

UNIVERSIDADE FEDERAL DE PELOTAS
Centro de Desenvolvimento Tecnológico
Programa de Pós-Graduação em Computação



Dissertação

**Algoritmo Auditável: O Papel dos Sistemas de Recomendação na Criação de
Filtros Bolha de Desinformação sobre as Urnas Eletrônicas no YouTube**

Yan Ballinhas Soares

Pelotas, 2023

Yan Ballinhas Soares

Algoritmo Auditável: O Papel dos Sistemas de Recomendação na Criação de Filtros Bolha de Desinformação sobre as Urnas Eletrônicas no YouTube

Dissertação apresentada ao Programa de Pós-Graduação em Computação do Centro de Desenvolvimento Tecnológico da Universidade Federal de Pelotas, como requisito parcial à obtenção do título de Mestre em Ciência da Computação.

Orientador: Prof. Dr. Luciano Volcan Agostini
Coorientadora: Prof^a. Dr^a. Raquel da Cunha Recuero

Pelotas, 2023

Universidade Federal de Pelotas / Sistema de Bibliotecas
Catalogação na Publicação

S676a Soares, Yan Ballinhas

Algoritmo auditável : o papel dos sistemas de recomendação na criação de filtros bolha de desinformação sobre as urnas eletrônicas no youtube / Yan Ballinhas Soares ; Luciano Volcan Agostini, orientador ; Raquel da Cunha Recuero, coorientador. — Pelotas, 2023.

210 f.

Dissertação (Mestrado) — Programa de Pós-Graduação em Computação, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, 2023.

1. Sistemas de recomendação. 2. Desinformação. 3. Urnas eletrônicas. 4. Youtube. I. Agostini, Luciano Volcan, orient. II. Recuero, Raquel da Cunha, coorient. III. Título.

CDD : 005

Dedico este trabalho a todos aqueles que deram a vida para que fosse possível o livre exercício democrático do sufrágio no Brasil.

Democracia para sempre!

AGRADECIMENTOS

Agradeço ao Claudio por me possibilitar escolher meu próprio caminho;
À Lucia, por me incentivar a buscar sempre mais;
Ao Theo, por me inspirar a ser uma pessoa cada vez melhor;
Ao Felipe, por ter sempre sido um Norte na minha vida;
Ao Gustavo, por nunca ter me deixado desistir;
E aos meus amigos, por tudo.

*O que sabemos é uma gota,
O que não sabemos é um oceano.*

—DITO ATRIBUÍDO POPULARMENTE A SIR ISAAC NEWTON

A heresia das heresias era o bom senso.

—GEORGE ORWELL, 1984

... Enganarmo-nos é o único privilégio humano frente a todos os outros organismos! Quem erra, chega à verdade! Sou ser humano precisamente porque erro. Ainda ninguém chegou a uma verdade qualquer sem antes se ter enganado catorze vezes, ou talvez cento e catorze, e isso é um mérito, neste sentido.

—DOSTOIÉVSKI, CRIME E CASTIGO

RESUMO

SOARES, Yan Ballinhas. **Algoritmo Auditável: O Papel dos Sistemas de Recomendação na Criação de Filtros Bolha de Desinformação sobre as Urnas Eletrônicas no YouTube**. Orientador: Luciano Volcan Agostini. Coorientadora: Raquel da Cunha Recuero. 2023. 210 f. Dissertação (Mestrado em Ciência da Computação) – Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, Pelotas, 2023.

Os Sistemas de Recomendação são uma ferramenta idealizada a partir da necessidade de navegação na massiva quantidade de dados disponíveis virtualmente. Sua atual implementação, utilizando de dados pessoais e subjetividades inerentes a cada usuário para o direcionamento de conteúdos, têm sido alvo de estudos e críticas em diversas frentes. Tendo tido comprovada influência em decisões populares como as eleições estadunidenses de 2016 e o referendo do Brexit, estes sistemas acabam por reforçar visões pré-estabelecidas de mundo, permitindo a criação de filtros bolha pelos quais os usuários consomem informações e moldam suas opiniões através de um viés de confirmação. Neste sentido, um usuário que consome conteúdos desinformativos acaba sendo direcionado a consumir ainda mais conteúdos desinformativos. Quando consideramos as Urnas Eletrônicas – um dos tópicos de maior fluxo informacional dos últimos anos e alvo de constantes campanhas desinformativas – a criação de filtros bolha de desinformação se torna ainda mais grave e uma ameaça à democracia. Levando em consideração que o YouTube é a segunda mídia social mais utilizada no Brasil, com um total de 142 milhões de usuários (quase 70% da população brasileira, segundo o Censo 2022), o presente trabalho buscou auditar de que maneira os Sistemas de Recomendação da plataforma atuam na criação de filtros bolha de desinformação, tanto para resultados de pesquisa como em vídeos recomendados, levando em consideração os diferentes perfis de consumo de conteúdos relacionados às Urnas Eletrônicas. Através da metodologia desenvolvida pelo autor, foi possível atestar a presença generalizada de desinformação em vídeos populares da plataforma acerca do tópico, além de um aumento geral de desinformação sendo recomendada em todos os perfis. Para o perfil de consumo desinformativo, foi possível comprovar a criação de um filtro bolha de desinformação, com o Sistema de Recomendação reduzindo consideravelmente os vídeos informativos recomendados e aumentando significativamente os vídeos desinformativos recomendados – chegando ao ponto de um a cada três vídeos recomendados apresentar potencial desinformativo.

Palavras-chave: Sistemas de Recomendação. Desinformação. Urnas Eletrônicas. YouTube.

ABSTRACT

SOARES, Yan Ballinhas. **Auditable algorithm: Recommender System's role in creating misinformation filter bubbles about electronic ballots on YouTube.**

Advisor: Luciano Volcan Agostini. 2023. 210 f. Dissertation (Masters in) – , Federal University of Pelotas, Pelotas, 2023.

Recommender Systems are a tool conceived out of the need of effective navigation through the vast amount of data available online. Its current implementation, abusing of user's personal data and inherent subjectivities for better recommendations has been the subject of studies and criticism on several fronts. With its proven influence on popular decision-making processes such as USA's 2016 Elections and the Brexit referendum, these systems end up reinforcing pre-established views of the world, allowing the creation of filter bubbles through which users consume information and shape their opinions with the influence of a confirmation bias. In this sense, users that consume misinformative contents end up being directioned to the consumption of even more misinformative contents. When we consider Brazilian Electronic Ballots – one of the topics with the greatest information fluxes over the last years and target of constant misinformative campaigns – the creation of misinformation filter bubbles poses an even greater threat to democracy. Considering that YouTube is the second most used social media in Brazil, with a total of 142 million users (almost 70% of the country's population, according to the 2022 Census), this study sought to audit how the platform's Recommender Systems contribute to the creation of misinformation filter bubbles, both in search results and recommended videos, simulating diverse content consumption profiles related to Electronic Ballots. Through the author's developed methodology, it was possible to confirm the widespread presence of misinformation in popular videos on the platform regarding this topic, as well as a general increase in recommended misinformation across all profiles. For the misinformation consumption profile, the creation of a misinformation filter bubble was proven, with the Recommender System significantly reducing the recommended informative videos and substantially increasing the recommended misinformative videos — to the point where one out of every three recommended videos presents potential misinformation.

Keywords: Recommender Systems. Misinformation. Electronic Ballots. YouTube.

LISTA DE FIGURAS

1	Exemplo de Sistema de Recomendação e seus componentes	28
2	Comparação entre uma matriz de (a) classificação e (b) filtragem colaborativa. Fonte: AGGARWAL (2016) – traduzido pelo autor.	33
3	Gráfico de cauda inversa representando a frequência itens avaliados em ordem decrescente. Fonte: SILVA (2020).	34
4	Arquitetura em alto nível de um SR baseado em conteúdo. Fonte: RICCI et al. (2010) – traduzido pelo autor.	36
5	3 "V"s do <i>Big Data</i> . Fonte: BU (2017) – traduzido pelo autor.	44
6	Representação de uma Rede Neural Profunda (RNP). Fonte: IBM (2020) – traduzido pelo autor.	45
7	Linha do Tempo do SR do YouTube. Fonte: VECTOR (2018) – traduzido pelo autor.	51
8	Arquitetura em alto nível do SR do YouTube. Fonte: (VECTOR, 2018) – traduzido pelo autor.	52
9	Sistema de Coleta de Dados do Usuário. Fonte: VECTOR (2018) – traduzido pelo autor.	53
10	Sistema de Geração de Candidato. Fonte: VECTOR (2018), – traduzido pelo autor.	54
11	Sistema de Ranqueamento. Fonte: VECTOR (2018) – traduzido pelo autor.	55
12	Tipos de Desordens Informativas. Fonte: WARDLE (2018) – traduzido pelo autor.	60
13	Evolução das Urnas Eletrônicas até 2018. Fonte: FERRÃO et al. (2019).	74
14	Etapas do voto impresso proposto pela PEC 135/2019. Fonte: SATIE (2021).	78
15	Invasão à Praça dos Três Poderes ocorrida em 8 de janeiro de 2023. Fonte: FERNANDES (2023).	81
16	Número de <i>posts</i> com <i>links</i> sobre desconfiança eleitoral no Facebook entre 2014 e 2020. Fonte: RUEDIGER; GRASSI (2020a).	94

17	Evolução no número de <i>posts</i> contendo desconfiança acerca do sistema eleitoral no mês de novembro de 2020 no Facebook. Fonte: RUEDIGER; GRASSI (2020b).	95
18	Termos mais frequentes nos <i>posts</i> analisados que incitam desconfiança acerca do sistema eleitoral no Facebook. Fonte: RUEDIGER; GRASSI (2020b).	96
19	Evolução das postagens com alegação de fraude nas urnas eletrônicas e defesa do voto impresso auditável no Facebook de novembro de 2020 à janeiro de 2022. Fonte: RUEDIGER et al. (2022).	98
20	Etapas da metodologia desenvolvida	100
21	Resultado da plataforma Google Trends na Pesquisa do YouTube Brasil acerca do Assunto “Urna Eletrônica” em 12 de setembro de 2022	102
22	Principais resultados obtidos no Google Trends sobre o assunto “Urna Eletrônica” em 12 de setembro de 2022	103
23	Página de pesquisa de vídeos do YouTube	105
24	Página de vídeo específico do YouTube	105
25	Funcionamento do <i>script</i> utilizado na Etapa 1	108
26	SERP-MS x Pares (Termo de Pesquisa x Algoritmo de Ordenação) dos resultados da Etapa 1	117
27	Distribuições percentuais de cada tipo de vídeo classificado para cada um dos quatro Algoritmos de Ordenação – Etapa 1 – Vídeos Recomendados	120
28	Distribuições percentuais de cada tipo de vídeo classificado para cada um dos cinco Termos de Pesquisa – Etapa 1 – Vídeos Recomendados	122
29	Comparação do SERP-MS médio entre o Perfil Desinformativo e a Etapa 1, utilizando como parâmetro a média entre os TPs para cada AO	126
30	Comparação do SERP-MS médio entre o Perfil Desinformativo e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP	127
31	Comparação da distribuição percentual média nos Vídeos Recomendados entre o Perfil Desinformativo e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP	128
32	Comparação do SERP-MS médio entre o Perfil Informativo e a Etapa 1, utilizando como parâmetro a média entre os TPs para cada AO	132
33	Comparação do SERP-MS médio entre o Perfil Informativo e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP	133
34	Comparação da distribuição percentual média nos Vídeos Recomendados entre o Perfil Informativo e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP	135
35	Comparação do SERP-MS médio entre o Perfil Misto e a Etapa 1, utilizando como parâmetro a média entre os TPs para cada AO	139

36	Comparação do SERP-MS médio entre o Perfil Misto e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP . . .	140
37	Comparação da distribuição percentual média nos Vídeos Recomendados entre o Perfil Misto e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP	141
38	Comparação do SERP-MS médio entre o Perfil Neutro e a Etapa 1, utilizando como parâmetro a média entre os TPs para cada AO . . .	146
39	Comparação do SERP-MS médio entre o Perfil Neutro e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP . . .	146
40	Comparação da distribuição percentual média nos Vídeos Recomendados entre o Perfil Neutro e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP	148
41	SERP-MS x Pares (Termo de Pesquisa x Algoritmo de Ordenação) dos resultados da Etapa 3 – Perfil Desinformativo	199
42	Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 4 Algoritmos de Ordenação – Etapa 3 – Vídeos Recomendados – Perfil Desinformativo	200
43	Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 5 Termos de Pesquisa – Etapa 3 – Vídeos Recomendados – Perfil Desinformativo	201
44	SERP-MS x Pares (Termo de Pesquisa x Algoritmo de Ordenação) dos resultados da Etapa 3 – Perfil Informativo	202
45	Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 4 Algoritmos de Ordenação – Etapa 3 – Vídeos Recomendados – Perfil Informativo	203
46	Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 5 Termos de Pesquisa – Etapa 3 – Vídeos Recomendados – Perfil Informativo	204
47	SERP-MS x Pares (Termo de Pesquisa x Algoritmo de Ordenação) dos resultados da Etapa 3 – Perfil Misto	205
48	Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 4 Algoritmos de Ordenação – Etapa 3 – Vídeos Recomendados – Perfil Misto	206
49	Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 5 Termos de Pesquisa – Etapa 3 – Vídeos Recomendados – Perfil Misto	207
50	SERP-MS x Pares (Termo de Pesquisa x Algoritmo de Ordenação) dos resultados da Etapa 3 – Perfil Neutro	208
51	Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 4 Algoritmos de Ordenação – Etapa 3 – Vídeos Recomendados – Perfil Neutro	209
52	Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 5 Termos de Pesquisa – Etapa 3 – Vídeos Recomendados – Perfil Neutro	210

LISTA DE TABELAS

1	Exemplo da matriz de avaliações $R = [r_{uj}]$ utilizando valores intervalares de 1 até 5	30
2	Matriz de avaliações de 4 usuários a 9 itens, a média das avaliações em comum do Usuário 1 com os demais e a métrica de similaridade do Usuário 1 com os demais utilizando Coeficiente de Pearson . . .	31
3	Matriz de avaliações de um usuário a 13 itens (filmes), sendo os itens 1-8 do <i>dataset</i> D_L (avaliados) e 9-13 do <i>dataset</i> D_U (sem avaliação), as palavras-chave referentes ao gênero do filme (1 se o filme for do gênero e 0 se não). As avaliações são positivas (1) ou negativas (-1).	38
4	Métrica de similaridade (Cosseno) entre todos os itens, os K-Vizinhos Mais Próximos (KNN) e a avaliação predita e normalizada para cada item.	38
5	Estudos analisados na revisão sistêmica de BANDY (2021) separados por domínio e categoria.	86
6	Trabalhos relacionados mais relevantes para o desenvolvimento da metodologia utilizada nesta Dissertação, sendo: TR1 – HUSSEIN; JUNEJA; MITRA (2020), TR2 – TOMLEIN et al. (2021), TR3 – PAPADAMOU et al. (2020), TR4 – ABUL-FOTTOUH; SONG; GRUZD (2020), TR5 – HUSSEIN; ELDARDIRY (2020), TR6 – JUNEJA; MITRA (2021)	86
7	Classificação dos vídeos em (HUSSEIN; JUNEJA; MITRA, 2020) .	88
8	Classificação dos vídeos em (PAPADAMOU et al., 2020)	90
9	Códigos retornados no processamento da tabela de <i>links</i> do TSE .	106
10	Anotação utilizada para classificar os vídeos extraídos	109
11	Número de vídeos da lista do TSE na classificação da Etapa 0 (setembro 2022 x junho 2023)	115
12	Percentual de vídeos classificados na Etapa 1 e SERP-MS dos pares Termo de Pesquisa x Algoritmo de Ordenação (Termo de Pesquisa – TP, Algoritmo de Ordenação – AO, Vídeos Informativos – VI, Vídeos Neutros – VN, Vídeos Desinformativos – VD, Pesquisa – P e Recomendados – R)	116

13	Percentual de vídeos classificados na Etapa 3 e SERP-MS dos pares Termo de Pesquisa x Algoritmo de Ordenação (Termo de Pesquisa – TP, Algoritmo de Ordenação – AO, Vídeos Informativos – VI, Vídeos Neutros – VN, Vídeos Desinformativos – VD, Pesquisa – P e Recomendados – R) – Perfil Desinformativo	125
14	Comparação entre as distribuições percentuais médias para cada TP entre a Etapa 1 e a Etapa 3 (Termo de Pesquisa – TP, Vídeos Neutros – VN, Vídeos Informativos – VI, Vídeos Desinformativos – VD) – Perfil Desinformativo (PD)	130
15	Percentual de vídeos classificados na Etapa 3 e SERP-MS dos pares Termo de Pesquisa x Algoritmo de Ordenação (Termo de Pesquisa – TP, Algoritmo de Ordenação – AO, Vídeos Informativos – VI, Vídeos Neutros – VN, Vídeos Desinformativos – VD, Pesquisa – P e Recomendados – R) – Perfil Informativo	131
16	Comparação entre as distribuições percentuais médias para cada TP entre a Etapa 1 e a Etapa 3 (Termo de Pesquisa – TP, Vídeos Neutros – VN, Vídeos Informativos – VI, Vídeos Desinformativos – VD) – Perfil Informativo (PI)	136
17	Percentual de vídeos classificados na Etapa 3 e SERP-MS dos pares Termo de Pesquisa x Algoritmo de Ordenação (Termo de Pesquisa – TP, Algoritmo de Ordenação – AO, Vídeos Informativos – VI, Vídeos Neutros – VN, Vídeos Desinformativos – VD, Pesquisa – P e Recomendados – R) – Perfil Misto	138
18	Comparação entre as distribuições percentuais médias para cada TP entre a Etapa 1 e a Etapa 3 (Termo de Pesquisa – TP, Vídeos Neutros – VN, Vídeos Informativos – VI, Vídeos Desinformativos – VD) – Perfil Misto (PM)	143
19	Percentual de vídeos classificados na Etapa 3 e SERP-MS dos pares Termo de Pesquisa x Algoritmo de Ordenação (Termo de Pesquisa – TP, Algoritmo de Ordenação – AO, Vídeos Informativos – VI, Vídeos Neutros – VN, Vídeos Desinformativos – VD, Pesquisa – P e Recomendados – R) – Perfil Neutro	144
20	Comparação entre as distribuições percentuais médias para cada TP entre a Etapa 1 e a Etapa 3 (Termo de Pesquisa – TP, Vídeos Neutros – VN, Vídeos Informativos – VI, Vídeos Desinformativos – VD) – Perfil Neutro	150
21	Avaliação do autor e dos avaliadores do Protocolo de Classificação	182
22	Resultados obtidos nas 10 iterações da Etapa 2	185
23	Resultados de <i>Search Engine Results Page – Misinformation Score</i> (SERP-MS)	195

LISTA DE ABREVIATURAS E SIGLAS

SR	Sistema de Recomendação
BD	<i>Big Data</i>
AM	Aprendizado de Máquina
IA	Inteligência Artificial
AP	Aprendizado Profundo
RNP	Redes Neurais Profundas
RGPD	Regulamento Geral de Proteção de Dados
MCI	Marco Civil da Internet
LGPD	Lei Geral de Proteção de Dados
UE	Urna Eletrônica
TRE	Tribunal Regional Eleitoral
PEC	Proposta de Emenda à Constituição
TI	Trabalho Individual
TR	Trabalho Relacionado
PC	Protocolo de Codificação
QP	Questão de Pesquisa
H	Hipótese
FC	Filtragem Colaborativa
BC	Baseado em Conteúdo
SERP-MS	<i>Search Engine Results Page – Misinformation Score</i>
TP	Termo de Pesquisa
AO	Algoritmo de Ordenação

SUMÁRIO

1	INTRODUÇÃO	17
1.1	Questões de Pesquisa e Hipóteses	24
1.2	Estrutura do Trabalho	25
2	SISTEMAS DE RECOMENDAÇÃO	26
2.1	Conceitos Principais sobre SRs	27
2.2	Modelos e Evolução	30
2.2.1	Modelos de Filtragem Colaborativa	30
2.2.2	Modelos Baseados em Conteúdo	35
2.2.3	Modelos de Filtragem Híbrida	39
2.3	Big Data, Aprendizado de Máquina e Redes Neurais Profundas	42
2.3.1	Privacidade em Sistemas de Recomendação e Regulamentação	47
2.4	O Sistema de Recomendação do YouTube	50
2.4.1	A Arquitetura do Sistema do YouTube	51
3	DESINFORMAÇÃO E URNAS ELETRÔNICAS	57
3.1	Classificando a Desinformação	59
3.2	Análise de Conteúdo	62
3.3	A Influência Social dos Sistemas de Recomendação e os Filtros Bolha de Desinformação	64
3.4	As Urnas Eletrônicas e o Voto no Brasil	68
3.4.1	A História do Sufrágio no Brasil	69
3.4.2	A Urna Eletrônica	74
3.4.3	Auditoria e Testes Públicos de Segurança (TPS)	76
3.4.4	O Voto Impresso e a Reunião com Embaixadores	77
3.4.5	A Lista de Links do TSE e a Atuação do Tribunal Durante as Eleições de 2022	79
3.5	Os Últimos Desdobramentos Acerca do Assunto - Pós-Eleições	80
3.5.1	O Projeto de Lei 2630/2020 - PL das Fake News	81
4	TRABALHOS RELACIONADOS	83
4.1	Sistemas de Recomendação e Filtros Bolha de Desinformação em Mídias Sociais	83
4.1.1	Trabalho Relacionado 1 - HUSSEIN; JUNEJA; MITRA (2020)	86
4.1.2	Trabalho Relacionado 2 - TOMLEIN et al. (2021)	87
4.1.3	Trabalho Relacionado 3 - PAPADAMOU et al. (2020)	89
4.1.4	Trabalho Relacionado 4 - ABUL-FOTTOUH; SONG; GRUZD (2020)	90
4.1.5	Trabalho Relacionado 5 - HUSSEIN; ELDARDIRY (2020)	91

4.1.6	Trabalho Relacionado 6 - JUNEJA; MITRA (2021)	92
4.2	Desinformação sobre Urnas Eletrônicas	93
4.2.1	Trabalho Relacionado 7 - RUEDIGER; GRASSI (2020a)	93
4.2.2	Trabalho Relacionado 8 - RUEDIGER; GRASSI (2020b)	94
4.2.3	Trabalho Relacionado 9 - RUEDIGER et al. (2022)	96
5	METODOLOGIA	99
5.1	Questões de Ética	100
5.2	Etapa 0 - Pré-processamento	101
5.3	Etapa 1 - Extração dos Vídeos Sem Perfil e Criação do Protocolo de Classificação	106
5.4	Etapa 2 - Teste de Homogeneidade das Recomendações	110
5.5	Etapa 3 - Extração dos Vídeos com os Quatro Perfis de Visualização	111
5.6	Métricas de Avaliação dos Resultados	112
5.6.1	<i>Search Engine Results Page – Misinformation Score</i> (SERP-MS)	112
5.6.2	Coeficiente de Determinação (R^2)	112
6	RESULTADOS	114
6.1	Etapa 0 - Pré-Processamento	115
6.2	Etapa 1 - Extração dos Vídeos Sem Perfil e Criação do Protocolo de Classificação	116
6.2.1	Resultados dos Vídeos de Pesquisa	117
6.2.2	Resultados dos Vídeos Recomendados	118
6.2.3	Comparação entre os Algoritmos de Ordenação	119
6.2.4	Comparação entre os Termos de Pesquisa	121
6.3	Etapa 3 - Extração dos Vídeos com os Quatro Perfis de Visualização	124
6.3.1	Perfil Desinformativo	125
6.3.2	Perfil Informativo	131
6.3.3	Perfil Misto	137
6.3.4	Perfil Neutro	145
7	CONCLUSÕES	152
	REFERÊNCIAS	161
APÊNDICE A	PROTOCOLO DE CLASSIFICAÇÃO	176
APÊNDICE B	TABELA DO TESTE DE VALIDAÇÃO DO PROTOCOLO DE CLASSIFICAÇÃO	182
APÊNDICE C	TABELA DE RESULTADOS DA ETAPA 2	185
APÊNDICE D	TABELA DE RESULTADOS DE <i>SEARCH ENGINE RESULTS PAGE – MISINFORMATION SCORE</i> (SERP-MS)	195
APÊNDICE E	GRÁFICOS DO SERP-MS E DISTRIBUIÇÃO PERCENTUAL DE VÍDEOS RECOMENDADOS POR ALGORITMO DE ORDENAÇÃO E TERMO DE PESQUISA PARA CADA PERFIL DA ETAPA 3	198

1 INTRODUÇÃO

Desenvolvidos a partir da necessidade de filtragem de informação na massiva quantidade de dados disponíveis *online*, os **Sistemas de Recomendação** (SRs) executam a tarefa de transformar dados sobre usuários e suas preferências em previsões de possíveis gostos e preferências futuras dentro de uma plataforma (LÜ et al., 2012).

Desde sua primeira implementação (um sistema utilizado em listas de *e-mails* para a recomendação de conteúdos que pudessem interessar ao usuário baseado nas preferências de outros usuários – o Tapestry, de 1992 (GOLDBERG et al., 1992)), os SRs passaram por uma série de iterações. A evolução destes Sistemas acompanhou o crescimento da internet, tendo os primeiros modelos sido baseados na avaliação de outros usuários (**filtragem colaborativa**) e na descrição de itens (**baseados em conteúdo**) e os atuais operando de maneira híbrida, através de métricas desenvolvidas com os modelos anteriores e com uma camada adicional **baseada em contexto** que utiliza dados de usuários extraídos e processados através de técnicas de Inteligência Artificial (IA) – **Big Data** (BD) e **Aprendizado de Máquina** (AM) (PORTUGAL; ALENCAR; COWAN, 2018).

Os SRs foram responsáveis por modificar a maneira que um usuário interage com plataformas e seus conteúdos na internet. Exemplos atuais do uso destes Sistemas vão desde a recomendação de novas músicas ou artistas em plataformas como Last.fm e Spotify até a recomendação de produtos a partir de compras pregressas em plataformas como a Amazon. Entretanto, apesar de seus inúmeros pontos positivos, como a possibilidade de um usuário conseguir encontrar o conteúdo (produto, vídeo, notícia, etc.) que melhor se adéque a seus interesses ou então auxiliando na possibilidade de serendipidade no encontro de novos conteúdos relevantes em meio a bilhões de possibilidades, os SRs são passíveis de influências não tão legítimas.

O problema começa em 2009, momento em que o Google começa seus esforços para personalizar os resultados de busca de usuários individualmente e se agrava em 2013, quando a empresa desenvolve um modelo de negócios extremamente lucrativo utilizando SRs: passa a implementar em sua plataforma de propagandas (o Google AdWords – atual Google Ads) algoritmos que utilizam dados pessoais do

usuário (como localização, tipo de dispositivo, etc.) de maneira automatizada para maior eficiência em seus anúncios veiculados, tornando-se precursor dos meios de **hiper-direcionamento de conteúdo** (WEAVER, 2022). Com o advento desse modelo, abre-se a possibilidade de direcionar determinados conteúdos a usuários específicos a partir de seus dados até então privados, de maneira a influenciá-los a clicar em determinado *link* ou consumir um determinado conteúdo de acordo com quem paga mais.

Não demorou muito até que outras **Big Techs** tenham corrido atrás da sua fatia de bolo nesse novo modelo de negócios que se provou muito bem sucedido. Atualmente, quatro das cinco maiores empresas de tecnologia do mundo utilizam alguma implementação deste padrão de conteúdo direcionado: Facebook (agora Meta), Microsoft, Google (Alphabet) e Amazon. A única que fica de fora dessa lista é a Apple (ZUBOFF, 2021).

O problema é que, para que tais recomendações sejam efetivas, são utilizadas técnicas de IA que sumariamente extraem informações e comportamentos de um usuário a fim de criar um **perfil** específico para ele (**profiling**), através do qual são geradas as recomendações. Em linhas gerais, o funcionamento atual dos SRs pode ser explicado da seguinte maneira:

Para direcionar, inicialmente as plataformas captam por rastreamento a matéria-prima – os dados das ações das pessoas – que forma suas características e seu comportamento nas redes, suas manchas de informação. Posteriormente, e às vezes concomitantemente, trocam esses dados colhidos entre elas (as plataformas). Com o pente fino da extração desses dados (emocionais, biométricos, georreferenciados etc.), a análise prediz padrões comportamentais e, como os rastros de dados provocam outras camadas de dados, é possível fazer correlação para influenciar as próximas ações do público-alvo, uma forma de modular essas pessoas escolhidas para inculir exatamente o que acham que devem inculir (PRADO, 2022).

Com a utilização desta técnica de extração e análise de dados, um novo fenômeno acaba surgindo: os **filtros bolha**, que, por definição, são o resultado individual de diferentes processos de busca de informação, percepção e seleção que culminam no usuário recebendo, dentro do universo de informações disponíveis, apenas uma parcela que se enquadra em suas noções pré-concebidas (GESCHKE; LORENZ; HOLTZ, 2019).

No livro que cunha o termo, de 2011, o autor Eli Pariser utiliza de um conceito que já era estudado no campo da Psicologia – o **viés da confirmação** – para explicar esse efeito:

A bolha dos filtros tende a amplificar drasticamente o viés da confir-

mação – de certa forma, é para isso que ela serve. O consumo de informações que se ajustam à nossas ideias sobre o mundo é fácil e prazeroso; o consumo de informações que nos desafiam a pensar de novas maneiras ou a questionar nossos conceitos é frustrante e difícil. É por isso que os defensores de uma determinada linha política tendem a não consumir a mídia produzida por outras linhas. Assim, um ambiente de informação baseado em indicadores de cliques favorecerá o conteúdo que corrobora nossas noções existentes sobre o mundo, em detrimento de informações que as questionam (PARISER, 2012).

Cabe destacar também que ocorre uma **assimetria de informação** quando consideramos os SRs da maneira que são implementados atualmente (em sua maioria, algoritmos proprietários completamente opacos, sem código aberto): o usuário não possui informações suficientes sobre quais parâmetros de seu perfil são efetivamente utilizados na geração de recomendações. Neste sentido, BOZDAG et al. (2014) destaca, inclusive, a dificuldade para a comprovação de estudos empíricos de filtros bolha em plataformas sociais, como o Facebook, por não permitir ou limitar a reprodução de resultados utilizando dados da plataforma.

Em estudos feitos sobre quais parâmetros em SRs utilizados em plataformas sociais influenciam mais diretamente na geração de recomendações – HUSSEIN; JUNEJA; MITRA (2020), PAPADAMOU et al. (2020), para citar alguns – o **histórico de uso** de um usuário em uma plataforma tem se mostrado, dentre outros, um dos mais influentes parâmetros na homogeneidade do conteúdo recomendado.

É possível constatar, portanto, o seguinte cenário: a maior parcela dos usuários de mídias sociais têm seus dados e comportamentos extraídos e processados a fim de gerar o direcionamento e recomendação de determinados conteúdos, sugeridos de acordo com parâmetros sobre os quais o usuário sequer tem conhecimento. O que ocorre, então, quando tais dados acabam por sugerir **conteúdos desinformativos** ou tendenciosos a uma parcela destes usuários?

Dois casos que tiveram maior destaque e foram estudados a fundo envolvendo a problemática dos filtros bolha gerados por SRs são as eleições norte-americanas de 2016 (escândalo da Cambridge Analytica), e o referendo popular do Brexit, tendo sido constatada a influência direta dos SRs e conteúdos direcionados nos resultados de ambos os pleitos (WHITTAKER et al., 2021).

Ainda mais grave que as influências políticas, os filtros bolha de SRs podem ser prejudiciais até mesmo à saúde: um estudo de 2021 aponta que o SR do YouTube no Brasil tende a recomendar mais vídeos com discursos pró-cloroquina (remédio com ineficiência comprovada contra a COVID-19) do que anti-cloroquina (SOARES et al., 2021).

Em ARENDT (2013) são elucidados alguns direitos humanos que impedem a con-

sagração de um estado totalitário. Dois deles são postos em cheque quando consideramos a atual implementação e uso dos SRs: o direito à informação e o direito à intimidade e à vida privada. O primeiro sendo cerceado com o bombardeio de conteúdos desinformativos recomendados a uma parcela significativa da sociedade e o segundo com a fronteira entre o ente público e privado sendo borrada com o amplo acesso e extração de dados dos mais variados de um indivíduo a fim de hiper-direcionar conteúdos.

Fica evidente então o poder destrutivo que tais Sistemas são capazes de infligir ao estado democrático de direito, reforçando visões de mundo que muitas vezes não condizem com a realidade e são retroalimentadas por conteúdos desinformativos que acabam direcionando o usuário a mais conteúdos desinformativos.

Quando pensamos esta problemática no contexto do cenário brasileiro, constatamos que nossa jovem democracia corre perigo. Já tendo ocupado em 2011 a 45ª posição no Índice de Democracia – relatório da Unidade de Inteligência do *The Economist*, que mensura e pontua diferentes aspectos democráticos em 167 países –, hoje o Brasil ocupa a 49ª posição (UNIT, 2021). Segundo o relatório anual do *Institute for Democracy and Electoral Assistance* (IDEA), o Brasil foi o país com o maior retrocesso democrático do mundo entre os anos de 2020 e 2021. Muito deste retrocesso ocorre em decorrência da má administração do governo federal no cenário pandêmico e também dos ataques do ex-presidente Jair Bolsonaro às Urnas Eletrônicas (UEs) e às Supremas Cortes do país (IDEA, 2021).

O Sistema Eleitoral consiste em um conjunto de processos e técnicas que permitem à população brasileira o exercício do voto e manifestação da sua soberania democrática, além de todas as etapas anteriores e posteriores às eleições, como protocolos de auditoria, apuração de pleitos e transmissão de votos. No centro deste complexo sistema – e alvo de constantes ataques – estão as **UEs**.

Empregada nas eleições brasileiras desde 1996, a UE tem história muito mais antiga na legislação nacional. A intenção inicial de estabelecer um sistema eletrônico vem do primeiro Código Eleitoral (Decreto 21.076, de 24/02/1932), sob o governo de Getúlio Vargas, que trazia a previsão de utilização de máquinas de votar, embora não levada a efeito, nos artigos 57 e 58 (TRE-SP, 2021).

As UEs são, portanto, a concretização de um longo anseio popular de eleições feitas de maneira limpa e acessível por todos. Modernizando o processo eleitoral frente ao resto do mundo, o Brasil conta com uma das apurações de votos mais rápidas e eficientes entre todas as nações democráticas desde o começo do século, disponibilizando os resultados do pleito poucas horas após seu encerramento (MACHADO FILHO, 2021).

Muitas foram as fraudes que marcaram as eleições brasileiras no período entre sua redemocratização e o ano 2000 (primeiro uso em todo território nacional das UEs),

tendo um dos episódios mais notáveis ocorrido nas eleições de 1994 no estado do Rio de Janeiro, quando ainda eram utilizadas cédulas para contagem de votos. No episódio em questão, pelo menos três pessoas responsáveis pela contagem dos votos foram presas quando se constatou que um a cada três boletins de urnas havia sido alterado durante a etapa de digitação dos votos para envio ao Tribunal Regional Eleitoral (TRE) do Rio de Janeiro. Por consequência, a contagem de votos e resultado final do pleito foram atrasados em semanas (FILHO, 2021).

Nos 22 anos de uso das UEs, não foram comprovadas fraudes nas eleições (ARREGUY, 2021). Não obstante, ainda que a maioria dos brasileiros mostre algum grau de confiança nas UEs – os números mais recentes apontam 42% com confiança elevada e 31% com confiança moderada, segundo pesquisa do instituto DataFolha (LINHARES, 2022) –, os 24% que alegam não confiar nas UEs não podem ser ignorados, colocando o modelo em evidência e sendo questionado por relevante parcela da população.

São recorrentes os debates em torno de modificações ao sistema em vigor das UEs, principalmente em anos eleitorais. Em particular, no ano de 2021, este debate foi exaltado em decorrência da tramitação da Proposta de Emenda a Constituição (PEC) 135/2019, que pautava sobre a implementação do voto impresso e foi rejeitada no plenário da Câmara dos Deputados no dia 10 de agosto daquele ano (CRISTALDO, 2021).

O debate se agrava, porém, quando a desinformação acerca das UEs passou a fazer parte da campanha do ex-presidente da República, Jair Bolsonaro, candidato às eleições em 2022: das 40 postagens no Facebook mais populares alegando fraude nas urnas entre novembro de 2020 e janeiro de 2022, 13 (32,5%) são da página do ex-presidente, sendo estas responsáveis por quase metade das interações nas postagens analisadas (RUEDIGER et al., 2022).

Nesse cenário, torna-se de fundamental importância o estudo dos fluxos de informação acerca das UEs em mídias sociais e o papel desempenhado pelos SRs. Em primeiro lugar, porque quase dois terços dos brasileiros (64%) utilizam redes sociais para consumo de notícias e, dentre todas as plataformas, o YouTube é a mais utilizada como fonte de informações – cerca de 43% de seus usuários a utilizam como meio para consumo de notícias (CARRO, 2022). Além disso, segundo LIMA (2022):

Desde que anunciou uma nova política para a redução da disseminação de informações enganosas sobre as eleições no Brasil, em março de 2022, o YouTube vem removendo vídeos com alegações infundadas sobre fraudes nas urnas eletrônicas. Estudos mostram, no entanto, que essa prática não é feita de forma contínua e transparente, além de apresentar alcance limitado em comparação com a quantidade de desinformação que circula na plataforma de vídeos. Em 18 de março,

dias antes do anúncio do YouTube, levantamento do professor Marcelo Alves, do Departamento de Comunicação da PUC-Rio, mostrou que ao menos 1.701 vídeos com questionamentos sobre a lisura das eleições no Brasil e com conteúdo golpista estavam disponíveis. Uma nova rodada da pesquisa constatou que até o final de maio apenas 75 deles haviam sido removidos pela empresa ou pelos próprios usuários que compartilharam o material. Ou seja, em cerca de dois meses e meio, as novas regras de moderação atingiram apenas 4,4% dos conteúdos analisados.

Apesar dos recentes esforços de plataformas como YouTube e Twitter para a redução de conteúdos desinformativos (principalmente após a explosão de informações em relação à pandemia de COVID-19 e também sobre a vacinação contra o Coronavírus) e a própria mudança de políticas do Facebook após o escândalo da *Cambridge Analytica* nas eleições estado-unidenses de 2016, as mídias sociais continuam tendo papel central na circulação de conteúdo desinformativo. Destaca-se, em particular, a baixa eficiência das políticas de moderação de conteúdos desinformativos compartilhados no Facebook: um estudo de Oxford analisando vídeos desinformativos sobre COVID-19 do YouTube compartilhados no Facebook chegou a conclusão que menos de 1% das publicações que continham *links* para tais vídeos com fatos distorcidos foram identificadas como desinformativas pelas políticas de moderação em vigor (KNU-UTILA et al., 2020).

Seguindo esta lógica e levando em conta que o YouTube atua como a principal fonte de informação de usuários de mídias sociais no Brasil (CARRO, 2022), é possível afirmar que uma significativa parcela da população utiliza de **vídeos** como fonte primária para consumo de notícias. Diversos trabalhos comprovam a criação de filtros bolha de conteúdos desinformativos em vídeos no YouTube em inglês – como HUSSEIN; JUNEJA; MITRA (2020) e PAPADAMOU et al. (2020) –, o que indica potencialmente que o mesmo deva ocorrer para o conteúdo em português sobre as UEs.

Levando em conta todos os pontos acima expostos e posicionando-se no encontro dos campos de saber da Computação e da Comunicação, o presente trabalho tem como objetivo compreender melhor os Sistemas de Recomendação e sua influência direta na criação de filtros bolha de conteúdos desinformativos. Propõe-se, portanto, o seguinte estudo: o papel dos SRs utilizados no YouTube na recomendação de conteúdo desinformativo sobre as UEs no Brasil, de acordo com o histórico de visualização de um determinado usuário, e a criação de filtros bolha desinformativos decorrentes de seu uso.

Para isso, foram desenvolvidos *scripts* em Python a fim de automatizar as buscas e extrações de vídeos de pesquisa e recomendados no YouTube, visto que a API desta plataforma conta com diversas limitações e não permite a simulação de perfis com

um histórico de uso específico, passo necessário na aplicação desta metodologia. A metodologia aplicada se dividiu em três etapas, além de uma pré-etapa (composta pela criação dos *scripts* e definição dos termos de pesquisa), conforme explicado abaixo.

Na **primeira etapa**, após ter sido feita a definição dos termos de pesquisa relevantes sobre o tema, foi extraído um *corpus* de vídeos, divididos entre vídeos obtidos diretamente na pesquisa e recomendações feitas a partir destes. Para cada um dos termos de pesquisa, são utilizados os quatro algoritmos de ordenação disponíveis para pesquisa no YouTube: **Relevância**, **Data de envio**, **Contagem de visualizações** e **Classificação**. A partir da análise destes vídeos, foram utilizados três classificadores propostos em um Protocolo de Classificação (posteriormente validado): **Sem potencial desinformativo** – para vídeos que abordam o tema e não contêm desinformação –, **Com potencial desinformativo** – para vídeos que abordam o tema e contêm desinformação – e **Outro** – para vídeos que não têm relação com o tema.

Na **segunda etapa**, o *corpus* de vídeos extraídos e já analisados da etapa anterior é utilizado para testar o número de vídeos necessários para gerar homogeneidade nas recomendações do YouTube.

Partindo destes resultados, na **terceira e última etapa**, foram criados quatro perfis com os mesmos parâmetros – idade, gênero, localização e informações já constatadas como relevantes aos SRs em outros trabalhos –, porém alimentados com históricos de visualização distintos, cada um consumindo vídeos apenas de uma das categorias supracitadas, além de um quarto perfil que tem seu histórico composto 50% por vídeos com potencial desinformativo e 50% de vídeos sem potencial desinformativo. Com estes perfis e históricos de uso construídos, foram gerados quatro novos *corpus* de vídeos utilizando o mesmo *script* da primeira etapa.

A partir dos novos dados gerados para cada perfil, foi realizada uma análise e categorização final, a fim de buscar atestar a influência dos SRs em vídeos do YouTube na criação de filtros bolha de conteúdos desinformativos, de acordo com o histórico de uso de um determinado usuário e tendo como objeto de pesquisa as UEs.

Uma análise inicial do discurso acerca das UEs em vídeos do Facebook foi realizada no Trabalho Individual (TI) do autor ao final do ano de 2021, a fim de verificar a possibilidade real de implementação de uma metodologia própria para o desenvolvimento desta Dissertação.

Neste primeiro estudo, foram analisados 638 vídeos utilizando os termos "urna eletrônica", "voto impresso" e "voto auditável". Destes, 11,91% dos vídeos foram classificados como sem potencial desinformativo, 16,14% como com potencial desinformativo e 71,94% como não relacionados (SOARES, 2021). É possível notar, portanto, o domínio do discurso desinformativo acerca do tema a frente do discurso que busca informar ou elucidar os usuários sobre as UEs em vídeos da plataforma.

A análise proposta torna-se imperativa para uma melhor compreensão do cenário brasileiro de 2022, levando em consideração a influência que vídeos do YouTube exercem como fonte de informação no país e os desdobramentos que o consumo de conteúdos desinformativos referentes às UEs poderiam levar naquele ano eleitoral, sendo alvo de constantes ataques do próprio ex-presidente da República e seus seguidores.

1.1 Questões de Pesquisa e Hipóteses

Este trabalho objetiva responder as seguintes **Questões de Pesquisa (QPs)** e propondo as seguintes **Hipóteses (Hs)**:

- **QP1** (Etapa 1) – Como é apresentado o conteúdo sobre as Urnas Eletrônicas em vídeos do YouTube? Este conteúdo tende a um potencial desinformativo?
 - **QP1.1** (Etapa 1 – Pesquisa) – Como este conteúdo é apresentado nos resultados de pesquisa? Algum termo de pesquisa ou algoritmo de ordenação leva à maior presença de desinformação nos vídeos de pesquisa?
 - **QP1.2** (Etapa 1 – Recomendações) – Como este conteúdo é apresentado nas Recomendações de vídeos? Algum termo de pesquisa ou algoritmo de ordenação leva à maior presença de desinformação nos vídeos recomendados?

***H1** - Tanto o termo de pesquisa utilizado para o encontro de resultados acerca de um tema como o algoritmo de classificação utilizado irão apresentar uma tendência maior ou menor à desinformação em seus resultados de pesquisa e também nas recomendações obtidas a partir dos vídeos de pesquisa.*

- **QP2** (Etapas 2 e 3) – Qual é a influência do histórico de visualização de um usuário no Sistema de Recomendação de vídeos do YouTube?
 - **QP2.1** (Etapa 2) – Qual é o número médio de vídeos que um usuário deve assistir acerca de um tema para que comece a haver homogeneidade em suas recomendações?
 - **QP2.2** (Etapa 3) – Os vídeos de pesquisa apresentam diferença de potencial desinformativo de acordo com o histórico de visualização de um usuário? Os termos de pesquisa e algoritmos de ordenação utilizados influenciam de alguma maneira particular nestes resultados?
 - **QP2.2** (Etapa 3) – Os vídeos recomendados apresentam diferença de potencial desinformativo de acordo com o histórico de visualização de um

usuário? Os termos de pesquisa e algoritmos de ordenação utilizados influenciam de alguma maneira particular nestes resultados?

H2 – Os resultados obtidos tanto na Pesquisa como nas Recomendações de vídeos de um usuário irão variar de acordo com o tipo de conteúdo presente em seu histórico de visualização.

- **QP3** (Etapa 3) – É possível atestar a criação de um filtro bolha de desinformação no Sistema de Recomendação do YouTube a partir do consumo de vídeos desinformativos por um usuário?

H3 – Usuários que consomem conteúdo desinformativo terão seus vídeos recomendados tendendo mais a conteúdos desinformativos quando comparados à usuários que consomem conteúdos informativos, mistos ou neutros em relação a um tema (filtro bolha).

1.2 Estrutura do Trabalho

O presente trabalho segue a seguinte estrutura: neste primeiro capítulo foi feita a Introdução do tema, oferecendo um *background* resumido da problemática em questão, introduzindo a metodologia a ser desenvolvida e apresentando as Questões de Pesquisa que o trabalho objetiva responder e Hipóteses levantadas pelo autor.

O Capítulo 2 será responsável por explicar os Sistemas de Recomendação a fundo, desde suas primeiras implementações até o estado atual, apontando os desdobramentos possíveis de seu uso. O próximo Capítulo (3) foca em explicar o fenômeno da Desinformação, além de suas categorias e formas de análise e elucida os pontos relativos ao tópico que será utilizado como estudo de caso (as Urnas Eletrônicas).

O Capítulo 4 apresenta os trabalhos relacionados com maior relevância para essa dissertação de mestrado. Os trabalhos relacionados são divididos em dois grupos: (i) trabalhos que atestam a criação de filtros bolha através dos sistemas de recomendação sistemas de recomendação e seu papel no consumo de conteúdos desinformativos em mídias sociais e plataformas de compra e (ii) trabalhos que analisam o discurso desinformativo acerca das urnas eletrônicas em plataformas como Twitter, Facebook e YouTube.

O Capítulo 5 detalha a Metodologia desenvolvida e aplicada pelo autor. Neste Capítulo serão explicadas todas as etapas desenvolvidas no estudo, desde a extração dos vídeos até a criação dos perfis distintos de utilização da plataforma e ferramentas para que sejam analisadas as categorias.

No Capítulo 6 serão expostos os resultados obtidos neste trabalho e serão debatidos os seus significados. Por fim, no Capítulo 7 são respondidas as Questões de Pesquisa, analisadas as Hipóteses propostas e apresentadas as conclusões do autor.

2 SISTEMAS DE RECOMENDAÇÃO

Com a popularização da internet ao final dos anos 90 e começo do novo século, a quantidade de conteúdo disponível no meio virtual e o número de usuários acessando esse conteúdo aumentou exponencialmente. No ano 2000, o total de domínios registrados na *web* era de cerca de 17 milhões. Hoje, este número passa de 1,9 bilhão (PROJECT, 2022). Já na quantidade de usuários, a *International Telecommunication Union* (ITU) estima que no ano 2000 apenas 7% da população mundial tinha acesso à internet, ao passo que em 2021 esse valor já era de 68% (ITU-D, 2021).

Para que essa massiva quantidade de informações fosse navegável de maneira eficiente, os Sistemas de Recomendação (SRs) foram desenvolvidos a fim de personalizar a navegação do usuário através da recomendação de publicações, notícias, vídeos, serviços e produtos que tenham a maior relevância possível. Os SRs passaram por uma série de iterações (Seção 2.2) até a consagração dos atuais modelos estado-da-arte, que operam de maneira híbrida, utilizando tecnologias de *Big Data*, mineração de dados e aprendizado de máquina.

Tendo sido inicialmente concebido como uma maneira de agregar utilidade ao usuário, relacionando usuários de um mesmo serviço com gostos similares (filtragem colaborativa) e, conseqüentemente, permitindo uma maior serendipidade nas suas descobertas individuais, os SRs acabaram se tornando parte essencial do modelo econômico da internet e por fim, acabaram subvertidos a um modelo de uso de dados neoliberal, ofuscando cada vez mais o limite entre os dados privados dos usuários e seu uso comercial (RICCI et al., 2010).

Pioneiro na recomendação de conteúdos hiper-direcionadas, o Google guiou a construção de SRs a partir do modelo do PageRank, que mais tarde foi utilizado como princípio básico do seu sistema de recomendação de propagandas, o Google Ads. Ao notar a possibilidade de rentabilidade a partir da captação destes dados de superávit comportamental dos usuários, as *Big Techs* rapidamente adotaram o mesmo modelo, presente atualmente em mídias sociais como Facebook, Instagram e Twitter, lojas virtuais como Amazon e serviços de *streaming* como Spotify e Netflix (AGGARWAL, 2016).

Além de toda a problemática envolvendo a privacidade dos usuários (Seção 2.3.1), os SRs também são vulneráveis a ataques e manipulações. Estudos demonstraram a vulnerabilidade de SRs à injeção de dados tendenciosos de usuários falsos (criados com o intuito de manipular o algoritmo) na recomendação final de usuários legítimos (FANG et al., 2018), além dos conhecidos casos de manipulação social através de SRs em eventos como as eleições estadunidenses de 2016 (Cambridge Analytica), a votação do Brexit e as eleições brasileiras de 2018. Estes pontos serão detalhados na Seção 3.3.

A fim de elucidar os princípios básicos que norteiam o funcionamento dos SRs, suas diversas implementações e evolução através do tempo, este Capítulo irá utilizar como principal base conceitual os livros AGGARWAL (2016) e RICCI et al. (2010).

2.1 Conceitos Principais sobre SRs

Segundo AGGARWAL (2016), o princípio fundamental dos SRs é que dependências significativas existem em atividades centralizadas entre **usuários** e **itens**. Essas dependências podem ser aprendidas através de uma orientação a dados de uma **matriz de avaliações** e o modelo resultante é utilizado para fazer **previsões a outros usuários**. Levando em conta este conceito básico, a Figura 1 ilustra um exemplo de SR com avaliações feitas em produtos de uma loja de eletrônicos.

Cabe salientar que os **itens** aqui especificados passíveis de recomendação por um SR podem ser de inúmeros tipos, não sendo restritos apenas a produtos, e vão variar de acordo com a plataforma em questão sendo analisada. Alguns exemplos: publicações (Facebook), notícias (Google News, G1), vídeos (YouTube, Vimeo), músicas (Spotify, Apple Music), filmes (Netflix, Amazon Prime), aplicativos (App Store, Play Store), entre outros.

Os quatro pilares que devem nortear as recomendações em um SR são:

- Relevância - Principal objetivo de um SR. Um SR que não faça a recomendação de conteúdos que sejam relevantes ao usuário não cumpre sua utilidade;
- Novidade - Um SR deve prezar sempre pela recomendação de itens que o usuário não tenha visto no passado;
- Serendipidade - A serendipidade, por definição, é a sensação de descoberta “casual” de algo novo, que não necessariamente está dentro da bolha de gostos do usuário. Em um SR, isto acontece na forma de recomendações que fujam do padrão de consumo/avaliação do usuário, mas que podem ser de interesse deste;
- Variedade - É importante haver um grau de variedade entre os itens recomendados a um usuário. Caso contrário, existe a possibilidade de uma recomendação

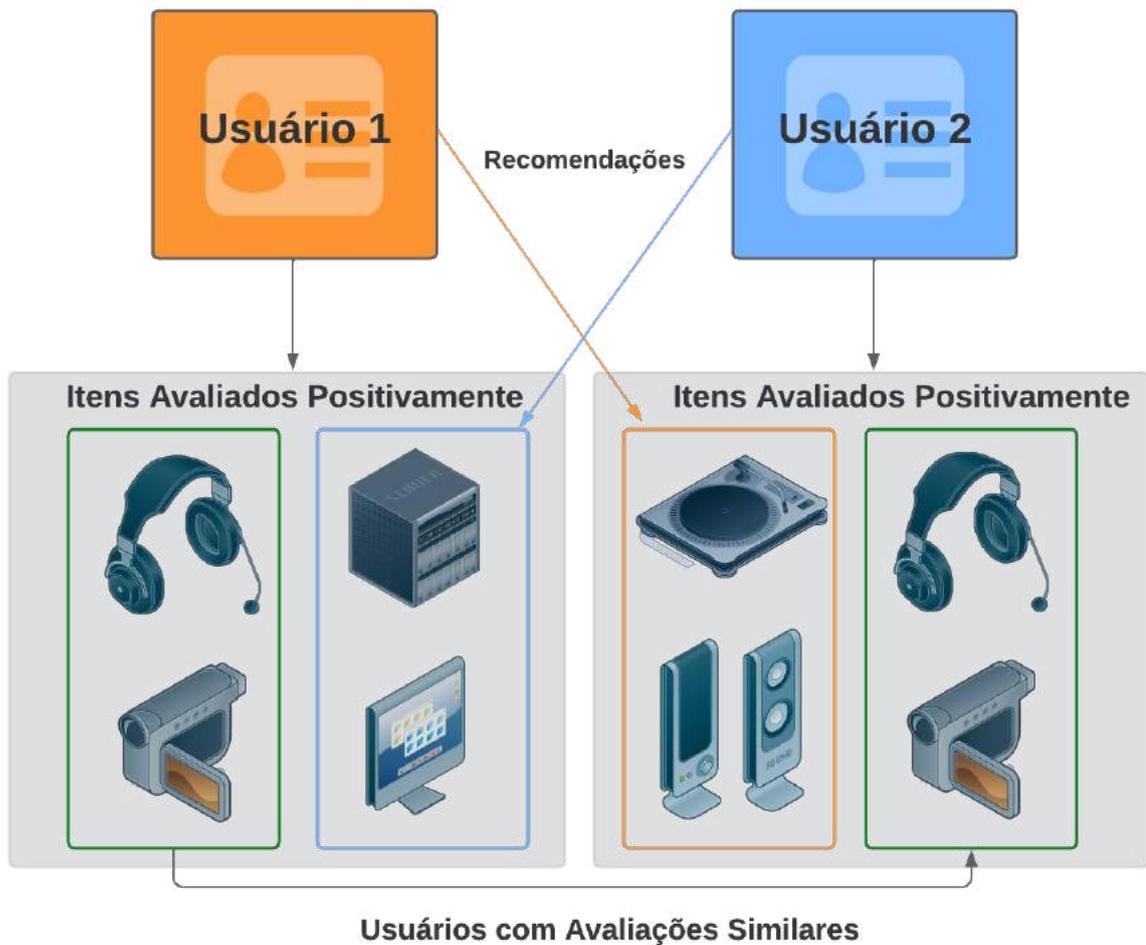


Figura 1 – Exemplo de Sistema de Recomendação e seus componentes

com impacto negativo ao usuário se repetir com as demais recomendações.

Uma plataforma que incorpora um SR pode considerar dois tipos de dados primários para o processamento de suas recomendações: avaliações **explícitas**, nas quais o usuário gera um determinado valor para um item de dentro da plataforma (por exemplo: a avaliação de itens comprados na Amazon em até cinco estrelas ou os botões de “gostei” e “não gostei” do YouTube) ou métricas **implícitas**, nas quais o usuário não determina um valor a um item, mas este é estipulado pela própria plataforma (por exemplo: o histórico de vídeos assistidos do YouTube ou itens comprados na Amazon, porém não avaliados). Além disso, os SRs atuais também levam em consideração uma série de **metadados** inerentes a cada usuário (geolocalização, idade, comportamento, etc.) para a geração de recomendações – estes conceitos serão aprofundados na Seção 5.

As **avaliações explícitas** de um usuário podem ser categorizadas em quatro tipos:

- Valores intervalares - Valor do tipo numérico discreto, estipulado dentro de um

conjunto de valores. Por exemplo: o sistema de avaliação de cinco estrelas implementado em plataformas como a Amazon. Esse tipo de valor numérico é o mais usual em SRs, pois consegue gerar um determinado peso (positivo ou negativo em relação ao item) à heurística implementada no algoritmo que processa as recomendações. Um caso particular desse tipo de valor é a atual implementação das reações do Facebook, onde o usuário pode escolher dentre sete tipos de reação a uma determinada publicação - podendo expressar, além do gosto, também tristeza ou raiva, por exemplo;

- Valores contínuos - Valor do tipo numérico real, geralmente implementado em forma de “barra deslizante”. Esse tipo de valor em avaliações é pouco usual pois dificulta o processo de definição de um peso para os infinitos valores possíveis;
- Valores binários - Possibilita que o usuário declare o gosto ou desgosto por determinado item, geralmente através de um botão de *like* e *dislike*, como no YouTube, por exemplo.
- Valores unários - Possibilitam que o usuário declare apenas se gostou de determinado item, não havendo a possibilidade de declarar explicitamente o fato de não ter gostado deste. Esse sistema é implementado no Instagram.

Além disso, os SRs podem ser implementados seguindo dois tipos de modelagens primárias:

- Modelo de predição - Baseado nas avaliações explícitas ou implícitas de determinado usuário e/ou de usuários similares, o SR pode fazer predição da avaliação que seria dada a um determinado item;
- Modelo de ranqueamento - Faz recomendações dos **top-k** (sendo **k** um número estipulado pela heurística implementada) itens para um usuário, não sendo necessária a predição exata da avaliação de um determinado item, mas sim a determinação dos itens mais relevantes para o usuário.

Estes conceitos são implementados de diferentes maneiras, de acordo com a plataforma em questão. Por exemplo: as recomendações de vídeos no YouTube são implementadas utilizando um modelo de ranqueamento (são apresentados os vídeos mais relevantes como recomendação a partir de um vídeo específico ou então a partir de um usuário específico) e valores binários para a avaliação de itens. Já na Amazon, os itens são avaliados com valores intervalares (uma a cinco estrelas) e as recomendações na página de um produto são apresentadas com modelos de predição (“compre com”) e modelos de ranqueamento (“produtos relacionados a este item” e “clientes que visualizaram este item também visualizaram”).

Tabela 1 – Exemplo da matriz de avaliações $R = [r_{uj}]$ utilizando valores intervalares de 1 até 5

Itens \Rightarrow	1	2	3	4	...	n
Usuários \Downarrow						
1	1		5	5	...	$[r_{1n}]$
2		2		4	...	$[r_{2n}]$
3	5	3	4	5	...	$[r_{3n}]$
4	5		3	1	...	$[r_{4n}]$
...	$[r_{...n}]$
m	$[r_{m1}]$	$[r_{m2}]$	$[r_{m3}]$	$[r_{m4}]$	$[r_{m...}]$	$[r_{mn}]$

2.2 Modelos e Evolução

Os SRs têm sua origem no problema fundamental de filtragem de informações em sistemas computacionais, que já era abordado em sistemas *offline* e é mencionado pela primeira vez ainda em 1982 por Peter Denning no artigo DENNING (1982), no qual o autor salientava a importância do recebimento das informações e o controle do processo de recuperação e filtragem destas para que a informação correta alcançasse o usuário que correto (CAZELLA; NUNES; REATEGUI, 2010).

Com o avanço da tecnologia e da pesquisa em SRs, diversos modelos e técnicas foram desenvolvidos. Considerada a primeira implementação de um SR, o *Tapestry* (GOLDBERG et al., 1992) foi um sistema de recebimento de mensagens que permitia que o usuário filtrasse em quais listas gostaria de se inscrever utilizando a avaliação feita por outros usuários – a primeira implementação de um sistema de **filtragem colaborativa**.

Dentre os modelos desenvolvidos, cabe o destaque de três, que serão detalhados nas Seções a seguir: modelos de **filtragem colaborativa** (2.2.1), modelos **baseados em conteúdo** (2.2.2) e modelos **híbridos** (2.2.3).

2.2.1 Modelos de Filtragem Colaborativa

Os SRs de filtragem colaborativa (FC) são baseados na premissa de que usuários similares exibem padrões comportamentais similares em relação à avaliações (FC baseada em usuários) e itens similares recebem avaliações parecidas (FC baseada em itens) (AGGARWAL, 2016).

Para a predição de valores de avaliação de um usuário em itens ainda não avaliados (e, conseqüentemente, a recomendação dos **top-k** itens), este modelo de SR irá utilizar uma matriz de avaliações R incompleta e altamente esparsa que relaciona m usuários (linhas) com n itens (colunas), sendo $[r_{uj}]$ a denotação de uma célula de avaliação qualquer de um usuário u em um item j . A Tabela 1 ilustra um exemplo da matriz de avaliações $R = [r_{uj}]$ utilizando valores intervalares de um até cinco.

Existindo esta matriz de avaliações, a técnica de filtragem colaborativa ocorre em

Tabela 2 – Matriz de avaliações de 4 usuários a 9 itens, a média das avaliações em comum do Usuário 1 com os demais e a métrica de similaridade do Usuário 1 com os demais utilizando Coeficiente de Pearson

Itens (<i>i</i>) ⇒ Usuários (<i>u</i>) ↓	1	2	3	4	5	6	7	8	9	Média de Avaliações	Métrica de Similaridade
1	1		5	5	3	4		2	4	3,428	1
2	4	2		4	2	5	1	3	5	3,833	0,368
3	5	3	4	5	3			1	3	3,5	0,2423
4	5		3	1	2	3	3	4		3	-0,7794

três passos (CAZELLA; NUNES; REATEGUI, 2010):

1. É feito o cálculo do peso de cada usuário em relação à similaridade ao usuário alvo (métrica de similaridade);
2. É selecionado um subconjunto de usuários com maiores similaridades (vizinhos) para considerar na predição;
3. São normalizadas as avaliações e computadas as predições ponderando as avaliações dos vizinhos com seus pesos.

A fim de exemplificar como é feita a predição de avaliações em um SR utilizando FC, os passos acima listados serão seguidos utilizando os valores presentes na Tabela 2 e calculando a similaridade do usuário um com os demais usuários.

Para o cálculo da métrica de similaridade (Passo 1), é necessário escolher um coeficiente de similaridade – sendo Cosseno e Pearson dois dos mais populares para este fim (SHEUGH; ALIZADEH, 2015). Aqui, será utilizado o coeficiente de Pearson, denotado por sim_{ab} (Fórmula 1 – adaptada de CAZELLA; NUNES; REATEGUI (2010) e SHEUGH; ALIZADEH (2015)):

$$sim_{ab} = \frac{\sum_{i \in I_{ab}} (r_{ai} - \bar{r}_a)(r_{bi} - \bar{r}_b)}{\sqrt{\sum_{i \in I_{ab}} (r_{ai} - \bar{r}_a)^2 \sum_{i \in I_{ab}} (r_{bi} - \bar{r}_b)^2}} \quad (1)$$

Essa fórmula deve ser aplicada iterativamente (somatório) considerando cada item *i* que tenha sido avaliado tanto pelo usuário *a* como pelo usuário *b* (conjunto I_{ab} de itens avaliados em comum). Considera-se aqui r_{ai} a avaliação do usuário *a* ao item *i*, r_{bi} a avaliação do usuário *b* ao item *i* e \bar{r}_a e \bar{r}_b a média aritmética simples de todas as avaliações feitas em itens em comum pelos usuários *a* e *b*, respectivamente. O cálculo da métrica de similaridade do usuário um com o usuário três ($\bar{r}_a = 3,33$ e $\bar{r}_b = 3,5$), por exemplo, se daria da seguinte maneira:

$$sim_{13} = \frac{(-2, 33) * 1, 5 + 1, 67 * 0, 5 + 1, 67 * 1, 5 + (-0, 33) * (-0, 5) + (-1, 33) * (-2, 5) + 0, 66 * (-0, 5)}{\sqrt{((-2, 33)^2 + 1, 67^2 + 1, 67^2 + (-0, 33)^2 + (-1, 33)^2 + 0, 66^2) * (1, 5^2 + 0, 5^2 + 1, 5^2 + (-0, 5)^2 + (-2, 5)^2 + (-0, 5)^2)}}$$

$$sim_{13} = 0, 2423$$

O segundo passo consiste na definição do subconjunto de vizinhos que mais se assemelham (têm maior métrica de similaridade) com o usuário alvo. Para isso, é definido um limiar mínimo de similaridade que deve ser atingido. Em CAZELLA; NUNES; REATEGUI (2010) este limiar é de 0,3. Para fins de aplicação neste exemplo, o limiar escolhido é de 0,2. Isto limita, portanto, o subconjunto de vizinhos do usuário um aos usuários dois e três.

Para que seja feita a predição do possível valor de avaliação que um usuário daria a um item não avaliado, baseando-se no subconjunto de seus vizinhos mais próximos (definidos no Passo 2), no Passo 3 é feito o cálculo da média ponderada das avaliações realizadas pelos vizinhos no item-alvo, seguindo a Fórmula 2 – adaptada de CAZELLA; NUNES; REATEGUI (2010) e ZENG et al. (2004).

$$p_{ai} = \bar{r}_a + \frac{\sum_{b=1}^n (r_{bi} - \bar{r}_b) * sim_{ab}}{\sum_{b=1}^n |sim_{ab}|} \quad (2)$$

A fórmula consiste no cálculo da média ponderada do desvio da média das avaliações dos vizinhos (ZENG et al., 2004). Neste caso, p_{ai} será o valor predito de avaliação de um usuário a (previamente escolhido na primeira etapa para os cálculos de similaridade e, por conseguinte, segunda etapa para a definição do subgrupo de seus vizinhos) a um item-alvo i . O somatório deverá ocorrer iterativamente para todos os usuários b pertencentes ao grupo de n vizinhos. Nesta fórmula, \bar{r}_a será a média dos valores avaliados pelo usuário em comum com todos os seus n vizinhos, \bar{r}_b será a média dos valores avaliados pelo usuário b da corrente iteração em comum com o usuário a , r_{bi} será a avaliação dada por este usuário ao item-alvo i e sim_{ab} é a métrica de similaridade calculada anteriormente do usuário a com o usuário b .

Para fins de exemplo, será descrito abaixo o cálculo da predição de avaliação do item 2 pelo usuário 1, considerando os usuários 2 ($b = 1$) e 3 ($b = 2$) como seu subgrupo de vizinhos ($n = 2$, $\bar{r}_a = 3$, $\bar{r}_{b=1} = 3, 833$, $\bar{r}_{b=2} = 3, 5$, $sim_{11} = 0, 368$, $sim_{12} = 0, 2423$):

$$p_{12} = 3 + \frac{((2 - 3,833) * 0,368) + ((3 - 3,5) * 0,2423)}{0,368 + 0,2423}$$

$$p_{12} = 1,6972$$

Este cálculo preditivo é aplicado em diversos itens-alvo de um usuário que não tenham sido avaliados previamente, levando em conta os itens que foram avaliados em comum por seus vizinhos. Com isto, é gerada uma lista dos **top-k itens** que serão recomendados para aquele usuário em particular. Considerando o padrão de avaliação de até cinco estrelas para um item empregado aqui, o item dois não entraria nos *top-k* itens indicados pelo SR ao usuário um.

O algoritmo empregado nesse modelo de SR pode ser interpretado como uma generalização do problema de Inteligência Artificial (IA) de classificação. Enquanto nas matrizes dos algoritmos de classificação as linhas de treinamento são separadas das linhas de teste e uma única coluna é considerada a coluna da variável dependente (a ser predito), nos SRs de FC os valores que serão preditos podem estar em qualquer célula da matriz de avaliações, não havendo distinção entre linhas de treinamento e de teste, conforme ilustrado na Figura 2 – considerando as células demarcadas como as contendo os valores faltantes que serão preditos.

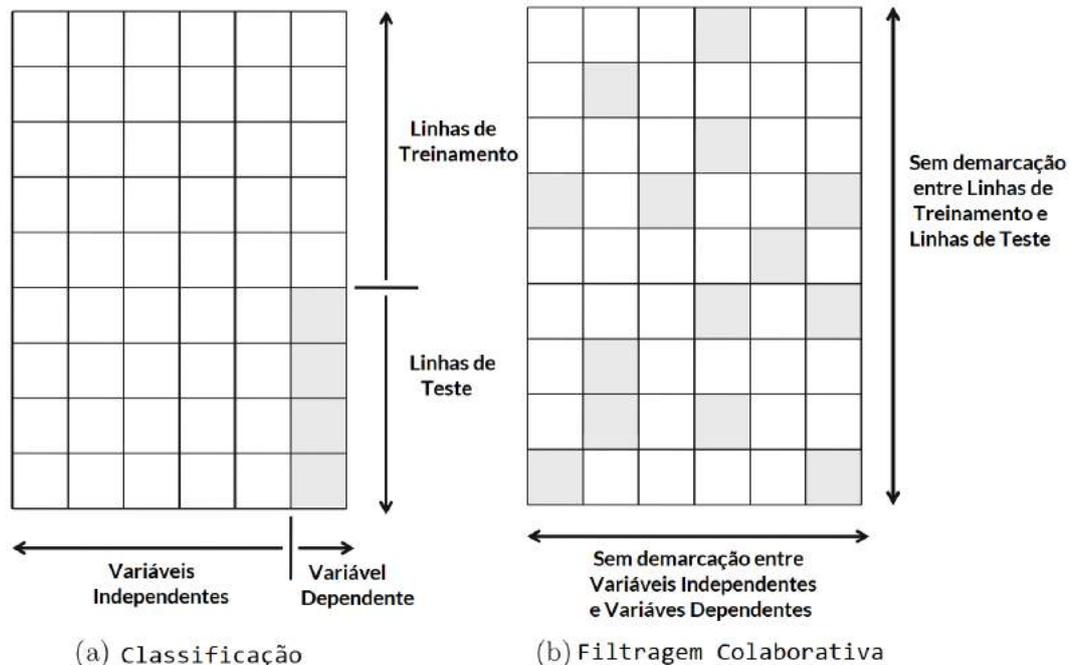


Figura 2 – Comparação entre uma matriz de (a) classificação e (b) filtragem colaborativa.

Fonte: AGGARWAL (2016) – traduzido pelo autor.

Este modelo de SR permite uma alta **serendipidade**, visto que usuários com gostos similares podem apresentar um item que fure a bolha de gostos de um usuário em particular. Porém, ele apresenta dois pontos negativos:

- Problema de *cold start* para novos usuários - Este modelo de SR não irá funcionar bem se houver poucas avaliações do usuário, visto que não será possível definir com precisão um grupo de vizinhos com gostos similares e, assim, calcular as recomendações corretas;
- Tendência em avaliações - Um modelo que se baseia em avaliações explícitas (como é o caso aqui) pode acabar gerando recomendações repetitivas e sem variedade, dada a tendência que esse tipo de avaliação apresenta: alguns itens mais populares são muito avaliados e, por consequência, têm uma tendência maior de ser recomendados e outros são menos avaliados e tem poucas chances de ser escolhidos pelo SR. Essa tendência é ilustrada na forma de **gráfico de cauda inversa**, conforme ilustrado na Figura 3.

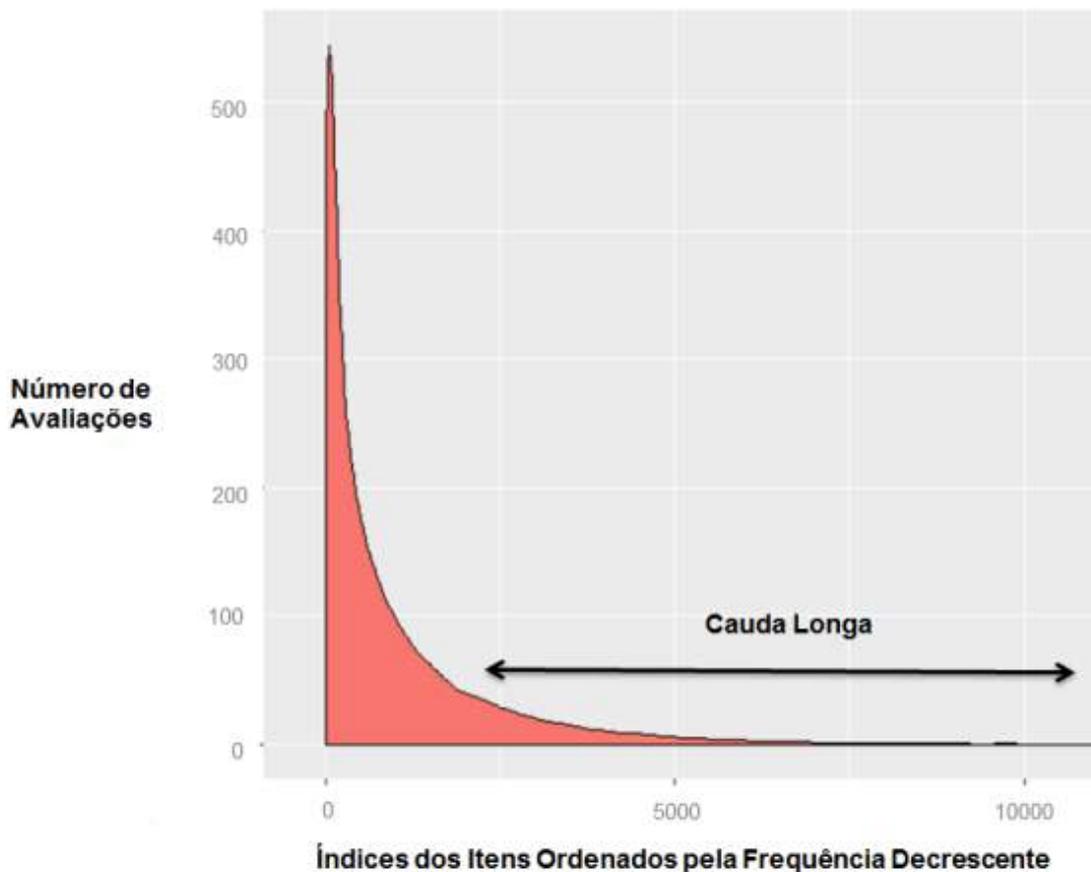


Figura 3 – Gráfico de cauda inversa representando a frequência itens avaliados em ordem decrescente.
Fonte: SILVA (2020).

2.2.2 Modelos Baseados em Conteúdo

Os modelos de SR baseados em conteúdo (BC) se diferenciam dos modelos de FC por pautar suas recomendações nos atributos dos **itens** avaliados (explícita ou implicitamente) e não na relação entre os **usuários**. Aqui são utilizadas as avaliações feitas pelo próprio usuário, o que é particularmente útil quando um item é novo e conta com poucas avaliações (AGGARWAL, 2016).

Ao combinar dados das descrições de itens como *tags*, gênero (para o caso de filmes/livros) e demais informações relevantes com o perfil de um usuário (manifestado através de seu histórico de uso em uma determinada plataforma com um modelo de aprendizagem baseado em itens visualizados e/ou avaliados), é possível prever itens similares que se adéquem ao gosto do usuário.

O perfil de um usuário neste tipo de SR pode ser representado, por exemplo, por um modelo de classificação/regressão baseado em um conjunto de documentos de treinamento que relacionam as descrições de itens (palavras-chave, gênero e outras informações) com *labels* compostas pelas avaliações do usuário (avaliações numéricas, *like*, *dislike*, etc.)

Este tipo de SR é particularmente útil para a recomendação de novos itens (com poucas ou nenhuma avaliação) a um usuário, visto que as informações relevantes de um item podem ser extraídas no momento em que é inserido na plataforma, possibilitando a recomendação a qualquer usuário com um SR BC. Isso elimina o problema de *cold start* para novos itens, mas não para novos usuários.

A geração das recomendações aqui é feita em três módulos, sendo os dois primeiros compostos por etapas que ocorrem *offline* (aprendizado e criação de modelo) e o último por etapas *online* (recomendação em tempo real) (RICCI et al., 2010). Uma arquitetura em alto nível é apresentada na Figura 4 e detalhada a seguir:

1. Analisador de conteúdo (pré-processamento e extração de *features*) - Transforma as propriedades não-estruturadas de um item (geralmente explicitadas de maneira textual) em representações vetoriais ou multi-dimensionais estruturadas utilizando palavras-chave. Aqui podem ser dados pesos diferentes para cada atributo do item, de acordo com a sua importância para recomendação de outros itens similares. Por exemplo: o gênero de um livro/filme tem mais importância que o autor/diretor na geração das palavras-chave;
2. Aprendiz de perfil (aprendizado de perfis de usuários baseado em conteúdo) - É feita a criação de um perfil de usuário com base nas representações já estruturadas dos itens, combinadas com as avaliações explícitas ou com comportamento implícito (histórico de uso) do usuário em uma plataforma. Estes dados obtidos de um perfil de usuário em conjunto com os atributos dos itens

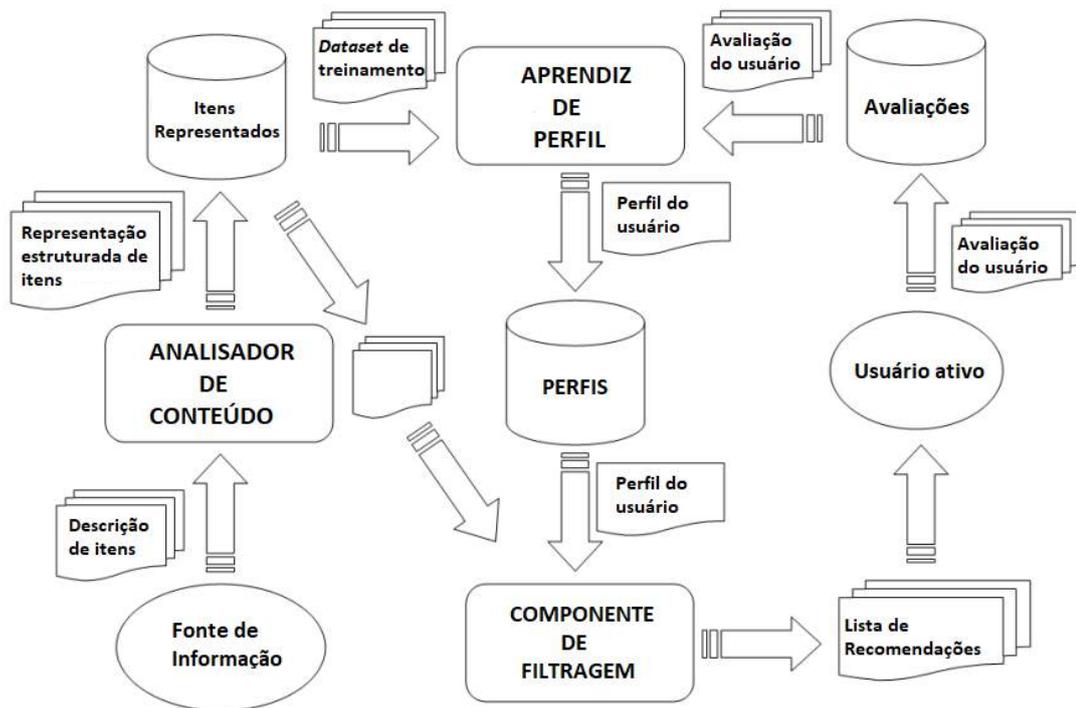


Figura 4 – Arquitetura em alto nível de um SR baseado em conteúdo.
 Fonte: RICCI et al. (2010) – traduzido pelo autor.

avaliados/utilizados são utilizadas como dados de treinamento para um modelo de aprendizado de máquina, de maneira similar a modelos de classificação/regressão;

3. Componente de filtragem (filtragem e recomendação) - O modelo de aprendizagem gerado na etapa anterior é utilizado em tempo real (*online*) para a geração de recomendações. Isso pode se dar na página inicial de uma plataforma ou combinando as informações da página de um item específico com o modelo criado a partir do perfil do usuário.

Visto que muitas vezes as descrições textuais contidas em itens são extensas (sinopses de livros/filmes, parágrafos inteiros em notícias), o primeiro módulo (analisador de conteúdo) utiliza de uma série de heurísticas para a ponderação de palavras-chave de maior relevância. Alguns exemplos de heurísticas utilizadas nesta etapa são o índice GINI, entropia e desvio padrão - similares aos utilizados em algoritmos de aprendizado de máquina para problemas de classificação (BATISTA, 2018).

Após a geração das representações estruturadas das palavras-chave de maior relevância, no módulo aprendiz de perfil é feita a combinação destas com o *feedback* explícito ou implícito do usuário (treinamento do modelo de previsão) e, após, é feita a filtragem e recomendação de itens. Nesse processo, são gerados dois *datasets* distintos para cada perfil de usuário: D_L e D_U . O primeiro é composto pelos itens avaliados

pelo usuário (*labeled*) e o segundo pelos demais itens não avaliados (*unlabeled*) e que poderão ser recomendados ao usuário em questão.

A partir dos itens avaliados (D_L), é gerado um modelo de treinamento (classificador) que servirá de base para as recomendações. Nesta etapa, podem ser implementados diferentes métodos. Para fins de exemplificação, será utilizado aqui o método mais simples para este fim: o **K-Vizinhos Mais Próximos** (KNN – *K-Nearest Neighbors*), que utiliza uma **métrica de similaridade** (*sim*), assim como os SRs de FC (2.2.1).

Como os SRs BC utilizam de conteúdos textuais para a geração de suas recomendações, a métrica de similaridade mais comum a ser utilizada é a de **Cosseno**, conforme a Equação 3. O uso do Cosseno neste caso é particularmente relevante por se tratar de uma métrica de similaridade adaptável a vetores de tamanho variável (FRIEDMAN et al., 2007).

$$sim(\bar{X}, \bar{Y}) = \frac{\sum_{i=1}^d x_i y_i}{\sqrt{\sum_{i=1}^d x_i^2} \sqrt{\sum_{i=1}^d y_i^2}} \quad (3)$$

Na equação acima, \bar{X} e \bar{Y} são, respectivamente, itens de D_L e D_U (já tendo passado pela extração de palavras-chave e, portanto, estruturados); d representa o número total de palavras-chave das representações estruturadas destes itens e i a iteração atual dentre estas palavras e x_i e y_i a frequência normalizada das palavras-chave nos itens \bar{X} e \bar{Y} , respectivamente. Quanto maior o valor do cosseno entre os dois itens, maior a sua similaridade.

Para cada item em D_U é feito o cálculo de *sim* com os itens de D_L . A partir disso, são estipulados os K-Vizinhos Mais Próximos, determinados a partir de um limiar mínimo pré-estabelecido. A **média das avaliações destes itens** é estipulada, portanto, como a avaliação do item em D_U . Isso é feito iterativamente e, por fim, é recomendado o item com a maior avaliação predita (no caso de modelos de predição) ou então são recomendados os itens com as maiores avaliações preditas (no caso de modelos de ranqueamento).

Como exemplo, temos a Tabela 3 ilustrando os dois *datasets* em um SR BC para recomendação de filmes, na qual os itens 1-8 representam o *dataset* de filmes avaliados (D_L) e os itens 9-13 representam os filmes ainda não avaliados (D_U) e, portanto, alvo das predições do SR.

Considera-se para este exemplo que a etapa de extração de palavras-chave resultou no **gênero do filme** como sendo a *feature* mais relevante para as recomendações. Os filmes são normalizados, portanto, como sendo de um gênero (1) ou não (0). O usuário pode avaliar de forma binária os filmes, sendo um utilizado para avaliações positivas e menos um para avaliações negativas.

Tabela 3 – Matriz de avaliações de um usuário a 13 itens (filmes), sendo os itens 1-8 do *dataset* D_L (avaliados) e 9-13 do *dataset* D_U (sem avaliação), as palavras-chave referentes ao gênero do filme (1 se o filme for do gênero e 0 se não). As avaliações são positivas (1) ou negativas (-1).

Itens ⇒ Palavras-Chave ↓	D_L								D_U				
	1	2	3	4	5	6	7	8	9	10	11	12	13
Comédia	1	0	0	1	1	1	1	1	0	1	0	1	0
Aventura	1	0	1	1	0	0	0	1	0	1	0	0	0
Terror	0	1	1	0	0	0	1	0	1	0	0	0	0
Romance	0	0	0	0	1	0	0	0	0	0	1	0	0
Suspense	0	1	0	1	0	0	0	0	1	0	0	0	1
Avaliação	1	1	1	-1	-1	1	1	-1	?	?	?	?	?

Tabela 4 – Métrica de similaridade (Cosseno) entre todos os itens, os K-Vizinhos Mais Próximos (KNN) e a avaliação predita e normalizada para cada item.

Itens	1	2	3	4	5	6	7	8	KNN	AP	AN
9	0	0	0,5	0,4	0	0	0,5	0	3, 7	1	1
10	1	0	0,5	0,8	0,5	0,7	0,5	1	1, 3, 4, 5, 6, 7, 8	0,15	0
11	0	0	0	0	0,7	0	0	0	5	-1	-1
12	0,7	0	0	0,6	0,7	1	0,7	0,7	1, 4, 5, 6, 7, 8	0	0
13	0	0,7	0	0,6	0	0	0	0	2, 4	0	0

O cálculo abaixo mostra a aplicação da fórmula do Cosseno para a avaliação de similaridade entre os itens 9 e 3 e a Tabela 4 mostra a similaridade calculada para todos os itens de D_U , seus K-Vizinhos Mais Próximos (KNN), a Avaliação Predita (AP – média das avaliações dos KNN) e a Avaliação Normalizada (AN – avaliações preditas acima do limiar estabelecido de 0,5) para cada item. Fica evidente, portanto, que o item nove seria recomendado.

$$sim(9, 3) = \frac{(0 * 0) + (0 * 1) + (1 * 1) + (0 * 0) + (1 * 0)}{\sqrt{(0^2 + 0^2 + 1^2 + 0^2 + 1^2)}\sqrt{(0^2 + 1^2 + 1^2 + 0^2 + 0^2)}}$$

$$sim(9, 3) = 0,5$$

Por fim, o componente de filtragem no SR BC será responsável por filtrar as recomendações de acordo com as avaliações preditas na etapa anterior e por adaptar o SR de acordo com as possíveis mudanças implicadas com o uso da plataforma, como, por exemplo, a inserção de novas avaliações por parte do usuário (aumento do *dataset* D_L) ou inserção de novos itens passíveis de recomendação na plataforma (aumento do *dataset* D_U).

Por depender unicamente das *features* extraídas de cada item, a **explicação** das

recomendações neste modelo de SR é simplificada quando comparada aos modelos de FC. É possível explicitar o motivo de um item estar sendo recomendado de acordo com as palavras-chave que apresentaram maior relevância para aquela recomendação. No exemplo desenvolvido acima, a recomendação do item nove poderia ser explicitada ao usuário como feita baseada em seu gosto por filmes de terror e suspense.

Por outro lado, assim como os SRs de FC, existem dois pontos negativos principais inerentes a este modelo de SR:

- Os SRs BC tendem a apresentar **baixa serendipidade e novidade** nos itens recomendados. Isso se deve ao fato da recomendação aqui depender apenas de características específicas de cada item, sem que seja levado em conta o fator social de descoberta. Isso acaba levando, em última instância, a recomendação de conteúdos repetitivos e que não furam a bolha de gostos do usuário;
- Apesar de resolver o problema de **cold start** para novos itens, este modelo não resolve o problema para **novos usuários** (quando ainda não há avaliações – implícitas ou explícitas). Isso implica na necessidade de apresentação e avaliação de itens por novos usuários, a fim de alimentar o modelo de treinamento e, assim, gerar recomendações.

2.2.3 Modelos de Filtragem Híbrida

Os modelos atuais de SRs empregam a filtragem híbrida, que combina aspectos dos SRs de filtragem colaborativa (FC), baseados em conteúdo (BC) e outras metodologias de filtragem, como:

- Filtragem baseada em conhecimento - Levam em consideração regras e informações específicas de um determinado domínio – como, por exemplo, considerando um serviço de recomendação no domínio de automóveis, utiliza de informações como modelos, ano, cor, etc. e gerencia como essas informações se relacionam entre si, através de regras associativas – e utilizam modelos de inferência para gerar recomendações (BURKE, 2002).

Nestes modelos, são utilizados grafos ou ontologias na representação dos conteúdos e dados de relevância, de maneira que estes sejam estruturados de forma associativa para geração de recomendações. Em um primeiro momento, o sistema constrói um perfil de usuário (da mesma maneira que os de filtragem colaborativa e baseada em conteúdo, explicitamente através de questões ou avaliações ou pelo próprio histórico de uso do usuário). Após, é feito o mapeamento de critérios e restrições do domínio para aquele usuário e, por fim, são aplicados

métodos de inferência para a recomendação de outros produtos/itens (FRIEDRICH; ZANKER, 2011).

Apesar de não utilizados vastamente, este tipo de filtragem ainda é encontrado em plataformas como Cookpad (recomendação de receitas) (GAUDANI, 2016) e Amadeus (recomendação de viagens) (PREZET, 2019).

- Filtragem baseada em popularidade - Método mais simples de filtragem, recomenda os itens mais populares em determinada plataforma, levando em conta que itens mais populares tendem a agradar um amplo número de usuários. Essa popularidade pode ser medida através de avaliações dos usuários, o número de visualizações ou vendas ou até a avaliação média de um determinado item (ponderando a média entre indicativos de aprovação e desaprovação) (ADOMAVICIUS; TUZHILIN, 2005).

Após a determinação dos itens mais populares, o SR ainda pode passar por uma etapa adicional de filtragem para personalizar as recomendações a um determinado usuário dentre todos os itens mais populares (utilizando FC ou BC, por exemplo). As plataformas sociais mais utilizadas atualmente (Instagram, Twitter, YouTube e Facebook) utilizam de maneira híbrida este tipo de filtragem.

- Filtragem demográfica - Funciona como uma espécie de extensão da FC, utilizando de informações como idade, gênero, geolocalização, ocupação, estado civil e outros para gerar recomendações. Essas informações podem ser coletadas através de um cadastro inicial do usuário ou de maneira indireta, devendo ser definida explicitamente através da Política de Privacidade e Termos de Uso da plataforma (Subseção 2.3.1) (ADOMAVICIUS; TUZHILIN, 2005).

Através destes dados, os usuários são segmentados em diferentes grupos. Dentro destes grupos, são analisadas possíveis semelhanças entre usuários e isso define a preferência de conteúdo/item que deve ser direcionado na recomendação àquele grupo (SU; KHOSHGOFTAAR, 2009). Atualmente, estes dados e grupos são definidos através de uma ampla gama de informações e técnicas de *Big Data* (Seção 5).

- Filtragem baseada em contexto - Similar à filtragem demográfica, mas leva em consideração dados contextuais do usuário, como dispositivo de acesso, hora do dia, localização do usuário em tempo real, eventos recentes de relevância ou até o próprio histórico de uso até aquele momento (ADOMAVICIUS; TUZHILIN, 2011).

Um exemplo padrão deste tipo de SR é o de recomendação de restaurantes enquanto um usuário está em determinado local próximo ao estabelecimento. Sis-

temas como o Google Maps (SOUTHERN, 2018) e TripAdvisor (NILASHI et al., 2018) utilizam essa metodologia.

A utilização de uma metodologia híbrida nos SRs busca solucionar os pontos negativos inerentes às metodologias utilizadas de maneira isolada. Aumentando a robustez, precisão e diversidade nas recomendações, estes SRs conseguem, no geral, solucionar os problemas de *cold start* e de baixa serendipidade dos SRs BC e, também, reduzem a criação de um gráfico de cauda inversa em decorrência das tendências em avaliações dos SRs de FC (BURKE, 2002).

Para fins de exemplificação, quando são colocadas em evidência três das maiores plataformas sociais da atualidade (Facebook, Twitter e Instagram), é possível notar a maneira como os SRs são utilizados de maneira híbrida nas mais diversas *features*, operando em cada uma delas.

No Facebook, o *Feed* de Notícias utiliza uma mistura de FC e BC, calculando uma pontuação de relevância para cada *post*, baseado em sinais emitidos pelo usuário, como por exemplo: se o conteúdo curtido por um usuário é um vídeo ou uma foto, se o usuário curtiu diversas publicações de uma determinada pessoa ou página, se o usuário costuma mais compartilhar ou comentar um determinado tipo de postagem, a localização do usuário, etc. Essa pontuação define quão alto no Feed estará um determinado *post* para um usuário (LADA; WANG; YAN, 2021).

A recomendação de Amigos é feita majoritariamente através de FC, com usuários com o maior número de amigos em comum sendo ranqueados no topo para a recomendação de amizades. São levados em consideração, também, usuários de grupos em comum, marcações em fotos e redes de contato (filtragem demográfica, utilizando informações de cidade de residência, trabalho ou escola do usuário) (FACEBOOK, 2023a). Já a recomendação de Páginas leva em consideração as Páginas já curtidas pelo usuário e os tópicos de interesse analisados pelo SR, utilizando também de páginas curtidas por usuários com gostos similares e amigos do usuário em questão (FC) (FACEBOOK, 2023b).

Já o Twitter utiliza fortemente de FC para a recomendação de usuários a seguir e tweets, fazendo uma análise comparativa de interesses em comum entre contas e também uma análise de conteúdo publicado pelo próprio usuário (BC) para a recomendação de *tweets*. Ele utiliza, também, de filtragem baseada em popularidade, com *tweets* populares sendo recomendados a usuários que possam ter interesse naquele conteúdo (TWITTER, 2023).

Além disso, o Twitter utiliza também de dados geográficos (filtragem demográfica e baseada em contexto) dos usuários para a recomendação de Tópicos em Alta e Eventos ocorrendo em regiões próximas ao usuário e que podem ser de relevância para este (CHENG; CAVERLEE; LEE, 2010).

No Instagram, conteúdos e perfis são recomendados utilizando uma mistura de FC e BC. No *Feed* do usuário, para a recomendação de novas contas a seguir, é levado em consideração o tipo de conteúdo com o qual o usuário interage mais e também contas com seguidores em comum. Por exemplo: usuários que seguem um grande número de contas comerciais de restaurantes, terão mais restaurantes recomendados (também levando em consideração informações demográficas e de contexto, como localização atual do usuário). Já as publicações recomendadas que aparecem neste *Feed* são geralmente impulsionadas (pagas) por páginas ou perfis que estejam de acordo com os gostos do usuário em questão (NAYEK; DAS, 2021).

Já no *Feed* "Explorar", é feita uma combinação de publicações populares com conteúdos curados através de filtragem BC, tendo a plataforma gerado uma lista de preferências do usuário principalmente através das publicações que o mesmo curte e o tipo de páginas com as quais mais interage e sendo esta atualizada constantemente na medida em que o usuário utiliza mais a rede (MEDVEDEV, 2019).

Em comum à todas estas plataformas, o sistema de direcionamento de propagandas também utiliza de um SR baseado nos interesses de cada usuário, seja pelo histórico de uso apresentado por este ou pela segmentação demográfica feita através de dados inerentes ao usuário ou contextuais, como geolocalização, idade, gênero, etc. Além disso, essas plataformas utilizam *cookies* para o monitoramento de atividades em sites de terceiros para direcionar de maneira ainda mais eficiente a publicidade incorporada em sua plataforma (maiores detalhes na Subseção 2.3.1) (MNEIMNEH, 2023).

Dada a proporção tomada pelas plataformas virtuais e as infinitas possibilidades de combinações de dados de relevância (*Big Data*) ao gerar uma recomendação a um usuário, atualmente todas as recomendações utilizam técnicas de Aprendizado de Máquina (AM) para uma maior eficiência. Estas técnicas serão explicadas na Seção a seguir.

2.3 *Big Data*, Aprendizado de Máquina e Redes Neurais Profundas

No contexto dos SRs, *Big Data* (BD) se refere ao conjunto de técnicas e ferramentas utilizadas para o processamento de uma quantidade massiva de dados de múltiplos usuários no intuito de gerar recomendações personalizadas mais eficientes a cada um. Este tipo de método utiliza de dados como histórico de uso de usuários de uma plataforma, suas preferências manifestadas através de suas atividades *online*, suas avaliações, interações e demais dados demográficos específicos a um usuário ou uma comunidade (ver Filtragem demográfica na Seção anterior).

O BD neste cenário se consolida principalmente através de três "V"s - Volume, Velocidade e Variedade (alguns autores utilizam quatro ou cinco "V"s, adicionando Ve-

racidade e Valor aos princípios de BD, mas para fins de simplificação e direcionamento do conteúdo, aqui serão abordados estes três) (Figura 5):

- Volume - Com a proliferação de dispositivos digitais e a abrangência do acesso a estes, os SRs têm acesso à grandes volumes de dados. Lidar com essa quantidade massiva de informações requer técnicas de armazenamento, processamento e análise escaláveis. A escalabilidade destes sistemas é crucial para garantir o processamento e geração de recomendações em tempo hábil (TAKÁCS et al., 2009);
- Velocidade - Tange a capacidade de processamento de dados em tempo real ou quase real, exigindo um processamento distribuído para processar e responder às solicitações de recomendação de uma plataforma no mínimo tempo possível e com os dados mais recentes e atualizados de um usuário. Isso é importante principalmente em comércios eletrônicos, pois converte-se diretamente na possibilidade de maior lucro à empresa com uma recomendação eficiente feita durante um processo de compra (JAMIY et al., 2015);
- Variedade - Os SRs precisam processar e analisar uma ampla gama de tipos de informações, podendo estas serem apresentadas de maneira estruturada (dados demográficos), semiestruturada (avaliações) ou não-estruturada (textos de avaliação de um item ou *posts* de uma rede, incluindo informações visuais). Para isso, é necessário que estes SRs tenham ferramentas para a análise de linguagem natural, análise de sentimentos e extração de recursos relevantes que influenciem na geração de uma recomendação mais precisa (JAMIY et al., 2015).

No cerne do processamento dessa quantidade massiva de dados está o Aprendizado de Máquina (AM), um conjunto de algoritmos de Inteligência Artificial (IA) que extrai padrões e conhecimentos a partir dos dados analisados, gerando modelos preditivos para a geração de recomendações. Para isso, estes algoritmos são treinados a partir do comportamento de um determinado usuário, sua interação com itens em uma plataforma e seus dados demográficos ou contextuais.

O principal algoritmo de AM utilizado nos SRs das plataformas virtuais atualmente é o de *deep learning*, ou Aprendizado Profundo (AP). Esta técnica utiliza de Redes Neurais Profundas (RNP), um tipo de Rede Neural composta de várias camadas capazes de aprender representações complexas e diversas de dados e gerar recomendações mais eficientes a partir de uma quantidade grande de dados de entrada, estabelecendo pesos de importância para cada um na geração de uma recomendação.

As RNPs utilizam camadas de neurônios matemáticos para processar dados. A informação é passada através de cada camada, com a saída da camada anterior fornecendo entrada para a próxima camada. A primeira camada em uma RNP é chamada



Figura 5 – 3 "V"s do *Big Data*.

Fonte: BU (2017) – traduzido pelo autor.

de camada de entrada, enquanto a última é chamada de camada de saída. Todas as camadas entre as duas são referidas como camadas ocultas. Cada camada é tipicamente um algoritmo simples e uniforme contendo um tipo de função de ativação (ACADEMY, 2022a). Uma representação de uma RNP pode ser vista na Figura 6.

No contexto dos SRs, cada uma das camadas de uma RNP funciona da seguinte maneira:

1. Camada de Entrada - Os dados de entrada em uma RNP de um SR podem incluir informações sobre o usuário, como histórico de uso, preferências declaradas, comportamento de navegação, dados demográficos, etc. e também informações sobre itens, como descrição, categoria, avaliações de usuários, etc. Estes dados são usualmente representados através de vetores de características. Essa etapa também faz a normalização e pré-processamento dos dados de entrada para que possam ser processados pelo restante do sistema. A escolha dos dados

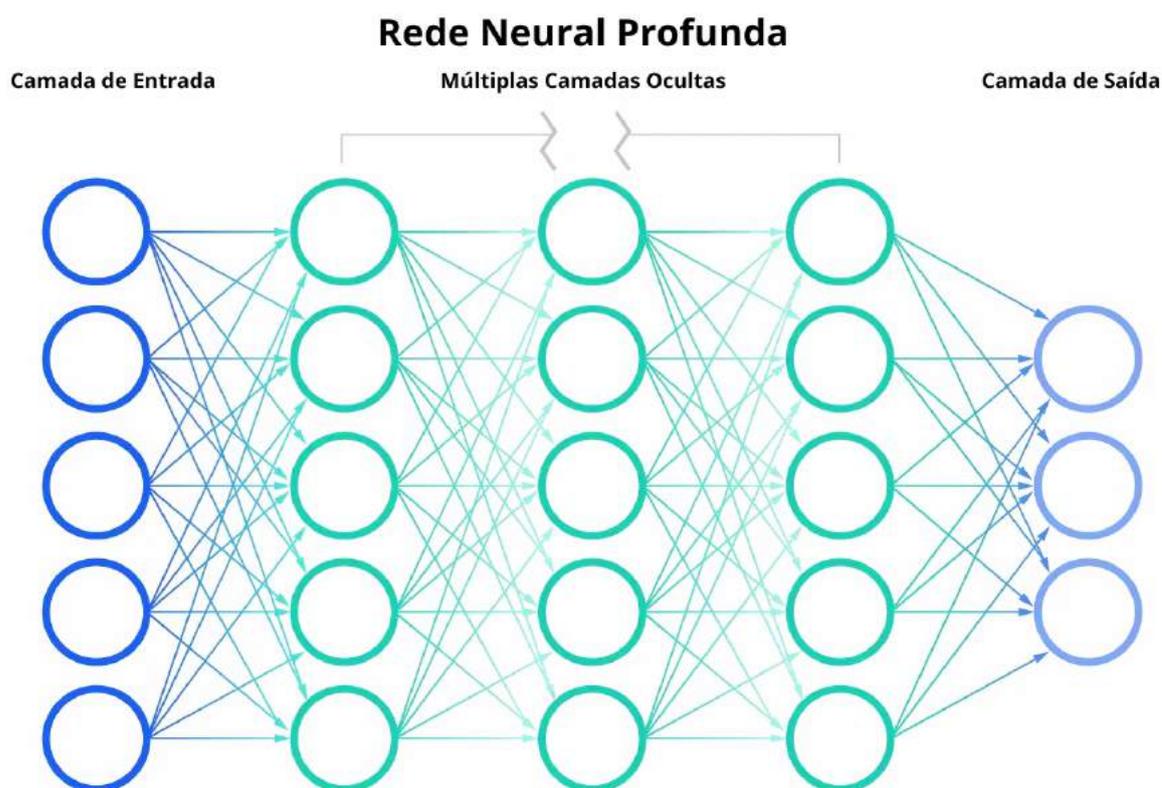


Figura 6 – Representação de uma Rede Neural Profunda (RNP).
 Fonte: IBM (2020) – traduzido pelo autor.

de entrada relevantes é um processo iterativo e, em muitos casos, automatizado (os problemas de privacidade dos usuários decorrentes disso são elucidados na Subseção 2.3.1) (ZHANG et al., 2019);

2. **Camadas Ocultas** - Cada neurônio dentro de uma camada oculta opera aplicando um algoritmo de transformação linear dos sinais de entrada que recebe seguido de uma função de ativação não-linear. A transformação linear envolve o cálculo do produto escalar entre os valores de entrada e o peso associado a cada conexão entre os neurônios, expresso através da Equação $z = \sum_i w_i x_i + b$, onde z representa a soma ponderada das entradas (ou a saída após a aplicação da transformação linear), w_i é o peso associado à conexão entre o neurônio e a entrada x_i e b representa o *bias* (viés), que permite ao neurônio aprender um valor constante (SHARMA, 2017).

A introdução da não-linearidade no processo é feita para possibilitar que o sistema aprenda relações complexas entre os dados de entrada. Para tanto, é utilizada uma função de ativação. A escolha da função de ativação varia de acordo com o tipo de recomendação sendo feita pelo SR e a relação de importância de cada um dos dados de entrada. Algumas das mais utilizadas são (ACADEMY,

2022b):

- (a) Função Sigmoidal - Transforma valores variáveis de entrada em um intervalo entre 0 e 1. É expressa através da expressão $f(z) = \frac{1}{1+e^{(-z)}}$, na qual e representa a base do logaritmo natural (número de Euler). Quando z é grande e positivo, $f(z)$ se aproxima de um e o inverso também ocorre, com um z grande e negativo aproximando o resultado da função de zero;
- (b) Função ReLU (*Rectified Linear Unit*) - Função utilizada pela sua simplicidade e eficiência computacional, retorna zero para valores negativos e mantém os valores positivos. É expressa por $f(z) = \max(0, z)$;
- (c) Função Tangente Hiperbólica (Tanh) - Mapeia os valores de entrada para um intervalo entre -1 e 1. Expressa pela fórmula $f(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$. Similar à Função Sigmoidal, porém apresenta uma saída simétrica em relação a zero;
- (d) Função *Softmax* - Funciona como uma espécie de extensão da Função Sigmoidal, testando a probabilidade de um sinal de entrada pertencer à uma determinada classe. É bastante utilizada na Camada de Saída, para a definição da saída final considerando um vetor de valores processados nas camadas ocultas anteriores. Considerando o vetor destes valores $z = (z_1, z_2, \dots, z_k)$, a fórmula da função é expressa por $f(z) = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}}$.

3. Camada de Saída - Consiste em um conjunto de neurônios que realizam uma transformação final dos dados processados nas camadas anteriores e pode gerar uma classificação binária, uma pontuação de relevância ou uma classificação multiclasse. Nessa Camada também são interpretados os resultados para a geração das recomendações finais a um usuário

Em uma Camada de Saída de classificação binária, é definido um limiar para a tomada de decisão de um item ser ou não recomendado. Na pontuação de relevância, o resultado pode ser interpretado de maneira direta, funcionando como uma pontuação para a recomendação de um item (itens que pontuam mais alto são recomendados primeiramente). Já em uma classificação multiclasse, o resultado é interpretado como a probabilidade de recomendação para cada classe (categorias de itens, por exemplo), sendo a classe com maior probabilidade a escolhida para recomendação (ZHANG et al., 2019).

As RNP são utilizadas em SRs principalmente pelos seguintes fatores (ZHANG et al., 2019):

- Transformação Não-linear: RNPs são capazes de modelar a não-linearidade presente em dados de entrada. Essa propriedade

possibilita a captura de padrões de interação usuário/item intrínsecas. RNPs são capazes de aproximar qualquer função contínua com uma precisão arbitrária, variando as combinações e escolhas de ativação. Essa propriedade permite um melhor reflexo da preferência de um usuário;

- **Aprendizado de Representação:** Em geral, uma grande quantidade de informações descritivas sobre itens e usuários estão disponíveis em uma aplicação do mundo real. Utilizar destas informações possibilita uma melhor compreensão dos itens e dos usuários, resultando portanto em um melhor SR. As vantagens do uso de RNP para o aprendizado de representação são duas: (i) reduz o esforço da construção manual das *features* relevantes em um SR, automatizando o processo de escolha a partir dos dados de entrada, com abordagens de aprendizado supervisionado ou não supervisionado e (ii) permite que os modelos de recomendação incluam informações de conteúdos heterogêneos, como texto, imagens, áudio e até vídeo;
- **Modelagem Sequencial:** RNPs mostram resultados promissores em um número de tarefas de modelagem sequencial, como processamento de linguagem natural, tradução de máquina, reconhecimento de fala, *chatbots* e outros. São amplamente aplicáveis e flexíveis na mineração de sequências estruturais em dados. Este tipo de modelagem é importante para a mineração das dinâmicas temporais de comportamento do usuário e de evolução de itens. Por exemplo: predição de um próximo item a ser adicionado em um carrinho de compras e recomendações baseadas em sessão são aplicações típicas;
- **Flexibilidade:** Com o advento de vários *frameworks* populares para o desenvolvimento de RNPs, é possível obter uma alta flexibilidade. A maioria das ferramentas para este fim são desenvolvidas de maneira modular e com uma comunidade ativa. Essa modulação permite, por exemplo, a combinação de diferentes modelos de RNPs para a criação de poderosos modelos híbridos que são capazes de capturar diferentes fatores e características.

2.3.1 Privacidade em Sistemas de Recomendação e Regulamentação

Embora o avanço dos sistemas de recomendação híbridos, que combinam técnicas de BD e AP, tenha possibilitado melhorias significativas na eficiência das recomendações geradas, é importante abordar uma série de problemas decorrentes do contexto social inerente ao uso desses sistemas. Os sistemas de recomendação baseados em

AP funcionam como caixas pretas, o que dificulta a verificação minuciosa do uso, processamento e armazenamento dos dados coletados de cada usuário para a geração de recomendações em diversas plataformas que utilizam esse tipo de sistema.

Na literatura relacionada a esse tema foi cunhado, inclusive, o conceito de "paradoxo de personalização-privacidade": a tensão entre a oferta de serviços altamente personalizados, como os SRs, e as preocupações dos usuários em relação à privacidade de seus dados pessoais. Ele envolve a ideia de que, embora a personalização possa melhorar a experiência do usuário, a coleta e o uso extensivos de dados pessoais podem levantar preocupações em relação à privacidade e ao controle sobre suas informações (AWAD; KRISHNAN, 2006).

Ao concordar com os Termos de Uso e Políticas de Privacidade de uma plataforma, os usuários acabam as vezes dando acesso a uma gama muito maior de dados pessoais que vão além dos estritamente necessários para o funcionamento devido das *features* da plataforma e dos SRs que ela implementa. Além disso, o usuário, que muito possivelmente não lê os Termos de maneira extensiva (o levantamento feito em CAKEBREAD (2019) aponta um percentual de 91% de usuários que sequer os leem), acaba concordando com o compartilhamento de seus dados com parceiros comerciais e anunciantes daquela plataforma, podendo resultar na exposição de informações pessoais a empresas que podem utilizá-las para fins de *marketing* direcionado, publicidade personalizada ou até mesmo vendê-las a outras organizações (MCDONALD; CRANOR, 2008).

Além disso, os *cookies* – pequenos arquivos textuais que são armazenados no navegador do usuário ao visitar um site e que registram informações sobre sua navegação, como suas configurações de preferência, histórico de uso, cliques, tempo passado em uma determinada página, etc. – são utilizados para coletar dados comportamentais do usuário e posteriormente aproveitados para aprimorar as recomendações geradas pelo SR da plataforma (CAHN et al., 2016).

Estes *cookies*, além de registrar as preferências de um usuário, podem coletar informações sensíveis, como detalhes de *login* e informações financeiras. Além do potencial compartilhamento destes dados com terceiros, há a possibilidade de vazamento destes dados em casos de brechas ou invasões destes sistemas (CAHN et al., 2016).

Dois casos de grande relevância global contribuíram significativamente para o debate sobre proteção de dados e privacidade. O primeiro caso é o da Cambridge Analytica, que envolveu a extração e uso indevido de dados de usuários para influenciar o comportamento virtual de multidões. Esse escândalo teve um impacto direto nas eleições presidenciais de 2016 nos Estados Unidos. O segundo caso é a votação do Brexit, na qual dados foram utilizados para influenciar a opinião pública no referendo sobre a saída do Reino Unido da União Europeia. Ambos os casos afetaram as deci-

sões democráticas de duas das maiores potências mundiais.

Fica evidente, portanto, a necessidade de regulamentação deste tipo de sistema. Neste sentido, legislações específicas ao redor do mundo foram criadas. Um exemplo relevante é o Regulamento Geral de Proteção de Dados (RGPD) da União Europeia, que entrou em vigor em maio de 2018. O RGPD estabelece um conjunto abrangente de regras e diretrizes para o tratamento de dados pessoais, com o objetivo de garantir a privacidade dos indivíduos e fortalecer o controle sobre suas informações.

O RGPD introduziu requisitos rigorosos para as organizações que coletam e processam dados pessoais, incluindo a necessidade de obter consentimento explícito dos usuários, a obrigação de fornecer transparência sobre como os dados são usados e protegidos, além de estabelecer direitos claros para os indivíduos, como o direito de acesso, retificação e exclusão de seus dados pessoais (União Europeia, 2016).

Por aqui, o Marco Civil da Internet (Lei nº 12.965/2014) estabelece diretrizes para o uso da Internet, incluindo princípios como a garantia da privacidade e a proteção dos dados pessoais dos usuários. Esse marco legal busca assegurar a liberdade de expressão e a privacidade dos indivíduos na era digital (BRASIL, 2014). Entretanto, cabe destacar que o MCI isenta as plataformas do conteúdo que é postado por seus usuários. Esse paradigma de isenção, baseado principalmente nas primeiras leis de regulamentação da internet dos anos 90, teve mudanças de entendimento significativas após os escândalos da Cambridge Analytica, votação do Brexit, a desinformação acerca da pandemia de COVID-19, entre outros fatores sociais de grande relevância no meio virtual, o que leva à atual tramitação do Projeto de Lei (PL) 2630/2020 (Ver Seção 3.5.1).

Além do MCI, a Lei Geral de Proteção de Dados (LGPD - Lei nº 13.709/2018), inspirada no RGPD da União Europeia, entrou em vigor no Brasil em setembro de 2020. A LGPD estabelece regras e diretrizes para o tratamento de dados pessoais por empresas e organizações, incluindo aquelas que operam sistemas de recomendação. Ela reforça a importância do consentimento informado dos usuários, a transparência na coleta e uso de dados, além da responsabilidade das empresas em proteger as informações pessoais dos usuários (BRASIL, 2018).

Apesar de toda a legislação criada a partir da preocupação sobre a privacidade e uso de dados de usuários, os SRs ainda apresentam outros desafios e problemas. Um deles é a criação de filtros bolha (ver Seção 3.3), nos quais os algoritmos de recomendação tendem a mostrar aos usuários apenas conteúdos semelhantes aos que eles já consomem, limitando sua exposição a diferentes perspectivas e ampliando a polarização de opiniões (PARISER, 2012). Além disso, os sistemas de recomendação também podem contribuir para a propagação de desinformação, uma vez que priorizam a maximização do engajamento e da retenção dos usuários, sem avaliar a veracidade ou a qualidade das informações recomendadas (DEL VICARIO et al., 2016). Esses proble-

mas ressaltam a necessidade contínua de reflexão e aprimoramento dos sistemas de recomendação, considerando não apenas a proteção da privacidade, mas também o impacto social e a responsabilidade ética envolvidos.

2.4 O Sistema de Recomendação do YouTube

Como o YouTube é o foco deste trabalho, essa seção discute um pouco mais o seu SR. O SR do YouTube é amplamente conhecido por seu papel crucial na personalização e apresentação de conteúdo aos usuários. Por meio de algoritmos complexos baseados em técnicas de AP, o YouTube busca oferecer recomendações personalizadas que cativem e mantenham o interesse dos usuários. No entanto, a opacidade desse sistema tem sido uma preocupação crescente. À medida que a plataforma coleta uma imensa quantidade de dados sobre os usuários, surgem questionamentos sobre como essas informações são utilizadas e como o algoritmo toma decisões de recomendação. A natureza fechada e complexa dos modelos de AP torna difícil para os usuários compreenderem como suas informações são processadas e como o conteúdo é selecionado para exibição.

Começemos, portanto, com as informações oferecidas pelo próprio YouTube: o SR do YouTube funciona de maneira híbrida, incorporando vários dos conceitos explicados previamente neste Capítulo. Ele mistura, em seu sistema proprietário, FC, BC, demográfica e contextual, operando através de RNPs para a geração de recomendações e que é aprimorada todos os dias através da interpretação de mais de 80 bilhões de sinais, dentre eles “cliques, tempo de visualização, respostas de enquetes, compartilhamentos, *likes*, e *dislikes*” (GOODROW, 2021).

Seu SR foi evoluindo com o tempo, como é possível ver na Figura 7. Começou primeiramente (em 2008) como um SR baseado em popularidade, recomendando simplesmente os vídeos mais populares da plataforma naquele momento. Foi evoluindo e utilizando de técnicas de AM e AP que levam em conta uma quantidade massiva de informações para a recomendação de vídeos. Vale frisar que, como é possível observar na Linha do Tempo, o SR passou por diversas mudanças para se adequar quanto à moderação de conteúdos sensíveis ou de cunho criminoso e, mais recentemente, conteúdos desinformativos (GOODROW, 2021):

"Utilizamos recomendações para limitar conteúdos de baixa qualidade de serem amplamente vistos desde 2011, quando construímos classificadores para identificar vídeos de cunho racista ou violento e prevenir que fossem recomendados. Então em 2015, notamos que conteúdos sensacionalistas estavam aparecendo nas páginas iniciais e os rebaixamos nas recomendações. Um ano depois, começamos a prever a potencialidade de um vídeo incluir menores em situações de risco e



Figura 7 – Linha do Tempo do SR do YouTube.
Fonte: VECTOR (2018) – traduzido pelo autor.

os removemos das recomendações. E em 2017, para garantir que nosso SR fosse justo com comunidades marginalizadas, começamos a avaliar o AM que alimenta nossos sistemas para garantir justiça entre estes grupos – como a comunidade LGBTQ+.

A ascensão da desinformação nos anos recentes nos levou a atentar nas maneiras que nosso SR inclui conteúdos desinformativos problemáticos e conteúdos “limitrofes” – isto é, conteúdo que chega perto, mas não viola nossas Orientações à Comunidade. Isto inclui vídeos de teorias conspiratórias (“o pouso na Lua foi falsificado”) ou conteúdo que espalha desinformação (“suco de laranja pode curar câncer”).

2.4.1 A Arquitetura do Sistema do YouTube

Quando o lado mais técnico do funcionamento deste SR é avaliado, as informações disponíveis são esparsas e pouco esclarecedoras. O YouTube tem sido discreto sobre os detalhes específicos de sua arquitetura de recomendação, mantendo-a em grande parte como um segredo comercial (caixa preta). Uma exceção a essa regra é o artigo COVINGTON; ADAMS; SARGIN (2016), escrito por engenheiros da equipe

do YouTube, que dá informações – ainda que em alto nível – do funcionamento interno de seu SR.

O Projeto “*The YouTube Video Recommendation System*”, da Universidade de Georgetown, compilou, em 2018, os principais estudos e artigos que buscam compreender e destrinchar o SR do YouTube (incluindo o artigo citado no parágrafo anterior) e a maneira que cada um de seus subsistemas funciona (VECTOR, 2018). A Figura 8 demonstra uma visão em alto nível do funcionamento da arquitetura deste SR. Cada um dos subsistemas será explicado na sequência.

O YouTube contou com quatro versões de arquiteturas de seu SR: a primeira era baseada em um grafo que fazia ligações de usuários-vídeos para a geração de recomendações (BALUJA et al., 2008). A segunda utilizava informações sobre a atividade do usuário na plataforma, recomendando vídeos similares aos visualizados previamente (DAVIDSON et al., 2010). A terceira também era baseada na atividade prévia, mas com um algoritmo mais robusto e eficiente (BENDERSKY et al., 2014). Já a quarta (e atual) versão utiliza de RNPs, resolvendo o problema de dados em larga escala, melhorando sua eficiência e eliminando ruídos nas recomendações (COVINGTON; ADAMS; SARGIN, 2016).



Figura 8 – Arquitetura em alto nível do SR do YouTube.

Fonte: (VECTOR, 2018) – traduzido pelo autor.

Desenvolvido através de técnicas de AP, a arquitetura atual do SR do YouTube opera através de três subsistemas: o de Coleta de Dados do Usuário, o de Geração de Candidato e o de Ranqueamento (Figura 8). Na Geração de Candidato, o aprendizado do SR é formulado como um problema de classificação multiclasse: dentre todos os vídeos existentes, prevê em quais deles o usuário pode ter engajamento. Já no Ranqueamento, o aprendizado é formulado como um problema de regressão logística ponderada: dado um par usuário/vídeo, prevê se o usuário irá engajar com aquele vídeo ou não (SMITH, 2021).

Na etapa de Coleta de Dados do Usuário (Figura 9), são extraídos dados sobre vídeos e sobre o usuário. Os dados sobre vídeos contêm: dados de conteúdo (duração, qualidade do vídeo, etc.), número de visualizações, avaliações, número de comentários e compartilhamentos, data de *upload* e meta-dados (título, descrição, etc.) (VECTOR, 2018).

Já os dados de usuários incluem suas avaliações, *likes* em vídeos, canais inscritos,

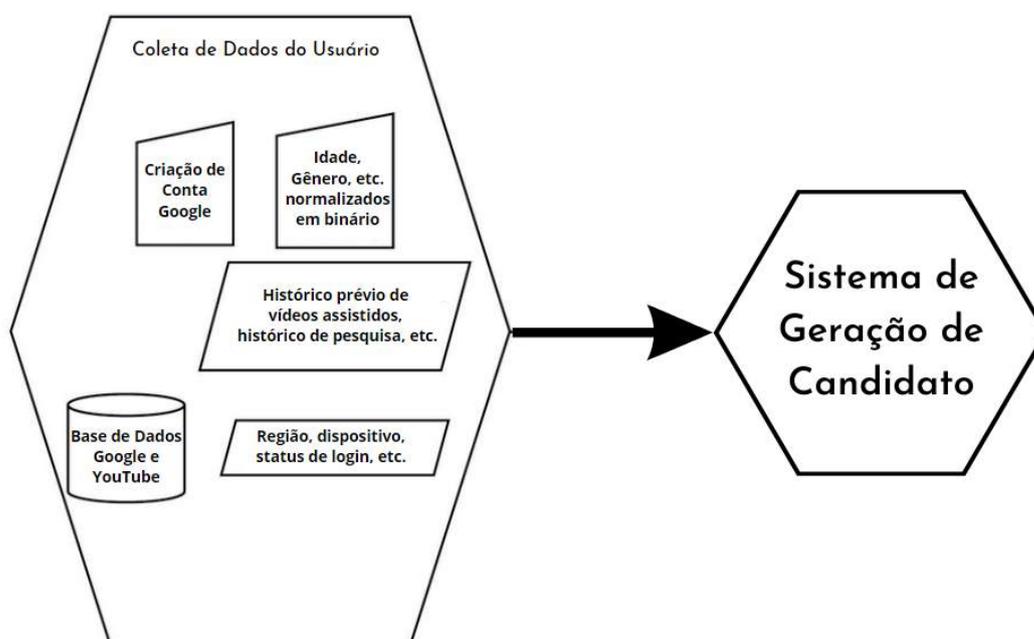


Figura 9 – Sistema de Coleta de Dados do Usuário.
 Fonte: VECTOR (2018) – traduzido pelo autor.

etc. (dados explícitos) e vídeos que o usuário começou a assistir, vídeos que assistiu mais que uma fatia pré-determinada, etc. (dados implícitos). Dados inerentes ao usuário também são considerados nesta etapa, como dados contextuais e demográficos (VECTOR, 2018).

Dados explícitos tendem a ser mais escassos (um usuário pode gostar de um vídeo e não apertar o botão de *like*, por exemplo), portanto os dados implícitos tendem a ter uma relevância maior nesta etapa: cliques de um usuário e tempo assistindo determinado vídeo tendem a ter maior ruído (usuários podem clicar sem querer em um vídeo), mas têm uma ordem de magnitude maior que a de dados explícitos. Na escala do YouTube, a quantidade de dados tende a ser mais importante que a qualidade dos dados (SMITH, 2021).

Um dado importante também nesta etapa é a sequência em que os vídeos são assistidos: o modelo do YouTube faz uma previsão de potenciais próximos vídeos a partir dos últimos termos pesquisados pelo usuário e seu histórico de visualização. Isso é feito utilizando dos últimos 50 vídeos assistidos pelo usuário e os 50 termos pesquisados no momento de treinamento da RNP (SMITH, 2021).

Os dados coletados na primeira etapa, são, então, transformados em um vocabulário fixo através de associações de alta dimensão e servem como entrada de uma RNP. O histórico de visualização do usuário, considerado aqui como “uma sequência de IDs de vídeo de tamanho variável” é mapeada para um “vetor denso”, que armazena todas as entidades através de associações. O histórico de pesquisa é transformado em uni-

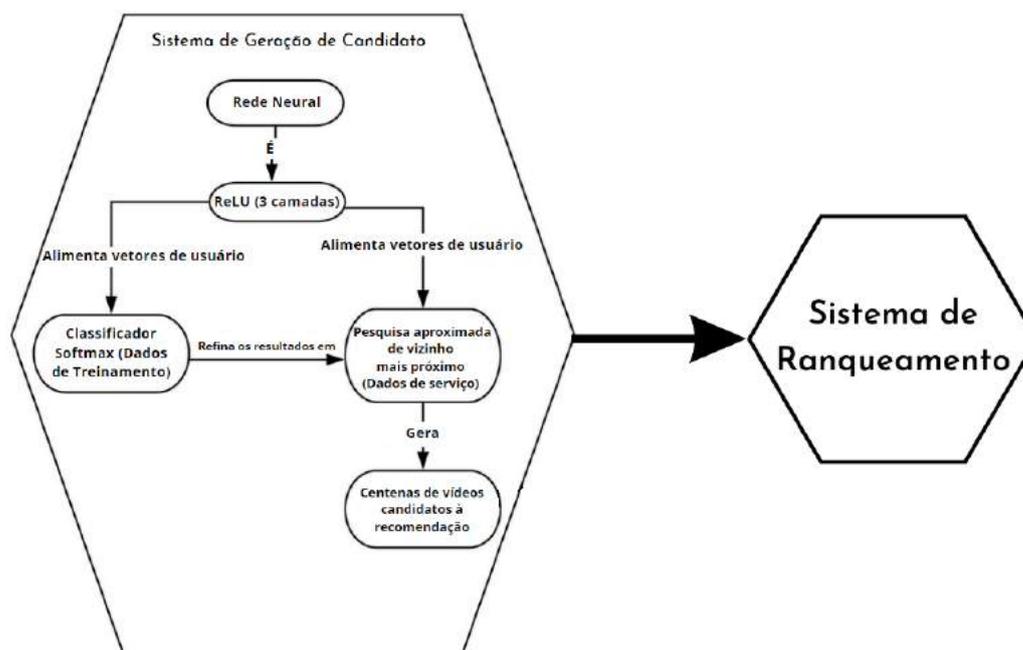


Figura 10 – Sistema de Geração de Candidato.
Fonte: VECTOR (2018), – traduzido pelo autor.

gramas e bigramas (sequências contíguas de um e dois itens, respectivamente), que são associados. Os dados associados de visualização e histórico são então ponderados e é gerado um “histórico de pesquisa denso sumarizado”. Como o processo de associação é feito na RNP, informações contínuas e categóricas podem ser inseridas, como gênero, estado de *login*, idade e demais informações do usuário, como localização informações do dispositivo. Estas informações são concatenadas na entrada da RNP após as informações de histórico de visualização e pesquisa (VECTOR, 2018).

As associações concatenadas passam por três camadas da RNP que utilizam da função de ativação ReLU (ver Seção 5), reduzindo o número de potenciais candidatos à recomendação a cada camada. Os dados de saída da RNP são distribuídos, então, em dados de treinamento e dados de serviço, com o primeiro auxiliando no refinamento de resultados para o segundo. Como camada de saída da RNP, é utilizada a função Softmax (ver Seção 5), tendo como saída uma distribuição multiclasse com a probabilidade de cada componente (vídeo). Esses dados são passados para os dados de serviço e a partir daí são calculados os vizinhos mais próximos ao usuário, para a geração de uma lista com centenas de vídeos com potencial para recomendação (FC) (VECTOR, 2018).

Além disso, para solucionar o problema de que SRs tendem a recomendar vídeos mais antigos, em decorrência do fato de serem treinados com dados históricos (não são gerados a todo momento), ele insere uma informação de tempo nos dados de treinamento. Por exemplo: se um dado de treinamento foi gerado na data atual, essa

informação é *setada* para “0” (mais relevante). Se o dado é mais antigo, este número vai sendo incrementado (menor relevância) (SMITH, 2021).

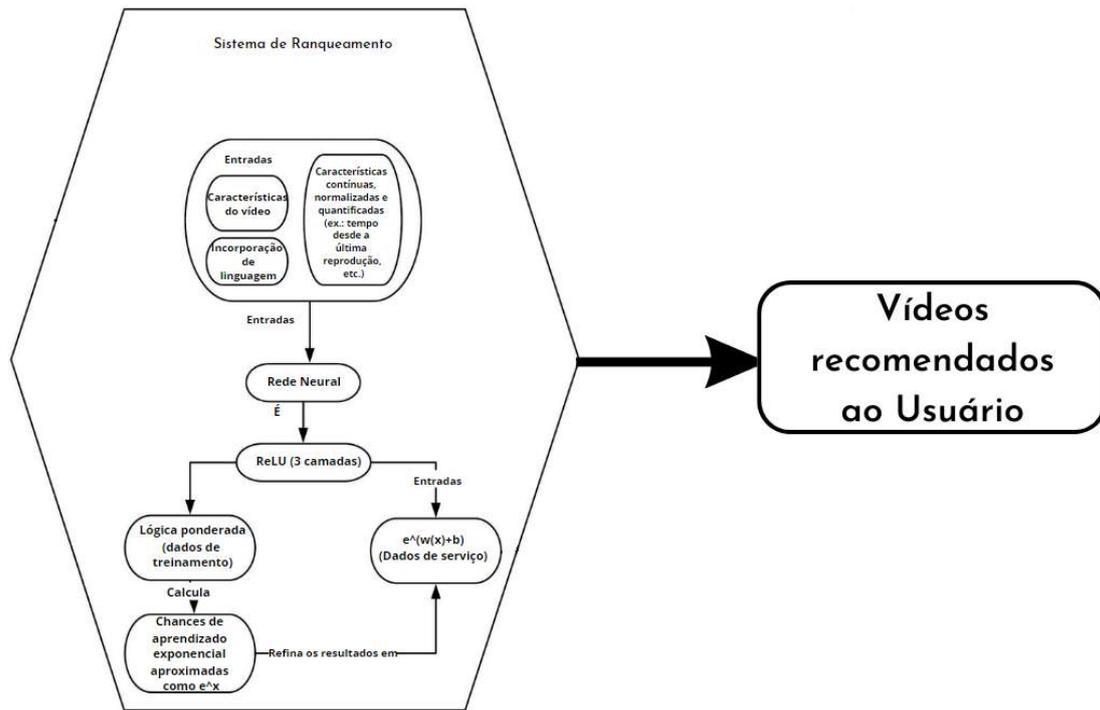


Figura 11 – Sistema de Ranqueamento.
Fonte: VECTOR (2018) – traduzido pelo autor.

Já a última etapa do SR, o Sistema de Ranqueamento (Figura 11), tem como entrada as centenas de vídeos de potencial recomendação da saída da Geração de Candidato e reduz essa lista para algumas dezenas, que são por fim apresentadas ao usuário (tanto em vídeos recomendados na sua página inicial como em vídeos específicos). Seu funcionamento é bastante similar ao da etapa anterior, com a diferença de que são utilizadas mais características dos vídeos (valores normalizados de tempo desde última visualização do usuário, número de impressões prévias sobre o vídeo, etc.) e é aplicada uma lógica nos dados de treinamento para estimar o potencial tempo de visualização de um determinado vídeo que aquele usuário terá, ranqueando mais alto os vídeos que apresentam uma probabilidade de maior tempo de visualização do usuário em questão (VECTOR, 2018).

Por fim, apesar destes *insights* no funcionamento do SR do YouTube, a maneira que cada um dos dados utilizados como entrada é ponderado para a geração das recomendações não é claro e, por se tratar de uma RNP, não pode ser claramente definido sequer pelos criadores do próprio sistema. Por isso, acaba sendo necessário estudos de auditoria destes sistemas para mensurar o impacto que eles têm, principalmente em fatores sociais, como o viés de conteúdos discriminatórios, a propagação de desinformação e a criação de filtros-bolha.

Nesse sentido, este trabalho pretende auditar este SR e a influência que ele exerce

na criação de filtros-bolha desinformativos acerca das Urnas Eletrônicas (UEs) no Brasil (ver Capítulo 5). Outros trabalhos relacionados também auditaram o SR do YouTube e de outras plataformas afim de averiguar o real impacto que os dados extraídos dos usuários têm nas recomendações geradas (ver Capítulo 4).

3 DESINFORMAÇÃO E URNAS ELETRÔNICAS

De histórias difamatórias antissemitas do Século XV na Europa até missivas de retribuição divina após o terremoto que atingiu Lisboa em 1755, as *fake news* têm circulado desde muito antes da invenção de jornais e mídias impressas (SCHUDSON; ZELIZER, 2017). Informações falsas e imprecisas permeiam a história da humanidade e dos meios de comunicação. No entanto, cabe destacar que o fenômeno da desinformação sofre um esvaziamento analítico quando é resumido à questão das *fake news* e confundido com outras formas de distorção dos fatos (MARTINS, 2020). É necessária uma análise aprofundada de todo o contexto e intenção na propagação de determinado conteúdo desinformativo.

Compreender o fenômeno da desinformação requer examinar não apenas o conteúdo em si, mas também o contexto social, político e cultural em que é disseminado. Além disso, é fundamental investigar as motivações por trás da criação e disseminação da desinformação, que podem envolver interesses políticos, econômicos ou ideológicos. Somente por meio de uma análise minuciosa e abrangente é possível entender a complexidade desse problema e desenvolver estratégias eficazes para combatê-lo. Segundo COOK; ECKER; LEWANDOWSKY (2015):

Existem diversas razões pelas quais a desinformação tem um potencial efeito prejudicial maior que a ignorância, isto é, a ausência de conhecimento. (i) A desinformação pode ser ativamente disseminada com a intenção de enganar. (...) Por exemplo: campanhas anticiência desinformam o público sobre questões que já obtiveram consenso na comunidade científica, como evolução biológica e a influência humana na mudança climática. Entretanto, a intenção de enganar não é mandatória. Por exemplo: a cobertura de uma notícia ocorrendo em tempo real necessita de atualizações regulares e correção de informações prévias. (...) (ii) Falsas crenças baseadas em desinformação são frequentemente mantidas com forte convicção, o que raramente é o caso com a ignorância. Por exemplo: pessoas que rejeitam a ciência climática também acreditam ser as melhores informadas sobre o assunto. (iii) A desinformação é muitas vezes imune à correção. Apesar de

claras retratações, desinformação e crenças falsas a ela associadas podem continuar influenciando o discurso e julgamento das pessoas. Essa influência contínua pode ser observada até quando as pessoas são expostas às retratações. Desinformação pode, portanto, afetar adversamente a tomada de decisões em sociedades democráticas que dependem de um público bem informado.

Indo na contramão da prometida Era da Informação do começo do século (CASTELLS, 2002), na qual testemunharíamos um acesso democrático e imparcial às fontes de notícias e meios de obtenção de conhecimento, vimos o avanço da lógica capitalista para com os meios de informação. Presenciamos a transformação do mercado das mídias de consumo em massa do século passado para o mercado hiper-individualizado do algoritmo nas redes, incluindo aqui os SRs.

Cunhado de capitalismo de vigilância (ZUBOFF, 2021), este novo modelo do capital, baseado amplamente na constatare utilização de dados coletados com a permissão dos próprios usuários (extração essa legitimada através de termos de compromisso e acordos de privacidade deliberadamente longos e complexos, não lidos por mais de 90% dos usuários (OBAR; OELDORF-HIRSCH, 2020)) para o direcionamento de propagandas e conteúdos sugeridos com base nestes dados.

Ressalta-se, porém, que a desinformação não nasceu com a internet nem com as redes sociais. Trata-se de um conjunto de práticas econômicas e políticas ligadas a um modelo irrigado pela lógica de acumulação econômica ampliada, de exploração do trabalho alienado e de dominação capitalista - algo que se fortalece de forma surpreendente em um contexto de crise. A digitalização, a interconexão mundial e as tecnologias associadas a elas intensificaram muito o potencial 'destrutivo' e desestabilizador da desinformação, mas estão longe de ser a razão da sua existência (WESTRUP et al., 2020).

As campanhas desinformativas encontram solo fértil na política e na ascensão da extrema-direita em nível global a partir de 2016, com a campanha de Donald Trump à presidência dos Estados Unidos. Mais tarde comprovado no escândalo da Cambridge Analytica, o então candidato republicano utilizou ativamente de propagandas direcionadas para desinformar determinadas parcelas dos votantes, além da utilização de *bots* para a produção de tais conteúdos (KAISER, 2020), o que possivelmente tenha levado à sua vitória contra a candidata democrata Hillary Clinton.

Surfando na onda das *fake news*, a campanha do então candidato à Presidência da República no Brasil Jair Bolsonaro se beneficiou da mesma estratégia: de 123 notícias falsas analisadas durante o período eleitoral, 104 (84,5%) beneficiavam o candidato (MACEDO, 2018). Até o final de seu mandato em 2022, ele utilizou reiteradamente de notícias falsas e desinformativas acerca de uma série de temas: de questões ambientais até o próprio sistema de votação que o elegeu.

O discurso desinformativo acerca das urnas eletrônicas (UEs) cresceu exponencialmente desde 2018, sendo especialmente presente em períodos eleitorais. Nas eleições de 2018, informações falsas sobre as UEs foram compartilhadas no Twitter e Facebook 3,4 milhões de vezes segundo a agência de *fact-checking* Aos Fatos. Já a agência Lupa informou que, somente no final de semana do primeiro turno desta eleição, informações consideradas inverídicas tiveram mais de 1,17 milhão de compartilhamentos no Facebook (BANDEIRA, 2018).

3.1 Classificando a Desinformação

Por se tratar de um tipo de classificação muitas vezes subjetiva, não há um consenso único e definitivo sobre a classificação de conteúdos de cunho desinformativo. Com parâmetros de análise que variam de acordo com o tema a ser estudado, é necessário fazer um aprofundamento do tipo de conteúdo, meio de propagação, a forma que está sendo comunicado, os agentes envolvidos na comunicação e também a intencionalidade destes, que muitas vezes pode ser explícita ou não. Segundo LEWANDOWSKY; ECKER; COOK (2017):

Para ser efetiva, a pesquisa científica acerca da desinformação deve ser considerada em um contexto político, tecnológico e social maior. O mundo da pós-verdade emergiu como um resultado de grandes tendências sociais, como o declínio de capital social e da confiança na ciência e o aumento da desigualdade social, polarização e fracionamento da mídia.

Em HABGOOD-COOTE (2019), há uma série de questionamentos relevantes que devem ser levados em conta no desenvolvimento de uma metodologia de classificação de conteúdo desinformativo e como o termo “*fake news*” por si só não consegue abranger todas essas nuances de maneira satisfatória:

- “*Fake News*” é um termo que pode ser aplicado apenas para notícias veiculadas *online* ou pode ocorrer em mídias tradicionais?
- Postagens individuais (feitas por indivíduos) podem ser consideradas “*Fake News*” ou apenas postagens feitas por veículos de comunicação no ambiente *online*? É possível aplicar o termo à mensagens do *WhatsApp*, por exemplo?
- Sátiras ou paródias de notícias podem ser consideradas “*Fake News*”?
- O termo “*Fake News*” se aplica a histórias completamente inverídicas, histórias parcialmente inverídicas ou histórias verdadeiras mas circuladas com intuito malicioso? Por exemplo, histórias verdadeiras no meio de um fluxo de informações

que não podem ser distinguidas entre falsas ou verdadeiras devem ser cunhadas da mesma maneira?

- O termo deve ser aplicado somente em informações circuladas com um certo tipo de intenção?
- O termo só pode ser aplicado em informações de ampla circulação?

Há também de ser considerada a barreira linguística inerente ao assunto nos trabalhos brasileiros: em inglês, há a distinção etimológica entre três tipos de desordens informativas - *misinformation*, *disinformation* e *malinformation*. A primeira é utilizada para se referir à conteúdos falsos ou incorretos propagados de maneira não intencional. A segunda é utilizada quando há intenção de desinformar ao se compartilhar determinada informação falsa. Já a última é relacionada a conteúdos que são expostos com a intenção de prejudicar, mas não são falsos. A Figura 12, extraída e traduzida de WARDLE (2018), expõe a diferença entre estes três tipos de desordem informativa.

Tipos de Desordens Informativas

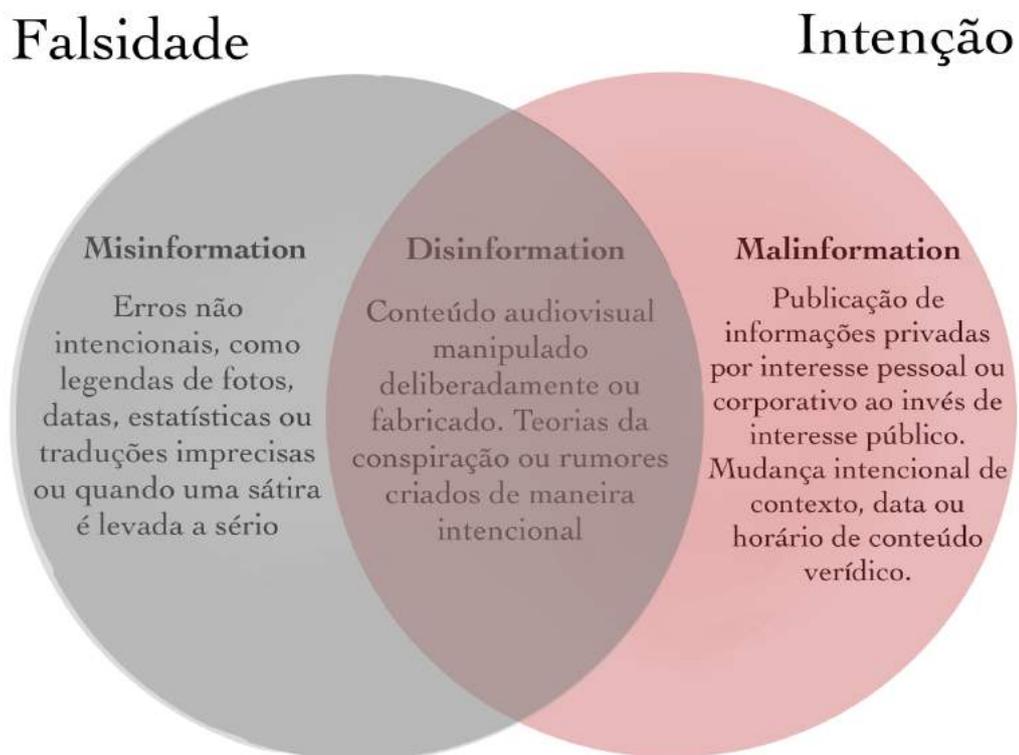


Figura 12 – Tipos de Desordens Informativas.
Fonte: WARDLE (2018) – traduzido pelo autor.

Apesar do termo “desinformação” por si só também não ser capaz de capturar todas as nuances sobre o conteúdo que define, é o mais apropriado na língua portuguesa para este tipo de análise, que visa compreender os fenômenos de informação que levam à crença em informações falsas ou distorcidas.

Diversos trabalhos apresentam diferentes modelos para a classificação dos conteúdos desinformativos. Este trabalho irá partir das categorias definidas por Claire Wardle em WARDLE (2017) e aprofundadas em WARDLE (2019) e irá detalhar o caminho metodológico utilizado a seguir. As sete categorias defendidas pela autora são:

1. Sátira ou paródia - Considerando os fenômenos desinformativos como um espectro (de menor a maior possibilidade de dano), esta categoria seria a primeira. O conteúdo satírico muitas vezes não tem o intuito de desinformar, sendo utilizado apenas como uma forma de humor.

O problema se encontra no amplo compartilhamento de conteúdos deste tipo, quando a origem do conteúdo pode vir a ser perdida e dá margem a interpretações errôneas do conteúdo sendo circulado, o que leva à interpretação de conteúdos absurdos ou fora de contexto como fatos sobre um determinado tópico. Há também a possibilidade do conteúdo originalmente humorístico ser utilizado por algum agente com o intuito de desinformar explicitamente;
2. Falsa conexão - Esta categoria é utilizada quando há uma desconexão entre o conteúdo sendo apresentado e o título/legenda (em caso de vídeos) ou manchetes (no caso de notícias). “*Clickbaits*”, por exemplo, se enquadram nesta categoria. O conteúdo apresentado não é necessariamente falso, mas ocorre distorção entre o fato apresentado e a maneira que é noticiado ou descrito, o que pode levar a interpretações errôneas;
3. Conteúdo enganoso - Categoria que pode ser considerada como a mais subjetiva ou difícil de ser enquadrada. Nela estão contidos fenômenos informacionais que utilizam de dados, estudos ou imagens de maneira arbitrária para a validação de uma opinião, sem levar em conta todo o contexto inerente da informação sendo usada.
4. Falso contexto - Utilização de imagens ou informações verdadeiras completamente fora de contexto. Esse tipo de desordem informativa é amplamente utilizado em conteúdos de cunho preconceituoso (racismo, homofobia, xenofobia, etc.), utilizando imagens de pessoas de determinado grupo social ou étnico e distorcendo as informações sobre o que de fato está ocorrendo na imagem.
5. Conteúdo impostor - Conteúdo manipulado que utiliza as *logos* ou o nome de instituições ou pessoas verdadeiras e com credibilidade para veicular fatos falsos

ou distorcidos. Por exemplo: um conteúdo completamente falso circulando como tendo sido reportado por uma emissora de televisão, como a Globo.

6. Conteúdo manipulado - Aqui enquadram-se montagens e conteúdos audiovisuais manipulados para insinuar a autoria de uma determinada pessoa ou alegar algo completamente falso. Uma técnica amplamente utilizada nesta categoria é a montagem utilizando imagens legítimas distintas para a insinuação de um fato falso.
7. Conteúdo fabricado - Conteúdo completamente falso e inventado circulado como informação verdadeira, sem que sejam apontadas fontes críveis ou, em muitos casos, sem menção nenhuma a autoria em conteúdos de grande circulação.

3.2 Análise de Conteúdo

Portanto, para a análise de informações tendo como base estes conceitos acerca de conteúdos desinformativos, foram desenvolvidas metodologias para diferentes fins:

Na análise de conteúdo (AC) e linguagem, são analisados padrões de desinformação contidos na linguagem utilizada em mensagens ou notícias, utilizando de técnicas como análise de sentimentos, detecção de polarização e identificação de informações falsas ou enganosas. Esta abordagem é amplamente utilizada para examinar textos e discursos com o objetivo de compreender e interpretar seu conteúdo, contexto e características linguísticas. Essa metodologia envolve a identificação e categorização de elementos textuais, como palavras-chave, temas, padrões de linguagem, sentimentos expressos e estratégias retóricas (KRIPPENDORFF, 2004).

Na análise de redes sociais busca-se estudar a estrutura e padrões de relacionamentos entre indivíduos, organizações, canais, etc. dentro de uma plataforma social. Para isso, é feita a análise da propagação de informações através das interações entre os nós da rede, como amigos, seguidores ou conexões e como a informação é replicada e alterada através destas ligações. Essa técnica é utilizada também para analisar comportamentos e influências sociais de entes sobre um determinado assunto e permitem uma análise da propagação de desinformação acerca de um assunto (BORGATTI; EVERETT; JOHNSON, 2018).

Como suporte a análise de conteúdo, a verificação de fatos (*fact-checking*) busca investigar a veracidade das informações apresentadas em notícias, discursos, declarações públicas e outros tipos de conteúdo. Essa abordagem busca identificar e corrigir erros, desinformação e informações enganosas, com o objetivo de fornecer ao público uma análise imparcial e precisa dos fatos (USCINSKI; BUTLER, 2013). No Brasil, temos uma série de agências de checagem de fatos que atuam para a contenção de conteúdos desinformativos de grande circulação no Brasil e fazem análises de maneira

imparcial, como a Agência Lupa e Aos Fatos.

Durante as Eleições de 2022, foi firmada inclusive uma parceria do TSE com as principais agências de checagem de fatos brasileiras para esclarecer conteúdos relativos ao processo eleitoral e sobre as UEs (ELEITORAL, 2022a). Informações verificadas por estas agências foram utilizadas na Metodologia do presente trabalho, detalhada no Capítulo 5.

Por fim, a análise de conteúdo desinformativo pode também ser feita de maneira automatizada através de técnicas de AM, treinadas a partir de conteúdos verificados previamente e atuando através do processamento de linguagem natural (podendo utilizar também transcrições de vídeos tanto para treinamento como para análise) (SHU et al., 2017).

Em CARDOSO SAMPAIO; LYCARIÃO (2021) é feita uma ampla análise histórica e a construção de um manual de aplicação da AC categorial quantitativa, utilizada neste trabalho. O autor constrói o manual a partir de diversas definições e aplicações de AC dos principais autores historicamente consolidados acerca do tema: Bardin, Bauer, Krippendorff e outros. Este livro será utilizado como base para as definições sobre o tema abaixo.

Segundo o autor, AC é uma técnica de pesquisa científica baseada em procedimentos sistemáticos, intersubjetivamente validados e públicos para criar inferências válidas sobre determinados conteúdos verbais, visuais ou escritos, buscando descrever, quantificar ou interpretar certo fenômeno em termos de seus significados, intenções, consequências ou contextos.

A AC deve ser feita, portanto, respondendo a três princípios epistemológicos fundamentais: replicabilidade, confiabilidade e validade. Segundo o autor:

Todos esses princípios são identificados não pela autoridade ou reputação da fonte, ou seja, dos autores ou instituição de origem do trabalho, mas pela capacidade que a AC tem de ser revisada pela comunidade científica em termos de consistência entre conceitos e os instrumentos metodológicos (validade), assim como em termos da precisão e acurácia dos codificadores ao utilizarem tal instrumento (confiabilidade). A própria oferta de condições para realizar essas revisões corresponde precisamente ao princípio de replicabilidade. Sem essa, a revisão não é possível e, portanto, o processo circular do desenvolvimento científico é interrompido.

Para seguir estes princípios, a técnica utilizada deve ser extensamente descritiva e passível de replicabilidade. Neste sentido, Krippendorff – um dos principais autores acerca do tema e responsável por uma das principais métricas de avaliação de confiabilidade de um protocolo de classificação (ver Seção 5) – define:

Técnicas devem ser confiáveis. Mais especificamente, técnicas de in-

investigação devem produzir resultados que sejam replicáveis. Isto é, pesquisadores trabalhando em diferentes contextos espaciais e temporais sob diferentes circunstâncias deveriam chegar aos mesmos resultados ao aplicar a mesma técnica aos mesmos dados. A replicabilidade é a forma mais importante de confiabilidade. A pesquisa também deve produzir resultados válidos, no sentido de que os esforços de pesquisa se mostrem abertos ao escrutínio cuidadoso e que as afirmações resultantes possam ser confrontadas diante de evidência disponibilizada independentemente (KRIPPENDORFF, 2004)

Levando em conta estes conceitos, para a análise do conteúdo desinformativo acerca das UEs, neste trabalho é feita uma AC categorial quantitativa. Para isso, foi desenvolvido um Protocolo de Classificação que define extensivamente o que deve ser considerado desinformação, utilizando como base análises feitas por agências de checagem de fatos e demais fontes confiáveis sobre o tema (na Seção 5 é detalhada a criação do Protocolo e também sua validação, utilizando o Alfa de Krippendorff).

3.3 A Influência Social dos Sistemas de Recomendação e os Filtros Bolha de Desinformação

Para que consigamos entender melhor a influência social dos SRs, cabe problematizar, em um primeiro momento, histórica e sociologicamente a relação que tem se desenvolvido entre a tecnologia, sociedade e política nas últimas décadas.

A tecnologia serviu em boa parte como base do movimento contracultural estadunidense que buscava diminuir o poder do Estado e do maquinário estatal, devolvendo ao povo sua capacidade de autonomia sobre seus gostos, escolhas e individualidade como um todo. Esta busca por individualidade decorrente do pós-segunda guerra e também do declínio do modelo fordista acabou sendo subvertida à lógica neoliberal de acúmulo de poder – agora em forma de dados – nas mãos de poucos (ZUBOFF, 2021).

O alerta sobre este acúmulo de poder e controle não é de agora. Em DELEUZE (1992) já se comentava:

Cada tipo de sociedade, evidentemente, pode-se fazer corresponder um tipo de máquina: as máquinas simples ou dinâmicas para as sociedades de soberania, as máquinas energéticas para as de disciplina, as cibernéticas e os computadores para as sociedades de controle.

Nos encontramos, portanto, no ponto-chave do capitalismo tardio (período do capital que compreende o final do século XX até os dias atuais, com a expansão de multinacionais e a globalização dos mercados (MANDEL, 1978)), no qual as *big techs* oferecem a tão sonhada emancipação do “sistema vigente” e das convenções sociais

previamente estabelecidas a passo que definem, por si só e muitas vezes contornando qualquer restrição legislativa, o padrão do comportamento que desejam dos usuários:

A Uber nos oferece serviços de transporte que se contrapõem ao setor existente dos táxis; o Airbnb nos ajuda a encontrar acomodações e evitar e evitar o setor hoteleiro; a Amazon se encarrega de vender livros sem passar pelas livrarias (...) não resta quase nenhuma restrição social, econômica ou política que o Vale do Silício não tenha se empenhado em romper (MOROZOV, 2019).

A simbiose desenvolvida entre a tecnologia e o capital se dá de maneira no mínimo paradoxal: enquanto a primeira prega a autonomia individual acima de qualquer coisa, oferecendo liberdade para expressão própria – em forma de *posts*, curtidas, compartilhamentos e diversas outras maneiras de gerar uma identidade virtual – desamarrando o indivíduo, portanto, das garras do *status quo* do que é imposto (seja pela mídia, pela política ou pela imposição do “politicamente correto”), serve ao mesmo tempo de maior fonte de superavit comportamental para o segundo que, em troca, utilizará dos dados gerados para controle social – com a única diferença de que o meio utilizado para moldar o comportamento, aqui, cabe no seu bolso.

Ao mesmo tempo em que esvaziam o debate acerca do Estado como agente do bem-estar social, aplicando a lógica neoliberal em tudo que tocam - “você pode, você consegue, você merece” - os tecnocratas do Vale do Silício atuam fora da legislação imposta a muito dos meios que condenam ou dos quais prometem emancipação, se utilizando da própria falta de transparência e possibilidade de análise que os SRs que desenvolveram possibilitam.

Na medida em que a autoconsciência e o senso crítico vão se esvaziando e a alienação aumentando, o usuário se torna fundamentalmente suscetível à qualquer influência que seja exercida sobre ele. Os SRs, principalmente nas mídias sociais, funcionando em grande parte através do engajamento dado pelos usuários e também pelo histórico de uso de cada um, acabam incentivando o viés de confirmação.

O viés de confirmação é a tendência de buscar, interpretar e lembrar informações de maneira a confirmar as próprias crenças e ideias preexistentes, enquanto se ignora ou descarta informações que contradizem essas crenças (PARISER, 2012).

O engajamento é uma métrica-chave para os SRs, uma vez que usuários engajados passam mais tempo em uma plataforma, interagem mais com o conteúdo e, conseqüentemente, geram mais dados valiosos para o sistema. Essa busca por engajamento acaba gerando uma dinâmica em que os conteúdos mais polêmicos, sensacionalistas, emocionalmente carregados ou impactantes têm maior probabilidade de serem recomendados, uma vez que tendem a gerar reações fortes e prolongar o tempo de interação dos usuários (SUI; HAWKINS; WANG, 2023). Partindo deste princípio, os sistemas de recomendação (SRs) têm o potencial de aumentar o engaja-

mento dos usuários ao direcionar conteúdos que reforçam seu viés de confirmação.

As emoções desempenham um papel importante nesse contexto. Conteúdos que despertam emoções intensas, sejam elas positivas ou negativas, têm maior probabilidade de serem compartilhados, comentados e curtidos, gerando um ciclo de retroalimentação para os SRs. Essas reações fortes podem ser exploradas pelos sistemas para aumentar o envolvimento dos usuários, mantendo-os conectados e alimentando a coleta de dados para melhorar ainda mais as recomendações.

Levando isso em consideração, é chamado de homofilia de conteúdo, o fenômeno em que os usuários são expostos principalmente a conteúdos que são semelhantes às suas preferências anteriores. Em termos computacionais, isso significa que nodos (conteúdos) similares dentro de um processo de AM tendem a ter laços relacionais, contribuindo com a estrutura de nodos individuais, ou seja, a conexão local entre os nodos de um sistema se conectam diretamente e têm uma relação de proximidade (ZHANG et al., 2016). No contexto do YouTube, por exemplo, isso cria um filtro bolha de informação, na qual as pessoas ficam presas em um conjunto limitado de perspectivas e têm menos acesso a visões divergentes ou informações que desafiam suas crenças.

O termo filtro bolha (ou bolha dos filtros), cunhado em PARISER (2012) é definido como um estado de isolamento intelectual resultante de buscas ou recomendações personalizadas nas quais um sistema determina quais informações o usuário gostaria de ver com base em informações previamente coletadas sobre o usuário, incluindo seu histórico de uso em uma plataforma. Esse processo se difere das mídias de propagação em massa e qualquer outro tipo de influência a qual tenhamos sido historicamente expostos enquanto sociedade de três maneiras relevantes:

Primeiro, estamos sozinhos na bolha. Um canal de TV a cabo que atenda a um interesse muito específico (o golfe, por exemplo) terá outros espectadores, com os quais compartilharemos um referencial. Agora, no entanto, cada pessoa está sozinha em sua bolha. Numa época em que as informações compartilhadas são a base para a experiência compartilhada, a bolha dos filtros é uma força centrífuga que nos afasta uns dos outros.

Segundo, a bolha dos filtros é invisível. Os espectadores de fontes de notícias conservadoras ou progressistas geralmente sabem que estão assistindo a um canal com determinada inclinação política. No entanto, a pauta do Google não é transparente. O Google não nos diz quem ele pensa que somos ou por que está nos mostrando o resultado que vemos. Não sabemos se as suposições que o site faz sobre nós estão certas ou erradas – as pessoas talvez nem imaginem que o site está fazendo suposições sobre elas. Minha amiga que recebeu informações

sobre a *British Petroleum* (BP) destinadas a investidores ainda não entendeu por quê, posto que não investe na bolsa de valores. Por não escolhermos os critérios que os sites usarão para filtrar os diversos assuntos, é fácil intuirmos que as informações que nos chegam através de uma bolha de filtros sejam imparciais, objetivas, verdadeiras. Mas não são. Na verdade, quando as vemos de dentro da bolha, é quase impossível conhecer seu grau de parcialidade.

Por fim, nós não optamos por entrar na bolha. Quando ligamos o canal Fox News ou lemos o jornal *The Nation*, estamos fazendo uma escolha sobre o tipo de filtro que usamos para tentar entender o mundo. É um processo ativo: nós conseguimos perceber de que modo as inclinações dos editores moldam a nossa percepção, como quando usamos óculos com lentes coloridas. Mas não fazemos esse tipo de escolha quando usamos filtros personalizados. Eles vêm até nós – e, por serem a base dos lucros dos sites que os utilizam, será cada vez mais difícil evitá-los.

Essa filtragem seletiva de conteúdo acaba sendo ainda mais danosa quando leva à formação de filtros bolha de desinformação, em que informações falsas ou distorcidas são recomendadas àqueles usuários que já possuem um histórico de consumo de conteúdo desinformativo. As plataformas de mídia social, ao priorizarem a maximização do engajamento, podem inadvertidamente reforçar visões extremistas e propagar desinformação, prejudicando o debate saudável, a diversidade de opiniões e a busca pela verdade.

É importante, porém, ressaltar que estudos mais recentes expandem o conceito de filtro bolha cunhado por Parisier, como exposto por LOIOLA (2018):

Como já vimos, há uma série de estudos sobre o tema que colocam em cheque, ou ao menos geram dúvidas quanto a alguns preceitos básicos no que diz respeito à polarização online, questionando alguns elementos-chave da teoria. Ainda que com certas limitações, e levando em conta as consideráveis diferenças entre as ideias ligadas a bolha, homofilia e balcanização, é possível observar que não há um consenso em torno de alguns dos seus preceitos básicos, e por isso devem ser levados em conta com cuidado.

E ainda que existam uma quantidade de estudos que mostram uma crescente polarização e a formação de câmaras de eco, também é necessário um certo nível de cautela, já que é difícil estabelecer tais relações de causalidade. Como se tratam de estudos com uma grande quantidade de variáveis envolvidas, é necessário um cuidado especial com elas, e com as suas correlações, elementos fundamentais para se analisar certos fenômenos com a devida cautela, especialmente quando eles são originais de outros países, com um ambiente político

que apresentam uma série de diferenças institucionais quando comparados com o Brasil.

Debate-se também sobre a familiaridade e envolvimento de um usuário com o tema que pesquisa: na percepção de uma possível ameaça, a escolha de informações que confirmam ideias pré-estabelecidas de usuário se acentua caso a pessoa tenha pouca experiência com o tópico em pauta. Já para pessoas com maior domínio do assunto, este fenômeno é limitado, tendo sido observado que as pessoas nesta ponta do espectro acabam procurando informações mais diversificadas (LIAO; FU, 2013).

Outro fator que leva a redução na diversidade das recomendações e, por conseguinte, a criação de filtros bolha é a presença de tópicos partidários específicos que podem se aproveitar da desinformação para gerar paixão ideológica alinhada com os princípios de determinado partido. Isso pode ser observado em uma série de estudos e acaba se estendendo inclusive à mídia tradicional, tendo sido constatado, por exemplo, nos Estados Unidos, a preferência por eleitores republicanos pelo canal Fox News (LOIOLA, 2018).

No Brasil, tal fenômeno se prolifera em canais como Jovem Pan e veículos como O Antagonista, abertamente porta-vozes da extrema-direita e que utilizam de pautas identitárias para um maior reconhecimento dos espectadores engajados com este viés político, mesmo que para isso utilizem de informações fora de contexto ou conteúdos explicitamente desinformativos.

Buscando comprovar a existência destes filtros bolha de desinformação em relação à UEs no YouTube, o presente trabalho explorou outras publicações que comprovaram a criação de tais filtros bolha para outros tópicos (ver Capítulo 4) e desenvolveu, a partir destes, uma Metodologia própria para a comprovação desse fenômeno (ver Capítulo 5).

3.4 As Urnas Eletrônicas e o Voto no Brasil

No Brasil, a soberania popular é exercida pelo sufrágio universal, voto direto e secreto, sendo facultativo para os maiores de 16 anos e menores de 18, assim como para os maiores de 70 anos e analfabetos. Contudo, o voto é obrigatório para os eleitores que tenham entre 18 e 70 anos (PAES, 2013).

O sufrágio é um direito público subjetivo, ou seja, um direito próprio da condição de cidadão, que inclui tanto o poder de escolha dos representantes quanto a possibilidade de concorrer aos cargos públicos eletivos. Quanto ao voto, embora seja obrigatório para uma determinada faixa da população, representa uma verdadeira conquista política para o povo brasileiro (PAES, 2013).

Sendo, portanto, o voto um direito e um dever cívico exercido pelos cidadãos para escolher seus representantes políticos e participar ativamente do processo democrá-

tico, a partir de 2000, o país adotou em nível o sistema de UEs como forma de garantir agilidade, segurança e transparência nas eleições.

As urnas eletrônicas são dispositivos eletrônicos de votação que substituíram o antigo método manual de cédulas de papel. Elas permitem que os eleitores escolham seus candidatos por meio de uma interface digital e registrem seus votos de forma rápida e precisa. O sistema de urnas eletrônicas tem sido amplamente utilizado no Brasil, tornando o processo eleitoral mais eficiente e confiável, garantindo a integridade e a contagem correta dos votos. Além disso, a tecnologia utilizada nas urnas eletrônicas busca oferecer segurança contra fraudes e garante o sigilo do voto, contribuindo para a lisura das eleições no país (COIMBRA, 2014).

No decorrer desta Seção serão apresentados detalhes históricos sobre o exercício do sufrágio no Brasil, além de todo o funcionamento das UEs, alvos de campanhas massivas de desinformação, principalmente em anos eleitorais.

3.4.1 A História do Sufrágio no Brasil

Para uma análise mais completa em torno da propagação de desinformação sobre as UEs, torna-se imperativa a análise da história do voto e todos os diferentes modos e legislações que regeram o exercício democrático no Brasil através do tempo.

Desde o começo da colonização do Brasil pelos portugueses em 1532 e durante todo o período colonial, orientado pelas Ordenações Filipinas, o poder era exercido centralmente pela Coroa e administrado regionalmente através das Câmaras e Conselhos (AZEVEDO, 2018), que contavam com a escolha de seus representantes (vereadores, juízes de paz e procuradores) através de eleições indiretas e em dois graus (eram escolhidos primeiramente os eleitores e estes elegiam, então, os representantes).

O sufrágio era praticado somente por homens acima de 25 anos, católicos, casados ou emancipados, que fossem proprietários de terras e “sem impurezas de sangue”. (NICOLAU, 2012). Ou seja, o voto era praticado unicamente pelos homens da casta da nobreza e os burgueses da época.

A partir de 1824, após a declaração de Independência e com a instauração do Brasil Império, Dom Pedro outorgou a primeira constituição brasileira, na qual constava a divisão do Poder Legislativo entre Câmara de Deputados (eleitos) e Senado (com os senadores escolhidos pelo Imperador dentro de lista tríplice indicada por eleições). O Brasil, dividido então em províncias, contava com um senador para cada dois deputados provinciais eleitos.

Durante este período, o voto era censitário (o eleitor deveria cumprir certos critérios) e a eleição era feita em quatro graus. No primeiro grau, poderiam votar homens livres acima de 25 anos ou casados e militares acima de 21 anos, que comprovassem renda anual acima de 100 mil réis. Estes eleitores elegiam os compromissários

(segundo grau).

Os compromissários elegiam, então, os eleitores da paróquia (terceiro grau), que então elegiam os eleitores da comarca (quarto grau). Todos estes deveriam comprovar uma renda mínima anual de 200 mil réis. Somente aos eleitores da comarca cabia o voto em deputados (que deveriam ter renda mínima de 400 mil réis anuais) e em senadores (que deveriam ter renda mínima de 800 mil réis anuais).

Apesar de contarem com uma participação popular mais ampla quando comparadas ao período colonial, as eleições neste período nada mais eram que uma maneira de sustentar as escolhas parlamentares do Imperador. Por esse motivo, ocorriam diversas fraudes eleitorais, facilitadas pelo voto que era feito em papel e assinado pelo eleitor. Algumas reformas eleitorais foram feitas nesse período visando a redução de fraudes, como o uso do título eleitoral (NICOLAU, 2002).

Com a Proclamação da República em 1889 foi iniciado o sistema presidencialista no país. Este período, chamado de Velha República, foi marcado pelas políticas do coronelismo e república do café-com-leite, com a alternância entre presidentes oligarcas mineiros e paulistas.

Apesar deste período ter marcado o fim do voto censitário, garantindo explicitamente o sufrágio direto e universal (BRASIL, 1891), ele ainda excluía mulheres, indígenas, analfabetos, mendigos e integrantes do clero (JAEGER, 2004). Podiam votar, portanto, somente os homens maiores de 21 anos.

Marcado por amplas fraudes eleitorais, o período regido pela primeira Constituição Republicana levou a cabo a política do coronelismo, que consistia no vasto poder dos coronéis no âmbito municipal, determinando quem seriam os representantes eleitos, através da coação de votos (voto de cabresto) e troca de benefícios e proteção oferecidas pelos militares, além de manipulação direta dos resultados eleitorais.

Dentre as fraudes ocorridas neste período, haviam os votos de pessoas que já haviam falecido, nomes de eleitores inventados, voto de pessoas que sequer estiveram presentes nas sessões eleitorais e outros tantos, chegando, inclusive, à eleição de Rodrigues Alves para a presidência em 1918 com 99% dos votos - um total de 386.467 votantes (JAEGER, 2004).

A Era Vargas, iniciada após a Revolução de 1930 com a ascensão de Getúlio Vargas ao poder, contou com a fundação da Justiça Eleitoral (JE) e a criação do primeiro Código Eleitoral (em 1932) (ZULINI; RICCI, 2020). Através desta legislação eleitoral, foi instituído o voto feminino e o voto secreto obrigatório em nível nacional para os maiores de 18 anos, além da criação do Tribunal Superior de Justiça Eleitoral (órgão que viria futuramente a se tornar o Tribunal Superior Eleitoral - TSE) e dos Tribunais Regionais Eleitorais (TREs).

Com o novo Código Eleitoral, o modelo de votação para Deputados, que até então era majoritário (distrital), passa a ser proporcional. No modelo majoritário, são eleitos

apenas os candidatos mais votados, independentemente de partido, de acordo com a região (distrito) na qual se candidatam. Já no modelo proporcional, os partidos políticos são levados em consideração para a ocupação dos cargos, não apenas os candidatos. Ou seja, um voto em um candidato corresponde a um voto em um partido, dando maiores chances de eleição de um determinado candidato de acordo com seu partido e, conseqüentemente, uma representatividade mais plural no poder legislativo.

“Como princípio representativo, o princípio proporcional considera terem as eleições, como principal função, a representação, no Parlamento, na medida do possível, de todas as forças sociais e grupos políticos da sociedade, na proporção de seu apoio eleitoral. (...) Para os que o defendem, esse princípio é mais justo e atende melhor ao imperativo democrático de dar voz às minorias, do que o princípio majoritário” (CINTRA, 2005).

Outro fato importante de ser destacado neste período foi o a exigência de registro dos candidatos antes do pleito, processo em que os partidos, alianças de partidos ou grupos de pelo menos cem eleitores deveriam registrar seus candidatos nos TREs com no mínimo cinco dias de antecedência ao dia das eleições (CANÊDO, 2003). Essa conquista, quando analisada junto ao voto secreto, vem no sentido oposto ao coronelismo e o voto de cabresto, dando à população a chance de se conscientizar de maneira mais clara na escolha de seus representantes.

Apesar das grandes conquistas neste período para o exercício democrático do voto e da representatividade plural de ideias no Parlamento com o Código Eleitoral (que inclusive fez a primeira menção à implementação de uma “máquina de votar” (VARGAS, 1932)), em 1937 o então presidente Vargas decreta estado de sítio, após alegações de uma série de revoltas no norte do país lideradas por movimentos militares alinhados a ideais fascistas. Intitulado de Estado Novo, este período teve um enorme retrocesso com o fechamento do Congresso Nacional e a extinção da Justiça Eleitoral, o Tribunal Superior e os Tribunais Regionais. Sob a ditadura de Vargas, o sufrágio não foi possível por quase 10 anos.

Este período ditatorial durou até 1945, quando, após o fim da Segunda Guerra e com a queda dos regimes autoritários na Europa, o Brasil sofreu diversas pressões internas e externas para a reimplantação do regime democrático. Mesmo com a promessa de eleições em dezembro daquele ano, Vargas foi deposto (BRAGA, 2016) e em seu lugar foi eleito Eurico Dutra.

Com o fim do regime totalitário do Estado Novo, o Brasil viveu um período de pleno exercício democrático por 19 anos, regido pela Constituição de 1946 e o Código Eleitoral de 1950, até o Golpe Militar de 1964. Durante este período cabe destacar a reinstauração da Justiça Eleitoral, que se tornou responsável por todo o processo eleitoral em nível nacional, estadual e municipal e o reconhecimento dos partidos políticos como unidades fundamentais no processo de voto, tornando-se obrigatório o registro

partidário para que um candidato pudesse concorrer a qualquer cargo.

A votação durante este período era feita através de cédula de papel oficial da Justiça Eleitoral depositada em urna velada (PORTO, 2002). Apesar das diversas fragilidades no sistema de votação, as eleições durante este período foram consideradas democráticas e limpas, ocorrendo casos de fraudes marginalizados.

Com o mundo imerso na Guerra Fria durante o começo dos anos 60, o Brasil, fazendo parte do bloco capitalista e sob a influência direta dos Estados Unidos, ficou à mercê do pânico generalizado aos movimentos de esquerda e à “ameaça comunista” que pairava no imaginário popular. Quem estava na presidência nessa época era João Goulart, presidente com tendências à esquerda do espectro político. Dentre as reformas que o presidente propunha, estava inclusive a possibilidade do sufrágio dos analfabetos em nível nacional (PIMENTEL, 2021). Com a soma de diversos fatores externos e internos e motivado pela “caça às bruxas” do comunismo, o Exército destituiu João Goulart.

Com o Golpe Militar em 1º de abril de 1964, mais uma vez o exercício do sufrágio é abalado, limitando o povo brasileiro às eleições diretas apenas de deputados federais, estaduais e vereadores. Período marcado pela censura e regido pelos Atos Institucionais (AIs), o cenário político brasileiro foi reformado ao bipartidarismo: de um lado, a Aliança Renovadora Nacional (Arena - partido dos militares) e, de outro, o Movimento Democrático Brasileiro (MDB), considerado de oposição (BRASIL, 1965).

Ainda que o sufrágio fosse extremamente limitado, em 1968 ele se tornou nulo: foi instaurado o Ato Institucional 5 (AI-5) (BRASIL, 1968), que fechou o Congresso Nacional e deu plenos poderes para os militares no controle da política nacional, engolfando o país, então, em uma Ditadura Militar. Cabe destacar que neste mesmo período foi criada a Lei de Segurança Nacional (CASTELLO BRANCO, 1967), que censurava por completo qualquer manifestação contra o governo militar.

Esse conturbado período só começou a ver seu fim após a abolição dos Atos Institucionais, em 1978, cedendo à pressão externa dos Estados Unidos (abalado após a Guerra do Vietnã) e interna dos movimentos que visavam os Direitos Humanos. Começa, então, o processo de Redemocratização sob o mando do presidente João Figueiredo.

Com os partidos políticos sendo novamente colocados no centro do debate político, os brasileiros voltaram a praticar o voto em 1982 nas eleições diretas para governadores, deputados estaduais, deputados federais e senadores. Marcou também este período o movimento Diretas Já, em 1983 e 1984 que exigia a volta da eleição direta para o cargo de presidente.

Sob o governo do presidente José Sarney, ainda eleito indiretamente, foi promulgada a Constituição Federal de 1988, que rege a democracia brasileira até os dias atuais. Nela é determinado que “todo poder emana do povo, que o exerce por meio

de representantes eleitos ou diretamente” (BRASIL, 1988), introduzindo novamente o sufrágio universal (e, finalmente, com a inclusão dos analfabetos) e as eleições diretas para todos os níveis dos Poderes Legislativo e Executivo.

Foi reduzida a idade mínima para o exercício do voto para 16 anos (com a obrigatoriedade permanecendo ainda somente após os 18 anos) e implementado o sistema eleitoral proporcional para Deputados e Vereadores e majoritário para Senadores, Prefeito, Governador e Presidente (CINTRA, 2005). Todos os cargos, com exceção dos Senadores, têm mandatos de quatro anos, com a possibilidade de uma única recondução no caso de Prefeito, Governador e Presidente e sem limite de reconduções para os Deputados e Vereadores. Os Senadores têm mandatos de oito anos, também sem limite de recondução.

Neste período que compreendeu 16 eleições diretas (sendo oito destas para presidente) desde a redemocratização do país, o sistema de votação passou por uma série de mudanças. A maior, cabe destacar, foi a implementação em nível nacional das UEs como sistema único de votação a partir das eleições de 2000. A UE será detalhada na Seção 3.4.2.

Quando analisamos a história do sufrágio no Brasil, conquistado em sua plenitude a duras penas através dos últimos quase 500 anos, podemos ver drásticas mudanças ocorrendo de tempos em tempos. Com essa história marcada por uma constante instabilidade na democracia popular, temos hoje o livre direito à escolha de nossos governantes.

Nos 33 anos desde a redemocratização, tivemos presidentes de esquerda, presidentes de direita e um Parlamento composto por uma pluralidade de ideias dentro de todo o espectro político. Passando por melhorias sempre que necessário, não estamos estagnados em um sistema político fixo, mas sim adaptável de acordo com as necessidades de determinado tempo para a manutenção da democracia e da soberania popular.

Aliando a baixa educação política no projeto educacional nacional (que culmina, enfim, em um desinteresse generalizado sobre a política em todos os níveis) a massiva presença da internet e das redes sociais na criação de um senso comum no espectro social (mais baseado em opiniões que fatos), está criado um campo fértil para a propagação de desinformação sobre as UEs.

A polarização política, no eterno embate de “nós contra eles”, amplia essa percepção e faz com que o grande parte da população acabe por acreditar em informações descabidas, propagadas sem base alguma na realidade. Com o sistema de votação não seria diferente, sendo um tema de ampla politização e divergência radical de opiniões entre os próprios representantes eleitos. Representantes eleitos estes que, inclusive, em muitos casos, colocam em cheque o próprio sistema que os elegeram, no intuito da criação de narrativas artificiais para justificativa de eventual derrota em

eleições futuras.

A história nos ensina que melhorias são sempre necessárias para o pleno exercício democrático através do sistema de votação vigente. Não podemos admitir, porém, que essas discussões sejam baseadas em conteúdo falso ou desinformativo.

3.4.2 A Urna Eletrônica

A implementação da UE como meio de votação oficial no processo eleitoral começa a partir de 1996, após as fraudes generalizadas que marcaram as eleições de 1994. A sua implementação é feita de maneira progressiva, sendo utilizada em 57 municípios nas eleições de 1996, em dois terços do eleitorado nacional em 1998, e em todo o território nacional a partir do ano 2000. (OLIVEIRA, 2021)

A Figura 13, extraída de (FERRÃO et al., 2019) sumariza alguns pontos da evolução histórica dos modelos de urna eletrônica brasileira até 2018. Na sua mais recente versão, a urna conta com um processador Intel Atom e a memória RAM foi ampliada para 500 MB (OLIVEIRA, 2021).

	Até 1998	Até 2004	Até 2018
Voto em Papel	<p>1985: Início do cadastro único e informatizado dos eleitores.</p> <p>1989: 1ª votação através de um computador (Brusque/SC).</p> <p>1994: Primeiro resultado das eleições via recursos computacionais da própria Justiça Eleitoral.</p> <p>1995: Apresentação de alguns dos primeiros protótipos da urna eletrônica (TRE/MG).</p>	<p>dos mesários, que as levavam ao TRE local ao final da votação. O registro de voto impresso foi removido da urna. O poder de processamento foi melhorado, tornando possível a inclusão de fotos de todos os candidatos. Cerca de 57% dos votantes utilizaram as urnas.</p> <p>2000: As urnas começaram a ser instaladas previamente no local de votação. Pela primeira vez, todos os eleitores votaram eletronicamente. Houve avanço quanto à acessibilidade, com a inclusão de áudio e a justificativa eleitoral através da própria urna.</p>	<p>A Ordem dos Advogados do Brasil, o Ministério Público e os representantes dos partidos políticos puderam participar da especificação e desenvolvimento dos programas utilizados na urna.</p> <p>2006: Inserção do leitor biométrico nas urnas.</p> <p>2008: A biometria foi utilizada pela primeira vez nas votações.</p>
Voto Eletrônico	<p>1996: Início da implantação das urnas eletrônicas no Brasil. As urnas possuíam um módulo de impressão externo (MIE) de voto para depositar na urna física e, devido ao baixo poder de processamento e capacidade de armazenamento, só mostravam as fotos dos principais cargos políticos. Cerca de 32% do total de votantes utilizaram as urnas.</p> <p>1998: Nas primeiras eleições, as urnas eram entregues nas casas</p>	<p>2002: O módulo de impressão externo foi novamente acoplado às urnas. Passou-se a utilizar o Sistema Operacional (SO) Windows CE, substituindo o SO VirtuOS.</p> <p>2004: O MIE foi descontinuado.</p>	<p>2009: Adição de <i>smart card</i> e <i>display</i> no terminal do mesário. Utilização do SO Linux nas urnas.</p> <p>2011 e 2013: Atualização das urnas com sensores biométricos de maior qualidade e um botão para ligar/desligar a urna, que antes era realizado por chave física.</p> <p>2018: Volta do MIE em 6% das urnas (escolhidas de forma aleatória). O Aplicativo e-Título (título de eleitor digital) oficial da Justiça Eleitoral foi disponibilizado.</p>

Figura 13 – Evolução das Urnas Eletrônicas até 2018.

Fonte: FERRÃO et al. (2019).

A preparação das UEs segue com rigor uma série de complexos processos com várias camadas de segurança. Além disso, são feitas auditorias das UEs antes e após o processo de votação - detalhados na Seção 3.4.3.

O processo que ocorre no dia da votação segue as seguintes etapas (adaptado de

(SENADO, 2018)):

1. A UE, que faz o registro dos votos dos eleitores, também apura os resultados da seção eleitoral, depois que a votação se encerra;
2. Quando termina a votação, a máquina emite o **Boletim de Urna (BU)**, um relatório impresso que contém o registro de todos os votos digitados nela. Cópias do BU são afixadas na seção eleitoral e enviadas a cartórios, para conferência dos interessados, e distribuídas aos partidos;
3. Esses dados de votação, que ficam gravados em um cartão de memória dentro de cada urna, também são gravados, criptografados, em uma **Mídia de Resultado (MR)** acoplada à máquina - uma espécie de pen drive exclusivo da Justiça Eleitoral, com assinatura digital;
4. O mesário presidente da seção eleitoral, após a votação, rompe o lacre da urna e retira a MR, entregando-a, em uma embalagem lacrada, a integrantes da junta eleitoral;
5. O dispositivo é encaminhado ao TRE do estado ou a um local próprio para transmissão dos dados ao *data center* do TRE. A transmissão é feita por uma rede de comunicação privada da JE. Em localidades de difícil acesso, como aldeias indígenas e comunidades ribeirinhas, a transmissão é por satélite;
6. Nos TREs, os dados passam por uma série de certificações de segurança antes de serem totalizados;
7. Primeiro, é verificada a assinatura digital, para conferir se aquele resultado foi mesmo gerado pela UE daquela seção eleitoral. Após a verificação, os dados são decifrados;
8. Somando os dados de todas as urnas, os TREs totalizam os votos e divulgam os resultados para deputado estadual e distrital, deputado federal, senador e governador;
9. Caso seja uma eleição presidencial, os dados são enviados ao TSE para totalização.

A transmissão dos votos armazenados nas MRs para o TRE local é feita através de uma conexão VPN (*Virtual Private Network*) configurada com uma mídia própria contendo o software *JE-Connect*, que permite que computadores locais sejam utilizados como transmissores dos resultados da votação da sessão eleitoral de maneira segura.

O modelo utilizado nas eleições de 2022 contaram com uma camada adicional de segurança: a certificação ICP-Brasil, que garante a execução na Urna somente de sistemas assinados pela Corte (POVO, 2021).

3.4.3 Auditoria e Testes Públicos de Segurança (TPS)

O processo de Auditoria de Funcionamento das Urnas Eletrônicas sob Condições Normais de Uso - nome dado ao processo de auditoragem das UEs que ocorre em todos os estados do Brasil - segue as seguintes etapas ((TSE, 2020)):

- **30 dias antes das eleições:** em cada TRE, é nomeada uma Comissão de Auditoria de Funcionamento das Urnas Eletrônicas. Essa Comissão é composta por um juiz de direito (que preside a Comissão) e, no mínimo, seis servidores da JE (sendo ao menos um da Corregedoria Regional Eleitoral, um da Secretaria Judiciária e um da Secretaria de Tecnologia da Informação). Um representante do Ministério Público é indicado pelo Procurador Regional Eleitoral para acompanhar os trabalhos da Comissão e, além disso, podem ser indicados para o mesmo fim representantes de partidos políticos, coligações, Ordem dos Advogados do Brasil (OAB), Congresso Nacional, STF, Corregedoria Geral da União (CGU), Polícia Federal (PF), Sociedade Brasileira de Computação (SBC), Conselho Federal de Engenharia e Agronomia e departamentos de Tecnologia da Informação (TI) das Universidades;
- **20 dias antes das eleições:** TREs divulgam onde será realizada a Auditoria e expede ofícios aos partidos políticos comunicando sobre horário e local onde será realizado sorteio das urnas a serem auditadas na véspera do pleito;
- **No dia anterior às eleições:** são sorteadas em **cerimônia pública** seções eleitorais de todo o país. As urnas sorteadas são retiradas, então, das suas sessões eleitorais e postas em uma sala vigiada pelos TREs. As urnas retiradas são substituídas por novas. São preenchidas cédulas de votação pelos representantes de partidos políticos e coligações e depositadas em uma urna lacrada;
- **No dia das eleições:** em cerimônia pública - transmitida até pelo YouTube por alguns TREs - todos os votos das cédulas preenchidas anteriormente são computados nas urnas e, ao final da votação e geração do BU, os dados emitidos pela urna são comparados com aqueles contidos nas células de votação. Além de todo esse processo de auditoria, o TSE também contrata uma empresa de auditoria terceirizada para acompanhar o procedimento em todo o país.

Além do processo de auditoria, desde 2009 - tornando-se obrigatório a partir de 2015 - são realizados os Testes Públicos de Segurança (TPS), que são testes organizados pelo TSE e dos quais podem participar **qualquer cidadão brasileiro acima de 18 anos** através inscrição em Edital próprio. Em 2021 ocorreu a sexta edição do TPS (TSE, 2019).

Nos TPS, partes do software das UEs são abertos para que possam ser testados exaustivamente por especialistas ou interessados em constatar a lisura dos programas utilizados nas urnas. Na edição de 2021 - ocorrida em novembro deste ano - foram executados 29 planos de ataques e foram identificadas cinco falhas (que não teriam o potencial de prejudicar os resultados das eleições) e que serão corrigidas antes do próximo pleito (D'AGOSTINO, 2021).

Nos últimos anos, além das eleições presidenciais de 2022, dois acontecimentos em específico deslançaram um grande fluxo de informações sobre as UEs, contando inclusive com campanhas desinformativas por parte de diversas autoridades: A PEC do Voto Impresso, em 2021, e a reunião do ex-presidente Jair Bolsonaro com embaixadores, em 2022. Esta última, inclusive, levou à inelegibilidade de oito anos do ex-presidente em decisão recente do TSE. Essas situações serão detalhadas a seguir.

3.4.4 O Voto Impresso e a Reunião com Embaixadores

A PEC 135/2019, proposta pela deputada Bia Kicis (PSL-DF) foi tramitada e rejeitada na Câmara dos Deputados no dia 10 de agosto de 2021 (CRISTALDO, 2021). A PEC propunha a alteração do Artigo 14 da CF, adicionando:

“§ 12 - No processo de votação e apuração das eleições, dos plebiscitos e dos referendos, independentemente do meio empregado para o registro do voto, é obrigatória a expedição de cédulas físicas conferíveis pelo eleitor, a serem depositadas, de forma automática e sem contato manual, em urnas indevassáveis, para fins de auditoria.”

O voto impresso já havia sido aprovado na mini-reforma eleitoral de 2015 e, posteriormente, foi declarada inconstitucional a impressão do voto digital em 2018 pelo STF (WESTIN, 2021). A pauta, amplamente defendida pela base aliada ao governo do ex-presidente Jair Bolsonaro, voltou aos holofotes naquele ano e contou com maciças campanhas por parte de deputados favoráveis à causa – dentre eles a própria deputada Bia Kicis, as deputadas Carla Zambelli e Caroline de Toni e o relator da PEC Filipe Barros.

A Figura 14 ilustra como seria a implementação deste voto impresso, caso a PEC fosse aprovada.

Cabe destacar que, em 2002 (no segundo pleito que teve as urnas eletrônicas como sistema único de votação), após o Módulo de Impressão Externo (MIE) ter sido acoplado às urnas eletrônicas novamente, foi feito um teste com a totalidade de urnas do Distrito Federal e do estado do Sergipe e parte das urnas em outros estados (6,18% do eleitorado brasileiro) utilizando um sistema de voto impresso, que depositava de maneira automática o voto em uma urna lacrada após confirmação explícita do eleitor.

Esse teste levou a uma série de problemas, tanto pelo desconhecimento do eleitor da necessidade da confirmação do voto impresso (situação observada em cerca de



Figura 14 – Etapas do voto impresso proposto pela PEC 135/2019.
Fonte: SATIE (2021).

60% das urnas que utilizaram o sistema no Rio de Janeiro), como por problemas mecânicos apresentados pelas urnas ou pelo MIE. No primeiro turno no Distrito Federal, o índice de quebra de urna foi de 5,3% (contra 1,41% da média nacional) e a média de seções que teve que passar para votação manual em decorrência dessas quebras foi de mais de 1% (contra 0,2% da média nacional) (TSE, 2017). Para efeito de comparação, no primeiro turno do pleito de 2020 foi necessária a troca de 3381 urnas, o que corresponde a 0,75% do contingente usado em nível nacional (DEPUTADOS, 2020).

No dia 18 de julho de 2022, o então presidente Jair Bolsonaro, apontado como derrotado nas pesquisas de intenção de voto para a corrida presidencial naquele momento, convocou embaixadores de todos os países para colocar em cheque a segurança do sistema eletrônico de votação e a imparcialidade das cortes superiores do país no julgamento de casos que o envolviam e envolviam o processo eleitoral (UOL, 2022).

As falas do ex-presidente usadas neste ataque foram falsas ou tendenciosas. O TSE emitiu uma nota após a reunião desmentindo 20 alegações feitas, dentre elas de que um hacker que invadiu o sistema do TSE teria liberdade de mexer no código-fonte das UEs, de que o Brasil, Butão e Bangladesh são os únicos países do mundo com o modelo de UEs utilizado no nosso pleito, entre outros (BETIM, 2022).

O vídeo da transmissão, feita na TV Brasil, permaneceu no ar no YouTube, Facebook e Instagram por mais de um mês, tendo sido retirado de circulação apenas em 23 de agosto, após decisão do TSE (RBA, 2022).

Estes acontecimentos, que não ocorreram de maneira isolada, fizeram parte de uma série de conteúdos desinformativos sobre as UEs em circulação e que foram encontrados reiteradamente nos vídeos analisados no presente trabalho.

3.4.5 A Lista de Links do TSE e a Atuação do Tribunal Durante as Eleições de 2022

Em 2020, o TSE lançou o Programa de Enfrentamento à Desinformação, que teve como objetivo o combate à desinformação relacionada ao processo eleitoral. Qualquer pessoa com conhecimento de mensagem, publicação ou notícia contendo desinformação sobre o processo eleitoral poderia encaminhar o conteúdo para o órgão, por meio da ouvidoria, e-mail ou qualquer outro canal de comunicação da instituição. Em junho deste ano o TSE lançou também o Sistema de Alerta de Desinformação Contra as Eleições. O sistema permite registrar casos de circulação ou disparo em massa e de “narrativas falsas que atentem contra a imagem de segurança e integridade do processo de votação e que possam afetar a normalidade do pleito eleitoral.” (SABENDO, 2022).

Em julho de 2022, a Fiquem Sabendo, agência independente especializada no acesso de dados através da Lei de Acesso à Informação, solicitou ao TSE uma série temporal com os *links* recebidos pelo Tribunal e determinados como potencial conteúdo desinformativo acerca do processo eleitoral. A tabela enviada pelo Tribunal conta com 1432 links de conteúdos do YouTube, 56 do Facebook, 25 do Instagram, 21 do Twitter, seis do Telegram e 39 de outras plataformas e pode ser acessada em <https://shorturl.at/fruNV> (acessado em 19/06/2023).

Essa lista de *links* serviu como base para a criação do Protocolo de Classificação do presente trabalho, explicado no Capítulo 5.

Tendo em vista, portanto, o imenso fluxo informacional acerca do processo eleitoral nas Eleições de 2022 e a decorrente leva de conteúdos desinformativos sobre o tema, potencializado muitas vezes pelos governantes no poder, o TSE adotou uma postura extraordinária para lidar com esta situação, aprovando em colegiado no dia 20 de outubro, a Resolução 23.714, que determina, dentre outros ELEITORAL (2022b):

(...)

Art. 2º É vedada, nos termos do Código Eleitoral, a divulgação ou compartilhamento de fatos sabidamente inverídicos ou gravemente descontextualizados que atinjam a integridade do processo eleitoral, inclusive os processos de votação, apuração e totalização de votos.

§ 1º Verificada a hipótese prevista no caput, o Tribunal Superior Eleitoral, em decisão fundamentada, determinará às plataformas a imediata remoção da URL, URI ou URN, sob pena de multa de R\$ 100.000,00 (cem mil reais) a R\$ 150.000,00 (cem e cinquenta mil reais) por hora

de descumprimento, a contar do término da segunda hora após o recebimento da notificação.

(...)

Art. 4º A produção sistemática de desinformação, caracterizada pela publicação contumaz de informações falsas ou descontextualizadas sobre o processo eleitoral, autoriza a determinação de suspensão temporária de perfis, contas ou canais mantidos em mídias sociais, observados, quanto aos requisitos, prazos e consequências, o disposto no art. 2º.

(...)

Art. 7º O disposto nesta Resolução não exclui a apuração da responsabilidade penal, do abuso de poder e do uso indevido dos meios de comunicação.

Fica evidente, portanto, a necessidade de apuração do papel e responsabilidade dos SRs na propagação e recomendação destes conteúdos desinformativos, principalmente durante o atípico período eleitoral ocorrido no ano passado.

3.5 Os Últimos Desdobramentos Acerca do Assunto - Pós-Eleições

Após as Eleições de 2022, finalizadas em 30 de outubro com a vitória em segundo turno do candidato Luís Inácio Lula da Silva (50,9% dos votos válidos) contra o candidato à reeleição Jair Messias Bolsonaro (49,1% dos votos válidos), aconteceram vários acontecimentos de interesse para este trabalho. Serão destacados brevemente aqui alguns deles, dada sua relação direta com as UEs e o fluxo informacional decorrente de todo o processo eleitoral do ano passado.

Diversas manifestações anti-democráticas aconteceram na frente de quartéis, com bloqueios de estradas e atos terroristas durante todo o mês de novembro e dezembro, com criminosos portando faixas pedindo intervenção militar e solicitando o código-fonte das UEs (disponível para consulta durante todo o ano antes das eleições e reiteradamente auditado através do TPS e de outras análises, realizadas extraordinariamente inclusive pelo próprio Exército), motivados em grande parte pela desconfiança relativa ao sistema impulsionada diversas vezes pelo ex-presidente (G1, 2022).

Ainda no mês de dezembro, bolsonaristas tentaram invadir a sede da Polícia Federal em Brasília no dia da diplomação do presidente eleito, em mais uma clara demonstração anti-democrática feita através de atos terroristas (UESLEI MARCELINO, 2022). No mesmo mês, após uma forte campanha virtual encabeçada pela *Sleeping Giants Brasil* (uma organização sem fins lucrativos que busca desmonetizar conteúdos desinformativos na internet) fez com que a Jovem Pan arquivasse pelo menos 376 vídeos

de conteúdos tendenciosos ou desinformativos sobre as UEs e a COVID-19 que permaneciam em circulação no canal da emissora (AMADO, 2022). Esta leva de vídeos, inclusive, recebeu uma classificação específica na Metodologia do presente trabalho (ver Capítulo 5).



Figura 15 – Invasão à Praça dos Três Poderes ocorrida em 8 de janeiro de 2023.
Fonte: FERNANDES (2023).

O maior ato terrorista, entretanto, ocorreu no dia 8 de janeiro do corrente ano. Bolsonaristas inconformados com a posse do presidente Lula invadiram a Praça dos Três Poderes, vandalizando e destruindo o Palácio do Planalto, o Supremo Tribunal Federal e o Congresso Nacional (ESTADÃO, 2023), como é possível observar na Figura 15. É possível traçar como uma raiz de toda esta problemática a reiterada desconfiança no sistema eleitoral e a massiva campanha de desinformação acerca do assunto, impulsionado em parte pelas próprias mídias sociais e os SRs, como o presente trabalho busca demonstrar. Os resultados encontrados encontram-se no Capítulo 6.

3.5.1 O Projeto de Lei 2630/2020 - PL das Fake News

Por fim, também como desdobramento dos fatos ocorridos durante o ano eleitoral de 2022 e a invasão do dia oito de janeiro, o governo eleito colocou como prioridade a votação do Projeto de Lei (PL) 2630/2020, do Senador Alessandro Vieira, que visa impor uma série de regulações e responsabilidades às *Big Techs* e o conteúdo que circula nas mídias sociais às quais são proprietárias (o Google com o YouTube, a Meta com o Facebook, o WhatsApp e o Instagram, e o Telegram, seu aplicativo de mensagens, etc.).

Apelidado de “PL das *Fake News*”, o Projeto cria a Lei Brasileira de Liberdade,

Responsabilidade e Transparência na Internet. O ponto principal é tornar obrigatória a regulamentação do conteúdo que circula nas plataformas digitais para que contas ou publicações com conteúdos considerados criminosos possam ser identificadas, excluídas ou sinalizadas (HENRIQUE, 2023). Essa nova regra vai de encontro com o primeiro entendimento (já ultrapassado) do Marco Civil da Internet, que não responsabiliza as empresas pelo conteúdo publicado nas redes (ver Seção 2.3.1).

Os principais pontos que cabem destaque em relação a este PL no contexto do presente trabalho são (FEDERAL, 2020):

1. Definição de *Fake News*: busca definir o que seriam as *Fake News*, conceituando-as como informações falsas ou enganosas, divulgadas de forma deliberada ou por negligência, com potencial de causar danos individuais ou coletivos, ou prejudicar processos eleitorais, a saúde pública, a segurança nacional ou a democracia;
2. Responsabilidade: busca impor às plataformas digitais, como redes sociais e aplicativos de mensagens, a responsabilidade pela identificação e remoção de conteúdos falsos. Elas devem adotar medidas para coibir a divulgação de desinformação e disponibilizar mecanismos para que os usuários possam denunciar conteúdos suspeitos;
3. Transparência: prevê a obrigatoriedade das plataformas digitais divulgarem informações sobre a identidade dos produtores de conteúdo patrocinado, além de tornarem transparentes os critérios de funcionamento de seus SRs;
4. Banco de dados de usuários: prevê a criação de um banco de dados com informações de usuários de redes sociais e aplicativos de mensagens, para fins de investigação criminal, dada a devida autorização judicial;
5. Proteção da privacidade e liberdade de expressão: Traz dispositivos para garantir a proteção da privacidade dos usuários e a preservação da liberdade de expressão, buscando evitar que ocorra censura ou violações dos direitos individuais.

No momento em que este trabalho é escrito, o PL – que havia sido colocada em pauta previamente e seria votado na Câmara dos Deputados – teve sua votação adiada, contando inclusive com campanhas irregulares e forte pressão das próprias *Big Techs* para barrar a criação da lei e chegando ao ponto de ser instalado inquérito para investigação do Google no STF após ser veiculada na página principal da plataforma, um artigo contrário à aprovação do PL (ROCHA, 2023).

4 TRABALHOS RELACIONADOS

Os Trabalhos Relacionados (TRs) que auxiliaram na criação da metodologia desenvolvida neste trabalho são divididos em duas categorias:

1. Trabalhos relacionados com a **proposta do trabalho**: trabalhos que atestam a criação de filtros bolha através dos SRs e seu papel no consumo de conteúdos desinformativos em mídias sociais e plataformas de compra;
2. Trabalhos relacionados com o **objeto de pesquisa**: trabalhos que analisam o discurso desinformativo acerca das UEs em plataformas como Twitter, Facebook e YouTube.

Assim, os trabalhos de interesse para essa dissertação de mestrado serão discutidos em duas seções distintas, de acordo com as duas categorias apresentadas acima.

4.1 Sistemas de Recomendação e Filtros Bolha de Desinformação em Mídias Sociais

Em 2017, seguindo as eleições estado-unidenses que colocaram Donald Trump no poder e no referendo que culminou com a saída do Reino Unido da União Europeia (Brexit), as mídias sociais passaram por um escrutínio ainda maior como possíveis responsáveis pela manipulação da opinião popular, sumariamente através de seus SRs e conteúdos hiper-direcionados.

Em DIFRANZO; GLORIA-GARCIA (2017), são apresentados estudos iniciais que apontam o papel do Facebook na criação de filtros bolha que reforçam ideologias políticas e visões de mundo pré-concebidas a um usuário, além de reforçar também o consumo de conteúdos conspiratórios sem possibilidade de verificação, com destaque para TUFKCI (2015) que refuta um estudo de 2015 feito pelo próprio Facebook diminuindo a influência que a plataforma teria na criação de filtros bolha e DEL VICARIO et al. (2016) que aponta a criação de duas comunidades completamente distintas e

segregadas dentro do ambiente do Facebook: uma que consome conteúdos científicos e outra que consome conteúdos conspiratórios, apontando a homogeneidade de conteúdo como o maior responsável por sua difusão – ou seja, os usuários de uma comunidade não consomem o conteúdo da outra. Nos anos seguintes, os estudos acerca deste fenômeno se intensificaram, colocando no centro da discussão o papel que os SRs desempenhavam no consumo de conteúdos desinformativos e criação de filtros bolha, especialmente em mídias sociais.

Em particular, os estudos de **auditação de algoritmos** (*algorithm audit*) lançaram luz sobre o funcionamento dos SRs e demais sistemas de caixa-preta (sem código aberto) utilizados em plataformas sociais proprietárias (Facebook, YouTube, Twitter, etc.) e as dinâmicas problemáticas que são desenvolvidas a partir destes. Apesar do termo “auditação” (ou “auditoria”) remeter ao processo de investigação para fins de responsabilização fiscal/econômica, os primeiros estudos de auditação em sentido amplo foram desenvolvidos por economistas do governo norte-americano para detectar discriminação racial em programas de habitação (SANDVIG et al., 2014). Este método, quando aplicado ao contexto computacional, é definido por RAJI; BUOLAMWINI (2019) da seguinte maneira:

Uma auditação de algoritmos envolve a coleta e análise de resultados de um algoritmo fixo ou modelo definido dentro de um sistema. Através da estimulação de uma população simulada de usuários, essas auditorias podem revelar padrões problemáticos em modelos de interesse.

BANDY (2021) faz uma revisão sistêmica dos trabalhos que utilizam auditação de algoritmos em sua metodologia, revisando 62 trabalhos e classificando as técnicas empregadas em cada um seguindo a definição proposta por SANDVIG et al. (2014):

- Auditação de código – Pesquisadores obtêm e analisam o código-fonte do algoritmo;
- Raspagem direta – Pesquisadores coletam dados diretamente através da API da plataforma ou através de outra consulta sistematizada;
- Contas de pesquisa ou “fantoques” – Pesquisadores coletam dados criando programas de computador que simulam usuários que então testam o algoritmo;
- Contas de interação ou “fantoques de transporte” – Similar ao método de fantoques, mas neste caso os usuários simulados afetam o sistema real e podem implicar efeitos aos usuários da plataforma;
- Contribuição colaborativa – Pesquisadores coletam dados contratando usuários da plataforma para testar o algoritmo.

A partir da revisão dos artigos foram propostas quatro categorias de problemas apontados pelos estudos de auditoria propostos por BANDY (2021) :

- Discriminação – O algoritmo trata de maneira desigual os usuários baseando-se em sua raça, idade, gênero, localização, *status* socioeconômico, etc.
- Distorção – O algoritmo apresenta mídia que distorce ou obscurece uma realidade subjacente, favorecendo uma ideologia política, hiper-personalizando recomendações ou até apresentando **conteúdo desinformativo** aos usuários;
- Exploração – O algoritmo utiliza de maneira inapropriada conteúdos de outras fontes e/ou informações sensíveis de usuários (muitas vezes sem o consentimento adequado);
- Erro de julgamento – O algoritmo faz classificações ou previsões incorretas, podendo levar aos outros três tipos de problemas.

A Tabela 5 totaliza o número de estudos revisados por BANDY (2021), dividindo-os por categoria e domínio. Os domínios encontrados aqui são: Busca (sistemas de pesquisa, como Google e Bing), Propaganda (sistemas de direcionamento de anúncios, como Google Ads), Recomendação (sistemas de recomendação de conteúdo, como os utilizados no YouTube e Facebook), Precificação (sistemas utilizados em comércios virtuais, como na Amazon), Visão (sistemas de identificação visual e processamento de imagens), Justiça Criminal (sistemas utilizados para previsão de reincidência criminal), Processamento de Linguagem (sistemas que processam linguagem natural ou fazem previsão de texto, como o Google Snippet) e Mapeamento (sistemas de mapas/localizações, como o Google Maps).

Observa-se que foram revisados sete trabalhos no domínio de Recomendação e na categoria de Distorção, sendo este cruzamento o foco do debate e da metodologia presentes nesta Dissertação.

Como este trabalho visa auditar o impacto dos SRs na criação de filtros bolha de desinformação acerca das UEs no YouTube, a metodologia mais adequada para este fim é o de contas de pesquisa, que permitem a simulação de usuários com diferentes perfis (neste caso, históricos de consumo de vídeos) e a análise de que maneiras estes perfis geram diferentes resultados de pesquisa e recomendações.

Tendo isso em vista, dois trabalhos relacionados dos sete envolvendo Sistemas de Recomendação no contexto de Distorção serviram como principais referências para o autor: HUSSEIN; JUNEJA; MITRA (2020) e JUNEJA; MITRA (2021). A partir destes trabalhos, foram encontrados outros quatro que utilizam o método de contas de pesquisa e são voltados para a análise dos SRs e filtros bolha de desinformação,

Tabela 5 – Estudos analisados na revisão sistêmica de BANDY (2021) separados por domínio e categoria.

Domínio	Discriminação	Distorção	Exploração	Erro de Julgamento	Total
Busca	5	18	2	-	25
Propaganda	3	2	3	4	12
Recomendação	1	7	-	-	8
Precificação	5	-	-	-	5
Visão	5	-	-	-	5
Justiça Criminal	1	-	-	3	4
Processamento de Linguagem	1	1	-	-	2
Mapeamento	-	1	-	-	1
Total	21	29	5	7	62

Tabela 6 – Trabalhos relacionados mais relevantes para o desenvolvimento da metodologia utilizada nesta Dissertação, sendo: TR1 – HUSSEIN; JUNEJA; MITRA (2020), TR2 – TOMLEIN et al. (2021), TR3 – PAPADAMOU et al. (2020), TR4 – ABUL-FOTTOUH; SONG; GRUZD (2020), TR5 – HUSSEIN; ELDARDIRY (2020), TR6 – JUNEJA; MITRA (2021)

Trabalho	Plataforma	Tema	Tipo de Conteúdo	Conteúdo Desinformativo (%)
TR1	YouTube	Teorias Conspiratórias	Audiovisual	15,36%
TR2	YouTube	Teorias Conspiratórias	Audiovisual	8,37%
TR3	YouTube	Conteúdo Pseudocientífico	Audiovisual	23,1%
TR4	YouTube	Vacinação	Audiovisual	19,98%
TR5	Amazon	Vacinação	Textual	5,54%
TR6	Amazon	Vacinação	Textual	10,47%

detalhados abaixo e expostos na Tabela 6. Estes trabalhos serão usados para comparação dos resultados dessa Dissertação (ver Capítulo 6) e discutidos em detalhes nas próximas seções.

O desenvolvimento da metodologia desta Dissertação (detalhada no Capítulo 4) leva em conta os diversos detalhes metodológicos implementados em cada um dos trabalhos relacionados expostos acima, que permitiram ao autor a construção de uma abordagem própria para a auditoria do SR utilizado no YouTube para os vídeos apresentados em sua pesquisa e recomendações, tendo como objeto de estudo a desinformação sobre as UEs.

4.1.1 Trabalho Relacionado 1 - HUSSEIN; JUNEJA; MITRA (2020)

O Trabalho Relacionado 1 (TR1) é apresentado em (HUSSEIN; JUNEJA; MITRA, 2020). Nesse trabalho são utilizados cinco tópicos que embasam uma significativa quantidade de teorias conspiratórias, para que seja feita a análise de conteúdo desinformativo em vídeos do YouTube sobre estes: o atentado de 11 de setembro, teoria da

Terra plana, chegada do homem à lua, movimento anti-vacina e teoria dos *chemtrails* (não muito conhecida no Brasil, essa teoria implica que os rastros deixados por aviões são, na verdade, agentes químicos ou biológicos utilizados para causar dano na saúde da população).

São criados, então, vários perfis com diferentes parâmetros (gênero, idade e localização) e extraídos os vídeos de pesquisa, relacionados e da página inicial de cada um dos perfis no YouTube. A partir da extração desses vídeos, são feitas classificações quanto a presença de conteúdo desinformativo nos vídeos resultantes e é feita uma análise da influência de cada um dos parâmetros no aumento ou redução da quantidade de conteúdo desinformativo. Após esta análise, é feito um outro experimento sobre a influência do **histórico de visualização** em perfis que consomem vídeos de cada uma das categorias estabelecidas.

Para a definição dos termos de pesquisa a serem utilizados na busca por vídeos destes tópicos, foi utilizado a plataforma **Google Trends**, comparando a relevância de cada um dos possíveis termos de pesquisa especificamente no YouTube e nas pesquisas gerais do Google. O total de termos de pesquisa definidos entre os cinco tópicos foi 49 e o total de vídeos coletados através dos diferentes perfis foi de 56475, sendo destes 2943 vídeos únicos.

Os autores definem, então, um **Protocolo de Codificação** (PC) de sete categorias para os vídeos extraídos e três classificações normalizadas: **Desmente** (Valor normalizado: -1), **Neutro** (Valor normalizado: 0) e **Promove** (Valor normalizado: 1). A definição dessas categorias foi feita após a análise de 25 vídeos selecionados aleatoriamente dentre os extraídos e múltiplas iterações e discussões entre os autores. As categorias, número de vídeos e classificações normalizadas correspondentes se encontram na Tabela 7. Os 79 vídeos que não estão presentes nesta Tabela são vídeos em outra língua, impossíveis de definição com as categorias estabelecidas ou então vídeos que foram removidos. Estes não foram considerados para fins do estudo.

Por fim, o trabalho conclui que a maior influência nos algoritmos de classificação e recomendação do YouTube advém do **histórico de visualização** dos usuários, com um aumento exponencial de conteúdo desinformativo sendo apresentado principalmente nos vídeos retornados diretamente pela **pesquisa** nos perfis que consomem conteúdo aqui categorizado como “Promove”. Ou seja, é comprovada a criação de um filtro de bolha nestes perfis.

4.1.2 Trabalho Relacionado 2 - TOMLEIN et al. (2021)

O Trabalho Relacionado 2 (TR2) é apresentado em TOMLEIN et al. (2021) e é uma versão atualizada do TR1, visando comparar os resultados e mudanças alegadas pelo YouTube em sua política de moderação de conteúdo desinformativo entre o período dos dois trabalhos (meio de 2019 – o mês de coleta do primeiro trabalho não é espe-

Tabela 7 – Classificação dos vídeos em (HUSSEIN; JUNEJA; MITRA, 2020)

Categoria	Número de Vídeos	Classificação Normalizada
Desmente conteúdos desinformativos acerca dos temas pesquisados	430	Desmente (-1)
Neutro e relacionado com os temas pesquisados	238	Neutro (0)
Promove ou busca justificar conteúdos desinformativos sobre os temas pesquisados	374	Promove (1)
Desmente conteúdos desinformativo acerca de outros temas desinformativos	64	Desmente (-1)
Neutro e relacionado com outros temas desinformativos	25	Neutro (0)
Promove ou busca justificar outros temas desinformativos	66	Promove (1)
Não relacionado com os temas pesquisados ou outros temas desinformativos	1667	Neutro (0)

cificado – a março de 2021), além de auditar também como é possível “furar” o filtro bolha desinformativo na plataforma.

Em comparação com o estudo de referência, este novo estudo aponta que apenas o tópico do movimento anti-vacina teve uma piora em quantidade de conteúdo desinformativo tanto nos resultados diretos de pesquisa como nas recomendações (*up-next* e 5 recomendações mais relevantes). Houve uma melhora no tópico da chegada do homem à lua (mais vídeos na categoria Desmente) nas recomendações, mesmo não havendo diferença significativa nos resultados de busca. Os demais tópicos não tiveram diferenças significativas nem na busca e nem nas recomendações.

O estudo atribui a piora nos resultados do movimento anti-vacina à massiva quantidade de desinformação relativa à vacina da COVID-19 e aponta que, mesmo com as mudanças de política de moderação implementadas pelo YouTube no período entre os dois estudos, não houve diferenças significativas na quantidade de desinformação sobre os tópicos em questão.

Por fim, o estudo consegue apontar a melhora (menos desinformação) nas recomendações e nos resultados de pesquisa quando o usuário assiste vídeos da categoria Desmente em todos os cinco tópicos, o que sugere que é possível “furar” o filtro bolha desinformativo através do consumo de vídeos informativos.

4.1.3 Trabalho Relacionado 3 - PAPADAMOU et al. (2020)

O Trabalho Relacionado 3 (TR3) é apresentado em PAPADAMOU et al. (2020), onde foi desenvolvido um algoritmo com técnicas de Inteligência Artificial (IA) para a classificação de vídeos no YouTube com conteúdo pseudocientífico acerca de quatro termos: COVID-19, anti-vacinação, anti-máscara e teoria da Terra plana. A partir deste algoritmo para a classificação dos conteúdos, é feita uma análise da influência do histórico de visualização do usuário na recomendação de mais ou menos conteúdos pseudocientíficos.

Para o treinamento do algoritmo de classificação com *deep learning*, é utilizado um *dataset* de vídeos extraídos através da API do YouTube com a pesquisa direta dos termos e os 10 primeiros vídeos relacionados a cada um dos vídeos retornados na pesquisa, totalizando 6600 vídeos únicos.

A classificação entre “Científico”, “Pseudocientífico” e “Irrelevante” é executada através de *crowdsourcing* na plataforma Appen (APPEN, 2021) por 992 classificadores, sendo 3 classificações por vídeo e utilizado como material de referência um PC desenvolvido pelos autores, levando em conta as particularidades de cada tema e o que é considerado como discurso pseudocientífico nestes. A classificação final é determinada, então, pela escolha feita pela maioria dentre as três classificações e os vídeos nos quais todos divergiram de opinião são descartados. Ao final, foram classificados então 5734 vídeos, conforme exposto na Tabela 8.

Tabela 8 – Classificação dos vídeos em (PAPADAMOU et al., 2020)

Tópico	#Científicos	#Pseudocientíficos	#Irrelevantes
COVID-19	607	368	721
Teoria da Terra plana	162	375	707
Anti-vacinação	363	394	1060
Anti-máscara	65	188	724
Total	1197	1325	3212

O trabalho prossegue, então, utilizando das classificações obtidas para a criação de um classificador próprio (baseado em RNP) e testa sua eficiência comparada com outros três algoritmos de IA: *Support Vector Machine (SVM)*, *Random Forest* e um algoritmo baseado em *Bidirectional Encoder Representations from Transformers (BERT)*.

Os algoritmos são treinados a partir dos seguintes metadados: transcrição do conteúdo do vídeo, título e descrição do vídeo, *tags* definidas pelo autor do vídeo, estatísticas do vídeo (número de visualizações, *likes*, etc.) e os 200 primeiros comentários, definidos pela métrica de relevância do YouTube. Ao final, o classificador baseado em RNP proposto pelos autores apresenta melhores resultados que os demais.

A partir deste classificador, os autores criam três perfis com históricos distintos de visualização: um usuário que consome vídeos científicos, outro que consome vídeos pseudocientíficos e outro que consome ambos. Para a simulação destes perfis (impossibilitada pela API do YouTube) e a extração dos vídeos retornados por cada instância (pesquisa e vídeos relacionados, além dos vídeos da página inicial do usuário), é criado um *script* utilizando **Selenium Bots**.

Ao final do trabalho, chega-se à conclusão de que o **histórico de usuário afeta significativamente nos algoritmos de recomendação** de vídeos do YouTube, sendo mais notável o aumento no número de vídeos pseudocientíficos nos resultados da pesquisa direta daqueles usuários que consomem conteúdo pseudocientífico. Ou seja, comprova-se aqui novamente a criação de um filtro de bolha que leva o usuário que consome conteúdo desinformativo a consumir ainda mais conteúdo desinformativo.

4.1.4 Trabalho Relacionado 4 - ABUL-FOTTOUH; SONG; GRUZD (2020)

O Trabalho Relacionado 4 (TR4), apresentando em ABUL-FOTTOUH; SONG; GRUZD (2020), fez a extração de 2122 vídeos sobre vacinas no YouTube em 9 de abril de 2019 (período pré-pandêmico), utilizando os 250 vídeos apresentados como mais relevantes na pesquisa e os primeiros 50 vídeos recomendados a partir de cada um deles.

Após, foram definidas três categorias para a codificação dos vídeos: pró-vacina, neutro e anti-vacina (desinformativo). Através da técnica de análise de redes soci-

ais, atestou-se que 67,45% dos vídeos recomendados eram pró-vacina, 19,98% anti-vacina e 15,27% neutros, além de ser atestado, também, que vídeos pró-vacina tinham uma tendência maior de recomendação.

Por fim, foi **comprovada a criação de filtros bolha** tanto para vídeos pró-vacina como anti-vacina, apresentando homofilia de conteúdo nas recomendações para ambos os casos através de usuários simulados consumindo cada tipo de conteúdo.

Cabe destacar aqui que uma pesquisa similar dos mesmos autores foi realizada em 2016 (SONG; GRUZD (2017)) e o cenário era o oposto: a vasta maioria dos vídeos (65%) eram anti-vacina, sendo estes massivamente recomendados a usuários no geral. Isso permite a percepção das mudanças de moderação de conteúdos relacionados à vacinação adotadas pelo YouTube em 2019, mas que englobou apenas conteúdos na língua inglesa, negligenciando conteúdos de outros idiomas (FERREIRA, 2020).

4.1.5 Trabalho Relacionado 5 - HUSSEIN; ELDARDIRY (2020)

O Trabalho Relacionado 5 (TR5), apresentado em HUSSEIN; ELDARDIRY (2020) é um estudo do mesmo autor do TR1. Este novo trabalho audita a pesquisa e as recomendações geradas pela Amazon com conteúdos desinformativos sobre vacinação e sua relação com o comportamento dos usuários. São analisados aproximadamente 526 mil itens de resultados de buscas e 182 mil itens recomendados, obtendo-se um total de 8566 itens únicos (não limitados apenas a livros, como estudos anteriores), classificados individualmente como pró-vacina, anti-vacina (desinformativo) ou neutro.

Para a seleção dos termos de pesquisa, o autor aqui utiliza duas fontes distintas: o Google Trends e o *autocomplete* da própria Amazon, utilizando “*vaccine*” e “*vaccination*” e salvando os resultados de frases sugeridas. Ao final são gerados 29 termos de pesquisa, categorizados de acordo com a expectativa que se tem em relação a cada um. Por exemplo: “*vaccine illusion*” é categorizado como possivelmente desinformativo (1) e “*vaccine*” como possivelmente neutro (0).

Após a coleta dos resultados de pesquisa dos termos e recomendações, foram analisados aleatoriamente 100 itens para o desenvolvimento do PC, com os possíveis valores sendo: -1 (pró-vacina), 0 (neutros), 1 (anti-vacina), 2 (não relacionados ao tópico), 3 (itens em outra língua que não inglês) e 4 (itens removidos). Na normalização, os itens da categoria 2 foram considerados neutros (0) e os das categorias 3 e 4 foram descartados, gerando ao final uma escala normalizada de três valores: -1, 0 e 1.

Seguindo uma metodologia similar ao TR1, após a categorização dos itens coletados, é realizada uma comparação da quantidade de itens desinformativos entre os cinco diferentes algoritmos de busca da Amazon (“destaques”, “preço: mais baixo a mais alto”, “preço: mais alto a mais baixo”, “avaliação média dos clientes” e “produtos mais novos”) e os termos de pesquisa utilizados. Além disso, a fim de mensurar a

influência do comportamento de um usuário nos resultados de pesquisa e recomendação, é construído o histórico de uso de usuários simulados através de quatro atividades: “pesquisa por item”, “navegação por item”, “adicionar item à lista de desejos” e “adicionar item ao carrinho”.

Por fim, o estudo conclui:

- Em relação às pesquisas: o **termo de pesquisa sendo utilizado influi diretamente na quantidade de desinformação encontrada** (os termos que haviam sido categorizados como possivelmente desinformativos acabaram apresentando maior quantidade de desinformação, como esperado) e a avaliação de usuários à itens tem um papel crucial influi diretamente em seu ranqueamento nos resultados de pesquisa. Ou seja, **itens desinformativos com melhores avaliações têm uma tendência a ranquear entre os primeiros resultados**, principalmente quando utilizados os algoritmos de pesquisa “destaques” e “avaliação média dos clientes”.
- Em relação ao comportamento do usuário: não houve mudança significativa no ranqueamento dos **resultados de pesquisa** após as atividades de consumo de conteúdo desinformativo, porém **houve alteração significativa (aumento de conteúdos desinformativos) nas recomendações da página inicial do usuário** quando realizava as ações “navegar por item”, “adicionar à lista de desejos” ou “adicionar ao carrinho” em itens categorizados como desinformativos.

4.1.6 Trabalho Relacionado 6 - JUNEJA; MITRA (2021)

O Trabalho Relacionado 6 (TR6) complementa o TR5, no sentido de aplicar a metodologia do trabalho de referência para a auditoria do Sistema de Recomendação e os algoritmos de busca da Amazon após o aumento de desinformação em relação às vacinas com a chegada da vacinação contra a COVID-19. Além disso, a metodologia é ampliada com a simulação de outras ações de usuários não inclusas no primeiro trabalho, para que possa ser analisada a criação do filtro bolha desinformativo a partir de outros padrões de uso da plataforma.

Em comparação com o trabalho anterior, houve um aumento de conteúdo desinformativo na análise inicial das buscas e recomendações: foram encontrados 10,47% de itens promovendo desinformação (contra 5,54% do TR5) e 8,99% que desmentem informações imprecisas (contra 6,36% do TR5). Outra diferença é que os algoritmos de busca que ranquearam melhor os itens desinformativos foram “avaliação média de usuários” e “preço: mais baixo a mais alto”. O aumento aqui é possivelmente justificado pela grande quantidade de desinformação em relação à vacinas contra a COVID-19 que não havia sido contemplada no trabalho anterior.

Em relação à personalização e criação de filtros bolha a partir do comportamento

de uso dos usuários, este estudo atestou uma maior influência para recomendações de conteúdo desinformativo quando um usuário realiza as ações “pesquisar + navegar por item” e “marcar todas as avaliações positivas mais relevantes como positivas” em itens categorizados como desinformativos, maior até que a ação “adicionar item ao carrinho” (que havia sido atestada como uma das mais influentes no TR5).

4.2 Desinformação sobre Urnas Eletrônicas

Os trabalhos aqui detalhados elucidaram uma maior compreensão acerca do fluxo de informações falsas ou distorcidas sobre as UEs na internet como um todo, em plataformas como Facebook, YouTube e Twitter.

Começando em 2020, a Diretoria de Análise de Políticas Públicas da Fundação Getúlio Vargas (DAPP - FGV) tem lançado uma série de Estudos sobre o tema “Digitalização e Esfera pública no Brasil”. Relacionados ao tema abordado aqui, destacam-se três destes trabalhos, elaborados em cooperação com o Tribunal Superior Eleitoral (TSE) e a Embaixada da Alemanha em Brasília: RUEDIGER; GRASSI (2020a), RUEDIGER; GRASSI (2020b) e RUEDIGER et al. (2022). Estes trabalhos foram definidos como Trabalhos Relacionados 7, 8 e 9 no escopo desta dissertação.

4.2.1 Trabalho Relacionado 7 - RUEDIGER; GRASSI (2020a)

No Trabalho Relacionado 7 (TR7), apresentado em (RUEDIGER; GRASSI, 2020a), é feita a análise de um *corpus* de aproximadamente 103 mil postagens entre os anos de 2014 e 2020 no Facebook e no YouTube, contendo *links* para notícias ou *posts* que incitam a desconfiança no sistema eleitoral, entre os que mais circulam em outras plataformas - *sites* da imprensa tradicional, *blogs* e portais de conteúdo partidariado ou não jornalístico.

A metodologia para enquadramento dos conteúdos textuais (não foram analisados imagens ou vídeos) considera a identificação de narrativas e eixos semânticos de “fraude nas urnas”, através de regras definidas em um caderno de referência metodológica próprio desenvolvido pelo autor (RUEDIGER, 2017).

Foram utilizadas a plataforma **CrowdTangle** e a API do YouTube para a extração de dados e metadados relevantes das postagens para fins de análise. A análise se dá, então, em três partes: circulação dos links nos diferentes anos, interação e alcance digitais do conteúdo e, por último, um enfoque específico nas eleições municipais de 2020.

A Figura 16 exhibe a quantidade de postagens com *links* contendo desinformação sobre as UEs no Facebook entre os anos de 2014 e 2020. Cabe destacar aqui que os picos ocorrem nos anos de eleições gerais (2014 e 2018), com ligeiras ascensões entre anos não eleitorais (2015, 2017 e 2019) e crescimento mais expressivo em eleições

municipais (2016 e 2020).

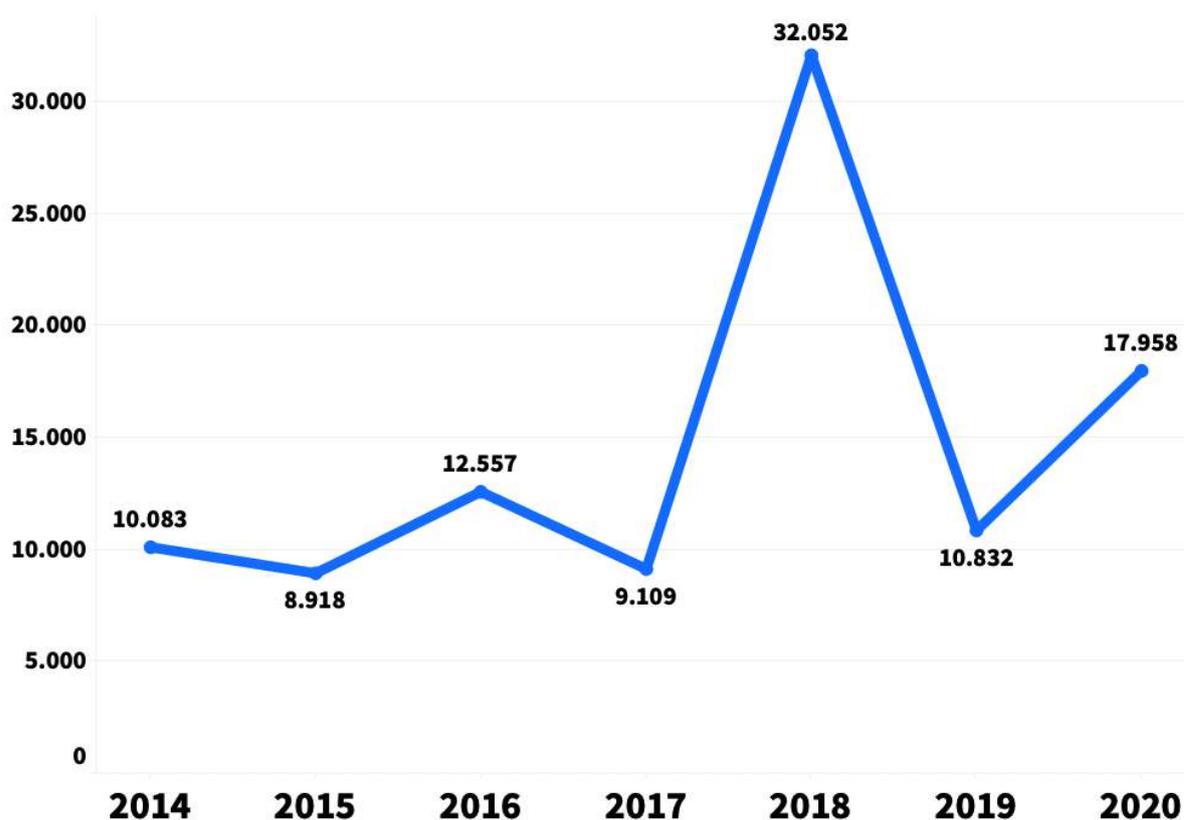


Figura 16 – Número de *posts* com *links* sobre desconfiança eleitoral no Facebook entre 2014 e 2020.

Fonte: RUEDIGER; GRASSI (2020a).

As interações e alcances das postagens apresentam comportamento similar. De maneira geral, **os anos eleitorais concentram picos de interações** a esses conteúdos ao longo dos meses de **setembro e outubro**, quando ocorrem as campanhas e votações. Já os anos não eleitorais exibem padrão de estabilidade em ocorrência e engajamento.

Nas eleições municipais de 2020 houve um menor engajamento nos *links* sobre fraude nas urnas, apesar de uma quantidade maior de *links* em circulação sobre o tema quando comparado com as eleições municipais anteriores. O autor não consegue definir o motivo específico desse menor engajamento, mas sugere uma possível mudança nas métricas de alcance desse tipo de conteúdo nas plataformas sociais.

4.2.2 Trabalho Relacionado 8 - RUEDIGER; GRASSI (2020b)

O Trabalho Relacionado 8 (TR8) é uma sequência do TR7. No TR8, apresentado em (RUEDIGER; GRASSI, 2020b), o autor se debruça especificamente sobre as eleições municipais de 2020 para a análise dos fluxos de informações a respeito da desconfiança nas UEs, fazendo agora a análise de um *corpus* de 1.426.687 *posts* do

Facebook, YouTube e Twitter entre os dias 1 e 30 de novembro de 2020, período que engloba os dois turnos das últimas eleições municipais.

O autor utiliza da mesma metodologia empregada no trabalho anterior e divide agora a análise em duas partes: a evolução das publicações durante o mês eleitoral e as redes de narrativas que amplificam o debate acerca da desconfiança nas urnas nos ambientes digitais.

Nesta análise se destaca a ampla presença da pauta do **voto impresso** nas eleições municipais, fato que era debatido geralmente nas eleições gerais. Isso se deve ao deslocamento do debate das eleições norte-americanas e as alegações de fraudes nas urnas por parte do eleitorado trumpista, replicado no discurso brasileiro acerca do tema. Outro fato de importante destaque é a **invasão do sistema do TSE**, ocorrida em 15 de novembro – data do primeiro turno das eleições (que não influenciou em nenhum dos resultados dos pleitos).

A Figura 17 ilustra a quantidade de *posts* obtidos na primeira parte da análise no Facebook e sua evolução a cada dia do mês de novembro. O autor destaca aqui três períodos com maior circulação deste conteúdo: entre 4 e 10 de novembro, motivado pela controvérsia causada por Trump sobre possível fraude eleitoral nas eleições estado-unidenses (22,8% dos *posts*); entre 15 e 20 de novembro, motivado principalmente pelo primeiro turno das eleições em questão e também a invasão aos sistemas do TSE (44,6% dos *posts*) e entre 28 e 30 de novembro, período das votações do segundo turno (7,1% dos *posts*).

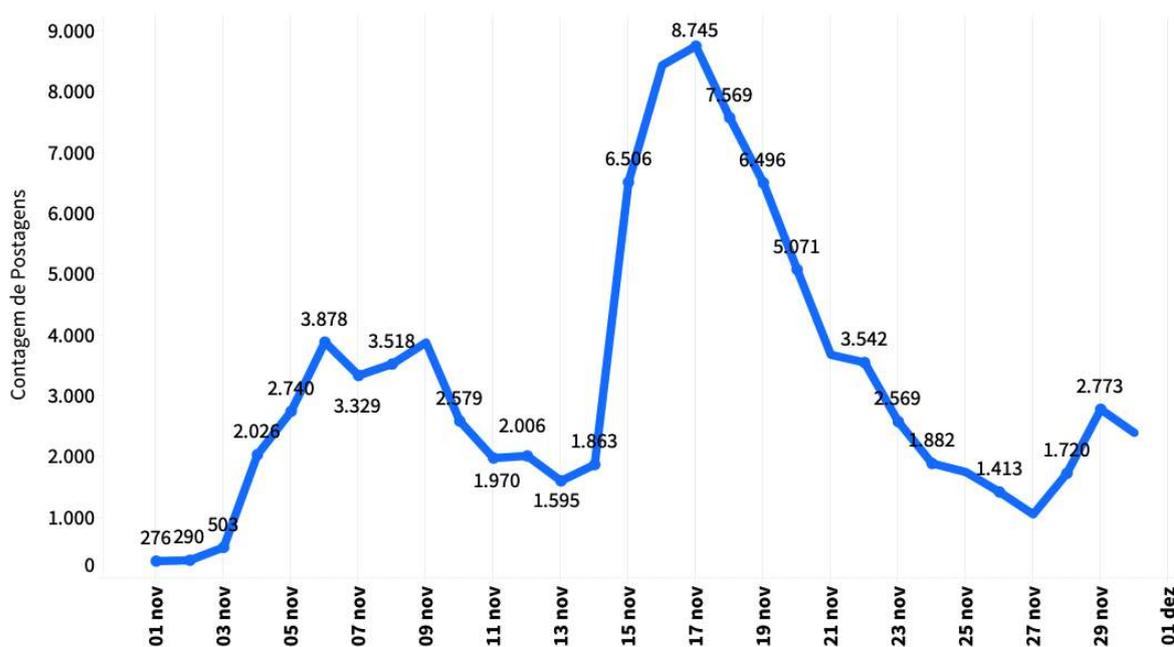


Figura 17 – Evolução no número de *posts* contendo desconfiança acerca do sistema eleitoral no mês de novembro de 2020 no Facebook.

Fonte: RUEDIGER; GRASSI (2020b).

Na segunda parte do trabalho é feita uma modelagem dos tópicos mais presentes no debate sobre desconfiança eleitoral nas três plataformas. A Figura 18 ilustra os termos mais presentes nos *posts* analisados do Facebook. Destacam-se aqui os termos **voto impresso** e **urna eletrônica**, termos de interesse para a presente Dissertação.

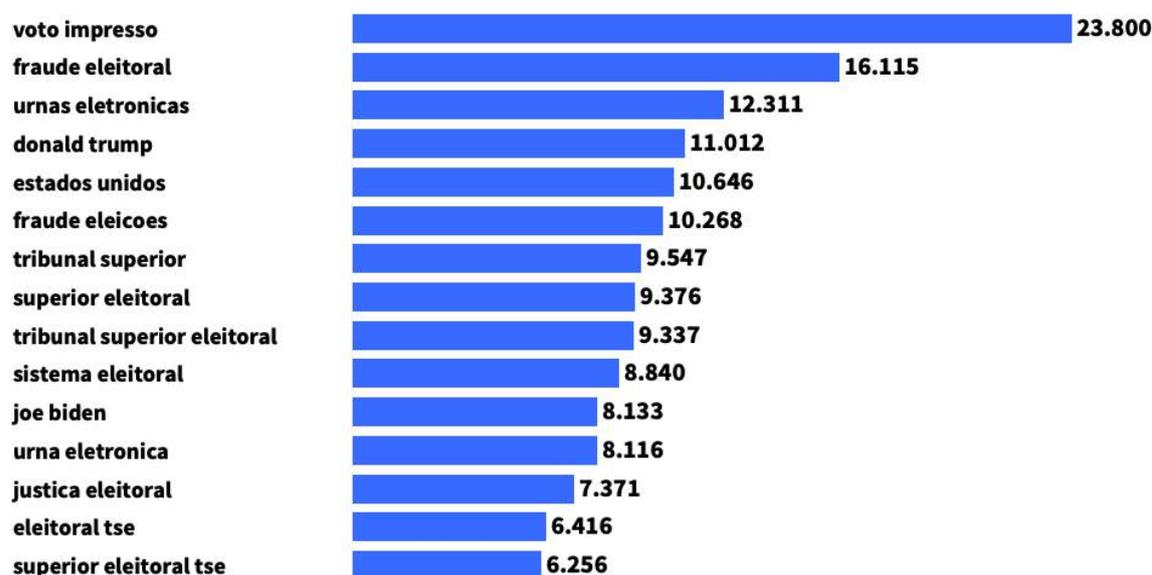


Figura 18 – Termos mais frequentes nos *posts* analisados que incitam desconfiança acerca do sistema eleitoral no Facebook.

Fonte: RUEDIGER; GRASSI (2020b).

4.2.3 Trabalho Relacionado 9 - RUEDIGER et al. (2022)

O Trabalho Relacionado 9 (TR9) apresenta uma última análise do autor do TR7 e do TR8. Este trabalho está apresentado em RUEDIGER et al. (2022) e o foco são 394.370 postagens no Facebook com acusações de fraude na urna eletrônica e defesa do “voto impresso auditável” (termo dado à PEC 135/2019, explicada em detalhe na Seção 3.4.4). A análise foi realizada ao longo de 15 meses: de 2 de novembro de 2020 à 18 de janeiro de 2022.

A análise é dividida em quatro partes: (1) a evolução que estas postagens tiveram no decorrer dos 15 meses, (2) quais contas foram as que geraram mais engajamento sobre o tema, (3) as postagens com maior número de interações e (4) as postagens contendo *links* e para onde estes apontavam. Resumidamente:

1. Durante os 15 meses analisados, houve uma média de 888 postagens diárias. Os meses com maior quantidade de postagens foram novembro de 2020 (eleições municipais – 12,7% do total de postagens), julho e agosto de 2021 (discussão acerca da PEC 135/2019 – 17,7% e 23,9% do total de postagens, respectivamente), sendo 10 de agosto de 2021 (dia da votação da PEC) o dia com o maior

número de postagens (responsável sozinho por 2,7% do total de postagens). A Figura 19 ilustra a evolução das postagens;

2. Do total de 111.748.306 interações nas postagens (compartilhamentos, comentários, curtidas e demais reações), doze contas se destacam com mais de 1 milhão de interações cada, sendo que 11 delas tinham ligação com o governo Bolsonaro ou eram alinhadas à agenda do conservadorismo de direita. A conta com maior número de interações foi da deputada bolsonarista Carla Zambelli. Além disso, 24 contas fizeram individualmente mais de mil postagens sobre o tema;
3. 40 postagens totalizaram 6.809.193 interações (mais de 6% do total), sendo 13 delas (32,5%) feitas pela página do ex-presidente Jair Bolsonaro – totalizando entre elas 3.227.981 interações (47,4% do total de interações dentre estas 40 postagens de maior engajamento) – sendo uma delas o vídeo de uma reportagem de 2009 que já foi considerada “fora de contexto” por agências de checagem;
4. 33,9% do total de postagens analisadas contêm *links*, tendo atraído 23,9% do total de interações. Dos 20 *links* externos (que levam à páginas fora do Facebook) com maior engajamento, o mais popular é a consulta pública no site do Senado sobre a PEC 135/2019 e quatro são *links* de vídeos do YouTube, com comentários de um advogado constitucionalista com críticas aos ministros do STF e defendendo o voto impresso.

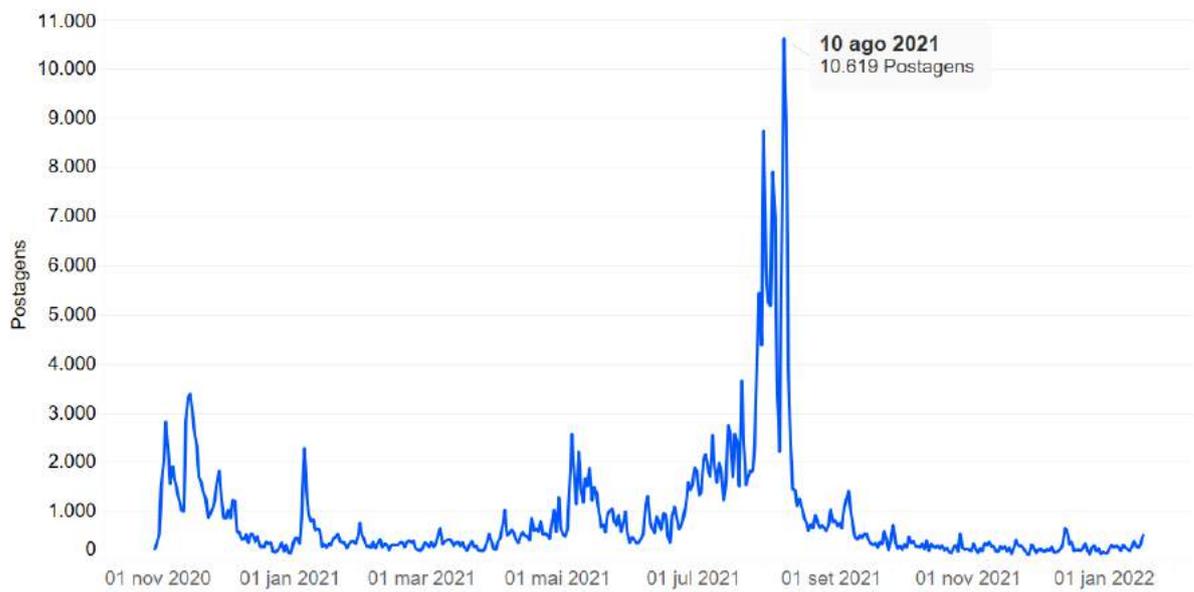


Figura 19 – Evolução das postagens com alegação de fraude nas urnas eletrônicas e defesa do voto impresso auditável no Facebook de novembro de 2020 à janeiro de 2022.

Fonte: RUEDIGER et al. (2022).

5 METODOLOGIA

A metodologia desenvolvida no presente trabalho foi construída a partir da revisão bibliográfica de estudos de auditoria de Sistemas de Recomendação, sejam eles de plataformas de compras (HUSSEIN; ELDARDIRY (2020)) ou plataformas de vídeo (HUSSEIN; JUNEJA; MITRA (2020), PAPADAMOU et al. (2020)). Esta metodologia visa medir a desinformação acerca das UEs a partir dos sistemas de pesquisa e recomendação do YouTube e consiste em quatro partes sequenciais, explicadas brevemente abaixo e detalhadas em Seções específicas na sequência. A Figura 20 ilustra um esquema geral de cada uma das etapas.

- Etapa 0 (Pré-processamento): são definidos cinco termos mais relevantes para o desenvolvimento da pesquisa, são desenvolvidos os *scripts* em Python utilizando Selenium Bots (SELENIUM, 2021) que serão responsáveis pela extração dos vídeos da página de pesquisa e dos vídeos relacionados a cada um destes, além dos *scripts* responsáveis pela construção dos históricos de cada perfil a ser testado na Etapa 3 e, por fim, é feito o processamento da lista de vídeos do TSE (Seção 3.4.5) para validar quais ainda estão no ar;
- Etapa 1 (Extração dos vídeos sem perfil e criação do Protocolo de Classificação): com os artefatos gerados na etapa anterior, o *script* de extração de vídeos é rodado para os cinco termos de pesquisa, utilizando os quatro algoritmos de ordenação do YouTube (detalhados na Seção 5.3) e a partir de uma seleção aleatória contendo vídeos da lista do TSE processada na Etapa 0 e vídeos extraídos nesta Etapa, é criado o Protocolo de Classificação, que é posteriormente validado.
- Etapa 2 (Teste de homogeneidade das recomendações): com as tabelas de vídeos já classificados da Etapa 1, roda-se um *script* de teste que utiliza um vídeo desinformativo como base e seleciona aleatoriamente outros vídeos classificados da mesma maneira para serem assistidos, sendo retornado ao vídeo base a cada novo vídeo assistido e calculado o Coeficiente de *Overlap* (detalhado

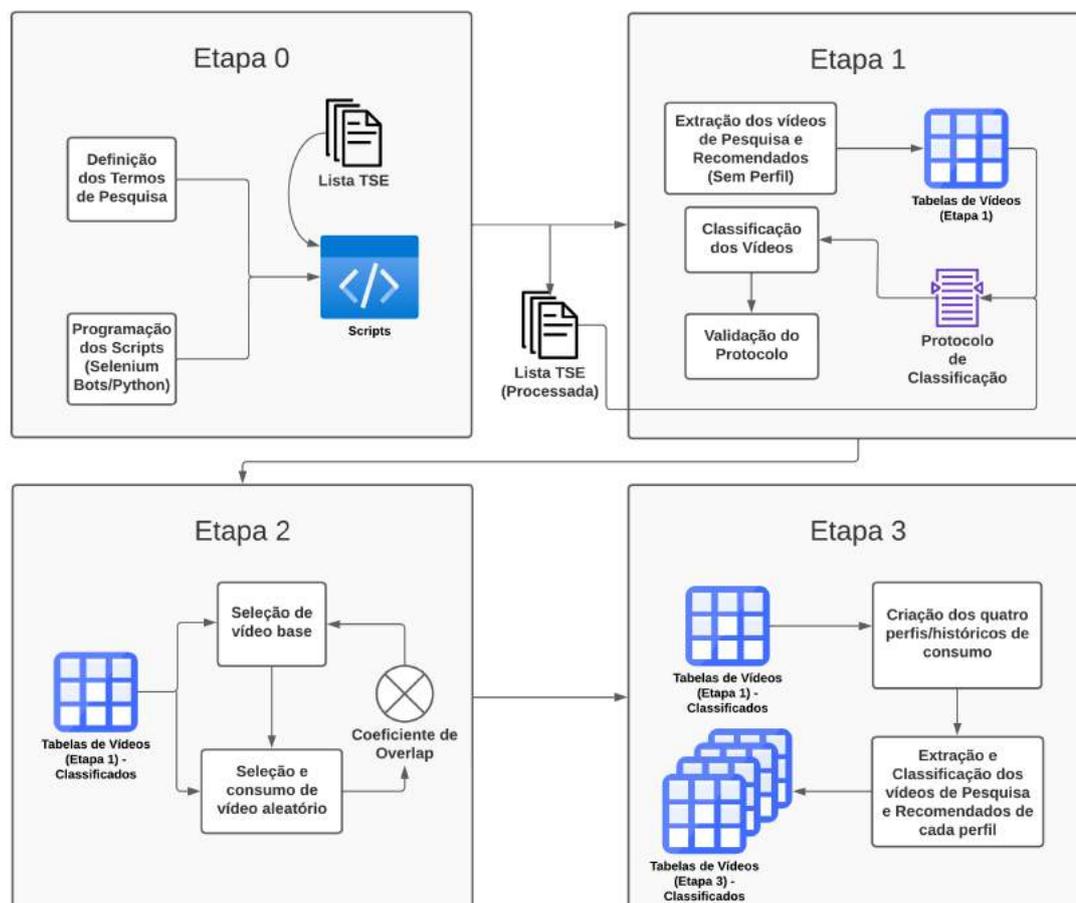


Figura 20 – Etapas da metodologia desenvolvida

na Seção 5.4) de suas recomendações, a fim de aferir a média de consumo de vídeos para começar a haver homogeneidade nas recomendações do SR;

- Etapa 3 (Extração dos vídeos com os quatro perfis de visualização e análise dos resultados): com os vídeos da Etapa 1 classificados e a média de vídeos necessários para homogeneização das recomendações calculados na Etapa 2, são criados quatro perfis distintos no YouTube, cada um com um histórico de consumo de conteúdo: informativo, desinformativo, misto e neutro. A partir destes, é repetida a extração de vídeos realizada na Etapa 1 para cada um dos perfis. Por fim, são gerados os resultados finais, apresentados e analisados no Capítulo 6.

5.1 Questões de Ética

O presente trabalho realiza a extração apenas de informações disponibilizadas publicamente através de vídeos de canais públicos no YouTube, devidamente indexados nos mecanismos de pesquisa e recomendação da rede, sem fazer nenhuma espécie

de *scraping* de dados pessoais ou que violam a privacidade de qualquer indivíduo ou grupo. Não é realizada, também, nenhuma alteração dos dados já presentes na rede, apenas é feita uma análise dos dados extraídos, de maneira que a Lei Geral de Proteção de Dados (LGPD) (BRASIL, 2018) é respeitada.

5.2 Etapa 0 - Pré-processamento

A primeira parte essencial de qualquer estudo de auditoria algorítmica em uma plataforma que envolve mecanismos de pesquisa e recomendações, como é o caso do YouTube, é a definição dos termos de pesquisa que consigam englobar diferentes aspectos relativos a um mesmo tema e que tenham a maior relevância possível no momento em que os dados são extraídos.

Dessa maneira, para que fossem definidos os cinco melhores termos de pesquisa que capturem as nuances acerca do assunto “Urnas Eletrônicas”, foi utilizada a ferramenta Google Trends (GOOGLE, 2023). Esta ferramenta permite mensurar o impacto e a relevância que um determinado tópico tem nas plataformas da empresa (Google Pesquisa, YouTube, Google Shopping, etc.) e também termos de pesquisa relacionados a este tópico que estão em relevância naquele momento.

O uso do Google Trends como ferramenta para definir termos de pesquisa em uma metodologia de pesquisa é respaldado por várias referências acadêmicas, oferecendo uma visão em tempo real das tendências de busca, permitindo que os pesquisadores identifiquem termos populares e relevantes para sua pesquisa. Em NUTI et al. (2014), por exemplo, é feita uma revisão sistemática sobre o uso da ferramenta em pesquisas voltadas para a área da saúde. A capacidade de rastrear padrões de busca que o serviço oferece pode ser aplicada em várias áreas de pesquisa, incluindo a abordada neste trabalho. Os Trabalhos Relacionados TR1 e TR3 utilizaram desta ferramenta, por exemplo.

Portanto, utilizando como base o assunto “Urna Eletrônica” e limitando o escopo do tempo de relevância para “últimos 90 dias” a partir de 12 de setembro de 2022, a localização para “Brasil” e a plataforma “Pesquisa do YouTube” (conforme Figura 21), foi possível observar alguns indicativos sobre o fluxo informacional acerca das UEs naquele momento.

A Figura 22 ilustra os principais achados desta etapa. Na Figura 22 é apresentado o pico de relevância nos três meses anteriores a busca ocorreu no dia 19 de julho de 2022. Esta data marca o dia após a reunião do ex-presidente Bolsonaro com embaixadores, onde houve um deliberado ataque às UEs e a membros da Suprema Corte do Brasil (ver Seção 3.4.4). A Figura 22 (b) e (c) apresentam os principais termos utilizados em pesquisas relacionadas que envolvem as UEs, Bolsonaro e também o humorista Danilo Gentili (que já fez esquetes de humor de alta circulação acerca

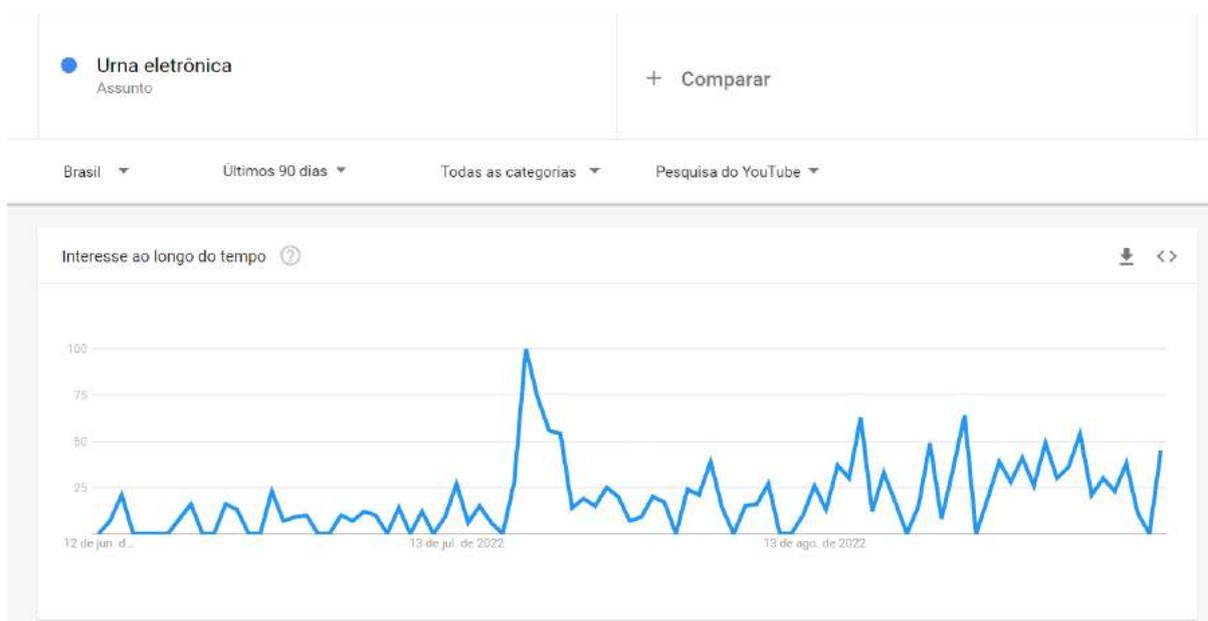


Figura 21 – Resultado da plataforma Google Trends na Pesquisa do YouTube Brasil acerca do Assunto “Urna Eletrônica” em 12 de setembro de 2022

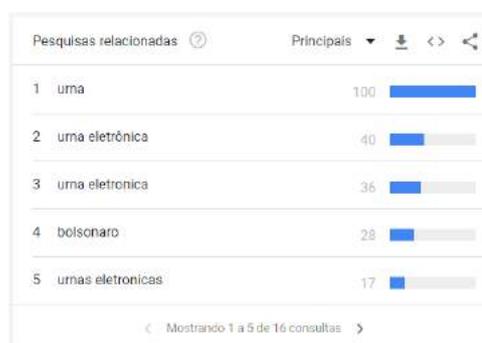
do assunto). Por fim, a Figura 22 (d) apresenta os principais termos com aumento repentino de popularidade que inclui a supracitada reunião de Bolsonaro com embaixadores, repete a relevância do humorista Danilo Gentili e apresenta o termo “nova urna eletronica 2022” e “som de urna eletronica” como termos relevantes acerca do assunto.

Em face destes resultados, os cinco termos de pesquisa escolhidos para a implementação desta metodologia foram:

1. “urna eletrônica” – dentre os termos mais relevantes, é o que melhor representa o objeto sendo estudado;
2. “urnas eletronicas” – extensão do termo acima;
3. “bolsonaro urnas eletrônicas” – termo que pode indicar a tendência apresentada em vídeos que envolvem o ex-presidente da República e suas opiniões acerca das UEs;
4. “bolsonaro embaixadores urnas” – termo de rápida ascensão de relevância que engloba vídeos sobre a reunião de Bolsonaro com embaixadores e seus desdobramentos;
5. “nova urna eletronica 2022” - termo de rápida ascensão de relevância que engloba vídeos que possam explicar as diferenças entre as UEs utilizadas nas eleições de 2022 das anteriores.



(a) Pico de relevância do assunto



(b) Principais pesquisas relacionadas - Página 1



(c) Principais pesquisas relacionadas - Página 2



(d) Pesquisas relacionadas com aumento repentino de popularidade

Figura 22 – Principais resultados obtidos no Google Trends sobre o assunto “Urna Eletrônica” em 12 de setembro de 2022

Os termos “urna” e “bolsonaro” não foram escolhidos por serem extremamente genéricos, o que poderia influenciar em resultados que não têm relação com o corrente tema de pesquisa. O termo “som da urna eletronica” não foi considerado por se tratar de algo muito específico e que não engloba a discussão geral sobre as UEs. Já os termos “danilo gentili” e “danilo gentili urna eletronica” foram descartados pela possibilidade de resultar na apresentação repetitiva de uns poucos vídeos em que o humorista ataca as UEs – sem agregar à discussão mais geral sobre o assunto.

Tendo definido os cinco termos de pesquisa e visando automatizar o processo de pesquisa e extração de vídeos do YouTube e considerando a limitação das APIs disponibilizadas para este fim (que não possibilitam, por exemplo, emular um perfil com determinado histórico de consumo de vídeos – fundamental para este trabalho – ou extrair os vídeos relacionados de um determinado vídeo específico), foram desenvolvidos *scripts* em Python (SOARES, 2023) utilizando *Selenium Bots*.

Selenium (SELENIUM, 2021) é um *framework* de código aberto com uma série de ferramentas de automação de testes para desenvolvimento *web*. Podendo ser implementado em C#, Ruby, Java, Python e JavaScript, é composto de três ecossistemas: uma IDE, que integra diversos *scripts* para testes automatizados diretamente em código, um *WebDriver*, que faz a conexão dos *scripts* desenvolvidos com diferentes navegadores (como Chrome e Firefox) e um *Grid*, voltado para clusterização de testes.

Neste trabalho foram desenvolvidos três *scripts* principais utilizando o Selenium WebDriver do Firefox: o *script* utilizado na Etapa 1, que é responsável pela extração dos vídeos de pesquisa e recomendados sem emular um perfil de consumo específico; o *script* usado na Etapa 2, responsável por mensurar a média de vídeos assistidos necessários para haver homogeneização das recomendações de vídeos; e o *script* usado na Etapa 3, que constrói quatro históricos de consumo distintos e repete a extração dos vídeos. O funcionamento individual de cada *script* será detalhado nas Seções a seguir e todos estão disponíveis no GitHub do autor: github.com/yanbss/YouTubeUrnas – junto dos resultados obtidos neste trabalho e também de *scripts* utilitários e demais arquivos desenvolvidos.

Todos os *scripts* desenvolvidos operam sobre a página de pesquisa de vídeos (Figura 23) e sobre as páginas de vídeos específicos (Figura 24). Cabe aqui uma observação do aviso da própria plataforma em relação aos conteúdos que envolvem UEs tanto na pesquisa como em vídeos acerca do assunto, direcionando o usuário a uma página da Justiça Eleitoral que explica o sistema.

Ainda nesta etapa de pré-processamento, é rodado um dos *scripts* utilitários para determinar quais vídeos da lista do TSE (ver Seção 3.4.5) foram removidos ou estão indisponíveis. Para isso, o *bot* visita cada um dos *links* do YouTube presentes na lista e analisa a disponibilidade do vídeo, retornando um dos códigos definidos na Tabela

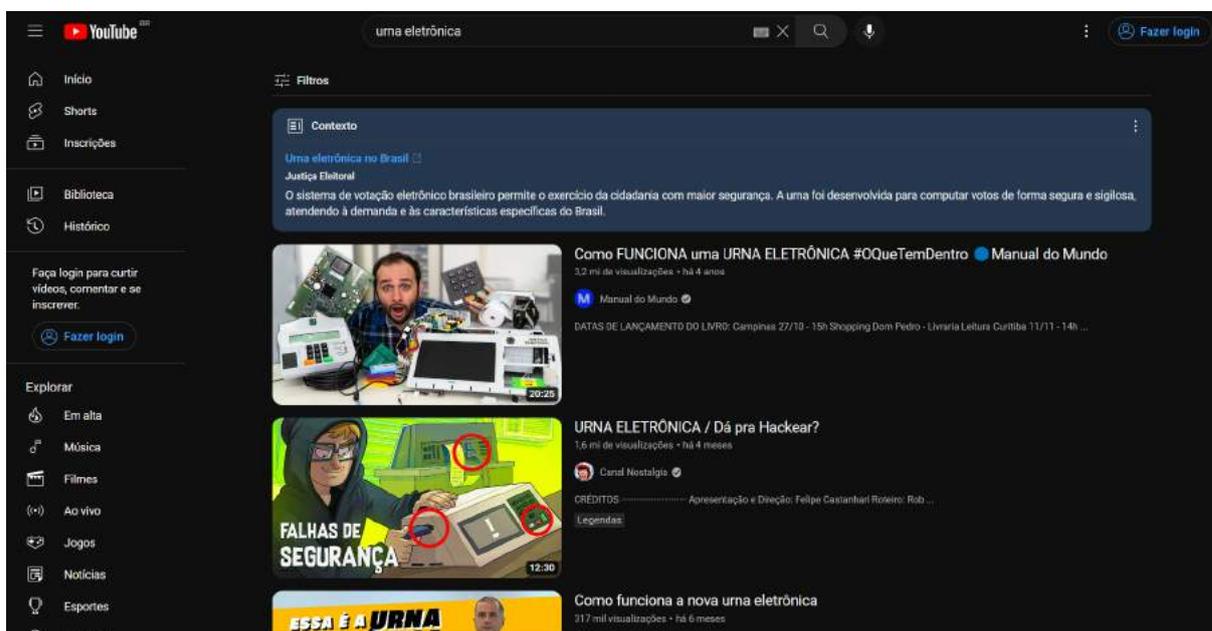


Figura 23 – Página de pesquisa de vídeos do YouTube

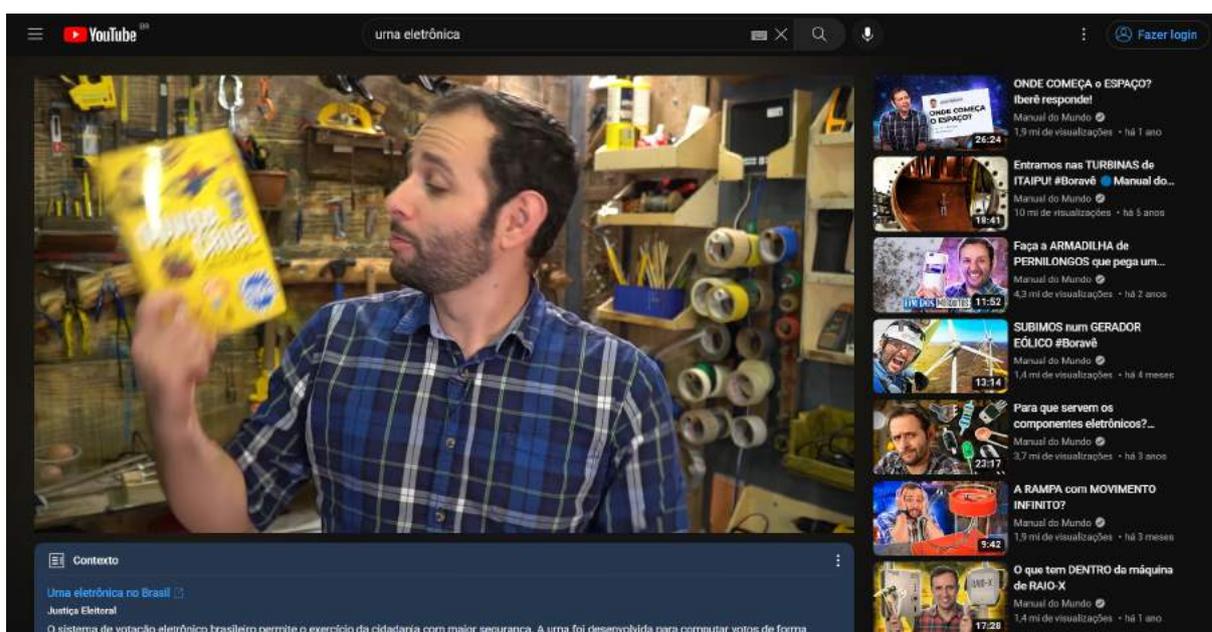


Figura 24 – Página de vídeo específico do YouTube

Tabela 9 – Códigos retornados no processamento da tabela de *links* do TSE

Código	Status do vídeo
0	Vídeo disponível (<i>online</i>)
1	Vídeo removido por violar as diretrizes da comunidade do YouTube
2	Vídeo indisponível ("Este vídeo não está mais disponível")
3	Gravação da transmissão ao vivo não está mais disponível
4	Vídeo privado (arquivado)
5	Vídeo indisponível ("Vídeo indisponível")
6	Vídeo de canal com conteúdo inadequado (<i>login</i> necessário)

9. Por fim, são geradas duas tabelas: uma contendo os vídeos ainda disponíveis (e que será utilizada na Etapa 1) e outra contendo os vídeos removidos, junto de seu código de *status*.

Cabe destacar que os códigos 2 e 5 (Vídeo indisponível) diferem na medida em que, vídeos que apresentam o código 2 foram removidos por alguma determinação externa ao canal que postou o vídeo (denúncia do vídeo ou moderação por parte do YouTube) e vídeos que apresentam o código 5 foram removidos por determinação do próprio canal (vídeo excluído pelo autor) ou pela exclusão do canal na plataforma.

5.3 Etapa 1 - Extração dos Vídeos Sem Perfil e Criação do Protocolo de Classificação

Com os *scripts* desenvolvidos e devidamente testados, foi realizada a primeira etapa de extração e processamento de vídeos: para cada um dos cinco termos de pesquisa definidos na Etapa 0, foram extraídos os primeiros vinte resultados de pesquisa (utilizando cada um dos quatro algoritmos de ordenação de resultados de pesquisa do YouTube – detalhados a seguir) e, de cada um destes vinte vídeos, são extraídos os primeiros dez vídeos recomendados. Neste processo de extração também é associado a cada vídeo o fator de Métrica de Popularidade (mp), calculado seguindo a Equação 4. Um funcionamento geral deste processo é exposto na Figura 25 e deta-

lhado a seguir.

$$\begin{aligned} \text{Métrica de Popularidade (mp)} = \text{número de visualizações} + \text{número de curtidas} \\ + \text{número de comentários} \end{aligned} \quad (4)$$

Os algoritmos de ordenação de resultados de pesquisa do YouTube operam da seguinte maneira:

- Relevância – Método padrão de classificação de pesquisa do YouTube, que ordena os resultados baseados em sua relevância em relação ao termo de pesquisa, levando em consideração fatores como o título, descrição e *tags* dos vídeos, assim como informações sobre os vídeos e seus autores;
- Data de envio – Ordena os resultados baseado na data de *upload*, com os vídeos mais novos aparecendo primeiro;
- Contagem de visualizações – Ordena os resultados baseado no número de visualizações, com os vídeos mais vistos aparecendo primeiro;
- Classificação – Ordena os resultados baseado na avaliação média dos vídeos, com os vídeos melhor avaliados aparecendo primeiro. A avaliação é determinada pelo número de *likes* e *dislikes* que um vídeo recebeu pelos usuários.

Conforme é possível observar na Figura 25, ao iniciar o *script*, é iterado um vetor contendo os 5 termos de pesquisa e, para cada termo de pesquisa, é iterado um vetor contendo os 4 algoritmos de ordenação. Para cada par (Termo de Pesquisa x Algoritmo de Ordenação), é efetuada a pesquisa e são carregados os resultados. A página é então rolada para baixo até que sejam carregados os primeiros 20 resultados daquele par.

É alocado, então, internamente ao *script*, uma lista contendo os *links* dos 20 vídeos carregados. Para cada um destes *links*, é feito o acesso à URL e calculado a métrica de popularidade (MP) daquele vídeo (seguindo a fórmula da Equação 4). Para cada primeiro vídeo acessado de cada par (Termo de Pesquisa x Algoritmo de Ordenação) é gerada uma tabela externa ao *script* (em forma de arquivos .csv) que irá conter os resultados de pesquisa e outra que irá conter os resultados de recomendação.

Para cada vídeo é inserido na tabela de pesquisa o *link* do vídeo e sua MP. Para cada vídeo são extraídos os 10 primeiros vídeos recomendados, sendo estes inseridos na tabela de recomendados (com o *link* de origem e o *link* do vídeo recomendado).

Esse processo é repetido até que todos os 20 vídeos da lista interna do *script* sejam acessados e seus dados sejam inseridos nas tabelas. Ao fim dos 20 vídeos

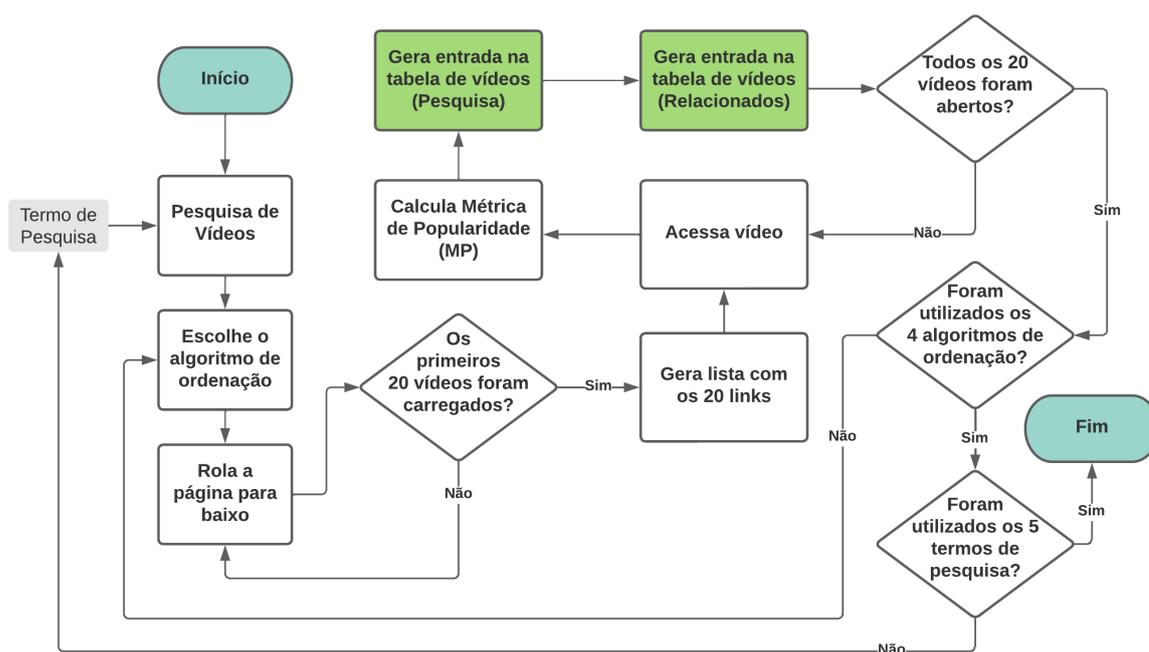


Figura 25 – Funcionamento do *script* utilizado na Etapa 1

do resultado de pesquisa, é iterado o próximo Algoritmo de Ordenação para aquele mesmo Termo de Pesquisa. Este processo é repetido até que todos os pares (Termo de Pesquisa x Algoritmo de Ordenação) sejam processados.

No fim deste processo são geradas 40 tabelas: 20 dos resultados de pesquisa e 20 de vídeos recomendados (5 termos de pesquisa x 4 algoritmos de ordenação). Cada tabela dos resultados de pesquisa contém 20 entradas únicas, ao passo que as tabelas de vídeos recomendados têm um número variado de entradas (até no máximo 200), visto que vídeos *shorts* do YouTube não contam com vídeos recomendados.

São escolhidos, então, 200 vídeos aleatoriamente – 100 da lista do TSE de vídeos que permanecem *online* após a Etapa 0 e 100 dos resultados dos vídeos de pesquisa extraídos nesta etapa. A partir destes vídeos, é feita a construção de um Protocolo de Classificação (disponível no Apêndice A) que divide os vídeos nas categorias detalhadas na Tabela 10.

Tendo finalizado a classificação de todos os vídeos desta primeira etapa de extração, foram selecionados 90 vídeos aleatoriamente (30 de cada uma das categorias) para a validação do Protocolo de Classificação com outros dois avaliadores. Para tal, cada um deveria classificar os 90 vídeos baseando-se no Protocolo de Classificação redigido pelo autor – sem consulta um ao outro ou ao próprio autor.

Para a validação, foi utilizada o Coeficiente Alfa de Krippendorff (KRIPPENDORFF, 2011) – medida estatística que avalia a concordância entre codificadores de um mesmo conjunto de dados – e foi obtido um grau de confiabilidade de $\alpha = 0,834$ entre o autor e o primeiro avaliador, $\alpha = 0,884$ entre o autor e o segundo avaliador e

Tabela 10 – Anotação utilizada para classificar os vídeos extraídos

Classificação	Descrição da Categoria	Classificação Normalizada
-1	Vídeo sobre ou contendo menções acerca das Urnas Eletrônicas, analisados seguindo o Protocolo de Classificação e sem potencial desinformativo	-1 (Informativo)
0	Vídeos sem relação ou menções às Urnas Eletrônicas	0 (Neutro)
1	Vídeo incluído na lista do TSE com links de vídeos com potencial desinformativo sobre as Urnas Eletrônicas	1 (Desinformativo)
2	Vídeos sobre ou contendo menções acerca das Urnas Eletrônicas, analisados seguindo o Protocolo de Classificação e com potencial desinformativo	1 (Desinformativo)
3	Vídeo Removido	Desconsiderado

$\alpha = 0,823$ entre o autor e ambos os avaliadores em conjunto, o que é considerado um resultado suficientemente confiável. A tabela contendo a classificação do autor e dos dois avaliadores para o conjunto de vídeos pode ser conferida no Apêndice B.

Este índice é utilizado em diversos trabalhos de Análise de Conteúdo para a validação dos dados classificados em determinadas categorias. Foi utilizado para a validação de protocolos de classificação de desinformação sobre COVID-19 em SOARES et al. (2021) e BIZZOTTO; MORLINO; SCHULZ (2022), e para a classificação de discurso de ódio em CINELLI et al. (2021), por exemplo. Segundo SAMPAIO; LYCARIÃO (2018):

Os valores dos testes de confiabilidade, geralmente, variam de “-1” a “1”, em que “1” indica uma concordância perfeita; “0” uma falta de confiabilidade por serem pareamentos aleatórios; e abaixo de zero que há uma discordância não aleatória ocorrendo. Ou seja, quando temos um índice negativo, os codificadores estão codificando o mesmo material a partir de compreensão significativamente e sistematicamente distintas uns dos outros. Qualquer valor acima de 0,9 é, em geral, considerado muito confiável e acima de 0,8 suficientemente confiável. Já valores entre 0,667 e 0,8 são considerados suficientes para variáveis experimentais (em aperfeiçoamento) no caso do alpha de Krippendorff.

5.4 Etapa 2 - Teste de Homogeneidade das Recomendações

A fim de se estabelecer o número médio de vídeos assistidos necessários para que se passe a ter homogeneidade nas recomendações apresentadas pelo SR do YouTube, nesta Etapa é realizado um teste baseado na bibliografia relacionada.

O teste segue o seguinte algoritmo apresentado abaixo e baseado em PAPADA-MOU et al. (2020):

- 1: Seja S um conjunto de 100 vídeos aleatórios classificados com potencial desinformativo
- 2: Seja V_{ref} um vídeo com potencial desinformativo selecionado aleatoriamente
- 3: Seja V_{refRec} o conjunto dos 10 primeiros vídeos recomendados de V_{ref}
- 4: $RH_{recs} \leftarrow \{V_{refRec}\}$
- 5: $S_{threshold} \leftarrow 0$ (Número de vídeos assistidos)
- 6: **Para cada** vídeo V em S **faça**:
- 7: Assista vídeo V
- 8: $W \leftarrow W + 1$
- 9: Extrai o conjunto dos 10 primeiros vídeos recomendados R de V_{ref}
- 10: Calcula o Coeficiente de Overlap O_{coef} entre R e RH_{recs}
- 11: **Se** $O_{coef} \geq S_{threshold}$ **então**
- 12: **retorna** W
- 13: **senão**
- 14: Adiciona R ao conjunto RH_{recs} retornado nas iterações prévias
- 15: **fim_se**
- 16: **fim_para**

O Coeficiente de *Overlap* (INMAN; JR, 1989) é uma medida de de similaridade que determina a sobreposição entre dois conjuntos finitos, sendo caracterizado como o tamanho da interseção entre os conjuntos divididos pelo tamanho do menor conjunto entre os dois. Sua fórmula é definida pela Equação 5.

$$Overlap(C1, C2) = \frac{|C1 \cap C2|}{\min(|C1|, |C2|)} \quad (5)$$

O Coeficiente de *Overlap* é utilizado neste trabalho para determinar o momento em que um vídeo base passa a ter suas recomendações similares à iteração anterior (quando o Coeficiente ultrapassa o limiar de 1.0).

Como o número de vídeos assistidos necessários não é exato, o teste foi rodado dez vezes e foi feita uma média dos resultados obtidos para atestar a homogeneidade das recomendações. Existe uma variabilidade natural nos resultados em decorrência de fatores como: (i) a variedade e diversidade dos tópicos relativos aos vídeos assistidos (os vídeos não falam necessariamente apenas sobre Urnas Eletrônicas, mas podem existir menções em seu conteúdo), (ii) a frequência de uso do usuário na pla-

taforma e (iii) a relevância da interação do usuário com o vídeo (como, por exemplo, se foi dado um *like* ou feito um comentário ou até se o usuário é inscrito no canal que postou o vídeo).

Com números variando de 7 até 77 vídeos assistidos, a média obtida foi de **33,5** vídeos assistidos necessários para que as recomendações passem a ter uma maior homogeneidade. Para fins de comparação, o resultado obtido no trabalho relacionado no qual o teste foi baseado foi de **22** vídeos, o que pode ser um primeiro indício de mudanças no funcionamento dos SRs do YouTube para que as recomendações sejam mais diversas.

A tabela contendo o Coeficiente obtido em cada uma das iterações do teste encontra-se no Apêndice C.

5.5 Etapa 3 - Extração dos Vídeos com os Quatro Perfis de Visualização

Por fim, são criados quatro perfis com as mesmas características demográficas (nome, gênero, idade e localização) e, com a utilização de um *script* auxiliar, é construído o histórico de visualização de cada um deles baseado em quatro modelos:

- Perfil Desinformativo – Histórico construído com os 100 vídeos classificados como **com potencial desinformativo** na Etapa 1 com maior Métrica de Popularidade (mp);
- Perfil Informativo – Histórico construído com os 100 vídeos classificados como **sem potencial desinformativo** na Etapa 1 com maior Métrica de Popularidade (mp);
- Perfil Misto – Histórico construído com os 50 primeiros vídeos utilizados nos perfis Desinformativo e Informativo, gerando um histórico composto metade por vídeos **com potencial desinformativo** e metade por vídeos **sem potencial desinformativo**;
- Perfil Neutro – Histórico construído com os 100 vídeos classificados como **neutros** na Etapa 1 com maior Métrica de Popularidade (mp).

Criados os perfis, é realizada para cada um deles, a extração dos vídeos de pesquisa e recomendados utilizando cada um dos termos de pesquisa e algoritmos de ordenação e os resultados da análise final são apresentados no Capítulo 6.

5.6 Métricas de Avaliação dos Resultados

5.6.1 Search Engine Results Page – Misinformation Score (SERP-MS)

Desenvolvida em HUSSEIN; JUNEJA; MITRA (2020) (TR1), a *Search Engine Results Page – Misinformation Score* (SERP-MS) é uma pontuação ponderada que varia de -1 a 1 e avalia o quão desinformativa é a lista de vídeos retornadas como resultado de uma pesquisa. Quanto mais próximo de 1 o valor resultante, mais desinformativo, e vice-versa (valores positivos são considerados majoritariamente desinformativos e valores negativos majoritariamente informativos).

A métrica considera que, quanto mais alto o ranqueamento do vídeo na lista de pesquisa (primeiros resultados), maior o peso de sua classificação e sua fórmula está ilustrada na Equação 6.

$$SERP - MS = \frac{\sum_{r=1}^n (x_i * (n - r + 1))}{\frac{n*(n+1)}{2}} \quad (6)$$

Na equação acima, n representa o número total de vídeos na lista de pesquisa retornada, r o número do vídeo atual (iterado de 1 até n no somatório) e x_i o valor associado ao vídeo atual (-1 para vídeos informativos, 0 para vídeos neutros e 1 para vídeos desinformativos).

Esta métrica foi utilizada nas Etapas 1 e 3 deste trabalho. O Apêndice D contém uma tabela única com todos os resultados de SERP-MS encontrados nas duas etapas e para cada um dos perfis.

5.6.2 Coeficiente de Determinação (R²)

O coeficiente de determinação (R²), é uma medida estatística utilizada para avaliar a qualidade de um modelo de regressão, incluindo as linhas de tendência. O cálculo do R² envolve comparar a variação dos valores observados com a variação dos valores previstos pelo modelo. Variando de 0 a 1, o R² fornece uma indicação de quanto a variabilidade dos dados é explicada pelo modelo de regressão. Em outras palavras, ele mede a proporção da variabilidade dos valores dependentes (variável de resposta) que é explicada pelas variáveis independentes (variáveis de entrada) do modelo (MARTINS, 2014).

Um valor de R² igual a 1 indica que todas as variações nos valores dependentes são perfeitamente explicadas pelas variáveis independentes do modelo. Por outro lado, um valor de R² igual a 0 indica que o modelo não explica nenhuma variação nos valores dependentes e é essencialmente inútil. A Fórmula para o cálculo de R² para um conjunto de dados de entrada está ilustrada na Equação 7.

$$R^2 = 1 - \frac{SQ_{res}}{SQ_{tot}} \quad (7)$$

Na equação acima, SQ_{res} representa a Soma dos Quadrados Residuais (diferença entre os valores observados e os valores previstos pelo modelo de regressão) e SQ_{res} representa a Soma dos Quadrados Totais (variação total dos valores observados em relação à média dos valores observados).

Esta métrica foi utilizada na Etapa 3, para comparar as linhas de tendência das médias de SERP-MS para cada Perfil em relação às médias obtidas na Etapa 1, considerando tanto a média entre os AOs como a média entre os TPs.

6 RESULTADOS

Aplicada a metodologia detalhada no Capítulo anterior, foram gerados resultados para cada uma das etapas e, ao final, para os quatro perfis com diferentes históricos de visualização. O número de vídeos analisados e classificados foi: 4.130 vídeos na Etapa 1, e, nos quatro perfis da Etapa 3 foram 3.870 vídeos com Perfil Desinformativo, 3.860 vídeos com Perfil Informativo, 3.860 vídeos com Perfil Misto e 3.640 vídeos com Perfil Neutro, totalizando 19.360 vídeos. A extração dos vídeos da Etapa 1 foi realizada em setembro de 2022 e os vídeos analisados entre setembro e outubro de 2022. A extração dos vídeos da Etapa 3 foi realizada em outubro de 2022 e os vídeos analisados entre novembro de 2022 e abril de 2023.

Neste Capítulo, estes resultados serão apresentados e discutidos para todas as Etapas (com exceção da Etapa 2, a qual já teve seus resultados apresentados diretamente no Capítulo anterior). Os resultados serão apresentados utilizando as Métricas expostas na Seção 5.6 e em termos de variação percentual (vídeos recomendados, comparação entre termos de pesquisa e algoritmos de ordenação). Para fins de simplificação da exibição dos resultados em tabelas e gráficos, se utilizará a seguinte abreviação:

- TP1 - Termo de Pesquisa 1: “bolsonaro embaixadores urnas”
- TP2 - Termo de Pesquisa 2: “bolsonaro urnas eletrônicas”
- TP3 - Termo de Pesquisa 3: “nova urna eletronica 2022”
- TP4 - Termo de Pesquisa 4: “urna eletrônica”
- TP5 - Termo de Pesquisa 5: “urnas eletronicas”
- AO1 - Algoritmo de Ordenação 1: Classificação
- AO2 - Algoritmo de Ordenação 2: Contagem de visualizações
- AO3 - Algoritmo de Ordenação 3: Data de envio
- AO4 - Algoritmo de Ordenação 4: Relevância

Tabela 11 – Número de vídeos da lista do TSE na classificação da Etapa 0 (setembro 2022 x junho 2023)

Código	Número de Vídeos (setembro 2022)	Número de Vídeos (junho 2023)	Diferença (%)
0	1.288	923	-28,34
1	17	22	+29,41
2	59	222	+276,21
3	8	8	0
4	14	179	+1.178,57
5	23	55	+139,13
6	1	1	0

Por fim, no Capítulo seguinte, serão respondidas as Questões de Pesquisa e debatidas as Hipóteses levantadas no Capítulo 1.

6.1 Etapa 0 - Pré-Processamento

A Tabela 11 apresenta os resultados encontrados ao rodar o *script* da Etapa 0 nos meses de setembro de 2022 (início da aplicação da metodologia) e junho de 2023 (fim da escrita do presente trabalho), com os respectivos códigos obtidos conforme explicação no Capítulo 5. Com esta Tabela, pretende-se avaliar o impacto das políticas implementadas pelo TSE para remoção de conteúdo desinformativo e demais desdobramentos após o pleito eleitoral de 2022.

Como é possível verificar na Tabela 11, houve uma diferença significativa na quantidade de vídeos da lista do TSE removidos entre setembro de 2022 e junho de 2023. Enquanto no período pré-eleitoral e anterior à série de acontecimentos detalhados nas Seções 3.4.5 e 3.5, 1.288 vídeos com potencial desinformativo continuavam disponíveis, **após todas as determinações do TSE e também pressão por parte de organizações como a *Sleeping Giants* Brasil para desmonetizar canais que espalham desinformação, atualmente (junho de 2023), 923 destes vídeos continuam online (uma diferença de 28,34%).**

É possível notar também que a vasta maioria dos vídeos que não se encontram mais disponíveis enquadram-se nas categorias dos códigos 2 (**Vídeo Removido por determinação externa ao canal autor – aumento de 276,21%**), 4 (**Vídeo Privado – Arquivado – aumento de 1178,57%**) e 5 (**Vídeo Removido por determinação do próprio autor ou pela exclusão do canal – aumento de 139,13%**) a passo que vídeos removidos com a *flag* de **violação das diretrizes do YouTube (código 1) tiveram um aumento de apenas 29,41%**.

Tabela 12 – Percentual de vídeos classificados na Etapa 1 e SERP-MS dos pares Termo de Pesquisa x Algoritmo de Ordenação (Termo de Pesquisa – TP, Algoritmo de Ordenação – AO, Vídeos Informativos – VI, Vídeos Neutros – VN, Vídeos Desinformativos – VD, Pesquisa – P e Recomendados – R)

TP	AO	SERP-MS	VI (%)		VN (%)		VD (%)	
			P	R	P	R	P	R
TP1	AO1	-0,26	55	13,77	15	75	30	8,16
TP1	AO2	0,09	50	18,53	0	68,78	50	10,24
TP1	AO3	-0,44	80	12,12	5	80,6	15	4,24
TP1	AO4	-0,76	85	27,69	10	58,46	5	11,28
TP2	AO1	0,45	36,84	8,49	0	77,7	63,15	10,45
TP2	AO2	-0,02	50	17,56	10	74,14	40	5,85
TP2	AO3	-0,04	50	5,8	5,55	87,74	44,4	3,22
TP2	AO4	-0,75	75	31,79	10	51,28	15	14,35
TP3	AO1	-0,89	95	42,94	0	55,88	5	1,17
TP3	AO2	-1	100	48,5	0	48,5	0	3
TP3	AO3	-1	100	31,57	0	67,89	0	0,5
TP3	AO4	-1	100	45,72	0	49,24	0	5,02
TP4	AO1	-0,75	89,47	22,75	0	75,13	10,52	2,11
TP4	AO2	-0,52	80	39,5	0	51,5	20	9
TP4	AO3	-1	100	20,67	0	78,21	0	1,11
TP4	AO4	-1	100	47,5	0	46	0	6,5
TP5	AO1	0,14	50	23	0	67,5	50	9,5
TP5	AO2	0,29	30	23	0	69	70	8
TP5	AO3	-0,54	85	34,91	0	60,35	15	4,73
TP5	AO4	-0,98	95	41,5	0	48	5	10,5
Média		-0,49	75,31	27,86	2,77	64,54	21,9	6,45

6.2 Etapa 1 - Extração dos Vídeos Sem Perfil e Criação do Protocolo de Classificação

A Tabela 12 apresenta todos os resultados encontrados na Etapa 1. Como é possível observar prontamente, os únicos quatro pares (Termo de Pesquisa x Algoritmo de Ordenação) que apresentaram um SERP-MS positivo (tendendo à desinformação) foram os pares (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 2 – Contagem de visualizações), (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação), (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação) e (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 2 – Contagem de visualizações). Nota-se também a **ordem de distribuição quantitativa entre os vídeos de Pesquisa é, em média, Informativo > Desinformativo > Neutro. Já entre os vídeos Recomendados é, em média, Neutro > Informativo > Desinformativo.** Os resultados serão explorados mais a fundo nas subseções a seguir.

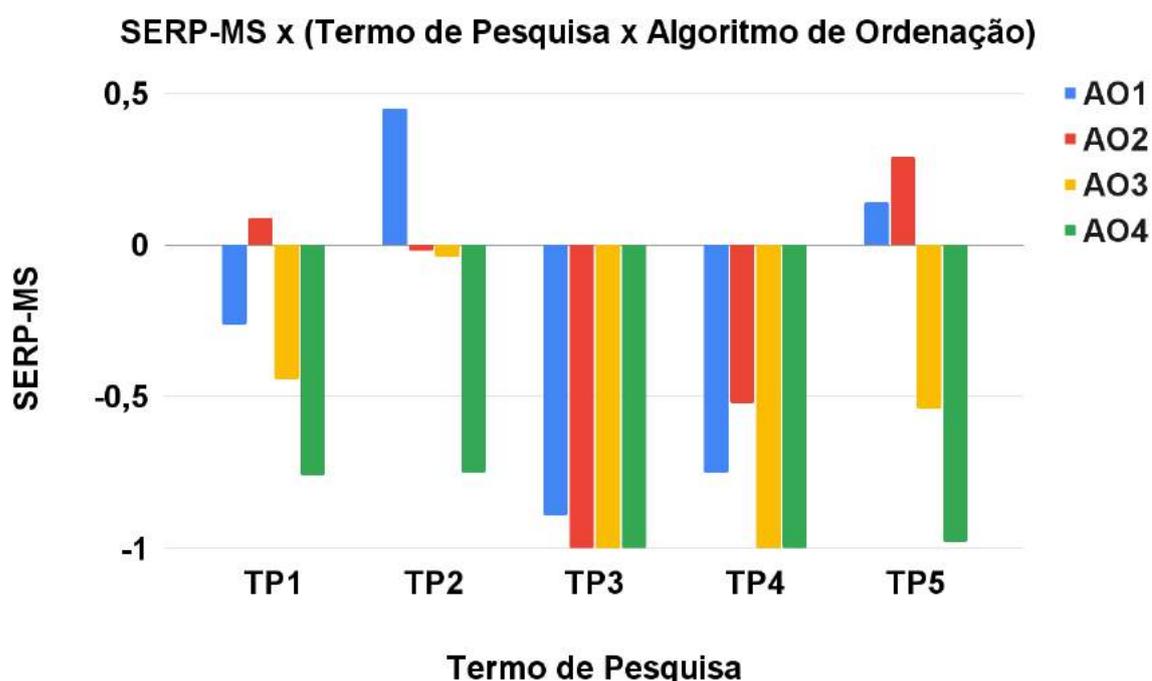


Figura 26 – SERP-MS x Pares (Termo de Pesquisa x Algoritmo de Ordenação) dos resultados da Etapa 1

Os valores em negrito na Tabela 12 simbolizam os valores que ficaram acima da média geral entre todos os Pares (Termo de Pesquisa x Algoritmo de Ordenação) para o valor referente à coluna em que se encontra.

6.2.1 Resultados dos Vídeos de Pesquisa

Como é possível observar na Figura 26 e na Tabela 12, a distribuição dos resultados de pesquisa, quando considerado o SERP-MS obtido entre os diferentes Pares (Termo de Pesquisa x Algoritmo de Ordenação), é majoritariamente informativa, com a média geral do SERP-MS de -0,49 e a distribuição percentual média de vídeos classificados em 75,31% informativos, 21,9% desinformativos e 2,77% neutros.

Os oito pares com um SERP-MS maior que a média (tendendo mais à desinformação que o valor médio) são (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 1 – Classificação: -0,26), (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 0,09), (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 3 – Data de envio: -0,44), (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação: 0,45), (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 2 – Contagem de visualizações: -0,02), (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 3 – Data de envio: -0,04), (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de

Ordenação 1 – Classificação: 0,14) e (Termo de Pesquisa 5 – “urnas eletronicas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 0,29). Os Termos de Pesquisa 3 e 4 não apresentaram um SERP-MS acima da média para nenhum Algoritmo de Ordenação.

Fica em evidência, portanto, a diferença exercida pela escolha do Termo de Pesquisa e Algoritmo de Ordenação na quantidade de desinformação presente nos resultados de pesquisa de vídeos do YouTube. **No caso em estudo, ambos os Termos de Pesquisa que continham “bolsonaro” em sua composição apresentaram a maior quantidade de desinformação.** Seguido destes, o Termo de Pesquisa 5 – “urnas eletronicas” (“urnas eletrônicas”) apresentou desinformação acima da média para vídeos quando utilizados os Algoritmos de Ordenação 1 e 2 e os Termos de Pesquisa 3 e 4 (“nova urna eletrônica 2022” e “urna eletrônica”) não apresentaram SERP-MS acima da média para nenhum Algoritmo de Ordenação, sendo o Termo de Pesquisa 3 – “nova urna eletrônica 2022” o menos desinformativo dentre todos.

Já em relação aos Algoritmos de Ordenação, a diferença é ainda mais preocupante, já que consistentemente entre todos os Termos de Pesquisa, o Algoritmo de Ordenação que apresentou maior quantidade de desinformação foi o Algoritmo de Ordenação 1 – Classificação – ou seja, o que prioriza os vídeos com a maior quantidade de *likes* e, em segundo lugar, o Algoritmo de Ordenação 2 – Contagem de visualizações, que ordena os resultados de pesquisa a partir dos vídeos com maior número de visualizações. **Portanto, em relação a vídeos de pesquisa sobre UEs, a desinformação se encontra mais presente entre os vídeos mais populares.**

É possível notar também que o **Algoritmo de Ordenação 4 – Relevância – Relevância (Algoritmo de Ordenação padrão quando um usuário realiza uma pesquisa de vídeo no YouTube), não apresentou um SERP-MS acima da média para nenhum Termo de Pesquisa.** Isso indica que, por padrão, os vídeos de pesquisa sobre as UEs no YouTube não tendem majoritariamente à desinformação.

6.2.2 Resultados dos Vídeos Recomendados

Já em relação aos vídeos recomendados, é possível notar ao se observar a Tabela 12 que a distribuição percentual média de vídeos classificados é majoritariamente neutra, sendo distribuída em 64,54% neutros, 27,86% informativos e 6,45% desinformativos.

Aqui, os nove pares que apresentam um percentual de vídeos desinformativos acima da média geral (6,45%) são (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 1 – Classificação: 8,16%), (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 10,24%), (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 4 – Relevância: 11,28%), (Termo de Pesquisa 2 – “bolsonaro

urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação: 10,45%), (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 4 – Relevância: 14,35%), (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 4 – Relevância: 6,5%), (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação: 9,5%), (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 8%), (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 4 – Relevância: 10,5%). **O Termo de Pesquisa 3 – “nova urna eletrônica 2022” foi o único Termo de Pesquisa que não apresentou percentual desinformativo acima da média nas suas recomendações com nenhum dos quatro Algoritmos de Ordenação, assim como o Algoritmo de Ordenação 3 – Data de envio não apresentou resultados acima da média para nenhum dos cinco Termos de Pesquisa.**

Quando é feita a interseção entre os pares que apresentam desinformação acima da média tanto na pesquisa como nas recomendações, é possível perceber os pares (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 1 – Classificação), (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 2 – Contagem de visualizações), (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação), (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação) e (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 2 – Contagem de visualizações).

Fica, portanto, mais uma vez evidente a diferença entre os Termos de Pesquisa e Algoritmos de Ordenação na quantidade de conteúdo desinformativo nas recomendações.

6.2.3 Comparação entre os Algoritmos de Ordenação

A Figura 27 e suas 4 sub-figuras apresentam os resultados obtidos das distribuições percentuais de cada categoria de vídeo (neutro, desinformativo e informativo) nos vídeos recomendados, colocando em evidência cada Algoritmo de Ordenação. Com esta Figura, busca-se comparar visualmente a diferença desta distribuição para cada um dos Algoritmos de Ordenação.

Como é possível observar nos gráficos comparativos da Figura 27, **o Algoritmo de Ordenação 3 – Data de envio leva a menor quantidade de desinformação consistentemente entre todos os Termos de Pesquisa.** O Termo de Pesquisa 3 – “nova urna eletrônica 2022”, que apresentou a menor quantidade de desinformação recomendada, apresenta aqui o menor valor de todos (Termo de Pesquisa 3 – “nova urna eletrônica 2022” x Algoritmo de Ordenação 3 – Data de envio: 0,5%).

O destaque maior aqui, porém, fica com o Algoritmo de Ordenação 4 – Relevância. **O único Algoritmo de Ordenação que não apresentou para nenhum Termo de Pesquisa um SERP-MS maior que a média nos resultados de pesquisa, apre-**

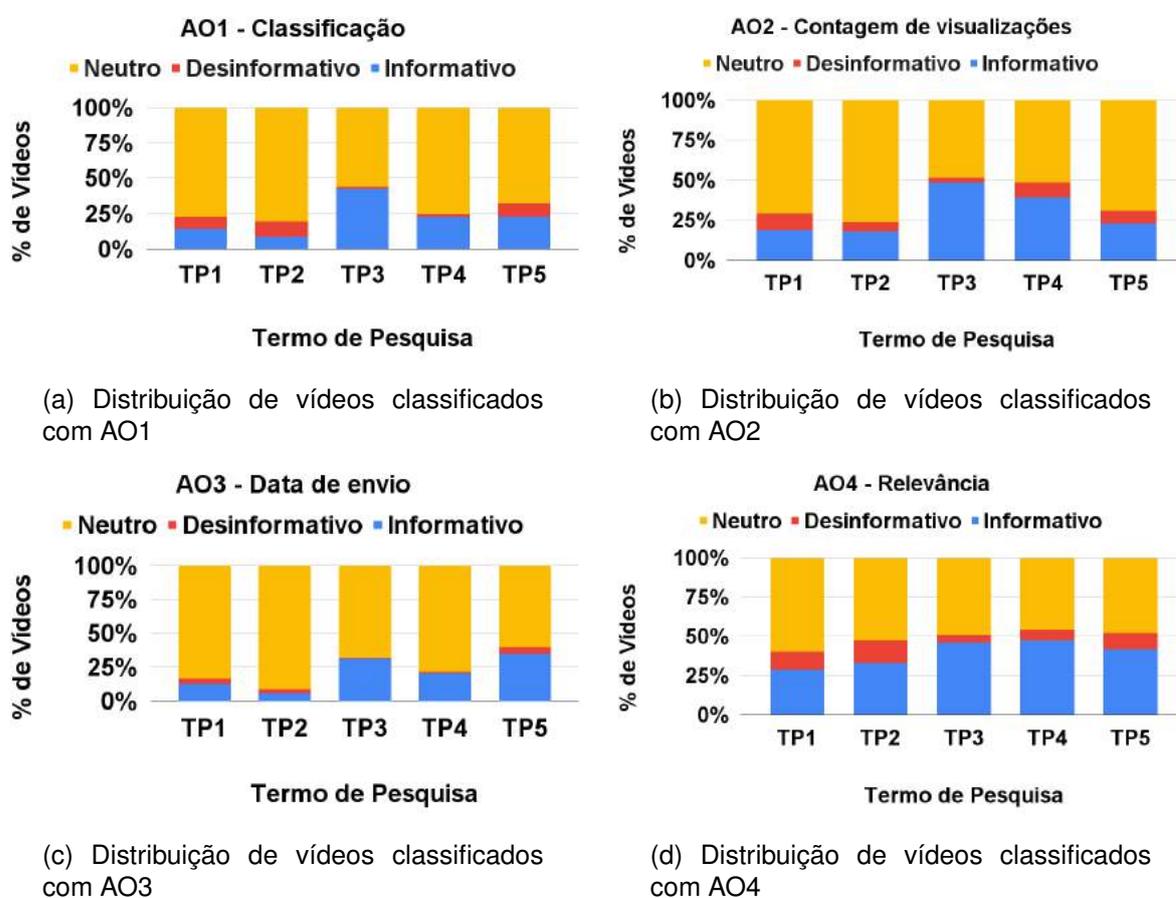


Figura 27 – Distribuições percentuais de cada tipo de vídeo classificado para cada um dos quatro Algoritmos de Ordenação – Etapa 1 – Vídeos Recomendados

senda, nos vídeos recomendados, quantidades de desinformação acima da média de forma consistente entre todos os Termos de Pesquisa (com exceção do Termo de Pesquisa 3 – “nova urna eletronica 2022”). Este Algoritmo de Ordenação também apresenta consistentemente grandes quantidades de conteúdo informativo recomendado, sendo que para os pares (Termo de Pesquisa 3 – “nova urna eletronica 2022” x Algoritmo de Ordenação 4 – Relevância), (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 4 – Relevância) e (Termo de Pesquisa 5 – “urnas eletronicas” x Algoritmo de Ordenação 4 – Relevância), a quantidade de vídeos recomendados informativos e desinformativos supera a quantidade de vídeos neutros. O único outro caso em que isto ocorre é no par (Termo de Pesquisa 3 – “nova urna eletronica 2022” x Algoritmo de Ordenação 2 – Contagem de visualizações). Conclui-se, portanto, que **mesmo que não apresente diretamente nos resultados de pesquisa quantidade de desinformação considerável, o Algoritmo de Ordenação padrão do YouTube é o que mais apresenta conteúdo informativo e desinformativo acerca de um tema nos vídeos recomendados, quando comparado aos outros Algoritmos de Ordenação.**

Já os Algoritmos de Ordenação que apresentaram maior quantidade de desinformação acima da média nos resultados de pesquisa (Algoritmo de Ordenação 1 – Classificação e Algoritmo de Ordenação 2 – Contagem de visualizações), apresentam aqui resultados variáveis. Para os Termos de Pesquisa com SERP-MS acima da média, há maior quantidade de vídeos desinformativos recomendados (**na interseção entre os dois conjuntos, estes são os únicos dois Algoritmos de Ordenação incluídos – ou seja, dentre os Algoritmos de Ordenação, são os que levam a maior desinformação tanto nos resultados de pesquisa como nos vídeos recomendados**), com destaque aqui para o par (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação) que é o único em que o percentual de vídeos desinformativos recomendados supera o de vídeos informativos: 10,45% contra 8,49%.

6.2.4 Comparação entre os Termos de Pesquisa

A Figura 28 e suas 5 sub-figuras apresentam os resultados obtidos das distribuições percentuais de cada categoria de vídeo (neutro, desinformativo e informativo) nos vídeos recomendados, colocando em evidência cada Termo de Pesquisa. Com esta Figura, busca-se comparar visualmente a diferença desta distribuição para cada um dos Termos de Pesquisa.

Como é possível observar na Figura 28, há evidente diferença também entre os Termos de Pesquisa na distribuição entre vídeos neutros, informativos e desinformativos. Como destacado anteriormente, o **Termo de Pesquisa 3 – “nova urna eletronica 2022” é o Termo de Pesquisa que apresenta menor quantidade de desinfor-**

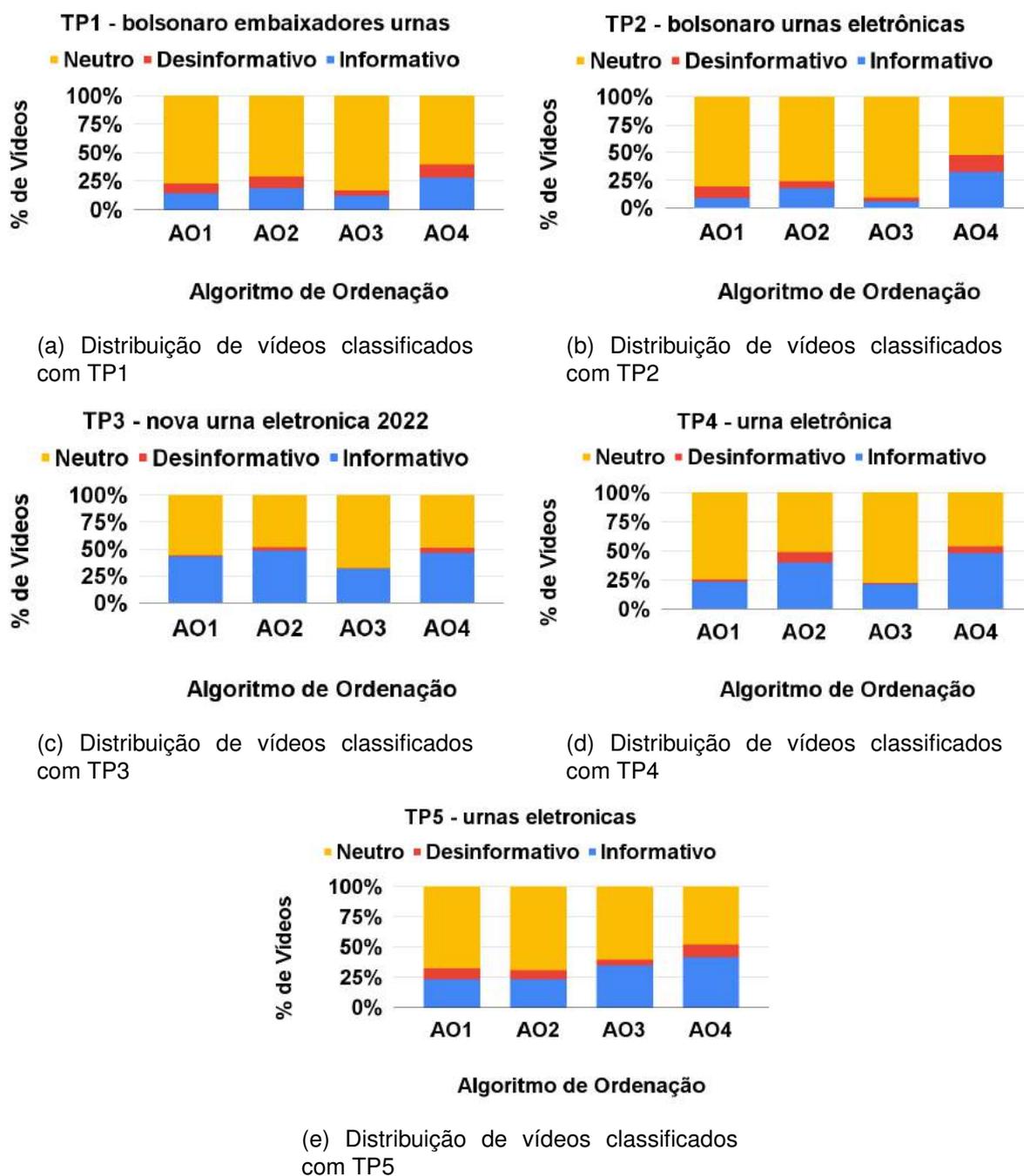


Figura 28 – Distribuições percentuais de cada tipo de vídeo classificado para cada um dos cinco Termos de Pesquisa – Etapa 1 – Vídeos Recomendados

mação em suas recomendações, com uma quantidade de vídeos informativos acima da média geral (27,86%) em todos os Algoritmos de Ordenação – (Termo de Pesquisa 3 – “nova urna eletronica 2022” x Algoritmo de Ordenação 1 – Classificação: 42,94%), (Termo de Pesquisa 3 – “nova urna eletronica 2022” x Algoritmo de Ordenação 2 – Contagem de visualizações: 48,5%), (Termo de Pesquisa 3 – “nova urna eletronica 2022” x Algoritmo de Ordenação 3 – Data de envio: 31,57%), (Termo de Pesquisa 3 – “nova urna eletronica 2022” x Algoritmo de Ordenação 4 – Relevância: 45,72%). Os únicos outros casos em que isto ocorre são nos pares (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 4 – Relevância: 31,79%), (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 2 – Contagem de visualizações: 39,5%), (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 4 – Relevância: 47,5%), (Termo de Pesquisa 5 – “urnas eletronicas” x Algoritmo de Ordenação 3 – Data de envio: 34,91%) e (Termo de Pesquisa 5 – “urnas eletronicas” x Algoritmo de Ordenação 4 – Relevância: 41,5%).

Já os Termos de Pesquisa 1 e 2, que apresentaram nos resultados de pesquisa a maior quantidade de desinformação, apresentam quantidade de vídeos neutros recomendados acima da média de forma consistente entre todos os Algoritmos de Ordenação (com exceção do Algoritmo de Ordenação 4 – Relevância, que, como demonstrado anteriormente, leva à maior quantidade de informação e desinformação nos vídeos recomendados entre todos os Termos de Pesquisa). Isso ocorre principalmente com o Algoritmo de Ordenação 3 – Data de envio: (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 3 – Data de envio: 80,6%) e (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 3 – Data de envio: 87,74%). **Isto pode indicar uma tentativa do SR de diversificação de conteúdo recomendado quando a carga informacional (informação e desinformação) acerca de um assunto é grande nos resultados de pesquisa para os Algoritmos de Ordenação 1, 2 e 3.**

O Termo de Pesquisa 4 – “urna eletrônica” apresenta um resultado variado: para os Algoritmos de Ordenação 1 e 3, a quantidade de desinformação é baixa: (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 1 – Classificação: 2,11%) e (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 3 – Data de envio: 1,11%). Nestes mesmos pares, a quantidade de vídeos recomendados neutros é bastante acima da média geral (64,54%): (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 1 – Classificação: 75,13%) e (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 3 – Data de envio: 78,21%). Em paralelo, a quantidade de vídeos informativos é acima da média geral (27,86%) entre os Algoritmos de Ordenação 2 e 4: (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 2 – Contagem de visualizações: 39,5%) e (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 4 – Relevância: 47,5%). **Cabe salientar aqui**

que o Termo de Pesquisa 4 – “urna eletrônica” (“urna eletrônica”) é o “termo padrão” acerca do assunto, o que pode indicar uma tendência maior do SR de tentativa de redução na recomendação de conteúdos desinformativos, tanto através da diversificação de conteúdo (maior quantidade de vídeos neutros nos Algoritmos de Ordenação 1 e 3), como na recomendação de fontes confiáveis (maior quantidade de vídeos informativos nos Algoritmos de Ordenação 2 e 4).

O mesmo não ocorre, porém, com o Termo de Pesquisa 5 – “urnas eletrônicas” (“urnas eletrônicas”). Aqui a quantidade de vídeos desinformativos recomendados é acima da média geral (6,45%) entre todos os Algoritmos de Ordenação, com exceção do Algoritmo de Ordenação 3 – Data de envio: (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação: 9,5%), (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 8%) e (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 4 – Relevância: 10,5%). Apesar de não ser o Termo de Pesquisa que apresenta maior quantidade de desinformação dentre os cinco, aqui ocorre um fenômeno interessante: nos pares onde a quantidade de vídeos informativos é abaixo da média, a quantidade de vídeos desinformativos é acima da média: (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação) e (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 2 – Contagem de visualizações), já nos pares onde a quantidade de vídeos informativos é acima da média, a quantidade de vídeos neutros é abaixo da média: (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 3 – Data de envio e Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 4 – Relevância). Ou seja, **dentre todos os Termos de Pesquisa, este é o que apresenta a distribuição geral mais equilibrada entre vídeos recomendados neutros, desinformativos e informativos.**

6.3 Etapa 3 - Extração dos Vídeos com os Quatro Perfis de Visualização

Nesta Seção serão expostos os resultados individuais de cada Perfil utilizado na Etapa 3 da Metodologia. Além disso, será feita uma comparação entre os SERP-MS médios (média entre todos os Algoritmos de Ordenação para cada Termo de Pesquisa e média entre todos os Termos de Pesquisa para cada Algoritmo de Ordenação) em cada Perfil com os resultados obtidos na Etapa 1. O mesmo será feito para a avaliação da distribuição percentual dos vídeos recomendados para cada Perfil. Para fins de clareza no texto e uma objetividade maior na explicação dos resultados, os gráficos individuais de todos os Pares (Termo de Pesquisa x Algoritmo de Ordenação) em termos de SERP-MS e comparação individual entre os Termos de Pesquisa e Algoritmos de Ordenação na distribuição dos vídeos recomendados estão presentes no Apêndice

Tabela 13 – Percentual de vídeos classificados na Etapa 3 e SERP-MS dos pares Termo de Pesquisa x Algoritmo de Ordenação (Termo de Pesquisa – TP, Algoritmo de Ordenação – AO, Vídeos Informativos – VI, Vídeos Neutros – VN, Vídeos Desinformativos – VD, Pesquisa – P e Recomendados – R) – Perfil Desinformativo

TP	AO	SERP-MS	VI (%)		VN (%)		VD (%)	
			P	R	P	R	P	R
TP1	AO1	-0,26	50	11,85	25	72,68	25	12,37
TP1	AO2	0	55	12,95	0	53,36	45	33,68
TP1	AO3	-0,44	83,3	12,66	0	80,38	16,6	6,96
TP1	AO4	-0,59	80	20,1	10	70,65	10	9,24
TP2	AO1	0,42	40	7,48	0	85,71	60	6,8
TP2	AO2	-0,06	50	12,29	16,6	64,8	33,33	22,9
TP2	AO3	-0,4	64,7	6,71	0	87,24	35,29	6,04
TP2	AO4	-0,69	80	24,73	10	58,06	10	17,2
TP3	AO1	-0,85	85	23,39	5	71,91	10	4,79
TP3	AO2	-0,84	95	31,55	0	57,14	5	11,3
TP3	AO3	-0,35	81,25	16	6,25	77,6	12,5	6,4
TP3	AO4	-1	100	59,39	0	32,99	0	7,61
TP4	AO1	-0,77	90	11,66	0	83,88	10	4,44
TP4	AO2	-0,48	78,95	21,64	0	64,18	21,05	14,18
TP4	AO3	-0,76	94,12	3,45	5,88	85,63	0	10,92
TP4	AO4	-1	100	47,23	0	47,74	0	5,02
TP5	AO1	0,2	40	17,36	5	72,45	55	10,18
TP5	AO2	-0,21	65	18,39	0	52,3	35	29,31
TP5	AO3	-0,64	88,8	27,89	5,55	63,26	5,55	8,84
TP5	AO4	-0,98	95	47,23	5	48,24	0	4,52
Média		-0,48	75,8	21,7	4,71	66,51	19,47	11,63
Diferença Média Etapa 1		+0,01	+0,49	-5,98	+1,94	+1,97	-2,43	+5,98

E.

Os valores em negrito nas Tabelas 13, 15, 17 e 17 representam os valores que ficaram acima da média geral entre todos os Pares (Termo de Pesquisa x Algoritmo de Ordenação) para o valor médio referente à coluna da Etapa 1 (Tabela 12) em que se encontra (para que fique em evidência os devidos aumentos e reduções em relação à Etapa 1 – ou seja, em comparação a um usuário sem *login* efetuado no YouTube e, por consequência, sem um histórico de uso).

6.3.1 Perfil Desinformativo

A Tabela 13 apresenta todos os resultados encontrados na Etapa 3 para o Perfil Desinformativo. Como é possível observar, apenas dois pares apresentaram um SERP-MS positivo (tendendo à desinformação): (Termo de Pesquisa 2 – “bolsonaro

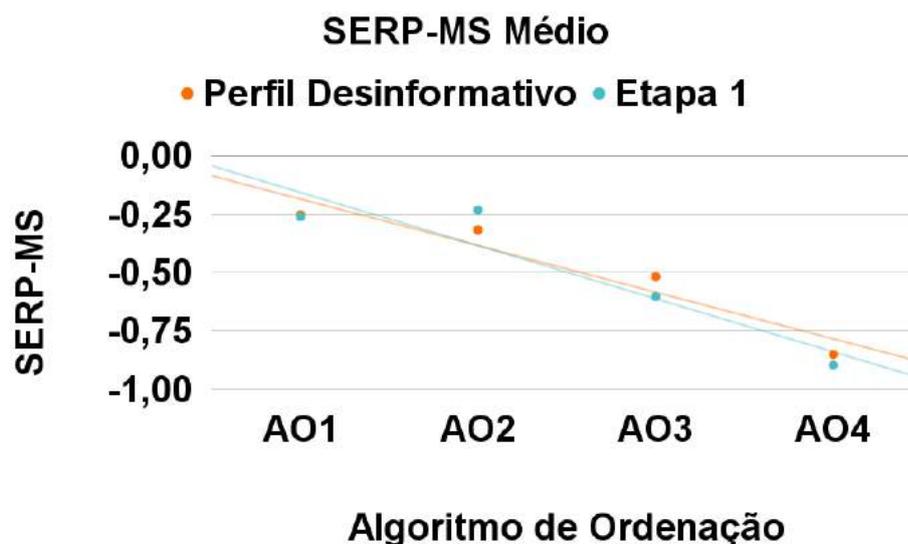


Figura 29 – Comparação do SERP-MS médio entre o Perfil Desinformativo e a Etapa 1, utilizando como parâmetro a média entre os TPs para cada AO

urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação: 0,42) e (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação: 0,2), o que apresenta uma redução de dois pares quando comparado à Etapa 1. Ocorre, porém, que quando comparado com o valor médio do SERP-MS da Etapa 1, dez pares apresentaram um valor acima da média da Etapa 1 (um aumento de dois pares). Observa-se também, prontamente, que a diferença média no SERP-MS foi de apenas +0,01 em relação à Etapa 1, ou seja, **em média, a diferença em relação à desinformação encontrada nos resultados de pesquisa é negligenciável.**

Já em relação à distribuição percentual dos vídeos recomendados, pode-se notar **um aumento geral na quantidade de vídeos desinformativos recomendados (nove pares acima da média na Etapa 1 x quatorze pares neste Perfil) e redução na quantidade de vídeos informativos recomendados (nove pares acima da média na Etapa 1 x cinco pares neste Perfil).** Os resultados serão detalhados nas Subseções a seguir.

6.3.1.1 Resultados dos Vídeos de Pesquisa

As Figuras 29 e 30 apresentam as médias de SERP-MS entre todos os TPs para cada Algoritmo de Ordenação e entre todos os Algoritmos de Ordenação para cada Termo de Pesquisa, respectivamente, do Perfil Desinformativo em comparação com os resultados obtidos na Etapa 1, além de apresentar as curvas de tendência de ambos. Com estas Figura, possibilita-se uma comparação visual dos resultados para os vídeos de pesquisa deste Perfil e da Etapa 1, colocando em evidência cada um dos Algoritmos de Ordenação e cada um dos Termos de Pesquisa.

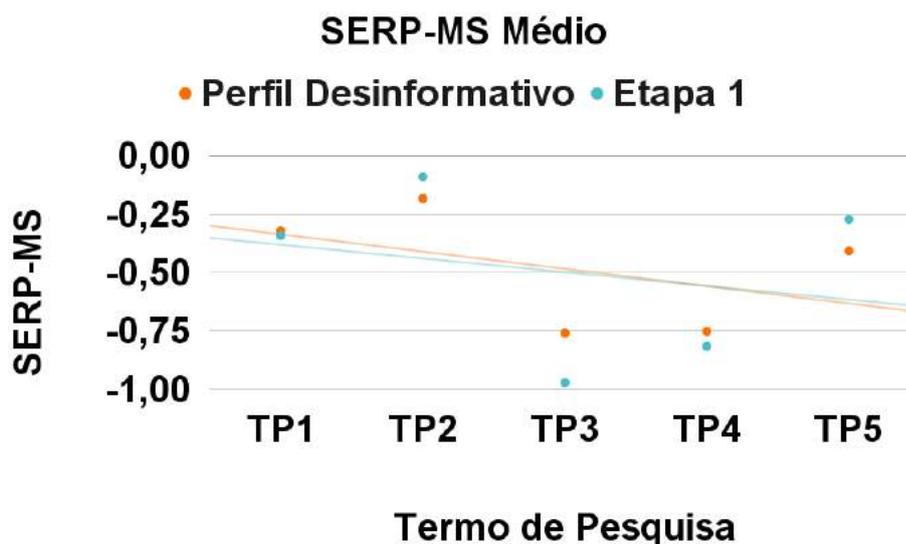


Figura 30 – Comparação do SERP-MS médio entre o Perfil Desinformativo e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP

Como é possível observar na Figura 29, quando são considerados os SERP-MS médios de cada um dos Algoritmos de Ordenação (média entre todos os Termos de Pesquisa para cada Algoritmo de Ordenação), o Perfil Desinformativo apresenta um SERP-MS médio maior que o da Etapa 1 para o Algoritmo de Ordenação 1 – Classificação (-0,252 x -0,262: +0,01), o Algoritmo de Ordenação 3 – Data de envio (-0,518 x -0,604: +0,086) e o Algoritmo de Ordenação 4 – Relevância (-0,852 x -0,898: +0,046). Já no Algoritmo de Ordenação 2 – Contagem de visualizações, a Etapa 1 apresenta um SERP-MS médio maior (-0,232 x -0,318: +0,086). Considerando as linhas de tendência, o R^2 calculado para a Etapa 1 é de 0,873 e o do Perfil Desinformativo é de 0,918 – uma diferença de 0,045.

Já ao observar a Figura 30, quando são considerados os SERP-MS médios de cada um dos Termos de Pesquisa (média entre todos os Algoritmos de Ordenação para cada Termo de Pesquisa), o Perfil Desinformativo apresenta um SERP-MS médio maior que o da Etapa 1 para o Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” (-0,3225 x -0,3425: +0,02), o Termo de Pesquisa 3 – “nova urna eletronica 2022” (-0,76 x -0,9725: +0,2125) e o Termo de Pesquisa 4 – “urna eletrônica” (-0,7525 x -0,8175: +0,065). Já a Etapa 1 apresenta um SERP-MS médio maior para o Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” (-0,09 x -0,1825: +0,0925) e o Termo de Pesquisa 5 – “urnas eletronicas” (-0,2725 x -0,4075: +0,135). Considerando as linhas de tendência, o R^2 calculado para a Etapa 1 é de 0,061 e o do Perfil Desinformativo é de 0,202 – uma diferença de 0,141.

Como observável em ambas as Figuras, as linhas de tendência se cruzam aproximadamente no ponto médio do gráfico. Isso, somado às diferenças de R^2 calcula-

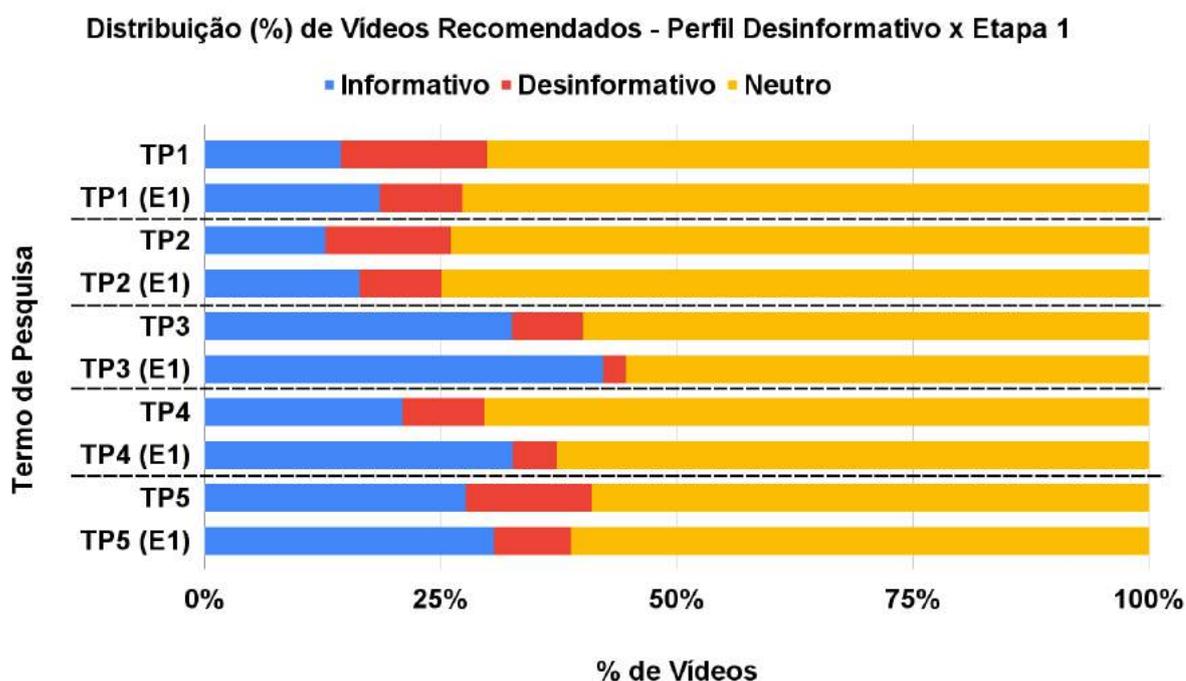


Figura 31 – Comparação da distribuição percentual média nos Vídeos Recomendados entre o Perfil Desinformativo e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP

das, levam à conclusão de que **em média, os resultados de pesquisa para o Perfil Desinformativo têm uma diferença negligenciável em relação à desinformação apresentada na Etapa 1. Ou seja, a presença de um histórico de visualização majoritariamente desinformativo não leva a uma presença maior de desinformação nos resultados de pesquisa** (o que se reflete no aumento de apenas 0,01 no SERP-MS médio em relação à Etapa 1, aumento de 0,49% na média de vídeos recomendados informativos, 1,94% para neutros e redução de 2,43% para desinformativos).

6.3.1.2 Resultados dos Vídeos Recomendados

A Figura 31 ilustra as distribuições percentuais de cada categoria de vídeo (informativo, neutro e desinformativo) entre os vídeos recomendados (levando em consideração a média entre todos os Algoritmos de Ordenação para cada Termo de Pesquisa) deste Perfil e os da Etapa 1. Com esta Figura, possibilita-se uma comparação visual dos resultados para os vídeos recomendados deste Perfil (barras superiores de cada divisória) e da Etapa 1 (barras inferiores de cada divisória), colocando em evidência cada um dos TPs.

Quanto à distribuição percentual de vídeos recomendados, é possível observar na Tabela 13 que entre as categorias ocorre a mesma ordem de distribuição observada na

Etapa 1: Neutros (66,51%) > Informativos (21,7%) > Desinformativos (11,63%). Mas **fica evidente a redução na quantidade de vídeos informativos recomendados (-5,98%) e o aumento da quantidade de vídeos desinformativos recomendados (+5,98%) – inclusive, essa redução e aumento ocorrem de maneira simétrica no Perfil Desinformativo.**

Aqui, os quatorze pares que apresentam um percentual de vídeos desinformativos acima da média da Etapa 1 (6,45%) são: (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 1 – Classificação: 12,37%), (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 33,68%), (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 3 – Data de envio: 6,96%), (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 4 – Relevância: 9,24%), (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação: 6,8%), (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 22,9%), (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 4 – Relevância: 17,2%), (Termo de Pesquisa 3 – “nova urna eletronica 2022” x Algoritmo de Ordenação 2 – Contagem de visualizações: 11,3%), (Termo de Pesquisa 3 – “nova urna eletronica 2022” x Algoritmo de Ordenação 4 – Relevância: 7,61%), (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 2 – Contagem de visualizações: 14,18%), (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 3 – Data de envio: 10,92%), (Termo de Pesquisa 5 – “urnas eletronicas” x Algoritmo de Ordenação 1 – Classificação: 10,18%), (Termo de Pesquisa 5 – “urnas eletronicas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 29,31%) e (Termo de Pesquisa 5 – “urnas eletronicas” x Algoritmo de Ordenação 3 – Data de envio: 8,84%). Ou seja, cada Termo de Pesquisa aparece em pelo menos dois Algoritmos de Ordenação com percentual maior que o percentual médio obtido por aquele par na Etapa 1, o que indica um **aumento generalizado na desinformação recomendada** (com destaque para o Termo de Pesquisa 1 – “bolsonaro embaixadores urnas”, que obteve uma média maior para todos os quatro Algoritmos de Ordenação).

Ao observar a Figura 31, este aumento generalizado na desinformação recomendada fica evidente. Quando é considerada a média entre os Algoritmos de Ordenação para cada Termo de Pesquisa, a distribuição percentual média apresenta-se como demonstrado na Tabela 14 (E1 - Etapa 1 e PD - Perfil Desinformativo). Fica evidente, portanto, que, **para um perfil com um histórico de visualização majoritariamente desinformativo, a recomendação de vídeos desinformativos aumenta entre todos os Termos de Pesquisa e Algoritmos de Ordenação (com um aumento médio de 5,98% e máximo de 6,8% – Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” – comparado a um usuário sem histórico), ao passo que a recomendação**

Tabela 14 – Comparação entre as distribuições percentuais médias para cada TP entre a Etapa 1 e a Etapa 3 (Termo de Pesquisa – TP, Vídeos Neutros – VN, Vídeos Informativos – VI, Vídeos Desinformativos – VD) – Perfil Desinformativo (PD)

TP (Etapa)	VN Média (%)	VI Média (%)	VD Média (%)
TP1 (E1)	72,7	18,5	8,7
TP1 (PD)	(-2,6)	(-4,1)	(+6,8)
TP2 (E1)	74,9	16,4	8,7
TP2 (PD)	(-0,9)	(-3,6)	(+4,5)
TP3 (E1)	55,4	42,2	2,4
TP3 (PD)	(+4,5)	(-9,6)	(+5,1)
TP4 (E1)	62,7	32,6	4,7
TP4 (PD)	(+7,7)	(-11,6)	(+3,8)
TP5 (E1)	61,2	30,6	8,2
TP5 (PD)	(-2,1)	(-2,9)	(+5)

Tabela 15 – Percentual de vídeos classificados na Etapa 3 e SERP-MS dos pares Termo de Pesquisa x Algoritmo de Ordenação (Termo de Pesquisa – TP, Algoritmo de Ordenação – AO, Vídeos Informativos – VI, Vídeos Neutros – VN, Vídeos Desinformativos – VD, Pesquisa – P e Recomendados – R) – Perfil Informativo

TP	AO	SERP-MS	VI (%)		VN (%)		VD (%)	
			P	R	P	R	P	R
TP1	AO1	-0,29	55,5	36,84	16,66	56,31	27,77	6,84
TP1	AO2	0	55	31,08	0	54,92	45	13,99
TP1	AO3	-0,51	92,85	34,37	0	62,5	7,14	3,12
TP1	AO4	-0,69	85	44,21	5	51,58	10	4,21
TP2	AO1	0,57	31,58	30,58	0	63,53	68,42	5,88
TP2	AO2	-0,15	50	19,67	27,77	62,84	22,22	17,49
TP2	AO3	-0,22	44,4	29,55	16,66	66,66	38,88	3,77
TP2	AO4	-0,68	75	45,21	15	46,8	10	7,98
TP3	AO1	-0,96	94,74	47,85	0	51,43	5,26	0,71
TP3	AO2	-0,83	95	34,61	0	58,33	5	7,05
TP3	AO3	-0,54	93,75	35,93	0	55,47	6,25	8,59
TP3	AO4	-1	100	61,5	0	36,5	0	2
TP4	AO1	-0,64	80	33,53	5	59,41	15	7,06
TP4	AO2	-0,58	84,21	32,54	0	52,38	15,79	15,08
TP4	AO3	-0,8	100	28,18	0	67,4	0	4,42
TP4	AO4	-0,85	80	41,71	20	55,08	0	3,2
TP5	AO1	0,41	25	35,75	5	63,13	70	1,11
TP5	AO2	-0,25	61,11	25,16	0	59,11	38,88	15,72
TP5	AO3	-0,66	88,23	43,66	0	52,11	11,76	4,22
TP5	AO4	-0,94	85	45,73	15	46,73	0	7,54
Média		-0,48	73,81	36,9	6,3	56,11	19,86	6,99
Diferença Média Etapa 1		+0,01	-1,5	+9,04	+3,53	-8,43	-2,04	+0,54

de vídeos informativos reduz (com uma redução média de 5,98% e máxima de 11,6% – Termo de Pesquisa 4 – “urna eletrônica” – comparado a um usuário sem histórico).

6.3.2 Perfil Informativo

A Tabela 15 apresenta todos os resultados encontrados na Etapa 3 para o Perfil Informativo. Como é possível observar, aqui mantém-se a quantidade de dois pares apresentando um SERP-MS positivo (tendendo à desinformação): (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação: 0,57) e (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação: 0,41), os mesmos dois pares que tiveram resultado positivo no Perfil Desinformativo. Portanto, novamente há uma redução de dois pares quando compa-

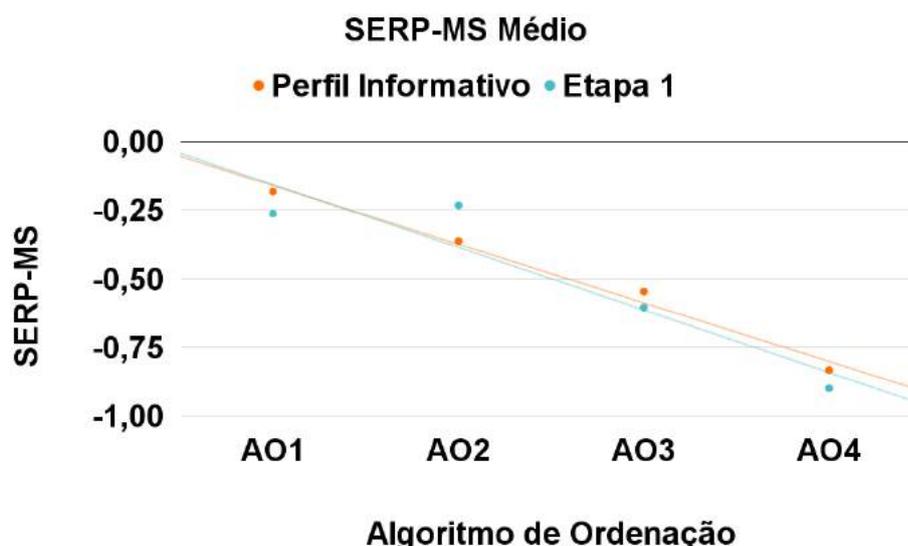


Figura 32 – Comparação do SERP-MS médio entre o Perfil Informativo e a Etapa 1, utilizando como parâmetro a média entre os TPs para cada AO

rado à Etapa 1. Ocorre, porém, que quando comparado com o valor médio do SERP-MS da Etapa 1, sete pares apresentaram um valor acima da média da Etapa 1 (uma redução de um par). Observa-se também, prontamente, que a diferença média no SERP-MS foi de apenas +0,01 em relação à Etapa 1 (mesma diferença observada no Perfil Desinformativo), ou seja, **em média, a diferença em relação à desinformação encontrada nos resultados de pesquisa é negligenciável.**

Já em relação à distribuição percentual dos vídeos recomendados, pode-se notar um **pequeno aumento na quantidade de vídeos desinformativos recomendados (nove pares acima da média na Etapa 1 x dez pares neste Perfil)** e um **aumento significativo na quantidade de vídeos informativos recomendados (nove pares acima da média na Etapa 1 x dezenove pares neste Perfil – ou seja, todos os pares com exceção do (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 25,16%))**. Os resultados serão detalhados nas Subseções a seguir.

6.3.2.1 Resultados dos Vídeos de Pesquisa

As Figuras 32 e 33 apresentam as médias de SERP-MS entre todos os Termos de Pesquisa para cada Algoritmo de Ordenação e entre todos os Algoritmos de Ordenação para cada Termo de Pesquisa, respectivamente, do Perfil Informativo em comparação com os resultados obtidos na Etapa 1, além de apresentar as curvas de tendência de ambos. Com estas Figura, possibilita-se uma comparação visual dos resultados para os vídeos de pesquisa deste Perfil e da Etapa 1, colocando em evidência cada um dos Algoritmos de Ordenação e cada um dos Termos de Pesquisa.

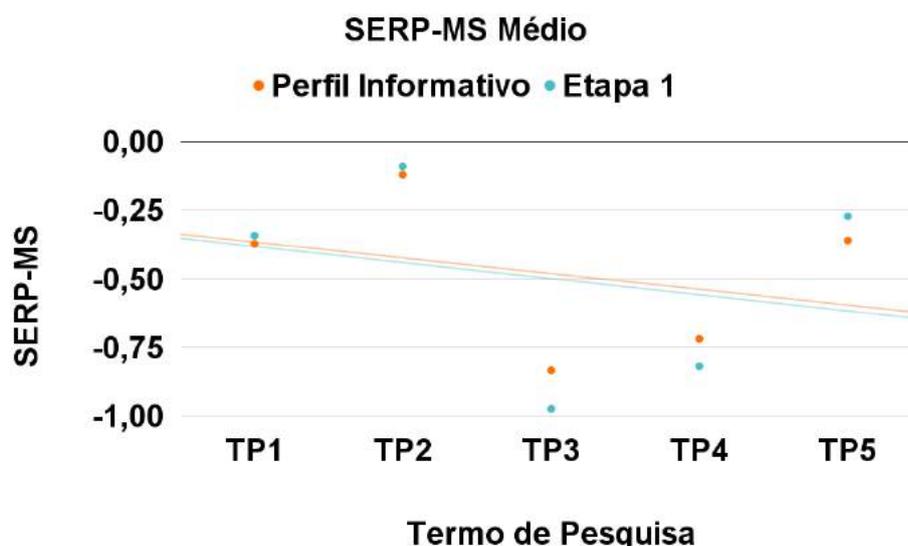


Figura 33 – Comparação do SERP-MS médio entre o Perfil Informativo e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP

Como é possível observar na Figura 32, quando são considerados os SERP-MS médios de cada um dos Algoritmos de Ordenação (média entre todos os Termos de Pesquisa para cada Algoritmo de Ordenação), o Perfil Informativo apresenta novamente um SERP-MS médio maior que o da Etapa 1 para o Algoritmo de Ordenação 1 – Classificação (-0,182 x -0,262: +0,08), o Algoritmo de Ordenação 3 – Data de envio (-0,546 x -0,604: +0,058) e o Algoritmo de Ordenação 4 – Relevância (-0,832 x -0,898: +0,066). Já no Algoritmo de Ordenação 2 – Contagem de visualizações, a Etapa 1 apresenta um SERP-MS médio maior (-0,232 x -0,362: +0,13). Considerando as linhas de tendência, o R^2 calculado para a Etapa 1 é de 0,873 e o do Perfil Informativo é de 0,986 – uma diferença de 0,113.

Já ao observar a Figura 33, quando são considerados os SERP-MS médios de cada um dos Termos de Pesquisa (média entre todos os Algoritmos de Ordenação para cada Termo de Pesquisa), o Perfil Informativo apresenta um SERP-MS médio maior que o da Etapa 1 apenas para o Termo de Pesquisa 3 – “nova urna eletronica 2022” (-0,8325 x -0,9725: +0,14) e o Termo de Pesquisa 4 – “urna eletrônica” (-0,7175 x -0,8175: +0,1). Já a Etapa 1 apresenta um SERP-MS médio maior para o Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” (-0,3425 x -0,3725: +0,0273), o Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” (-0,09 x -0,12: +0,03) e o Termo de Pesquisa 5 – “urnas eletronicas” (-0,2725 x -0,36: +0,0875). Considerando as linhas de tendência, o R^2 calculado para a Etapa 1 é de 0,061 e o do Perfil Informativo é de 0,097 – uma diferença de 0,036.

Como observável em ambas as figuras, as linhas de tendência correm paralelamente, sem uma diferença significativa em qualquer ponto. Isso, somado às dife-

renças de R^2 calculadas, levam à conclusão de que **em média, os resultados de pesquisa para o Perfil Informativo têm uma diferença negligenciável em relação à desinformação apresentada na Etapa 1. Ou seja, a presença de um histórico de visualização majoritariamente informativo não leva a uma presença maior de desinformação nos resultados de pesquisa** (o que se reflete no aumento de apenas 0,01 no SERP-MS médio em relação à Etapa 1, uma redução de 1,5% na média de vídeos recomendados informativos, um aumento de 3,53% para neutros e uma redução de 2,04% para desinformativos).

Outro fenômeno que cabe destacar aqui é a inversão em relação ao Perfil Desinformativo: a diferença entre as linhas de tendência é maior para os Algoritmos de Ordenação que para os Termos de Pesquisa. Ou seja, mesmo tendo um aumento de desinformação negligenciável em relação à Etapa 1, **os Algoritmos de Ordenação apresentam aqui uma influência maior na quantidade de desinformação nos resultados de pesquisa, ao passo que no Perfil Desinformativo, os Termos de Pesquisa apresentam uma influência maior.**

6.3.2.2 Resultados dos Vídeos Recomendados

A Figura 34 ilustra as distribuições percentuais de cada categoria de vídeo (informativo, neutro e desinformativo) entre os vídeos recomendados (levando em consideração a média entre todos os Algoritmos de Ordenação para cada Termo de Pesquisa) deste Perfil e os da Etapa 1. Com esta Figura, possibilita-se uma comparação visual dos resultados para os vídeos recomendados deste Perfil (barras superiores de cada divisória) e da Etapa 1 (barras inferiores de cada divisória), colocando em evidência cada um dos Termos de Pesquisa.

Quanto à distribuição percentual de vídeos recomendados, é possível observar na Tabela 15 que ocorre a mesma ordem de distribuição categórica observada na Etapa 1 e no Perfil Desinformativo: Neutros (56,11%) > Informativos (36,9%) > Desinformativos (6,99%). Mas **fica evidente o aumento na quantidade de vídeos informativos recomendados (+9,04%) e um pequeno aumento da quantidade de vídeos desinformativos recomendados (+0,54%).**

Aqui, os dez pares que apresentam um percentual de vídeos desinformativos acima da média da Etapa 1 (6,45%) são: (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 1 – Classificação: 6,84%), (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 13,99%), (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 17,49%), (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 4 – Relevância: 7,98%), (Termo de Pesquisa 3 – “nova urna eletrônica 2022” x Algoritmo de Ordenação 2 – Contagem de visualizações: 7,05%), (Termo de Pesquisa 3 – “nova urna

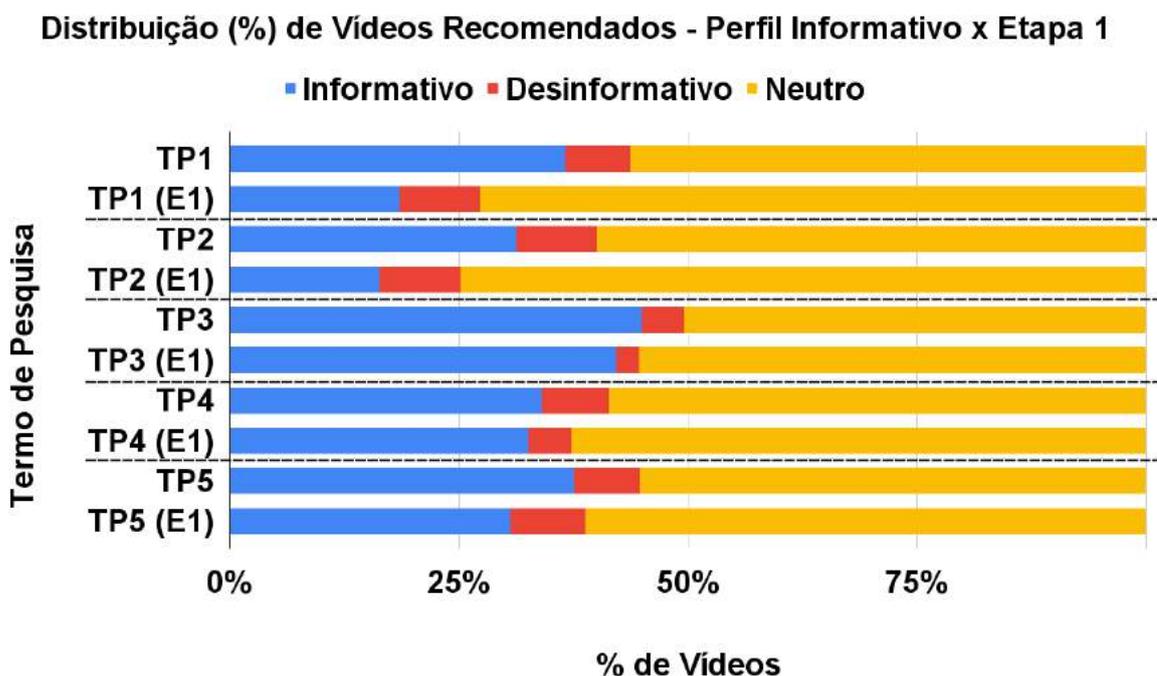


Figura 34 – Comparação da distribuição percentual média nos Vídeos Recomendados entre o Perfil Informativo e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP

eletronica 2022” x Algoritmo de Ordenação 3 – Data de envio: 8,59%), (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 1 – Classificação: 7,06%), (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 2 – Contagem de visualizações: 15,08%), (Termo de Pesquisa 5 – “urnas eletronicas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 15,72%) e (Termo de Pesquisa 5 – “urnas eletronicas” x Algoritmo de Ordenação 4 – Relevância: 7,54%). Ou seja, todos os Termos de Pesquisa e Algoritmos de Ordenação aparecem em pelo menos um par com percentual maior que o percentual médio obtido por aquele par na Etapa 1, o que indica um **aumento generalizado na desinformação recomendada, mesmo que menor quando comparado ao Perfil Desinformativo.**

O destaque maior aqui, porém, fica por conta dos vídeos informativos recomendados: todos os pares apresentaram valores acima da média obtida na Etapa 1 (27,86%), com exceção do (Termo de Pesquisa 5 – “urnas eletronicas” x Algoritmo de Ordenação 2 – Contagem de visualizações), que obteve um valor próximo (25,16%). Ou seja, **o aumento de vídeos informativos recomendados supera muito o aumento de vídeos desinformativos.**

Ao observar a Figura 34, este aumento generalizado na informação recomendada fica evidente. Quando é considerada a média entre os Algoritmos de Ordenação para cada Termo de Pesquisa, a distribuição percentual média apresenta-se como demons-

Tabela 16 – Comparação entre as distribuições percentuais médias para cada TP entre a Etapa 1 e a Etapa 3 (Termo de Pesquisa – TP, Vídeos Neutros – VN, Vídeos Informativos – VI, Vídeos Desinformativos – VD) – Perfil Informativo (PI)

TP (Etapa)	VN Média (%)	VI Média (%)	VD Média (%)
TP1 (E1)	72,7	18,5	8,7
TP1 (PI)	56,3 (-16,4)	36,6 (+18,1)	7 (-1,7)
TP2 (E1)	74,9	16,4	8,7
TP2 (PI)	60 (-14,9)	31,3 (+10,9)	8,8 (+0,1)
TP3 (E1)	55,4	42,2	2,4
TP3 (PI)	50,4 (-5)	45 (+2,8)	4,6 (+2,2)
TP4 (E1)	62,7	32,6	4,7
TP4 (PI)	58,6 (-4,1)	34 (+1,4)	7,4 (+2,7)
TP5 (E1)	61,2	30,6	8,2
TP5 (PI)	55,3 (-5,9)	37,6 (+7)	7,1 (-1,1)

trado na Tabela 16 (E1 - Etapa 1 e PI - Perfil Informativo). Fica evidente, portanto, que, **para um perfil com um histórico de visualização majoritariamente informativo, a recomendação de vídeos informativos aumenta entre todos os Termos de Pesquisa e Algoritmos de Ordenação (com um aumento médio de 9,04% e máximo de 18,1% – Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” – comparado a um usuário sem histórico), a passo que a recomendação de vídeos desinformativos também aumenta, porém em menor grau (com um aumento médio de 5,98% e máxima de 2,7% – Termo de Pesquisa 4 – “urna eletrônica” – comparado a um usuário sem histórico).**

É interessante destacar que aqui, mais uma vez, ocorre o oposto entre o Perfil Desinformativo e o Perfil Informativo: **o Termo de Pesquisa com o maior aumento médio de vídeos desinformativos recomendados no Perfil Desinformativo (TP1: 6,8%) é o que apresenta o maior aumento médio de vídeos informativos recomendados no Perfil Informativo (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas”: 18,1%).** Outro fenômeno que ocorre é que **o Termo de Pesquisa com a maior redução de vídeos informativos recomendados no Perfil Desinformativo (Termo de Pesquisa 4 – “urna eletrônica”: 11,6%) é o que apresenta, no Perfil Informativo, o menor aumento no percentual de vídeos informativos recomendados (Termo de Pesquisa 4 – “urna eletrônica”: 1,4%) e maior aumento percentual de vídeos desinformativos recomendados (Termo de Pesquisa 4 – “urna eletrônica”: 2,7%).**

6.3.3 Perfil Misto

A Tabela 17 apresenta todos os resultados encontrados na Etapa 3 para o Perfil Misto. Como é possível observar, aqui mantém-se a quantidade de dois pares apresentando um SERP-MS positivo (tendendo à desinformação): (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação: 0,34) e (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação: 0,28), os mesmos dois pares que tiveram resultado positivo no Perfil Desinformativo e no Perfil Informativo. Portanto, novamente existe uma redução de dois pares quando comparado à Etapa 1. Ocorre, novamente, como no Perfil Informativo, que quando comparado com o valor médio do SERP-MS da Etapa 1, sete pares apresentaram um valor acima da média da Etapa 1 (uma redução de um par). Observa-se também, prontamente, que a diferença média no SERP-MS foi de apenas -0,01 em relação à Etapa 1, ou seja, **em média, a diferença em relação à desinformação encontrada nos resultados de pesquisa é negligenciável.**

Já em relação à distribuição percentual dos vídeos recomendados, pode-se notar uma **redução geral na quantidade de vídeos desinformativos recomendados (nove pares acima da média na Etapa 1 x sete pares neste Perfil)** (mesmo havendo

Tabela 17 – Percentual de vídeos classificados na Etapa 3 e SERP-MS dos pares Termo de Pesquisa x Algoritmo de Ordenação (Termo de Pesquisa – TP, Algoritmo de Ordenação – AO, Vídeos Informativos – VI, Vídeos Neutros – VN, Vídeos Desinformativos – VD, Pesquisa – P e Recomendados – R) – Perfil Misto

TP	AO	SERP-MS	VI (%)		VN (%)		VD (%)	
			P	R	P	R	P	R
TP1	AO1	-0,18	55,5	21,05	16,66	75,26	27,77	3,68
TP1	AO2	0	55	10,71	0	55,61	45	33,67
TP1	AO3	-0,12	92,3	22,43	0	76,92	7,69	0,64
TP1	AO4	-0,66	85	26,88	5	69,35	10	3,76
TP2	AO1	0,34	36,84	17,61	0	75,47	63,16	6,92
TP2	AO2	-0,16	47,36	11,38	26,31	61,07	26,31	27,54
TP2	AO3	-0,48	70,59	14,65	5,88	84,71	23,53	0,64
TP2	AO4	-0,74	90	50,79	0	39,68	10	9,52
TP3	AO1	-0,92	90	48,66	0	50,66	10	0,66
TP3	AO2	-0,83	95	23,53	0	56,47	5	20
TP3	AO3	-0,49	88,23	31,39	0	64,96	11,76	3,65
TP3	AO4	-1	100	68,18	0	27,77	0	4,04
TP4	AO1	-0,8	85	36,78	0	61,49	15	1,72
TP4	AO2	-0,54	80	18,52	0	65,18	20	16,3
TP4	AO3	-0,83	85	12,16	10	87,16	5	0,67
TP4	AO4	-1	100	60,71	0	36,73	0	2,55
TP5	AO1	0,28	40	15,55	5	79,44	55	5
TP5	AO2	-0,25	66,66	12,2	0	55,81	33,33	31,98
TP5	AO3	-0,79	84,21	31,43	0	67,86	15,79	0,71
TP5	AO4	-0,99	95	58,88	0	35,02	5	6,09
Média		-0,5	77,08	29,67	3,44	61,33	19,57	8,99
Diferença Média Etapa 1		-0,01	+1,77	+1,81	+0,67	-3,21	-2,33	+2,54

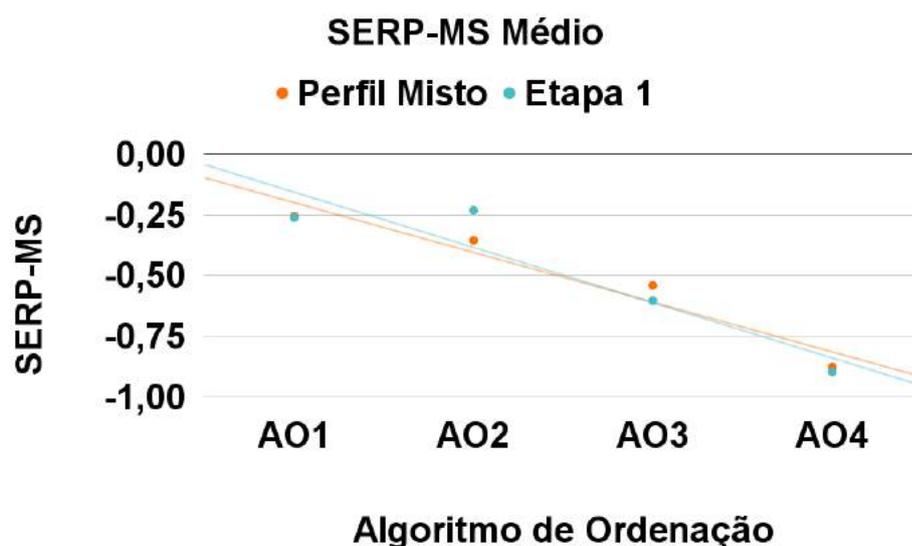


Figura 35 – Comparação do SERP-MS médio entre o Perfil Misto e a Etapa 1, utilizando como parâmetro a média entre os TPs para cada AO

um aumento médio de 2,54% quando considerados todos os pares) e uma **pequena redução geral na quantidade de vídeos informativos recomendados (nove pares acima da média na Etapa 1 x oito pares neste Perfil)** (mesmo havendo um aumento médio de 1,81% quando considerados todos os Pares). Os resultados serão detalhados nas subseções a seguir.

6.3.3.1 Resultados dos Vídeos de Pesquisa

As Figuras 35 e 36 apresentam as médias de SERP-MS entre todos os Termos de Pesquisa para cada Algoritmo de Ordenação e entre todos os Algoritmos de Ordenação para cada Termo de Pesquisa, respectivamente, do Perfil Misto em comparação com os resultados obtidos na Etapa 1, além de apresentar as curvas de tendência de ambos. Com estas Figura, possibilita-se uma comparação visual dos resultados para os vídeos de pesquisa deste Perfil e da Etapa 1, colocando em evidência cada um dos Algoritmos de Ordenação e cada um dos Termos de Pesquisa.

Como é possível observar na Figura 35, quando considerados os SERP-MS médios de cada um dos Algoritmos de Ordenação (média entre todos os Termos de Pesquisa para cada Algoritmo de Ordenação), o Perfil Misto apresenta, novamente, um SERP-MS médio maior que o da Etapa 1 para o Algoritmo de Ordenação 1 – Classificação (-0,256 x -0,262: +0,006), o Algoritmo de Ordenação 3 – Data de envio (-0,542 x -0,604: +0,062) e o Algoritmo de Ordenação 4 – Relevância (-0,878 x -0,898: +0,02) – os mesmos Algoritmos de Ordenação encontrados nos Perfis Desinformativo e Informativo. Já no Algoritmo de Ordenação 2 – Contagem de visualizações, a Etapa 1 apresenta um SERP-MS médio maior (-0,232 x -0,356: +0,124). Considerando as

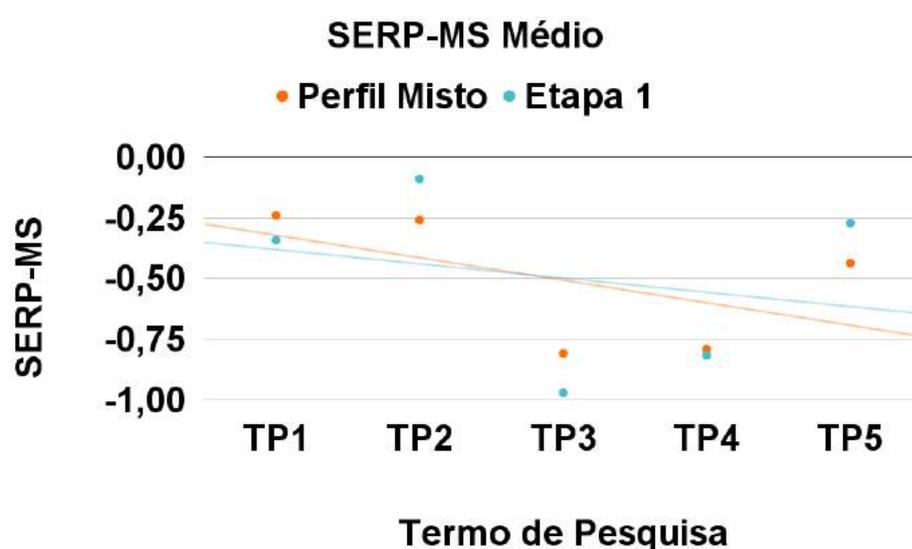


Figura 36 – Comparação do SERP-MS médio entre o Perfil Misto e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP

linhas de tendência, o R^2 calculado para a Etapa 1 é de 0,873 e o do Perfil Misto é de 0,937 – uma diferença de 0,1.

Já ao observar a Figura 36, quando considerados os SERP-MS médios de cada um dos Termos de Pesquisa (média entre todos os Algoritmos de Ordenação para cada Termo de Pesquisa), o Perfil Misto apresenta um SERP-MS médio maior que o da Etapa 1 para o Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” (-0,24 x -0,3425: +0,1025), o Termo de Pesquisa 3 – “nova urna eletrônica 2022” (-0,81 x -0,9725: +0,1625) e o Termo de Pesquisa 4 – “urna eletrônica” (-0,7925 x -0,8175: +0,025) – os mesmos Termos de Pesquisa encontrados no Perfil Desinformativo. Já a Etapa 1 apresenta um SERP-MS médio maior para o Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” (-0,09 x -0,26: +0,17) e o Termo de Pesquisa 5 – “urnas eletrônicas” (-0,2725 x -0,4375: +0,165). Considerando as linhas de tendência, o R^2 calculado para a Etapa 1 é de 0,061 e o do Perfil Misto é de 0,277 – uma diferença de 0,216.

Como observável em ambas as figuras, as linhas de tendência se cruzam aproximadamente no ponto médio do gráfico. Isso, somado às diferenças de R^2 calculadas, levam à conclusão de que **em média, os resultados de pesquisa para o Perfil Misto têm uma diferença negligenciável em relação à desinformação apresentada na Etapa 1. Ou seja, a presença de um histórico de visualização dividido entre vídeos informativos e desinformativos não leva a uma presença maior de desinformação nos resultados de pesquisa** (o que se reflete na redução de apenas 0,01 no SERP-MS médio em relação à Etapa 1, aumento de 1,5% na média de vídeos recomendados informativos, aumento de 0,67% para neutros e redução de 2,33% para

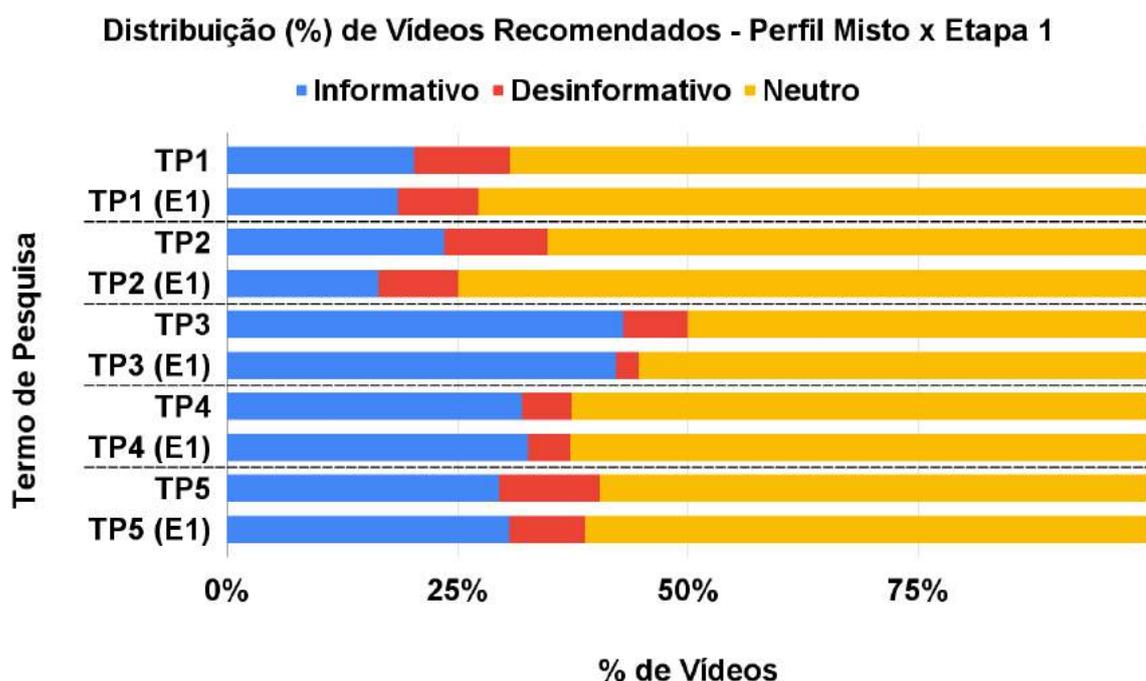


Figura 37 – Comparação da distribuição percentual média nos Vídeos Recomendados entre o Perfil Misto e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP

desinformativos).

Observa-se também que **o comportamento do Perfil Misto nos resultados de pesquisa tende mais ao comportamento observado no Perfil Desinformativo** (diferença maior entre R^2 s das linhas de tendência nos Termos de Pesquisa) que no Perfil Informativo (diferença maior entre R^2 s das linhas de tendência nos Algoritmos de Ordenação), mesmo com uma diferença média negligenciável da quantidade de desinformação quando comparado à Etapa 1.

6.3.3.2 Resultados dos Vídeos Recomendados

A Figura 37 ilustra as distribuições percentuais de cada categoria de vídeo (informativo, neutro e desinformativo) entre os vídeos recomendados (levando em consideração a média entre todos os Algoritmos de Ordenação para cada Termo de Pesquisa) deste Perfil e os da Etapa 1. Com estas Figura, possibilita-se uma comparação visual dos resultados para os vídeos recomendados deste Perfil (barras superiores de cada divisória) e da Etapa 1 (barras inferiores de cada divisória), colocando em evidência cada um dos Termos de Pesquisa.

Quanto à distribuição percentual de vídeos recomendados, é possível observar na Tabela 17 que ocorre a mesma ordem de distribuição categórica observada na Etapa 1 e nos Perfis Desinformativo e Informativo: Neutros (61,33%) > Informativos

(29,67%) > Desinformativos (8,99%). **Neste Perfil ocorre um pequeno aumento na quantidade de vídeos informativos recomendados quando comparado à Etapa 1 (1,81%) e um aumento um pouco maior na quantidade de vídeos desinformativos recomendados (2,54%).**

Aqui, os sete pares que apresentam um percentual de vídeos desinformativos acima da média da Etapa 1 (6,45%) são: (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 33,67%), (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação: 6,92%), (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 27,54%), (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 4 – Relevância: 9,52%), (Termo de Pesquisa 3 – “nova urna eletronica 2022” x Algoritmo de Ordenação 2 – Contagem de visualizações: 20%), (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 2 – Contagem de visualizações: 16,3%) e (Termo de Pesquisa 5 – “urnas eletronicas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 31,98%). **Destaca-se aqui o Algoritmo de Ordenação 2 – Contagem de visualizações por apresentar de maneira constante entre todos os Termos de Pesquisa com um elevado percentual de vídeos desinformativos recomendados (média de 25,9% entre todos os Termos de Pesquisa) e o Algoritmo de Ordenação 3 – Data de envio por apresentar os menores percentuais de vídeos desinformativos recomendados (média de apenas 1,27% entre todos os Termos de Pesquisa). Isso indica um comportamento mais variado no Perfil Misto quando comparado aos Perfis Informativo e Desinformativo.**

Ao observar a Figura 37, este comportamento variado fica evidente. Quando considerada a média entre os Algoritmos de Ordenação para cada Termo de Pesquisa, a distribuição percentual média apresenta-se como demonstrado na Tabela 18 (E1 - Etapa 1 e PM - Perfil Misto). Observa-se aqui o aumento percentual de vídeos desinformativos recomendados para todos os Termos de Pesquisa (mesmo que em menor grau quando comparado com o Perfil Desinformativo), o aumento de vídeos informativos recomendados para o Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” (1,8%), o Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” (7,2%) e o Termo de Pesquisa 3 – “nova urna eletronica 2022” (0,7%) e uma redução de vídeos informativos recomendados para o Termo de Pesquisa 4 – “urna eletrônica” (0,6%) e o Termo de Pesquisa 5 – “urnas eletronicas” (1,1%).

Fica evidente, portanto, que **para um perfil com um histórico de visualização dividido entre vídeos informativos e desinformativos, a recomendação de vídeos desinformativos aumenta entre todos os Termos de Pesquisa e Algoritmos de Ordenação (com um aumento médio de 2,54% e máximo de 4,7% – Termo de Pesquisa 3 – “nova urna eletronica 2022” – comparado a um usuário sem histó-**

Tabela 18 – Comparação entre as distribuições percentuais médias para cada TP entre a Etapa 1 e a Etapa 3 (Termo de Pesquisa – TP, Vídeos Neutros – VN, Vídeos Informativos – VI, Vídeos Desinformativos – VD) – Perfil Misto (PM)

TP (Etapa)	VN Média (%)	VI Média (%)	VD Média (%)
TP1 (E1)	72,7	18,5	8,7
TP1 (PM)	(-3,4)	(+1,8)	(+1,7)
TP2 (E1)	74,9	16,4	8,7
TP2 (PM)	(-9,7)	(+7,2)	(+2,5)
TP3 (E1)	55,4	42,2	2,4
TP3 (PM)	(-5,4)	(+0,7)	(+4,7)
TP4 (E1)	62,7	32,6	4,7
TP4 (PM)	(-0,1)	(-0,6)	(+0,6)
TP5 (E1)	61,2	30,6	8,2
TP5 (PM)	(-1,7)	(-1,1)	(+2,7)

Tabela 19 – Percentual de vídeos classificados na Etapa 3 e SERP-MS dos pares Termo de Pesquisa x Algoritmo de Ordenação (Termo de Pesquisa – TP, Algoritmo de Ordenação – AO, Vídeos Informativos – VI, Vídeos Neutros – VN, Vídeos Desinformativos – VD, Pesquisa – P e Recomendados – R) – Perfil Neutro

TP	AO	SERP-MS	VI (%)		VN (%)		VD (%)	
			P	R	P	R	P	R
TP1	AO1	-0,29	55,5	14,83	16,66	81,87	27,77	3,3
TP1	AO2	0	55	12,37	0	73,19	45	14,43
TP1	AO3	-0,26	85,71	7,79	0	87,66	14,28	4,54
TP1	AO4	-0,63	75	13,37	15	81,28	10	5,35
TP2	AO1	0,36	42,1	15,86	0	79,31	57,89	4,83
TP2	AO2	-0,2	50	10,73	27,77	71,75	22,22	17,51
TP2	AO3	-0,51	68,42	20,77	10,53	76,15	21,05	3,08
TP2	AO4	-0,76	90	33,87	0	58,6	10	7,53
TP3	AO1	-0,93	94,74	17,69	0	81,63	5,26	0,68
TP3	AO2	-0,83	95	28,85	0	62,18	5	8,97
TP3	AO3	-0,4	81,25	15,25	0	79,66	18,75	5,08
TP3	AO4	-1	100	42,93	0	47,98	0	9,09
TP4	AO1	-0,95	95	14,88	0	83,93	5	1,19
TP4	AO2	-0,53	78,95	20	0	63,85	21,05	16,15
TP4	AO3	-0,73	94,12	17,89	0	77,89	5,88	4,21
TP4	AO4	-1	100	45,83	0	48,21	0	5,95
TP5	AO1	-0,18	55	12,24	5	82,99	40	4,76
TP5	AO2	-0,37	70,59	16,25	0	64,37	29,41	19,37
TP5	AO3	-0,85	89,47	15,62	0	76,04	10,53	8,33
TP5	AO4	-0,98	95	41,33	0	53,06	5	5,61
Média		-0,55	78,5	20,92	3,75	71,58	17,7	7,5
Diferença Média Etapa 1		-0,06	+3,19	-6,94	+0,98	+7,04	-4,2	+1,05

rico), ao passo que a recomendação de vídeos informativos aumenta ou diminui, a depender do Termo de Pesquisa e Algoritmo de Ordenação em evidência (com um aumento médio de 1,81% e máximo de 7,2% – Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” – comparado a um usuário sem histórico).

Conclui-se, portanto, que, dentre todos os Perfis, o Perfil Misto é o que se comporta de maneira mais similar a um usuário sem histórico de visualização, apresentando as menores variações em relação à Etapa 1 quando consideradas em conjunto as variações nos vídeos informativos e desinformativos recomendados.

6.3.4 Perfil Neutro

A Tabela 19 apresenta todos os resultados encontrados na Etapa 3 para o Perfil Neutro. Como é possível observar, aqui, diferente dos três perfis anteriores, há apenas um par apresentando um SERP-MS positivo (tendendo à desinformação): (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação: 0,36), enquanto o par (Termo de Pesquisa 5 – “urnas eletrônicas” x Algoritmo de Ordenação 1 – Classificação: -0,18) passa a ter um valor negativo. Portanto, temos uma redução de três pares quando comparado à Etapa 1. A quantidade de pares com SERP-MS acima da média em relação à Etapa 1 é de oito pares, ou seja, a mesma quantidade de Pares da Etapa 1. Aqui, diferentemente dos Perfis anteriores, a diferença média do SERP-MS é de -0,06 em relação à Etapa 1, ou seja, **em média, a diferença em relação à desinformação encontrada nos resultados de pesquisa é pequena, mas não negligenciável.**

Já em relação à distribuição percentual dos vídeos recomendados, pode-se notar uma **pequena redução geral na quantidade de vídeos desinformativos recomendados (nove pares acima da média na Etapa 1 x oito pares neste Perfil)** (mesmo havendo um aumento médio de 1,05% quando consideramos todos os Pares) e uma **significativa redução geral na quantidade de vídeos informativos recomendados (nove pares acima da média na Etapa 1 x cinco pares neste Perfil)** (que se reflete na redução média de 6,94% na quantidade média de vídeos informativos recomendados, quando comparado a Etapa 1). Os resultados serão detalhados nas Subseções a seguir.

6.3.4.1 Resultados dos Vídeos de Pesquisa

As Figuras 38 e 39 apresentam as médias de SERP-MS entre todos os Termos de Pesquisa para cada Algoritmo de Ordenação e entre todos os Algoritmos de Ordenação para cada Termo de Pesquisa, respectivamente, do Perfil Neutro em comparação com os resultados obtidos na Etapa 1, além de apresentar as curvas de tendência de ambos. Com estas Figuras, possibilita-se uma comparação visual dos resultados para os vídeos de pesquisa deste Perfil e da Etapa 1, colocando em evidência cada um dos Algoritmos de Ordenação e cada um dos Termos de Pesquisa.

Como é possível observar na Figura 38, quando são considerados os SERP-MS médios de cada um dos Algoritmos de Ordenação (média entre todos os Termos de Pesquisa para cada Algoritmo de Ordenação), o Perfil Neutro apresenta um SERP-MS médio maior que o da Etapa 1 apenas para os Algoritmos de Ordenação 3 (-0,55 x -0,604: +0,054) e 4 (-0,874 x -0,898: +0,024) – diferentemente dos três perfis anteriores, nos quais o Algoritmo de Ordenação 1 – Classificação também era incluído neste conjunto. Já no Algoritmo de Ordenação 1 – Classificação (-0,262 x -0,398: +0,136) e no Algoritmo de Ordenação 2 – Contagem de visualizações (-0,232 x -0,386:

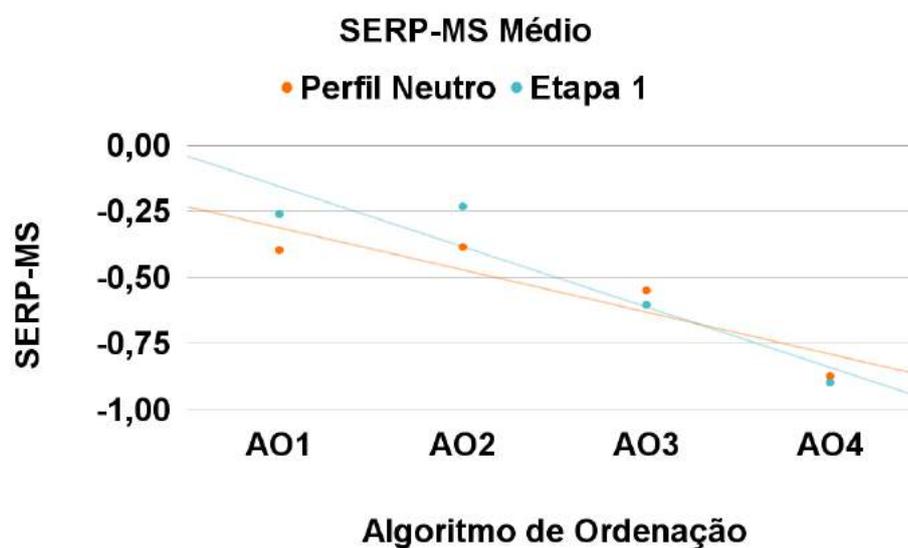


Figura 38 – Comparação do SERP-MS médio entre o Perfil Neutro e a Etapa 1, utilizando como parâmetro a média entre os TPs para cada AO

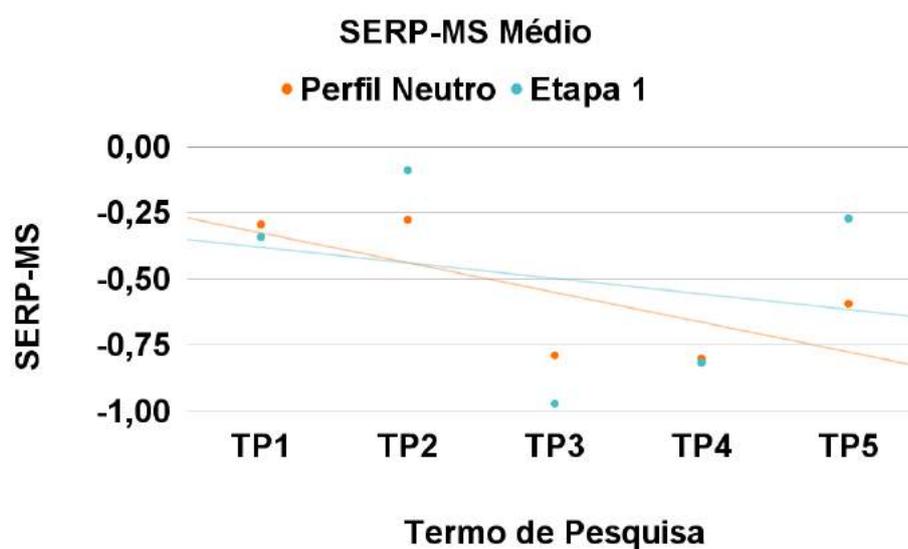


Figura 39 – Comparação do SERP-MS médio entre o Perfil Neutro e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP

+0,154), a Etapa 1 apresenta um SERP-MS médio maior. Considerando as linhas de tendência, o R^2 calculado para a Etapa 1 é de 0,873 e o do Perfil Neutro é de 0,818 – uma diferença de 0,055.

Já ao observar a Figura 39, quando são considerados os SERP-MS médios de cada um dos Termos de Pesquisa (média entre todos os Algoritmos de Ordenação para cada Termo de Pesquisa), o Perfil Neutro apresenta um SERP-MS médio maior que o da Etapa 1 para o Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” (-0,295 x -0,3425: +0,047), o Termo de Pesquisa 3 – “nova urna eletrônica 2022” (-0,79 x -0,9725: +0,1825) e o Termo de Pesquisa 4 – “urna eletrônica” (-0,8025 x -0,8175: +0,015) – os mesmos Termos de Pesquisa encontrados nos Perfis Desinformativo e Misto. Já a Etapa 1 apresenta um SERP-MS médio maior para o Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” (-0,09 x -0,2775: +0,1875) e o Termo de Pesquisa 5 – “urnas eletrônicas” (-0,2725 x -0,595: +0,3225). Considerando as linhas de tendência, o R^2 calculado para a Etapa 1 é de 0,061 e o do Perfil Neutro é de 0,482 – uma diferença de 0,421.

Como observável em ambas as Figuras, as linhas de tendência **não** se cruzam no ponto médio do gráfico. Isso, somado às diferenças de R^2 calculadas, levam à conclusão de que **em média, os resultados de pesquisa para o Perfil Neutro têm uma diferença não negligenciável em relação à desinformação apresentada na Etapa 1. Ou seja, a presença de um histórico de visualização composto por vídeos que não têm ligação com o assunto sendo pesquisado leva a uma redução na quantidade de desinformação nos resultados de pesquisa, em maior ou menor grau, dependendo do referencial utilizado** (o que se reflete na redução de 0,06 no SERP-MS médio em relação à Etapa 1, um aumento de 3,19% na média de vídeos recomendados informativos, um aumento de 0,98% para neutros e uma redução de 4,2% para desinformativos).

No Perfil Neutro, a escolha do Termo de Pesquisa tem um relevante impacto na quantidade de desinformação apresentada nos resultados de pesquisa, quando comparado à Etapa 1 (o que fica explícito com a diferença de 0,421 entre os R^2 das curvas de tendência). Apesar de apresentar o mesmo número de pares com SERP-MS acima da média em relação à Etapa 1 (oito pares), **o percentual de vídeos desinformativos nos resultados de pesquisa é consideravelmente menor quando comparado à Etapa 1 (21,9% x 17,7%: -4,2%)**. Isso fica ainda mais claro dependendo do Termo de Pesquisa utilizado, com o Termo de Pesquisa 5 – “urnas eletrônicas” apresentando a maior variação em comparação à Etapa 1 dentre todas as variações de todos os perfis, com uma redução de 0,3225.

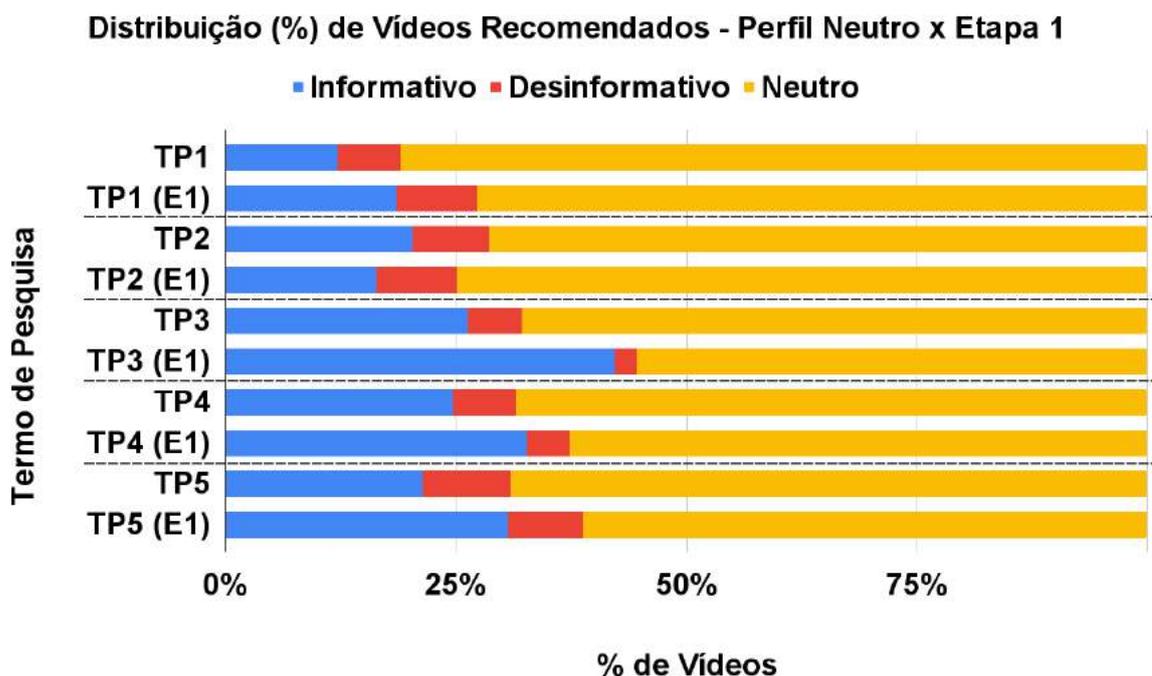


Figura 40 – Comparação da distribuição percentual média nos Vídeos Recomendados entre o Perfil Neutro e a Etapa 1, utilizando como parâmetro a média entre os AOs para cada TP

6.3.4.2 Resultados dos Vídeos Recomendados

A Figura 40 ilustra as distribuições percentuais de cada categoria de vídeo (informativo, neutro e desinformativo) entre os vídeos recomendados (levando em consideração a média entre todos os Algoritmos de Ordenação para cada Termo de Pesquisa) deste Perfil e os da Etapa 1. Com esta Figura, possibilita-se uma comparação visual dos resultados para os vídeos recomendados deste Perfil (barras superiores de cada divisória) e da Etapa 1 (barras inferiores de cada divisória), colocando em evidência cada um dos Termos de Pesquisa.

Quanto à distribuição percentual de vídeos recomendados, é possível observar na Tabela 19 ocorre a mesma ordem de distribuição categórica observada na Etapa 1 e nos Perfis Desinformativo e Informativo: Neutros (71,58%) > Informativos (20,92%) > Desinformativos (7,5%). **Neste Perfil ocorre uma considerável redução na quantidade de vídeos informativos recomendados quando comparado à Etapa 1 (6,94%) e um pequeno aumento na quantidade de vídeos desinformativos recomendados (1,05%) – o segundo menor aumento entre todos os Perfis, atrás apenas do Perfil Informativo.**

Aqui, os oito pares que apresentam um percentual de vídeos desinformativos acima da média da Etapa 1 (6,45%) são: (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 14,43%),

(Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 17,51%), (Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” x Algoritmo de Ordenação 4 – Relevância: 7,53%), (Termo de Pesquisa 3 – “nova urna eletronica 2022” x Algoritmo de Ordenação 2 – Contagem de visualizações: 8,97%), (Termo de Pesquisa 3 – “nova urna eletronica 2022” x Algoritmo de Ordenação 4 – Relevância: 9,09%), (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 2 – Contagem de visualizações: 16,15%), (Termo de Pesquisa 5 – “urnas eletronicas” x Algoritmo de Ordenação 2 – Contagem de visualizações: 19,37%) e (Termo de Pesquisa 5 – “urnas eletronicas” x Algoritmo de Ordenação 3 – Data de envio: 8,33%). **Destaca-se novamente o Algoritmo de Ordenação 2 – Contagem de visualizações, por apresentar-se de maneira constante entre todos os Termos de Pesquisa com um elevado percentual de vídeos desinformativos recomendados (média de 15,29% entre todos os Termos de Pesquisa – menor que no Perfil Misto) e o Algoritmo de Ordenação 1 – Classificação, por apresentar os menores percentuais de vídeos desinformativos recomendados (média de apenas 2,95% entre todos os Termos de Pesquisa). Isso indica um comportamento mais variado no Perfil Neutro quando comparado aos Perfis Informativo e Desinformativo, porém com uma tendência menor de recomendação de vídeos informativos quando comparado ao Perfil Misto.**

Ao observar a Figura 40, este comportamento fica evidente. Quando é considerada a média entre os Algoritmos de Ordenação para cada Termo de Pesquisa, a distribuição percentual média apresenta-se como demonstrado na Tabela 20 (E1 - Etapa 1 e PN - Perfil Neutro). Observa-se aqui o aumento percentual de vídeos neutros recomendados para todos os Termos de Pesquisa, com exceção do Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” (o único Termo de Pesquisa onde ocorre um aumento no percentual médio de vídeos informativos recomendados) e a redução considerável de vídeos informativos recomendados para o Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” (6,4%), o Termo de Pesquisa 3 – “nova urna eletronica 2022” (16%), o Termo de Pesquisa 4 – “urna eletrônica” (7,9%) e o Termo de Pesquisa 5 – “urnas eletronicas” (9,2%), além de um pequeno aumento médio dos vídeos desinformativos recomendados (sendo o maior aumento do Termo de Pesquisa 3 – “nova urna eletronica 2022”: 3,6%).

Fica evidente, portanto, que **para um perfil com um histórico de visualização composto majoritariamente por vídeos que não tem relação com um determinado tópico, a recomendação de vídeos informativos diminui entre quase todos os Termos de Pesquisa e Algoritmos de Ordenação (com uma redução média de 6,94% e máxima de 16% – Termo de Pesquisa 3 – “nova urna eletronica 2022” – comparado a um usuário sem histórico), ao passo que a recomendação de vídeos desinformativos aumenta ou diminui, a depender do Termo de Pesquisa e**

Tabela 20 – Comparação entre as distribuições percentuais médias para cada TP entre a Etapa 1 e a Etapa 3 (Termo de Pesquisa – TP, Vídeos Neutros – VN, Vídeos Informativos – VI, Vídeos Desinformativos – VD) – Perfil Neutro

TP (Etapa)	VN Média (%)	VI Média (%)	VD Média (%)
TP1 (E1)	72,7	18,5	8,7
TP1 (PN)	(+8,3)	(-6,4)	(-1,8)
TP2 (E1)	74,9	16,4	8,7
TP2 (PN)	(-3,4)	(+3,9)	(-0,5)
TP3 (E1)	55,4	42,2	2,4
TP3 (PN)	(+12,5)	(-16)	(+3,6)
TP4 (E1)	62,7	32,6	4,7
TP4 (PN)	(+5,8)	(-7,9)	(+2,2)
TP5 (E1)	61,2	30,6	8,2
TP5 (PN)	(+7,9)	(-9,2)	(+1,3)

Algoritmo de Ordenação em evidência (com um aumento médio de 1,05% e máximo de 3,6% – Termo de Pesquisa 3 – “nova urna eletrônica 2022” – comparado a um usuário sem histórico).

Conclui-se, portanto, que, apesar das recomendações do Perfil Neutro tenderem a ser majoritariamente de tópicos que não têm relação com o tópico em pesquisa – como nos demais Perfis – quando comparado a um usuário sem perfil, a tendência de que vídeos informativos que tenham relação com o tópico pesquisado sejam recomendados é menor.

7 CONCLUSÕES

Neste Capítulo final serão expostas as respostas para cada uma das Questões de Pesquisa propostas na Seção 1.1, além das avaliações para cada uma das Hipóteses de acordo com os resultados encontrados e expostos no Capítulo anterior. Por fim, serão feitas as considerações finais do autor.

A primeira questão de pesquisa apresentada neste trabalho foi: *Como é apresentado o conteúdo sobre as Urnas Eletrônicas em vídeos do YouTube? Este conteúdo tende a um potencial desinformativo?*

A resposta para a última questão é **não**. Nos resultados de pesquisa, a distribuição de vídeos é Informativo (Média 75,31%) > Desinformativo (Média 21,9%) > Neutro (Média 2,77%) e nas Recomendações, os vídeos são distribuídos em Neutro (Média 64,54%) > Informativo (Média 27,86%) > Desinformativo (Média 6,45%), como detalhado a seguir.

Outra questão de pesquisa, derivada da primeira, foi: *Como este conteúdo é apresentado nos resultados de pesquisa? Algum termo de pesquisa ou algoritmo de ordenação leva à maior presença de desinformação nos vídeos de pesquisa?*

Nesse caso, a média geral de SERP-MS, quando considerados todos os pares (Termo de Pesquisa x Algoritmo de Ordenação), é de -0,49 (majoritariamente informativa). Existe uma diferença significativa na quantidade de desinformação de acordo com o Termo de Pesquisa e Algoritmo de Ordenação escolhidos. O Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” (“bolsonaro embaixadores urnas”) e o Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas” (“bolsonaro urnas eletrônicas”) são os mais desinformativos dentre todos os termos. Por outro lado, o Algoritmo de Ordenação 1 – Classificação (Classificação) e Algoritmo de Ordenação 2 – Contagem de visualizações (Contagem de Visualizações) são os mais desinformativos dentre os algoritmos. Ou seja, para vídeos de pesquisa, a desinformação se encontra mais presente entre os vídeos mais populares e que contém o termo “bolsonaro” em sua composição. O Algoritmo de Ordenação 4 – Relevância (Relevância) não apresentou SERP-MS acima da média para nenhum dos cinco Termos de Pesquisa, o que mostra uma possível filtragem do próprio YouTube quanto à conteúdos desinformativos nos resultados

de pesquisa em seu Algoritmo de Ordenação padrão.

Uma última questão de pesquisa também derivada da primeira questão, foi: *Como este conteúdo é apresentado nas Recomendações de vídeos? Algum termo de pesquisa ou algoritmo de ordenação leva à maior presença de desinformação nos vídeos recomendados?*

Considerando um usuário sem perfil (e, conseqüentemente, sem histórico de visualização), a recomendação de vídeos é majoritariamente neutra através de todos os pares (Termo de Pesquisa x Algoritmo de Ordenação), sendo na maioria dos casos (16/20) superior ao percentual de vídeos informativos somado ao de vídeos desinformativos – com exceção dos Pares (Termo de Pesquisa 3 – “nova urna eletronica 2022” x Algoritmo de Ordenação 2 – Contagem de visualizações), (Termo de Pesquisa 3 – “nova urna eletronica 2022” x Algoritmo de Ordenação 4 – Relevância), (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 4 – Relevância) e (Termo de Pesquisa 5 – “urnas eletronicas” x Algoritmo de Ordenação 4 – Relevância) –, o que apresenta uma tendência à diversidade nas recomendações do SR do YouTube quando não há um histórico de uso atrelado ao usuário.

Nos vídeos recomendados, quando considerados os percentuais médios entre os AOs, o Algoritmo de Ordenação 3 – Data de envio (Data de Envio) apresenta consistentemente entre todos os Termos de Pesquisa a menor quantidade de desinformação recomendada. Já o Algoritmo de Ordenação 4 – Relevância (Relevância), que não apresentou nenhum SERP-MS maior que a média nos resultados de pesquisa, apresenta quantidades de desinformação acima da média nos vídeos recomendados para todos os Termos de Pesquisa e também um percentual alto de conteúdo informativo de maneira consistente. Ou seja, o Algoritmo de Ordenação padrão do YouTube, mesmo apresentando a menor desinformação entre os Algoritmos de Ordenação nos resultados de pesquisa, é o que mais apresenta conteúdo informativo e desinformativo acerca do tema pesquisado, quando comparado aos demais Algoritmos de Ordenação. Além disso, o Algoritmo de Ordenação 1 – Classificação (Classificação) e o Algoritmo de Ordenação 2 – Contagem de visualizações (Contagem de Visualização) são os únicos dois que apresentam tanto SERP-MS quanto percentual de vídeos desinformativos recomendados acima da média, o que é particularmente preocupante, visto que são os dois Algoritmos de Ordenação que ranqueiam nas primeiras posições os vídeos mais populares.

Considerando as conclusões acima, referentes à primeira questão de pesquisa e suas questões derivadas, retomamos a primeira hipótese investigada neste trabalho e apresentada no início dessa dissertação, qual seja:

Tanto o termo de pesquisa utilizado para o encontro de resultados acerca de um tema como o algoritmo de classificação utilizado irão apresentar uma tendência maior ou menor à desinformação em seus resultados de pesquisa e também nas recomen-

dações obtidas a partir dos vídeos de pesquisa.

A hipótese acima foi **validada**, como demonstrado nas respostas às questões de pesquisa acima. O uso de diferentes Termos de Pesquisa e de diferentes Algoritmos de Ordenação implicam em variações na quantidade de desinformação, tanto nos resultados de pesquisa quanto nos vídeos recomendados. Os vídeos classificados como mais populares (Algoritmo de Ordenação 1 – Classificação e Algoritmo de Ordenação 2 – Contagem de visualizações), quando selecionados nos resultados de pesquisa, levam também à maior desinformação nos vídeos recomendados, quando comparados aos demais Algoritmos de Ordenação. A desinformação ainda se faz mais presente quando são considerados o Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” e o Termo de Pesquisa 2 – “bolsonaro urnas eletrônicas”, que contam com a palavra “bolsonaro” em sua composição. O Termo de Pesquisa 3 – “nova urna eletronica 2022” (“nova urna eletronica 2022”) é o que menos apresenta desinformação em todos os cenários.

Uma segunda questão de pesquisa foi proposta: *Qual é a influência do histórico de visualização de um usuário no Sistema de Recomendação de vídeos do YouTube?*

Nesse caso, foi demonstrado que diferentes históricos de visualização irão gerar diferentes variações na quantidade de vídeos recomendados informativos, neutros e desinformativos acerca de um tema, como detalhado a seguir.

Uma questão de pesquisa derivada foi: *Qual é o número médio de vídeos que um usuário deve assistir acerca de um tema para que comece a haver homogeneidade em suas recomendações?*

Conforme apresentado nesta dissertação, o número médio de vídeos obtidos através das 10 iterações executadas na Etapa 2 deste trabalho é de 33,5 vídeos, não sendo possível determinar um número exato de vídeos necessários para que as recomendações comecem a apresentar homogeneidade (a variabilidade foi de 7 a 77 vídeos assistidos). Isso pode ocorrer por diferenças inerentes aos vídeos utilizados como base e também por conta dos vídeos escolhidos aleatoriamente nesta Etapa.

Outra questão de pesquisa que também derivou da segunda questão apresentada mais acima, foi: *Os vídeos de pesquisa apresentam diferença de potencial desinformativo de acordo com o histórico de visualização de um usuário? Os termos de pesquisa e algoritmos de ordenação utilizados influenciam de alguma maneira particular nestes resultados?*

O trabalho demonstrou que, dentre os quatro perfis analisados na Etapa 3, apenas o Perfil Neutro (PN) apresentou resultados de variação de SERP-MS não negligenciáveis, tendo uma pequena redução na desinformação média apresentada nos resultados de pesquisa. Para os demais Perfis (Desinformativo – PD, Informativo – PI e Misto – PM), a diferença média nos resultados de pesquisa foi negligenciável em relação à Etapa 1 (sem perfil).

Nos perfis Desinformativo, Informativo e Misto, a variação média do SERP-MS entre todos os pares foi de apenas 0,01 (com aumento no PD e PI e redução no PM). No Perfil Desinformativo, a escolha do Termo de Pesquisa teve maior impacto na variação dos resultados de pesquisa. No Perfil Informativo esta variação foi mais influenciada pela escolha de Algoritmo de Ordenação. O Perfil Misto, em termos gerais, se aproximou mais dos resultados do Perfil Desinformativo do que do Perfil Informativo nos vídeos de pesquisa.

Já no Perfil Neutro, a presença de um histórico de visualização composto por vídeos que não têm ligação com o assunto pesquisado leva a uma redução na quantidade de desinformação nos resultados de pesquisa. Aqui, como no Perfil Desinformativo, a escolha do Termo de Pesquisa teve um impacto maior na quantidade de desinformação apresentada nos resultados de pesquisa. Quando comparado à Etapa 1, o percentual de vídeos desinformativos nos resultados de pesquisa é consideravelmente menor (uma redução média de 4,2% no número de vídeos e chegando até uma redução de 0,3225 no SERP-MS médio quando considerado o TP5).

Mais uma questão de pesquisa surgiu sobre esse tema: *Os vídeos recomendados apresentam diferença de potencial desinformativo de acordo com o histórico de visualização de um usuário? Os termos de pesquisa e algoritmos de ordenação utilizados influenciam de alguma maneira particular nestes resultados?*

Nesse caso, a resposta é que, nos vídeos recomendados, diferentemente dos vídeos de pesquisa, a diferença entre os quatro Perfis é bastante significativa e fica evidente. Mesmo havendo um aumento percentual médio (em menor ou maior escala) entre todos os perfis na quantidade de vídeos desinformativos recomendados, o comportamento observado entre os quatro diverge.

No Perfil Desinformativo, há uma redução média percentual significativa nos vídeos informativos recomendados (-5,98%) ao passo que, simetricamente, há o mesmo aumento médio percentual significativo nos vídeos desinformativos recomendados (+5,98%). Nos pares deste perfil, cada Termo de Pesquisa aparece em pelo menos dois Algoritmos de Ordenação com percentual médio desinformativo maior do que aquele obtido por um usuário sem histórico. O aumento médio máximo de vídeos desinformativos recomendados ocorre no Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” – 6,8% e o par (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 2 – Contagem de visualizações) chega a apresentar 33,68% de vídeos desinformativos recomendados – basicamente, um a cada três vídeos.

No Perfil Informativo, há um aumento médio percentual muito significativo nos vídeos informativos recomendados (+9,04%), ao passo que há um pequeno aumento na média percentual de vídeos desinformativos recomendados (+0,54%). Todos os pares deste perfil, com exceção do (Termo de Pesquisa 5 x Algoritmo de Ordenação 2 – Contagem de visualizações), apresentaram uma média percentual de vídeos in-

formativos recomendados acima daqueles obtidos por um usuário sem histórico. O aumento médio máximo de vídeos informativos recomendados ocorre novamente no Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” – 18,1% e o par (Termo de Pesquisa 3 – “nova urna eletronica 2022” x Algoritmo de Ordenação 4 – Relevância) chega a apresentar 61,5% de vídeos informativos recomendados – basicamente, um a cada dois vídeos.

No Perfil Misto, há um pequeno aumento médio percentual nos vídeos informativos recomendados (1,81%) e um aumento um pouco maior na quantidade de vídeos desinformativos recomendados (2,54%). Aqui, o Algoritmo de Ordenação 2 – Contagem de visualizações apresenta-se de maneira consistente entre todos os Termos de Pesquisa com um elevado percentual de vídeos desinformativos recomendados (média de 25,9% entre todos os Termos de Pesquisa) e o AO3 apresenta os menores percentuais de vídeos desinformativos recomendados (média de apenas 1,27% entre todos os Termos de Pesquisa). O aumento médio máximo de vídeos desinformativos ocorre no Termo de Pesquisa 3 – “nova urna eletronica 2022” – 4,7% e o par (Termo de Pesquisa 3 – “nova urna eletronica 2022” x Algoritmo de Ordenação 4 – Relevância) chega a apresentar 68,18% de vídeos informativos recomendados, enquanto o par (Termo de Pesquisa 1 – “bolsonaro embaixadores urnas” x Algoritmo de Ordenação 2 – Contagem de visualizações) chega a apresentar 33,67% de vídeos desinformativos recomendados, o que reflete traços do comportamento observado tanto no Perfil Desinformativo como no Perfil Informativo – e, portanto, esse é o perfil que se comporta de maneira mais similar à um usuário sem histórico;

Por fim, no Perfil Neutro há uma redução média percentual significativa nos vídeos informativos recomendados (6,94%) ao passo que há um pequeno aumento na média percentual de vídeos desinformativos recomendados (1,05%). Aqui, novamente, o Algoritmo de Ordenação 2 – Contagem de visualizações apresenta-se constantemente entre todos os Termos de Pesquisa com um elevado percentual de vídeos desinformativos recomendados (média de 15,29% entre todos os Termos de Pesquisa) e, diferentemente do Perfil Misto, aqui o Algoritmo de Ordenação 1 – Classificação apresenta os menores percentuais de vídeos desinformativos recomendados (média de apenas 2,95% entre todos os Termos de Pesquisa). Há uma redução geral na quantidade de vídeos informativos recomendados (com a redução máxima aparecendo no Termo de Pesquisa 3 – “nova urna eletronica 2022” – 16%) e um pequeno aumento geral na quantidade de vídeos desinformativos recomendados (com o aumento máximo aparecendo também no Termo de Pesquisa 3 – “nova urna eletronica 2022” – 3,6%). Aqui, observa-se uma clara redução nos vídeos informativos recomendados quando o par com o maior percentual é o (Termo de Pesquisa 4 – “urna eletrônica” x Algoritmo de Ordenação 4 – Relevância) com 45,83% – uma redução de 15,67% em relação ao máximo no Perfil Informativo e de 22,35% em relação ao máximo no Perfil Misto.

Neste ponto, com as discussões acima, estamos aptos a discutir a segunda hipótese apresentada nessa dissertação, que foi:

Os resultados obtidos tanto na Pesquisa como nas Recomendações de vídeos de um usuário irão variar de acordo com o tipo de conteúdo presente em seu histórico de visualização.

Essa hipótese foi **parcialmente validada**, visto que, para quase todos os históricos de visualização, a variação nos resultados de pesquisa em relação a um usuário sem histórico foi negligenciável. Já nas recomendações, houve significativa mudança de acordo com o histórico do usuário em comparação a um usuário sem histórico. Usuários com um perfil de consumo majoritariamente desinformativo têm um aumento significativo de desinformação e redução significativa de informação recomendados,. Por outro lado, usuários com um perfil de consumo majoritariamente informativo, têm um aumento drástico de informação recomendada e um pequeno aumento de desinformação também. Já para um usuário com histórico misto entre vídeos informativos e desinformativos, há pequenas diferenças em relação a um usuário sem histórico. Por fim, um usuário que consome vídeos apenas sobre temas não relativos ao objeto de pesquisa em questão tende a ser recomendado com menos vídeos informativos do que um usuário sem histórico.

Por fim, a última questão de pesquisa proposta neste trabalho foi: *É possível atestar a criação de um filtro bolha de desinformação no Sistema de Recomendação do YouTube a partir do consumo de vídeos desinformativos por um usuário?*

Nesse caso a resposta é **sim**. Na medida em que um usuário que consome conteúdo desinformativo acaba recebendo em suas recomendações mais conteúdo desinformativo e menos conteúdo informativo, pode-se observar a criação de um filtro bolha de desinformação, conforme análise da hipótese final deste trabalho, apresentada abaixo.

Usuários que consomem conteúdo desinformativo terão seus vídeos recomendados tendendo mais a conteúdos desinformativos quando comparados à usuários que consomem conteúdos informativos, mistos ou neutros em relação a um tema (filtro bolha).

Levando em consideração as definições apontadas para os filtros bolha no Capítulo 3, essa hipótese foi **parcialmente validada** neste trabalho, visto que, para todos os históricos de visualização, houve pelo menos um pequeno aumento no percentual médio de vídeos desinformativos recomendados.

A hipótese acaba sendo parcialmente validada, pois, mesmo não havendo um mecanismo eficiente para que o filtro criado seja furado, ainda há uma forte recomendação de vídeos informativos ou neutros, mesmo com o aumento generalizado de vídeos desinformativos sendo recomendados para um Perfil Desinformativo e, simetricamente, tendo uma redução de vídeos informativos recomendados no mesmo. O

viés do usuário pode agir de maneira a cair em uma câmara de eco nas recomendações geradas de acordo com seu histórico, o que vai de encontro com as definições estabelecidas em PARISER (2012), LOIOLA (2018) e LIAO; FU (2013). Levando em consideração que o tema em questão (Urnas Eletrônicas) tornou-se um tópico extremamente identitário da extrema-direita durante a vigência do último governo e no momento em que este trabalho foi realizado, é possível dizer que há uma forte tendência na criação de um filtro bolha para usuários deste espectro nesta configuração.

O Perfil Desinformativo mostra o resultado mais grave dentre os quatro perfis, apresentando um aumento médio de 5,98% de vídeos desinformativos recomendados, ao passo que reduz de maneira simétrica o percentual médio de vídeos informativos em suas recomendações. Esse comportamento apresentado pelo SR leva à criação de um filtro bolha de desinformação, chegando ao ponto de um a cada três vídeos recomendados ser desinformativo quando utilizado o Termo de Pesquisa “bolsonaro embaixadores urnas” com o Algoritmo de Ordenação de Contagem de Visualizações. Na mesma medida, não encontrou-se para este perfil um mecanismo eficiente para que fosse furada tal bolha, com evidente redução geral de vídeos informativos sendo recomendados.

Destaca-se, também, o fato de que, dentre todos os Perfis, os dois Algoritmos de Ordenação que apresentaram maior percentual médio de desinformação recomendada foram os que ranqueiam primeiramente os vídeos mais populares, tanto pelo número de *likes* quanto pelo número de visualizações, o que aponta para uma ampla presença de desinformação generalizada entre vídeos populares na plataforma. Para o Perfil Desinformativo, o SR constantemente recomendava vídeos de canais específicos, como Jovem Pan e Gazeta do Povo, sendo estes alguns dos que apresentaram maior métrica de popularidade e desinformação generalizada nas classificações feitas na Etapa 1. Mesmo havendo o arquivamento e desmonetização de boa parte dos vídeos da Jovem Pan contendo desinformação sobre as UEs, fica notável aqui o impacto que os vídeos tiveram no reforço de visões anti-democráticas e que, por fim, podem ter culminado em ataques como o de 8 de janeiro de 2023.

Cabe salientar, porém, que o filtro bolha de conteúdo informativo parece atuar de maneira ainda mais forte quando observados os resultados do Perfil Informativo, onde 19 dos 20 pares obtiveram um percentual de vídeos informativos recomendados acima da média daqueles obtidos na Etapa 1 – mesmo este perfil apresentando também um pequeno aumento percentual na média de vídeos desinformativos recomendados entre todos os pares. Quando utilizado o Termo de Pesquisa “nova urna eletrônica 2022” e o Algoritmo de Ordenação de Relevância (padrão do YouTube), mais da metade dos vídeos recomendados são informativos sobre o tema.

O Perfil Misto acaba por demonstrar o equilíbrio de um “meio-termo” entre o Perfil Informativo e o Perfil Desinformativo, com aumentos percentuais em ambas as frentes:

informativa e desinformativa. Acaba, portanto, não podendo ser atestada a criação de um filtro bolha, apresentando, inclusive, um comportamento similar a um usuário que não tem um histórico de visualização associado. O Perfil Neutro tende a recomendar menos vídeos informativos, mas também não tem um aumento significativo nos vídeos desinformativos recomendados e, portanto, não pode se atestar a criação de um filtro bolha para este.

Este trabalho teve como objetivo principal investigar o impacto que os Sistemas de Recomendação podem exercer nas esferas política, social e democrática. Mesmo tendo sido idealizados como ferramentas de facilitação e de direcionamento de conteúdos no massivo acervo presente online em inúmeras plataformas, sua atual implementação, utilizando de dados pessoais e subjetividades inerentes a cada usuário, acabam por reforçar visões de mundo pré-estabelecidas e que, muitas vezes, são baseadas em desinformação – principalmente após o aumento exponencial de conteúdos desinformativos sendo utilizados como arma de manipulação para a manutenção do poder entre aqueles ocupando cargos da mais alta escala administrativa em muitos países.

No Brasil, foi possível observar uma aderência popular massiva a campanhas que buscavam desacreditar as instituições democraticamente estabelecidas, o equilíbrio entre os poderes e, de maneira ainda mais gravosa, a única solução cientificamente comprovada como saída de uma situação de calamidade pública, no caso da pandemia. É evidente a influência que as mídias sociais exerceram em tais campanhas, impulsionadas, inclusive, por aqueles no poder. As plataformas permitiram, e de maneira ainda mais gravosa, lucraram, com a viralização de conteúdos que despertam o que há de mais primal no ser humano. Abandonou-se a civilidade por parte significativa da população em prol do capital na mão de poucos.

As Urnas Eletrônicas, alvos de uma das maiores campanhas desinformativas já vistas em território brasileiro, demonstraram vez e vez novamente que servem como um sólido pilar da nossa democracia. Deram voz à legítima vontade do povo e embasaram a alternância de poder que é saudável em um país democrático. No último pleito, as Urnas Eletrônicas foram escrutinadas reiteradamente de maneira não usual, inclusive pelas Forças Armadas, a partir de um esforço do governo federal para tentar deslegitimar as urnas, como forma desesperada para manter o seu poder.

Passada a maior ameaça enfrentada desde o Golpe Militar de 1964, a democracia brasileira mais uma vez sobreviveu. Mesmo que o filtro bolha de desinformação, demonstrado no presente trabalho para o caso da plataforma YouTube, manipulando uma significativa massa popular, as instituições não ruíram. Fica aqui registrada a admiração do autor pelo singular trabalho exercido pelas Supremas Cortes deste país – em especial o Tribunal Superior Eleitoral – para a manutenção do estado democrático de direito, principalmente durante o período do pleito eleitoral de 2022.

A Computação, principalmente em nosso atual contexto social, não atua de forma isolada. Este trabalho serve, também, como tentativa do autor de ampliar o horizonte de possibilidades de estudos interdisciplinares a partir do olhar da Ciência da Computação, tentando entender, para além da teoria, o real impacto exercido pela tecnologia desenvolvida por egressos de cursos desta natureza. Mais do que linhas de código e paradigmas de programação, um cientista da computação acaba por atuar diretamente no que hoje é o filtro pelo qual a humanidade interpreta fatos: a tecnologia.

Em trabalhos futuros, pode-se ampliar o debate feito no presente trabalho para além das Urnas Eletrônicas, aplicando a Metodologia desenvolvida aqui para que possa ser auditado o Sistema de Recomendação do YouTube com relação a outros temas com ampla presença de desinformação na plataforma: vacinas, teorias conspiratórias, etc. Além disso, pode-se desenvolver uma técnica de classificação automática, a partir dos dados apresentados aqui, com modelos de Aprendizado de Máquina, para analisar transcrições de vídeos na plataforma e detectar a presença de desinformação de maneira automatizada. As possibilidades são muitas.

Por fim, fica registrado aqui que, na data em que é finalizada a primeira redação desta dissertação (29/06/2023), o ex-presidente da república Jair Bolsonaro tornou-se inelegível por oito anos em decorrência dos ataques proferidos contra as Urnas Eletrônicas, que foram objeto de investigação deste trabalho.

REFERÊNCIAS

ABUL-FOTTOUH, D.; SONG, M. Y.; GRUZD, A. Examining algorithmic biases in YouTube's recommendations of vaccine videos. **International Journal of Medical Informatics**, v.140, p.104–175, 2020.

ACADEMY, D. S. **Deep Learning Book - Capítulo 3 – O Que São Redes Neurais Artificiais Profundas ou Deep Learning?** Acessado em: 24/05/2023, <https://www.deeplearningbook.com.br/o-que-sao-redes-neurais-artificiais-profundas/>.

ACADEMY, D. S. **Deep Learning Book - Capítulo 8 – Função de Ativação.** Acessado em: 24/05/2023, <https://www.deeplearningbook.com.br/funcao-de-ativacao/>.

ADOMAVICIUS, G.; TUZHILIN, A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. **IEEE Transactions on Knowledge and Data Engineering**, v.17, n.6, p.734–749, 2005.

ADOMAVICIUS, G.; TUZHILIN, A. Context-Aware Recommender Systems. , p.217–253, 2011.

AGGARWAL, C. C. **Recommender Systems - The Textbook.** 1-498p.

AMADO, G. Pressão comercial faz Jovem Pan arquivar 376 vídeos sobre Covid e urna. **Metrópoles**, 2022.

APPEN. **Confidence to Deploy AI with World-Class Training Data.** Acessado em: 09/12/2021, <https://appen.com/>.

ARENDDT, H. **Origens do totalitarismo:** antissemitismo, imperialismo, totalitarismo.

ARREGUY, J. Nunca houve fraude comprovada nas eleições com urna eletrônica no Brasil. **Uol Confere**, 2021. Disponível em: <https://noticias.uol.com.br/confere/ultimas-noticias/2021/07/08/eleicoes-brasil-fraude.htm>. Acessado em: 07/12/2021.

AWAD; KRISHNAN. The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to Be Profiled Online for Personalization. **MIS Quarterly**, v.30, n.1, p.13, 2006.

AZEVEDO, A. G. **A História do Direito ao Voto no Brasil**. Trabalho de Conclusão de Curso - Faculdade de Direito Professor Jacy de Assis - Universidade Federal de Uberlândia.

BALUJA, S. et al. Video suggestion and discovery for youtube. In: WORLD WIDE WEB, 17., 2008. **Proceedings**....

BANDEIRA, O. O COMBATE À DESINFORMAÇÃO NA INTERNET: O QUE FAZER DAQUI PRA FRENTE? **Congresso em Foco**, 2018. Disponível em: <https://congressoemfoco.uol.com.br/temas/midia/o-combate-a-desinformacao-na-internet-o-que-fazer-daqui-para-frente/>. Acessado em: 10/12/2021.

BANDY, J. Problematic machine behavior: A systematic literature review of algorithm audits. **Proceedings of the acm on human-computer interaction**, v.5, n.CSCW1, p.1–34, 2021.

BATISTA, M. R. S. **A utilização de algoritmos de aprendizado de máquina em problemas de classificação**. 2018. Tese (Doutorado em Ciência da Computação) — Universidade de São Paulo.

BENDERSKY, M. et al. Up next. In: ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 20., 2014. **Proceedings**....

BETIM, F. TSE desmente alegações sem provas de Bolsonaro contra urnas eletrônicas. **Jota**, 2022.

BIZZOTTO, N.; MORLINO, S.; SCHULZ, P. J. Misinformation in Italian Online Mental Health Communities During the COVID-19 Pandemic: Protocol for a Content Analysis Study. **JMIR Res Protoc**, v.11, n.5, p.e35347, May 2022.

BORGATTI, S. P.; EVERETT, M. G.; JOHNSON, J. C. **Analyzing Social Networks**.

BOZDAG, E.; GAO, Q.; HOUBEN, G.-J.; WARNIER, M. Does offline political segregation affect the filter bubble? An empirical analysis of information diversity for Dutch and Turkish Twitter users. **Computers in human behavior**, v.41, p.405–415, 2014.

BRAGA, S. Estado Novo. **Fundação Getúlio Vargas; Centro de Pesquisa e Documentação de História Contemporânea do Brasil. Fatos e imagens: artigos ilustrados de fatos e conjunturas do Brasil**. Disponível em: <https://goo.gl/I9TQLQ>. Acesso em 08/10/2021, v.22, 2016.

BRASIL. CONSTITUIÇÃO DA REPÚBLICA DOS ESTADOS UNIDOS DO BRASIL (DE 24 DE FEVEREIRO DE 1891). **Congresso Nacional Constituinte**, 1891.

BRASIL. ATO INSTITUCIONAL Nº 2, DE 27 DE OUTUBRO DE 1965. **Diário Oficial da União**, 1965.

BRASIL. ATO INSTITUCIONAL Nº 5, DE 13 DE DEZEMBRO DE 1968. **Diário Oficial da União**, 1968.

BRASIL. CONSTITUIÇÃO DA REPÚBLICA FEDERATIVA DO BRASIL DE 1988. **Diário Oficial da União**, 1988.

BRASIL. **LEI Nº 12.965, DE 23 DE ABRIL DE 2014 - Marco Civil da Internet**. Lei nº 12.965/2014. Disponível em: <http://www.planalto.gov.br/ccivil_03/ato2011-2014/2014/lei/l12965.htm > .

BRASIL. **Lei Nº 13.709, DE 14 DE AGOSTO DE 2018 - Lei Geral de Proteção de Dados Pessoais (LGPD)**. Lei nº 13.709/2018. Disponível em: <http://www.planalto.gov.br/ccivil_03/ato2015-2018/2018/lei/L13709.htm > .

BU, C.-S. Understanding the 3 Vs of Big Data - Volume, Velocity and Variety. **Coforge**, 2017.

BURKE, R. **User Modeling and User-Adapted Interaction**, v.12, n.4, p.331–370, 2002.

CAHN, A.; ALFELD, S.; BARFORD, P.; MUTHUKRISHNAN, S. An Empirical Study of Web Cookies. In: INTERNATIONAL CONFERENCE ON WORLD WIDE WEB, 25., 2016. **Proceedings...**

CAKEBREAD, C. You're not alone, no one reads terms of service agreements. **Business Insider**, June 2019.

CANÊDO, L. B. Aprendendo a votar. **História da Cidadania**, p.517, 2003.

CARDOSO SAMPAIO, R.; LYCARIÃO, D. **Análise de conteúdo categorial**: manual de aplicação.

CARRO, R. Brazil | Reuters Institute digital news report 2022. **Reuters Institute for the Study of Journalism**, 2022. Disponível em: <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022/brazil>. Acessado em: 01/08/2022.

CASTELLO BRANCO, H. Decreto-Lei nº 314, de 13 de Março de 1967. **Diário Oficial da União**, 1967.

CASTELLS, M. Local and global: Cities in the network society. **Tijdschrift voor economische en sociale geografie**, v.93, n.5, p.548–558, 2002.

CAZELLA, S. C.; NUNES, M.; REATEGUI, E. A Ciência da Opinião: Estado da arte em Sistemas de Recomendação. **André Ponce de Leon F. de Carvalho; Tomasz Kowaltowski..(Org.). Jornada de Atualização de Informática-JAI**, p.161–216, 2010.

CHENG, Z.; CAVERLEE, J.; LEE, K. You are where you tweet. In: ACM INTERNATIONAL CONFERENCE ON INFORMATION AND KNOWLEDGE MANAGEMENT, 19., 2010. **Proceedings...**

CINELLI, M. et al. Online Hate: Behavioural Dynamics and Relationship with Misinformation. **CoRR**, v.abs/2105.14005, 2021.

CINTRA, A. O. Majoritário ou proporcional? **O SISTEMA POLÍTICO BRASILEIRO: Continuidade ou Reforma?**, p.17, 2005.

COIMBRA, R. C. M. Por que a urna eletrônica é segura. **Revista Eletrônica da EJE**, v.6, 2014.

COOK, J.; ECKER, U.; LEWANDOWSKY, S. Misinformation and how to correct it. **Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource**, p.1–17, 2015.

COVINGTON, P.; ADAMS, J.; SARGIN, E. Deep Neural Networks for YouTube Recommendations. In: ACM CONFERENCE ON RECOMMENDER SYSTEMS, 10., 2016. **Proceedings...**

CRISTALDO, H. Câmara dos Deputados rejeita PEC do Voto Impresso. **Agência Brasil**, 2021. Disponível em: <https://agenciabrasil.ebc.com.br/politica/noticia/2021-08/camara-dos-deputados-rejeita-pec-do-voto-impresso>. Acessado em: 07/12/2021.

D'AGOSTINO, R. Presidente do TSE, Barroso diz que teste de segurança nas urnas não mostrou risco à eleição. **G1**, 2021. Disponível em: <https://g1.globo.com/politica/noticia/2021/11/29/presidente-do-tse-barroso-diz-que-teste-de-seguranca-nas-urnas-nao-mostrou-risco-a-eleicao.ghtml>. Acessado em: 10/12/2021.

DAVIDSON, J. et al. The YouTube video recommendation system. In: ACM CONFERENCE ON RECOMMENDER SYSTEMS, 2010. **Proceedings...**

DEL VICARIO, M. et al. The spreading of misinformation online. **Proceedings of the National Academy of Sciences**, v.113, n.3, p.554–559, 2016.

DELEUZE, G. Controle e devir. *Futur Antérieur*, n. 1, primavera, 1990, entrevista a Toni Negri. **Conversações**, 1992.

DENNING, P. J. ACM president's letter: electronic junk. **Communications of the ACM**, v.25, n.3, p.163–165, 1982.

DEPUTADOS, R. C. dos. TSE contabiliza 3,3 mil urnas com defeito, de um total de 400 mil. **Câmara dos Deputados**, 2020. Disponível em: <https://www.camara.leg.br/noticias/707746-tse-contabiliza-33-mil-urnas-com-defeito-de-um-total-de-400-mil/>. Acessado em: 06/12/2021.

DIFRANZO, D.; GLORIA-GARCIA, K. Filter bubbles and fake news. **XRDS: Crossroads, The ACM Magazine for Students**, v.23, n.3, p.32–35, 2017.

ELEITORAL, T. S. **Eleições 2022**: TSE renova parceria com agências de checagem. Acesso em: 01/06/2023, Disponível em: <https://encurtador.com.br/cpNT6>.

ELEITORAL, T. S. **RESOLUÇÃO Nº 23.714, DE 20 DE OUTUBRO DE 2022**. RESOLUÇÃO Nº 23.714. Disponível em: <https://www.tse.jus.br/legislacao/compilada/res/2022/resolucao-no-23-714-de-20-de-outubro-de-2022>.

ESTADÃO. 8 de janeiro: O que se sabe sobre os ataques golpistas em Brasília após um mês da invasão. **Estadão**, 2023.

FACEBOOK. **De onde vêm as sugestões de Pessoas que você talvez conheça no Facebook**. Acessado em: 23/05/2023, https://www.facebook.com/help/163810437015615/?helpref=uf_share.

FACEBOOK. **Saiba mais e gerencie o conteúdo sugerido no seu Feed do Facebook**. Acessado em: 23/05/2023, https://www.facebook.com/help/485502912850153/?helpref=related_articles.

FANG, M.; YANG, G.; GONG, N. Z.; LIU, J. Poisoning Attacks to Graph-Based Recommender Systems. In: ANNUAL COMPUTER SECURITY APPLICATIONS CONFERENCE, 34., 2018, New York, NY, USA. **Proceedings...** p.381–392. (ACSAC '18).

FEDERAL, S. **Projeto de Lei nº 2630, de 2020**. Projeto da Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/141944emendas>.

FERNANDES, A. Invasão em Brasília completa um mês; prejuízo chegou a R\$ 21 milhões, e 1,4 mil foram presos. **R7 Brasília**, 2023.

FERRÃO, I. et al. Urnas Eletrônicas no Brasil: linha do tempo, evolução e falhas e desafios de segurança. **Revista Brasileira de Computação Aplicada**, v.11, n.2, p.1–12, maio 2019.

FERREIRA, F. **Antivaccine videos slip through YouTube's advertising policies, new study finds.**

FILHO, W. H. Caos nas eleições de 1994 mostra como era apuração de votos antes da urna eletrônica. **O Globo**, 2021. Disponível em: <https://blogs.oglobo.globo.com/blog-do-acervo/post/cedulas-de-papel-como-o-juiz-luiz-fulx-combateu-fraude-nas-eleicoes-de-1994-e-foi-ameacado-de-morte.html>. Acessado em: 07/12/2021.

FRIEDMAN, M.; LAST, M.; MAKOVER, Y.; KANDEL, A. Anomaly detection in web documents using crisp and fuzzy-based cosine clustering methodology. **Information Sciences**, v.177, n.2, p.467–475, 2007.

FRIEDRICH, G.; ZANKER, M. A Taxonomy for Generating Explanations in Recommender Systems. **AI Magazine**, v.32, n.3, p.90–98, Jun. 2011.

G1. Bolsonaroistas radicais fazem manifestação com intenções golpistas no feriado de 15 de novembro. **G1**, 2022.

GAUDANI, H. Personalized Recipe Recommendation System using Hybrid Approach. **International Journal of Innovative Research in Computer and Communication Engineering**, v.5, p.192–197, 06 2016.

GESCHKE, D.; LORENZ, J.; HOLTZ, P. The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers. **British Journal of Social Psychology**, v.58, n.1, p.129–149, 2019.

GOLDBERG, D.; NICHOLS, D.; OKI, B. M.; TERRY, D. Using Collaborative Filtering to Weave an Information Tapestry. **Commun. ACM**, v.35, n.12, p.61–70, dec 1992.

GOODROW, C. **On YouTube's Recommendation System.** Acesso em: 30/05/2023, Disponível em: <https://blog.youtube/inside-youtube/on-youtubes-recommendation-system/>.

GOOGLE. **Google Trends.** Disponível em: <<https://trends.google.com.br/trends/>>.

HABGOOD-COOTE, J. Stop Talking About Fake News! **Inquiry: An Interdisciplinary Journal of Philosophy**, v.62, n.9-10, p.1033–1065, 2019.

HENRIQUE, L. PL das Fake News: os 10 pontos principais para entender o projeto de lei. **Politize!**, 2023.

HUSSEIN, E.; ELDARDIRY, H. Investigating Misinformation in Online Marketplaces: An Audit Study on Amazon. **arXiv preprint arXiv:2009.12468**, 2020.

HUSSEIN, E.; JUNEJA, P.; MITRA, T. Measuring Misinformation in Video Search Platforms: An Audit Study on YouTube. **Proceedings of the ACM on Human-Computer Interaction**, v.4, n.CSCW1, p.1–27, May 2020.

IBM. **AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference?** Acessado em: 24/05/2023, <https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks>.

IDEA, I. The Global State of Democracy 2021. Building Resilience in a Pandemic Era. , 2021.

INMAN, H. F.; JR, E. L. B. The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities. **Communications in Statistics - Theory and Methods**, v.18, n.10, p.3851–3874, 1989.

ITU-D, I. T. U. D. S. **World Telecommunication/ICT Indicators Database**. Disponível em: <<https://www.itu.int/en/ITU-D/Statistics/Pages/publications/wtid.aspx>>.

JAEGER, G. E. R. **História do voto no Brasil**. 158p. Trabalho de Conclusão de Curso (Graduação) - Centro de Ciências Jurídicas, Universidade do Vale do Rio dos Sinos.

JAMIY, F. E.; DAIF, A.; AZOUAZI, M.; MARZAK, A. **The potential and challenges of Big data - Recommendation systems next level application**. Disponível em: <<https://arxiv.org/abs/1501.03424>>.

JUNEJA, P.; MITRA, T. Auditing e-commerce platforms for algorithmically curated vaccine misinformation. In: OF THE 2021 CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2021. **Proceedings**. . . . p.1–27.

KAISER, B. **Manipulados**: como a Cambridge Analytica e o Facebook invadiram a privacidade de milhões e botaram a democracia em xeque.

KNUUTILA, A. et al. Covid-related misinformation on YouTube: The spread of misinformation videos on social media and the effectiveness of platform policies. **COMPROM Data Memo**, n.6, 2020.

KRIPPENDORFF, K. **Content Analysis**: An Introduction to Its Methodology. 2nd.ed.

KRIPPENDORFF, K. Computing Krippendorff's alpha-reliability. , 2011.

LADA, A.; WANG, M.; YAN, T. **How does News Feed predict what you want to see?** Acessado em: 23/05/2023, <https://tech.facebook.com/engineering/2021/1/news-feed-ranking/>.

LEWANDOWSKY, S.; ECKER, U. K. H.; COOK, J. Beyond misinformation: Understanding and coping with the “post-truth” era. **Journal of Applied Research in Memory and Cognition**, v.6, n.4, p.353–369, Dec. 2017.

LIAO, Q. V.; FU, W.-T. Beyond the Filter Bubble: Interactive Effects of Perceived Threat and Topic Involvement on Selective Exposure to Information. In: SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2013, New York, NY, USA. **Proceedings**. . . . p.2359–2368. (CHI '13).

LIMA, S. YouTube remove só 4,4% dos vídeos com desinformação contra urna eletrônica. **Estadão**, 2022. Acessado em: 20/07/2022.

LINHARES, C. Datafolha: 73% confiam nas urnas; índice recua em meio a ofensiva de Bolsonaro. **Folha de S. Paulo**, 2022. Disponível em: <https://www1.folha.uol.com.br/poder/2022/05/datafolha-73-confiam-nas-urnas-indice-recua-em-meio-a-ofensiva-de-bolsonaro.shtml>. Acessado em: 28/07/2022.

LOIOLA, D. F. E. Recomendado Para Você: o impacto do algoritmo do YouTube na formação de bolhas. , 2018.

LÜ, L. et al. Recommender systems. **Physics reports**, v.519, n.1, p.1–49, 2012.

MACEDO, I. DAS 123 FAKE NEWS ENCONTRADAS POR AGÊNCIAS DE CHECAGEM, 104 BENEFICIARAM BOLSONARO. **Congresso em Foco**, 2018. Disponível em: <https://congressoemfoco.uol.com.br/area/pais/das-123-fake-news-encontradas-por-agencias-de-checagem-104-beneficiaram-bolsonaro/>. Acessado em: 10/12/2021.

MACHADO FILHO, W. P. **Vulnerabilidade da urna eletrônica**: mecanismos de segurança e transparência para garantia da lisura no processo de votação.

MANDEL, E. **Late Capitalism**.

MARTINS, H. **A Desinformação em Meio à crise social**. Disponível em: <https://diplomatie.org.br/a-desinformacao-em-meio-a-crise-social/>.

MARTINS, M. E. G. Coeficiente de correlação amostral. **Revista de Ciência Elementar**, v.2, n.2, p.69, 2014.

MCDONALD, A. M.; CRANOR, L. F. The Cost of Reading Privacy Policies. **I/S: A Journal of Law and Policy for the Information Society**, v.4, n.3, p.540–565, 2008.

MEDVEDEV, I. Powered by AI: Instagram’s Explore recommender system. **Instagram Engineering**, 2019.

MNEIMNEH, A. Surveillance Advertising. **The Flaw**, 2023.

MOROZOV, E. **Big Tech: A Ascensão dos Dados e a Morte da Política.**

NAYEK, J. K.; DAS, R. Evaluation of Famous Recommender Systems: A Comparative Analysis. **DigitalCommons@University of Nebraska - Lincoln**, 2021.

NICOLAU, J. **Eleições no Brasil: do Império aos dias atuais.**

NICOLAU, J. M. **História do voto no Brasil.**

NILASHI, M. et al. Travelers decision making using online review in social network sites: A case on TripAdvisor. **Journal of Computational Science**, v.28, p.168–179, 2018.

NUTI, S. V. et al. The use of google trends in health care research: a systematic review. **PloS one**, v.9, n.10, p.e109583, 2014.

OBAR, J. A.; OELDORF-HIRSCH, A. The biggest lie on the Internet: ignoring the privacy policies and terms of service policies of social networking services. **Information, Communication & Society**, v.23, n.1, p.128–147, 2020.

OLIVEIRA, F. M. d. A transparência e a auditoria da urna eletrônica: a soberania popular materializada na legitimidade do voto. , 2021.

PAES, J. P. L. O sufrágio e o voto no Brasil: direito ou obrigação? **Revista Eletrônica da EJE**, v.3, 2013.

PAPADAMOU, K. et al. "It is just a flu": Assessing the Effect of Watch History on YouTube's Pseudoscientific Video Recommendations. **arXiv preprint arXiv:2010.11638**, 2020.

PARISER, E. **O filtro invisível: o que a internet está escondendo de você.**

PIMENTEL, P. Democracia e reformas: uma análise comparativa entre o discurso de João Goulart na Central do Brasil e o pensamento de Alexis de Tocqueville e John Stuart Mill. **Politeia - História e Sociedade**, v.19, n.2, p.263–273, jan. 2021.

PORTO, W. C. **O voto no Brasil: da colônia à 6a. República.**

PORTUGAL, I.; ALENCAR, P.; COWAN, D. The use of machine learning algorithms in recommender systems: A systematic review. **Expert Systems with Applications**, v.97, p.205–227, 2018.

POVO, G. do. TSE anuncia novo modelo de urnas eletrônicas para as eleições de 2022. **Gazeta do Povo**, 2021. Disponível em: <https://www.gazetadopovo.com.br/republica/breves/tse-novo-modelo-urnas-eletronicas-eleicoes-2022/>. Acessado em: 10/12/2021.

PRADO, M. **Fake News e Inteligência Artificial**: O poder dos algoritmos na guerra da desinformação.

PREZET, H. **The next level of personalization with the Amadeus Travel Platform**. Acessado em: 23/05/2023, <https://amadeus.com/en/insights/blog/next-level-personalization-amadeus-travel-platform>.

PROJECT, R. T. S. Internet Live Stats. , 2022. Disponível em: <https://www.internetlivestats.com/>. Acessado em: 02/05/2022.

RAJI, I. D.; BUOLAMWINI, J. Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products. In: AAAI/ACM CONFERENCE ON AI, ETHICS, AND SOCIETY, 2019., 2019. **Proceedings...** p.429–435.

RBA, R. TSE manda retirar das redes vídeo com mentiras de Bolsonaro a embaixadores sobre urnas. **Rede Brasil Atual**, 2022.

RICCI, F.; ROKACH, L.; SHAPIRA, B.; KANTOR, P. B. **Recommender Systems Handbook**. 1st.ed.

ROCHA, M. Google nega abuso e pede ao STF fim de inquérito sobre PL das Fake News. **Folha de São Paulo**, 2023.

RUEDIGER, M. A. **Nem tão simples assim**: o desafio de monitorar políticas públicas nas redes sociais. Acessado em: 09/12/2021, https://observademocraciadigital.org/wp-content/uploads/2020/01/nem_tao_simples_assim_fgvdapp.pdf.

RUEDIGER, M. A. et al. Desinformação on-line e contestação das eleições. , 2022.

RUEDIGER, M. A.; GRASSI, A. Desinformação on-line e eleições no Brasil: a circulação de links sobre desconfiança no sistema eleitoral brasileiro no Facebook e no YouTube (2014-2020). , 2020.

RUEDIGER, M. A.; GRASSI, A. O ecossistema digital nas eleições municipais de 2020 no Brasil: o buzz da desconfiança no sistema eleitoral no Facebook, YouTube e Twitter. , 2020.

SABENDO, F. TSE recebe ao menos 1,5 mil links denunciados como desinformação eleitoral - Don't LAI to Me 83. **Don't LAI to me - Fiquem Sabendo**, 2022.

SAMPAIO, R.; LYCARIÃO, D. Eu quero acreditar! Da importância, formas de uso e limites dos testes de confiabilidade na Análise de Conteúdo. **Revista de Sociologia e Política**, v.26, n.66, p.31–47, June 2018.

SANDVIG, C.; HAMILTON, K.; KARAHALIOS, K.; LANGBORT, C. Auditing algorithms: Research methods for detecting discrimination on internet platforms. **Data and discrimination: converting critical concerns into productive inquiry**, v.22, p.4349–4357, 2014.

SATIE, A. Entenda o projeto de voto impresso que será analisado por comissão da Câmara. **CNN Brasil**, 2021. Disponível em: <https://www.cnnbrasil.com.br/politica/entenda-o-projeto-de-voto-impresso-que-sera-analisado-por-comissao-da-camara/>. Acessado em: 10/12/2021.

SCHUDSON, M.; ZELIZER, B. Fake News in context. **AA. VV., Understanding and Addressing the Disinformation Ecosystem**, p.1–4, 2017.

SELENIUM. **Overview | Selenium**. Acessado em: 02/11/2021, <https://www.selenium.dev/documentation/overview/>.

SENADO, A. Veja aqui o Funcionamento da Urna Eletrônica. **Agência Senado**, 2018. Disponível em: <https://www12.senado.leg.br/noticias/infograficos/2018/09/veja-aqui-o-funcionamento-da-urna-eletronica>. Acessado em: 10/12/2021.

SHARMA, A. **Understanding Activation Functions in Neural Networks**. Acessado em: 24/05/2023, <https://medium.com/the-theory-of-everything/understanding-activation-functions-in-neural-networks-9491262884e0>.

SHEUGH, L.; ALIZADEH, S. H. A note on pearson correlation coefficient as a metric of similarity in recommender system. In: **AI & ROBOTICS (IRANOPEN)**, 2015., 2015. **Anais....** p.1–6.

SHU, K. et al. Fake News Detection on Social Media: A Data Mining Perspective. **ACM SIGKDD Explorations Newsletter**, v.19, n.1, p.22–36, Sept. 2017.

SILVA, L. A. C. da. Modelo Bayesiano Heterocedástico de Fatorização Probabilística de Matrizes. , 2020.

SMITH, J. Breaking Down YouTube’s Recommendation Algorithm. **Towards Data Science**, January 2021.

SOARES, F. B. et al. Desinformação sobre o Covid-19 no WhatsApp: a pandemia enquadrada como debate político. **Ciência da Informação em Revista**, v.8, n.1, p.74–94, jun. 2021.

SOARES, F.; SALGUEIRO, I.; BONOTO, C.; VINHAS, O. YouTube as a source of information about unproven drugs for Covid-19: The role of the mainstream media and recommendation algorithms in promoting misinformation. , 09 2021.

SOARES, Y. Infodemia x Democracia: Análise da Desinformação Sobre o Sistema de Votação Brasileiro em Vídeos do Facebook. , p.81, 12 2021.

SOARES, Y. **YouTubeUrnas**. Disponível em: <<https://github.com/yanbss/YouTubeUrnas>>.

SONG, M. Y.-J.; GRUZD, A. Examining sentiments and popularity of pro-and anti-vaccination videos on YouTube. In: OF THE 8TH INTERNATIONAL CONFERENCE ON SOCIAL MEDIA & SOCIETY, 2017. **Proceedings**. . . . p.1–8.

SOUTHERN, M. G. **Google Maps to Deliver More Personal Recommendations Based on Machine Learning**. Acessado em: 23/05/2023, <https://www.searchenginejournal.com/google-maps-deliver-personal-recommendations-based-machine-learning/252127/#close>.

SU, X.; KHOSHGOFTAAR, T. M. A Survey of Collaborative Filtering Techniques. **Advances in Artificial Intelligence**, v.2009, p.1–19, Oct. 2009.

SUI, M.; HAWKINS, I.; WANG, R. When falsehood wins? Varied effects of sensational elements on users' engagement with real and fake posts. **Computers in Human Behavior**, v.142, p.107654, 2023.

TAKÁCS, G.; PILÁSZY, I.; NÉMETH, B.; TIKK, D. Scalable Collaborative Filtering Approaches for Large Recommender Systems. **Journal of Machine Learning Research**, v.10, p.623–656, 03 2009.

TOMLEIN, M. et al. An Audit of Misinformation Filter Bubbles on YouTube: Bubble Bursting and Recent Behavior Changes. In: FIFTEENTH ACM CONFERENCE ON RECOMMENDER SYSTEMS, 2021. **Anais**. . . . p.1–11.

TRE-SP. No papel desde 1932, no real desde 1996: a história da urna eletrônica. **TRE-SP**, 2021.

TSE. **Série Voto Impresso**: primeira experiência com impressão do voto foi nas eleições de 2002. Disponível em: <<https://www.tse.jus.br/imprensa/noticias-tse/2017/Fevereiro/serie-voto-impresso-primeira-experiencia-com-impresao-do-voto-foi-nas-eleicoes-de-2002>>. Acesso em: 28/09/2021.

TSE, T. S. E. Você sabe o que é o Teste Público de Segurança do Sistema Eletrônico de Votação? **Tribunal Superior Eleitoral - TSE**, 2019. Disponível em: <https://www.tse.jus.br/imprensa/noticias-tse/2019/Outubro/voce-sabe-o-que-e-o-teste-publico-de-seguranca-do-sistema-eletronico-de-votacao>. Acessado em: 10/12/2021.

TSE, T. S. E. Veja como é feita a auditoria de funcionamento das urnas eletrônicas. **Tribunal Superior Eleitoral - TSE**, 2020. Disponível em: <https://www.tse.jus.br/imprensa/noticias-tse/2020/Dezembro/veja-como-funciona-a-auditoria-de-funcionamento-das-urnas-eletronicas>. Acessado em: 10/12/2021.

TUFEKCI, Z. How facebook's algorithm suppresses content diversity (modestly) and how the newsfeed rules your clicks. **Recuperado de: goo. gl/QI1XEd**, 2015.

TWITTER. **About Twitter's account suggestions**. Acessado em: 23/05/2023, <https://help.twitter.com/en/using-twitter/account-suggestions#:~:text=We%20may%20make%20suggestions%20based,Learn%20more>.

UESLEI MARCELINO, M. C. M. Bolsonaristas tentam invadir sede da PF em Brasília após prisão de indígena apoiador do presidente. **UOL**, 2022.

UNIT, T. E. I. Democracy Index 2020: In sickness and in health? **The Economist**, p.75, 2021. Obtido através de cadastro pelo link: <https://www.eiu.com/n/campaigns/democracy-index-2020/>. Acesado em: 08/12/2021.

União Europeia. **Regulamento (EU) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016, relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados**. Acesso em: 27/05/2023. Disponível em: <<https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:32016R0679from=EN>>.

UOL. New York Times: diplomatas 'ficaram abalados' após reunião com Bolsonaro. **Eleições 2022**, 2022.

USCINSKI, J. E.; BUTLER, R. W. The Epistemology of Fact Checking. **Critical Review**, v.25, n.2, p.162–180, June 2013.

VARGAS, G. DECRETO Nº 21.076, DE 24 DE FEVEREIRO DE 1932 - Código Eleitoral. **Diário Oficial da União**, 1932.

VECTOR, T. **Architecture - The YouTube Recommendation System**. Acesso em: 27/05/2023, Disponível em: <https://youtuberecommends.2018.cctp506.georgetown.domains/architecture/>.

WARDLE, C. Fake news. It's complicated. **First Draft**, v.16, p.1–11, 2017.

WARDLE, C. The need for smarter definitions and practical, timely empirical research on information disorder. **Digital Journalism**, v.6, n.8, p.951–963, 2018.

WARDLE, C. First draft's essential guide to understanding information disorder. **First draft**, 2019.

WEAVER, B. E. The History of Google Ads 20 Years in the Making (Infographic). **Instapage by Postclick**, 2022. Disponível em: <https://instapage.com/blog/google-adwords-infographic>. Acessado em: 30/07/2022.

WESTIN, R. Entenda a polêmica em torno da PEC do voto impresso. **Senado Federal**, 2021. Disponível em: <https://www12.senado.leg.br/noticias/infomaterias/2021/06/entenda-a-polemica-em-torno-da-pec-do-voto-impresso>. Acessado em: 10/12/2021.

WESTRUP, A. C. et al. **Desinformação**: crise política e saídas democráticas para as fake news.

WHITTAKER, J.; LOONEY, S.; REED, A.; VOTTA, F. Recommender systems and the amplification of extremist content. **Internet Policy Review**, v.10, n.2, p.1–29, 2021.

ZENG, C.; XING, C.-X.; ZHOU, L.-Z.; ZHENG, X.-H. Similarity measure and instance selection for collaborative filtering. **International Journal of Electronic Commerce**, v.8, n.4, p.115–129, 2004.

ZHANG, D.; YIN, J.; ZHU, X.; ZHANG, C. Homophily, Structure, and Content Augmented Network Representation Learning. In: IEEE 16TH INTERNATIONAL CONFERENCE ON DATA MINING (ICDM), 2016., 2016. **Anais...** p.609–618.

ZHANG, S.; YAO, L.; SUN, A.; TAY, Y. Deep Learning Based Recommender System. **ACM Computing Surveys**, v.52, n.1, p.1–38, Feb. 2019.

ZUBOFF, S. **A era do capitalismo de vigilância**.

ZULINI, J. P.; RICCI, P. O Código Eleitoral de 1932 e as eleições da Era Vargas: um passo na direção da democracia? **Estudos Históricos (Rio de Janeiro)**, v.33, p.600–623, 2020.

Apêndices

APÊNDICE A – Protocolo de Classificação

PROTOCOLO DE CLASSIFICAÇÃO

Algoritmo auditável: o papel dos Sistemas de Recomendação na criação de filtros bolha de desinformação sobre as Urnas Eletrônicas no YouTube

Yan Ballinhas Soares

ybsoares@inf.ufpel.edu.br

1 INTRODUÇÃO

O protocolo de classificação aqui descrito visa orientar sobre a categorização de vídeos do YouTube sobre o sistema eletrônico de votação brasileiro (Urnas Eletrônicas – UEs) baseado no seu potencial desinformativo. Para esta análise, foram extraídos vídeos da plataforma utilizando os termos de pesquisa "urna eletrônica", "urnas eletrônicas", "nova urna eletrônica 2022", "bolsonaro urnas eletrônicas" e "bolsonaro embaixadores urnas", determinados através de uma análise prévia no Google Trends acerca do assunto "Urna Eletrônica".

A criação deste protocolo é baseada numa lista de 1289 vídeos com potencial desinformativo sobre as urnas compilado pelo Tribunal Superior Eleitoral (TSE) (disponível em [1]) processados pelo autor em uma etapa prévia para a remoção de *links* duplicados e vídeos indisponíveis. A partir desta lista, foram analisados aleatoriamente 50 vídeos para uma melhor compreensão acerca do fluxo informacional sobre as UEs na plataforma e o que pode vir a caracterizar um potencial desinformativo nestes.

Foram definidas três categorias distintas levando em conta o conteúdo dos vídeos: **Sem potencial desinformativo** (-1), **Com potencial desinformativo** (1) e **Neutro** (0).

Para vídeos com mais de 25 minutos, apenas os 25 minutos iniciais devem ser considerados.

2 DESINFORMAÇÃO SOBRE AS URNAS ELETRÔNICAS

A análise de conteúdos relacionados ao sistema eletrônico de votação brasileiro não é tarefa simples. Alvo de constantes ataques, a segurança das urnas eletrônicas já foi reiteradamente comprovada através dos Testes Públicos de Segurança (TPS), realizados em 2009, 2012, 2016, 2017, 2019 e 2021 [2], além de diversas investigações de supostas fraudes já desmentidas.

Estes ataques vêm sendo mais constantes desde o ano passado, com a tramitação da PEC 135/2019 que visava a implementação do voto impresso. A PEC foi votada no dia 5 de agosto e rejeitada na Câmara dos Deputados [3], mas o fluxo de informações falsas ou distorcidas acerca das UEs continua em circulação até agora.

Um argumento desinformativo que tem tomado força durante o corrente ano foi de que a invasão hacker ao sistema administrativo do TSE, ocorrida entre abril e setembro de 2018 e que ainda está em investigação pela Polícia Federal [4], deflagraria uma vulnerabilidade do sistema eleitoral. O argumento não tem respaldo algum, visto que já foi comprovado que a invasão possibilitou acesso apenas a arquivos que **não têm ligação com as urnas e eram utilizados apenas na administração interna do TSE.**

Este foi, inclusive, um dos argumentos utilizados pelo Presidente da República (PR) Jair Bolsonaro em uma reunião com embaixadores de vários países ocorrida em julho deste ano. Na reunião, o PR utilizou deste e outros argumentos infundados para atacar a segurança das urnas eletrônicas. O próprio TSE emitiu em seguida uma nota rebatendo 20 pontos desta reunião [5].

Apesar de diversas alegações sem credibilidade alguma e vídeos de ampla circulação que alegam comprovações de fraudes carecendo de fontes ou qualquer embasamento real, nunca foi verificada uma fraude no sistema eletrônico de votação brasileiro desde sua primeira implementação [6].

Na dúvida se algum conteúdo em um vídeo é falso ou desinformativo, é sempre importante pesquisar sobre as alegações que estão sendo feitas no vídeo. Existem amplas bases de dados sobre o tema de eleições e urnas eletrônicas em agências de checagem, como a Agência Lupa [7] [8], Aos Fatos [9], Fato ou Fake [10] e Projeto Comprova [11]. Ao analisar os vídeos, deve-se atentar, além do emprego dos termos "urna eletrônica" e "urnas eletrônicas", os termos "sistema eleitoral", "processo eleitoral", "sistema

de votação", "máquina de votar", "voto impresso" e "voto auditável". De maneira objetiva, os conteúdos com potencial desinformativo são definidos abaixo, separados por tópicos.

3 SOBRE A REUNIÃO COM EMBAIXADORES

3.1 Com potencial desinformativo (1)

Vídeos reproduzindo partes ou a integridade das falas desinformativas sobre as urnas eletrônicas na reunião de Bolsonaro com os embaixadores, **sem que sejam corrigidas as informações expostas**. O mesmo se aplica a vídeos que comentam sobre os tópicos falados na reunião, **sem esclarecer** que as informações passadas pelo Presidente foram desinformativas.

Vídeos que reproduzem as falas desinformativas do PR sobre as urnas eletrônicas na reunião e só mencionam "mentiras" ou "desinformação" em seu título, **sem que sejam claramente refutadas as informações no próprio vídeo**.

3.2 Sem potencial desinformativo (-1)

Vídeos que mostram partes da reunião com os embaixadores, mas **corrigem claramente as alegações feitas** ou então **mostram a fala feita pelo TSE após a reunião, refutando as informações prestadas pelo PR**.

4 SOBRE O VOTO IMPRESSO

4.1 Com potencial desinformativo (1)

Vídeos com insinuações de que a implementação do voto impresso seria o **equivalente ao uso de cédulas de votação**, como antigamente. A proposta da PEC 135/19 do voto impresso (rejeitada em agosto de 2021) utilizaria um sistema de impressão e depósito automático do comprovante de votação, além de medidas de segurança como o uso de QR code. **Não é o mesmo que uma cédula de papel e o eleitor não levaria o comprovante para a casa**.

4.2 Sem potencial desinformativo (-1)

Vídeos explicando como seria feita a implementação do voto impresso, caso aprovada a PEC 135/19 ou que debatam a validade da proposta, **sem utilizar de argumentos desinformativos ou alegações de fraudes**.

5 SOBRE AS ELEIÇÕES E URNAS ELETRÔNICAS NO GERAL

5.1 Com potencial desinformativo (1)

Vídeos com falas que apresentam **dados conspiratórios das eleições de 2014 e 2018**, que já foram desmentidos e provados. Dois clássicos exemplos aqui são de que o PR atual teria ganhado em 2018 no primeiro turno e de que nas eleições de 2014 o resultado teria sido fraudado, tendo como base as diversas oscilações entre primeiro e segundo lugar durante a contagem dos votos [12].

Vídeos que utilizam do argumento da **invasão hacker** ao sistema do TSE em 2020, alegando que colocaria em cheque a segurança das urnas eletrônicas e do sistema de votação [13].

Vídeos apresentando o argumento de que **a contagem de votos é feita pela empresa Oracle (terceirizada)**. O supercomputador cedido para o processamento dos votos é da Oracle, a contagem não é feita pela empresa [14].

Vídeos com insinuações de que a **contagem de votos é feita "à portas fechadas"** ou de **maneira secreta por servidores do TSE**. São convidados representantes dos partidos no processo de contagem e a apuração é feita de maneira pública em tempo real [15].

Vídeos que informam que as **sugestões das Forças Armadas** para melhorias das urnas e do processo de votação foram "rechaçadas" ou não foram consideradas pela Justiça Eleitoral. 10 sugestões das forças armadas foram acatadas e todas foram devidamente analisadas e encaminhadas [16].

Vídeos que mostram o corte de um programa de televisão no qual é informado que as **urnas eletrônicas foram hackeadas em duas horas** numa convenção de hackers. As urnas neste caso eram as urnas americanas e não têm nada a ver com o processo eleitoral brasileiro [17].

Vídeos em que é dito que a **votação eletrônica é utilizada só em 3 países**. Votação eletrônica é utilizada parcial ou totalmente em vários países, não só 3 [18].

Vídeos que falam que no caso de falha e substituição de urna, os **votos já computados são perdidos** [19].

Vídeos com **falas deslocada do passado** (principalmente de candidatos) colocando dúvida sobre o sistema eleitoral sendo utilizadas como argumento para duvidar da lisura eleitoral atual, **sem levar em conta as inúmeras mudanças** que ocorreram nas urnas eletrônicas e no processo eleitoral como um todo desde o momento da declaração sendo mostrada.

5.2 Sem potencial desinformativo (-1)

Vídeos que informam que 39% das urnas não passaram por testes públicos, de acordo com levantamento dos membros das Forças Armadas que compõem a Comissão de Transparência das Eleições. De fato, os modelos UE2020 (39% dos que serão utilizados nas eleições de 2022) não haviam sido testados no último TPS. O TSE explicou que isto não ocorreu porque os modelos foram recebidos após o final dos testes, em dezembro de 2021. Mesmo assim, estes modelos foram enviados e testados por pesquisadores da USP, UFPE e Unicamp e passaram normalmente nos testes [20].

6 SOBRE MANIFESTAÇÕES E O ARTIGO 142

6.1 Com potencial desinformativo (1)

Vídeos que embasam a **invocação do Artigo 142 da Constituição** (uso das Forças Armadas para garantia dos poderes constitucionais) com o pretexto de fraude nas eleições ou não confiança na lisura do processo de votação [21].

7 ATENÇÃO ESPECÍFICA A VÍDEOS ENVOLVENDO COMENTARISTAS

7.1 Com potencial desinformativo (1)

Vídeos que colocam comentaristas a opinar sobre as urnas eletrônicas ou o processo de votação e é utilizada de alguma fala desinformativa sem que esta seja **claramente desmentida ou corrigida**.

7.2 Sem potencial desinformativo (-1)

Quando o comentarista em questão é **claramente corrigido** e os fatos corretos são apresentados, de maneira que não fique uma dúvida sobre a lisura do processo eleitoral ou das urnas eletrônicas.

8 OBSERVAÇÕES FINAIS

Não se deve entrar no mérito de apurar desinformações sobre outros temas ou discursos que **não tenham a ver com o sistema de votação/urnas eletrônicas**. Muitas falas de comentaristas fazem uma série de acusações sobre outros temas. Estas não devem ser julgadas.

Para ser considerado com/sem potencial desinformativo, o vídeo deve falar **total ou parcialmente sobre as URNAS ELETRÔNICAS ou SISTEMA ELETRÔNICO DE VOTAÇÃO**. Vídeos sobre as eleições como um todo, programas de governo, candidaturas, etc. são considerados neutros (0).

Pesquisas eleitorais, **vídeos sobre os ministros do STF/TSE sem que mencione a urna ou o sistema eletrônico de votação** também são considerados neutros (0).

Discussões técnicas visando melhorias no sistema são absolutamente necessárias. Portanto, vídeos que contemplem este tipo de conteúdo (por exemplo: vídeos explicando a maneira que o voto impresso seria implementado), sem que sejam utilizados argumentos desinformativos sobre supostas fraudes são categorizados como **Sem potencial desinformativo** (-1).

Vídeos que explicam o funcionamento de votos brancos, nulos ou de legenda são considerados neutros (0), visto que não têm relação direta com as urnas eletrônicas, apenas explicam como funcionam os mecanismos de anulação do exercício do sufrágio.

REFERÊNCIAS

- [1] Fiquem Sabendo. Tse recebe ao menos 1,5 mil links denunciados como desinformação eleitoral - don't lai to me 83. *Don't LAI to me - Fiquem Sabendo*, 2022. Disponível em: <https://fiquemsabendo.substack.com/p/tse-ja-pediu-remocao-de-ao-menos>. Acessado em 17/09/2022.
- [2] Tribunal Superior Eleitoral. Teste público de segurança. *Justiça Eleitoral*, 2022. Disponível em: <https://www.justicaeleitoral.jus.br/tps/>. Acessado em 17/09/2022.
- [3] Heloisa Cristaldo. Câmara dos deputados rejeita pec do voto impresso. *Agência Brasil*, 2021. Disponível em: <https://agenciabrasil.ebc.com.br/politica/noticia/2021-08/camara-dos-deputados-rejeita-pec-do-voto-impresso>. Acessado em 17/09/2022.
- [4] Fabio Serapião. Entenda inquérito do ataque hacker ao tse utilizado por bolsonaro para questionar urnas. *Folha de São Paulo*, 2022. Disponível em: <https://www1.folha.uol.com.br/poder/2022/08/entenda-inquerito-do-ataque-hacker-ao-tse-utilizado-por-bolsonaro-para-questionar-urnas.shtml>. Acessado em 17/09/2022.
- [5] Redação Oeste. Tse rebate 20 pontos expostos por bolsonaro a embaixadores. *Revista Oeste*, 2022. Disponível em: <https://revistaoeste.com/politica/eleicoes-2022/tse-rebate-20-pontos-expostos-por-bolsonaro-a-embaxadores/>. Acessado em 17/09/2022.
- [6] G1. Urna eletrônica é segura e não há evidência de fraude, diz associação de peritos da polícia federal. *G1*, 2021. Disponível em: <https://g1.globo.com/politica/noticia/2021/08/05/urna-eletronica-e-segura-e-nao-ha-evidencia-de-fraude-diz-associacao-de-peritos-da-policia-federal.ghtml>. Acessado em 05/11/2021.
- [7] Agência Lupa. Checagem de fatos: Eleições 2018. *Agência Lupa*, 2018. Disponível em: <https://piaui.folha.uol.com.br/lupa/category/eleicoes-2018/>. Acessado em 05/11/2021.
- [8] Agência Lupa. Checagem de fatos: Eleições 2020. *Agência Lupa*, 2020. Disponível em: <https://piaui.folha.uol.com.br/lupa/category/eleicoes-2020/>. Acessado em 05/11/2021.
- [9] Aos Fatos. Aos fatos - valorize o que é real: "urna eletrônica". *Aos Fatos*, 2021. Disponível em: <https://www.aosfatos.org/search/?q=urna+eletr%C3%B4nica>. Acessado em 05/11/2021.
- [10] Fato ou Fake. G1 - fato ou fake. *Fato ou Fake*, 2021. Disponível em: <https://g1.globo.com/fato-ou-fake/>. Acessado em 05/11/2021.
- [11] Projeto Comprova. Projeto comprova - eleições. *Projeto Comprova*, 2021. Disponível em: <https://projetocomprova.com.br/?filter=eleicoes>. Acessado em 05/11/2021.
- [12] Projeto Comprova. Não há provas de fraudes em 'apagões' nas eleições de 2014 e 2018, ao contrário do que afirma post. *Estadão*, 2022. Disponível em: <https://politica.estadao.com.br/blogs/estadao-verifica/apagoes-eleicoes-2014-2018-fraude-urnas/>. Acessado em 18/09/2022.
- [13] Tribunal Superior Eleitoral. Fato ou boato: ataques hackers não afetaram a segurança das eleições municipais de 2020. *Justiça Eleitoral*, 2021. Disponível em: <https://www.tse.jus.br/comunicacao/noticias/2021/Junho/fato-ou-boato-ataques-hackers-nao-afetaram-a-seguranca-das-eleicoes-municipais-de-2020>. Acessado em 18/09/2022.
- [14] Tribunal Superior Eleitoral. Nota de esclarecimento sobre nuvem para contabilizar votos. *Justiça Eleitoral*, 2020. Disponível em: <https://www.tse.jus.br/comunicacao/noticias/2020/Novembro/nota-de-esclarecimento-sobre-nuvem-para-contabilizar-votos>. Acessado em 18/09/2022.

- [15] Tribunal Superior Eleitoral. Fato ou boato: é falso que a apuração das eleições seja feita de forma secreta por servidores do tse. *Justiça Eleitoral*, 2021. Disponível em: <https://www.tse.jus.br/comunicacao/noticias/2021/Julho/fato-ou-boato-e-falso-que-a-apuracao-das-eleicoes-seja-feita-de-forma-secreta-por-servidores-do-tse>. Acessado em 18/09/2022.
- [16] Lucas Rocha. Tse acolhe 10 sugestões de militares para o sistema eleitoral; saiba quais são. *Brasil de Fato*, 2022. Disponível em: <https://www.brasildefato.com.br/2022/06/13/tse-acolhe-10-sugestoes-de-militares-para-o-sistema-eleitoral-saiba-quais-sao>. Acessado em 18/09/2022.
- [17] Pedro Prata. Trecho de programa jornalístico é cortado e retirado de contexto para atacar urnas eletrônicas. *Estadão*, 2021. Disponível em: <https://politica.estadao.com.br/blogs/estadao-verifica/programa-sem-contexto-urnas-hackeadas/>. Acessado em 18/09/2022.
- [18] International Institute for Democracy and Electoral Assistance (International IDEA). Is e-voting currently used in any elections with emb. *International Institute for Democracy and Electoral Assistance (International IDEA)*, 2022. Disponível em: <https://www.idea.int/data-tools/question-view/742>. Acessado em 18/09/2022.
- [19] Ana Heloisa de Aragão Bastos. E quando a urna quebra durante a votação? *Justiça Eleitoral*, 2015. Disponível em: https://www.tse.jus.br/o-tse/escola-judiciaria-eleitoral/publicacoes/revistas-da-eje/artigos/revista-eletronica-eje-n-2-ano-5/copy_of_or_-_que_-_a_-_urna_-_eletronica_-_e_-_segura. Acessado em 18/09/2022.
- [20] Tribunal Regional Eleitoral AL. Universidades validam nova urna e códigos-fonte dos sistemas eleitorais. *Justiça Eleitoral*, 2022. Disponível em: <https://www.tre-al.jus.br/comunicacao/noticias/2022/Agosto/universidades-validam-nova-urna-e-codigos-fonte-dos-sistemas-eleitorais>. Acessado em 18/09/2022.
- [21] Câmara dos Deputados. Câmara emite parecer esclarecendo que artigo 142 da constituição não autoriza intervenção militar. *Agência Câmara de Notícias*, 2020. Disponível em: <https://www.camara.leg.br/noticias/667144-camara-emite-parecer-esclarecendo-que-artigo-142-da-constituicao-nao-autoriza-intervencao-militar/>. Acessado em 18/09/2022.

APÊNDICE B – Tabela do teste de validação do Protocolo de Classificação

Tabela 21 – Avaliação do autor e dos avaliadores do Protocolo de Classificação

Link	Autor	Avaliador 1	Avaliador 2
https://www.youtube.com/watch?v=HFcIWMnV1vM	1	1	1
https://www.youtube.com/watch?v=6fQNPPDQcyg	1	1	1
https://www.youtube.com/watch?v=_acUB4b9UyY	1	1	-1
https://www.youtube.com/watch?v=jzobSBH4wx8	-1	-1	-1
https://www.youtube.com/watch?v=jyT-gVeqRBA	0	0	0
https://www.youtube.com/watch?v=g9a6oZ7Wgdl	0	0	0
https://www.youtube.com/watch?v=cvgPonvurtM	-1	1	-1
https://www.youtube.com/watch?v=dAD5br-qO34	0	0	-1
https://www.youtube.com/shorts/oidnt1M44Pw	1	1	-1
https://www.youtube.com/watch?v=zmw2ObsnIFE	-1	-1	-1
https://www.youtube.com/watch?v=ipD5ALMyf2s	-1	-1	-1
https://www.youtube.com/watch?v=OW4SSO0aCWU	-1	-1	-1
https://www.youtube.com/watch?v=m-MZ65o8yEM	1	1	1
https://www.youtube.com/watch?v=YEuWmHXG3EY	0	1	0
https://www.youtube.com/watch?v=P0T-2dCLlx8	-1	1	-1
https://www.youtube.com/watch?v=cxYI5UYAnq8	0	0	0
https://www.youtube.com/watch?v=jzobSBH4wx8	-1	1	-1
https://www.youtube.com/watch?v=UfuGrYC2Nik	0	0	0
https://www.youtube.com/watch?v=PYZNi_eq18I	-1	-1	-1
https://www.youtube.com/watch?v=HFcIWMnV1vM	1	1	1
https://www.youtube.com/watch?v=eP8VXPnyULc	1	1	1
https://www.youtube.com/watch?v=_267qod4iXg	-1	-1	-1
https://www.youtube.com/watch?v=C8hfZVB8Do8	1	1	1
https://www.youtube.com/watch?v=Y6nrJpAsHkA	1	1	1
https://www.youtube.com/shorts/D-CcqxMwb0s	-1	1	-1
https://www.youtube.com/watch?v=O7qugPOOnbl	-1	-1	-1
https://www.youtube.com/shorts/hvZgcfvby_s	1	1	1
https://www.youtube.com/shorts/Z_eCEj2GLjw	-1	0	-1
https://www.youtube.com/watch?v=VO0kej14-dE	1	1	1

https://www.youtube.com/watch?v=i36mLoKMobQ	1	1	1
https://www.youtube.com/watch?v=dfPPu9xXcSQ	-1	-1	-1
https://www.youtube.com/watch?v=C8hfZVB8Do8	1	1	1
https://www.youtube.com/watch?v=BksNm7fDfyg	0	0	0
https://www.youtube.com/watch?v=oBiNzV8utV8	-1	-1	-1
https://www.youtube.com/watch?v=xkLt6hdal9Q	0	0	0
https://www.youtube.com/watch?v=Zx28GpYtOXc	1	1	1
https://www.youtube.com/watch?v=eUAS60YvRNY	-1	-1	-1
https://www.youtube.com/watch?v=eP8VXPnyULc	1	1	1
https://www.youtube.com/watch?v=Zx28GpYtOXc	1	1	1
https://www.youtube.com/watch?v=OSmFVLR9594	0	0	0
https://www.youtube.com/watch?v=GJI9RgqDIAI	-1	-1	-1
https://www.youtube.com/watch?v=ggLk6DwYFPY	-1	-1	-1
https://www.youtube.com/watch?v=Qr3BWKWixpk	0	0	0
https://www.youtube.com/watch?v=7AdArwt4EJs	-1	-1	1
https://www.youtube.com/watch?v=L4gbKspo4kY	0	0	0
https://www.youtube.com/watch?v=yAE1dVrVYC4	0	0	0
https://www.youtube.com/watch?v=gtYqrfMK2Tc	-1	-1	-1
https://www.youtube.com/watch?v=Occa5lusj3U	-1	-1	-1
https://www.youtube.com/watch?v=oWMqxf_xbAg	0	0	0
https://www.youtube.com/watch?v=_yPWaPv7ur4	1	1	1
https://www.youtube.com/watch?v=nR0cG4ntCKA	-1	-1	-1
https://www.youtube.com/watch?v=oBiNzV8utV8	-1	1	-1
https://www.youtube.com/watch?v=7hepX8Y7A-M	1	-1	-1
https://www.youtube.com/watch?v=p4D7OyXs9tw	0	0	0
https://www.youtube.com/shorts/hvZgcfvby_s	1	1	1
https://www.youtube.com/watch?v=SuDOY98H6wE	0	0	0
https://www.youtube.com/watch?v=lGhAf2ENtJs	1	1	1
https://www.youtube.com/watch?v=OSmFVLR9594	0	0	0
https://www.youtube.com/watch?v=eOlwlGo9g5M	0	0	0
https://www.youtube.com/watch?v=xkLt6hdal9Q	0	0	0
https://www.youtube.com/shorts/Ng-ILNHpysE	-1	-1	-1
https://www.youtube.com/shorts/hvZgcfvby_s	1	1	1
https://www.youtube.com/watch?v=B9WtfQizHD8	0	0	0
https://www.youtube.com/watch?v=PORTmOLhswW	-1	-1	-1
https://www.youtube.com/watch?v=acoTO_URqWo	0	0	0
https://www.youtube.com/shorts/o1XXk736Mc8	-1	0	-1
https://www.youtube.com/watch?v=1XQVLdzlrU0	0	0	0
https://www.youtube.com/watch?v=lrfKL54bwAc	1	1	1

https://www.youtube.com/watch?v=pTDjLBeaVEc	1	1	-1
https://www.youtube.com/watch?v=5yh_Ilij0ZE	1	1	1
https://www.youtube.com/watch?v=9nrR1HMmy2k	0	0	0
https://www.youtube.com/watch?v=TQ5-YPliw2g	1	1	1
https://www.youtube.com/watch?v=HooyxC4w2OI	0	0	0
https://www.youtube.com/watch?v=p1a_5nYwV1Q	0	0	0
https://www.youtube.com/watch?v=SuDOY98H6wE	0	0	0
https://www.youtube.com/watch?v=yAE1dVrVYC4	0	0	0
https://www.youtube.com/watch?v=jzobSBH4wx8	-1	-1	-1
https://www.youtube.com/watch?v=oBiNzV8utV8	-1	1	-1
https://www.youtube.com/watch?v=S87CD7S-L5E	-1	-1	-1
https://www.youtube.com/watch?v=Y9lclPkAilQ	1	1	1
https://www.youtube.com/watch?v=yWNhVXHz7vg	0	0	0
https://www.youtube.com/watch?v=m-MZ65o8yEM	1	1	1
https://www.youtube.com/watch?v=WzAJXvKAxxM	1	1	-1
https://www.youtube.com/watch?v=flHB9KISu6A	0	0	0
https://www.youtube.com/watch?v=Az58JIQ4v9w	-1	-1	-1
https://www.youtube.com/watch?v=62GYGqcRuos	1	1	1
https://www.youtube.com/watch?v=eOlwlGo9g5M	0	0	0
https://www.youtube.com/watch?v=_yPWaPv7ur4	1	1	1
https://www.youtube.com/watch?v=O7qugPOOnbl	-1	-1	-1
https://www.youtube.com/watch?v=xkLt6hdal9Q	0	0	0

APÊNDICE C – Tabela de Resultados da Etapa 2

Tabela 22 – Resultados obtidos nas 10 iterações da Etapa 2

Iteração	Vídeos Assistidos	Coefficiente de Overlap
1	0	0.0
1	1	0.7
1	2	0.3
1	3	0.6
1	4	0.5
1	5	0.3
1	6	0.7
1	7	0.3
1	8	0.4
1	9	0.9
1	10	0.7
1	11	0.5
1	12	0.5
1	13	0.8
1	14	0.9
1	15	0.8
1	16	0.6
1	17	0.8
1	18	0.9
1	19	0.8
1	20	0.6
1	21	0.8
1	22	0.7
1	23	0.8
1	24	0.8
1	25	0.9
1	26	0.8
1	27	0.8
1	28	0.7

1	29	0.6
1	30	0.5
1	31	0.6
1	32	0.5
1	33	0.8
1	34	0.6
1	35	0.5
1	36	0.8
1	37	0.8
1	38	0.9
1	39	0.9
1	40	0.8
1	41	0.8
1	42	0.9
1	43	0.8
1	44	0.9
1	45	0.8
1	46	0.9
1	47	0.5
1	48	1.0
2	0	0.0
2	1	0.3
2	2	0.3
2	3	0.4
2	4	0.5
2	5	0.5
2	6	0.1
2	7	0.6
2	8	0.3
2	9	0.7
2	10	0.7
2	11	0.3
2	12	1.0
3	0	0.0
3	1	0.1
3	2	0.1
3	3	0.4
3	4	0.3
3	5	0.4

3	6	0.9
3	7	0.8
3	8	0.8
3	9	0.7
3	10	0.6
3	11	0.7
3	12	0.8
3	13	0.7
3	14	0.6
3	15	0.7
3	16	0.6
3	17	0.5
3	18	0.8
3	19	0.6
3	20	0.8
3	21	0.6
3	22	0.9
3	23	0.7
3	24	0.4
3	25	0.7
3	26	0.7
3	27	0.8
3	28	0.8
3	29	1.0
4	0	0.0
4	1	0.3
4	2	0.5
4	3	0.4
4	4	0.7
4	5	0.4
4	6	0.5
4	7	0.5
4	8	0.5
4	9	0.7
4	10	0.6
4	11	0.8
4	12	0.3
4	13	0.4
4	14	0.4

4	15	0.4
4	16	0.6
4	17	0.5
4	18	0.8
4	19	0.6
4	20	0.4
4	21	0.5
4	22	0.4
4	23	0.7
4	24	0.5
4	25	0.6
4	26	0.6
4	27	0.3
4	28	0.7
4	29	0.8
4	30	0.7
4	31	0.5
4	32	0.5
4	33	0.5
4	34	0.3
4	35	0.8
4	36	0.6
4	37	0.9
4	38	0.5
4	39	0.6
4	40	0.7
4	41	0.6
4	42	0.7
4	43	0.8
4	44	0.6
4	45	0.9
4	46	0.9
4	47	0.7
4	48	0.6
4	49	0.8
4	50	0.9
4	51	0.9
4	52	1.0
5	0	0.0

5	1	0.1
5	2	0.4
5	3	0.4
5	4	0.1
5	5	0.5
5	6	0.3
5	7	1.0
6	0	0.0
6	1	0.8
6	2	0.6
6	3	0.7
6	4	0.0
6	5	0.2
6	6	0.7
6	7	0.4
6	8	0.3
6	9	0.7
6	10	0.2
6	11	0.7
6	12	0.5
6	13	0.7
6	14	0.7
6	15	0.6
6	16	0.6
6	17	1.0
7	0	0.0
7	1	0.2
7	2	0.4
7	3	0.1
7	4	0.6
7	5	0.4
7	6	0.6
7	7	0.4
7	8	0.2
7	9	0.5
7	10	0.7
7	11	0.4
7	12	0.9
7	13	0.8

7	14	0.8
7	15	0.3
7	16	0.4
7	17	0.5
7	18	0.8
7	19	0.7
7	20	0.6
7	21	0.8
7	22	0.8
7	23	0.7
7	24	0.5
7	25	0.7
7	26	0.7
7	27	0.6
7	28	0.7
7	29	0.8
7	30	0.8
7	31	0.5
7	32	0.8
7	33	0.9
7	34	0.5
7	35	0.8
7	36	0.7
7	37	0.8
7	38	0.7
7	39	0.9
7	40	0.7
7	41	0.6
7	42	0.7
7	43	0.5
7	44	0.7
7	45	0.7
7	46	0.8
7	47	0.9
7	48	0.7
7	49	0.7
7	50	0.9
7	51	0.9
7	52	0.7

7	53	0.9
7	54	0.7
7	55	0.7
7	56	0.9
7	57	0.7
7	58	0.6
7	59	0.6
7	60	0.6
7	61	0.7
7	62	0.8
7	63	0.9
7	64	0.8
7	65	0.7
7	66	0.8
7	67	0.6
7	68	0.8
7	69	0.8
7	70	0.8
7	71	0.8
7	72	0.9
7	73	0.9
7	74	0.9
7	75	0.3
7	76	0.9
7	77	1.0
8	0	0.0
8	1	0.5
8	2	0.6
8	3	0.6
8	4	0.8
8	5	0.5
8	6	0.5
8	7	0.6
8	8	0.8
8	9	0.5
8	10	0.5
8	11	0.6
8	12	0.5
8	13	0.6

8	14	0.5
8	15	0.5
8	16	0.4
8	17	0.4
8	18	0.7
8	19	0.7
8	20	0.7
8	21	0.6
8	22	0.6
8	23	0.8
8	24	0.5
8	25	0.9
8	26	0.8
8	27	0.7
8	28	0.6
8	29	0.6
8	30	0.6
8	31	0.8
8	32	0.4
8	33	0.7
8	34	0.8
8	35	0.8
8	36	0.9
8	37	0.4
8	38	0.5
8	39	0.7
8	40	0.7
8	41	0.7
8	42	0.9
8	43	0.8
8	44	0.7
8	45	0.8
8	46	0.6
8	47	0.9
8	48	0.7
8	49	0.6
8	50	0.7
8	51	0.8
8	52	1.0

9	0	0.0
9	1	0.2
9	2	0.5
9	3	0.3
9	4	0.4
9	5	0.6
9	6	0.2
9	7	0.4
9	8	0.3
9	9	0.6
9	10	0.6
9	11	0.5
9	12	1.0
10	0	0.0
10	1	0.2
10	2	0.6
10	3	0.7
10	4	0.1
10	5	0.5
10	6	0.7
10	7	0.3
10	8	0.8
10	9	0.7
10	10	0.7
10	11	0.6
10	12	0.7
10	13	0.7
10	14	0.9
10	15	0.9
10	16	0.9
10	17	0.9
10	18	0.7
10	19	0.7
10	20	0.9
10	21	0.8
10	22	0.9
10	23	0.6
10	24	0.8
10	25	0.6

10	26	0.6
10	27	0.5
10	28	0.6
10	29	1.0

APÊNDICE D – Tabela de Resultados de *Search Engine Results Page* –
Misinformation Score (SERP-MS)

Tabela 23 – Resultados de *Search Engine Results Page* – *Misinformation Score* (SERP-MS)

Perfil	Termo	Algoritmo	SERP-MS
Etapa 1	bolsonaro embaixadores urnas	Classificação	-0,27
Etapa 1	bolsonaro embaixadores urnas	Contagem	0,10
Etapa 1	bolsonaro embaixadores urnas	Data	-0,45
Etapa 1	bolsonaro embaixadores urnas	Relevância	-0,77
Etapa 1	bolsonaro urnas eletrônicas	Classificação	0,45
Etapa 1	bolsonaro urnas eletrônicas	Contagem	-0,02
Etapa 1	bolsonaro urnas eletrônicas	Data	-0,04
Etapa 1	bolsonaro urnas eletrônicas	Relevância	-0,75
Etapa 1	nova urna eletronica 2022	Classificação	-0,90
Etapa 1	nova urna eletronica 2022	Contagem	-1,00
Etapa 1	nova urna eletronica 2022	Data	-1,00
Etapa 1	nova urna eletronica 2022	Relevância	-1,00
Etapa 1	urna eletrônica	Classificação	-0,75
Etapa 1	urna eletrônica	Contagem	-0,52
Etapa 1	urna eletrônica	Data	-1,00
Etapa 1	urna eletrônica	Relevância	-1,00
Etapa 1	urnas eletronicas	Classificação	0,14
Etapa 1	urnas eletronicas	Contagem	0,30
Etapa 1	urnas eletronicas	Data	-0,54
Etapa 1	urnas eletronicas	Relevância	-0,98
Desinformativo	bolsonaro embaixadores urnas	Classificação	-0,27
Desinformativo	bolsonaro embaixadores urnas	Contagem	0,00
Desinformativo	bolsonaro embaixadores urnas	Data	-0,44
Desinformativo	bolsonaro embaixadores urnas	Relevância	-0,60
Desinformativo	bolsonaro urnas eletrônicas	Classificação	0,42
Desinformativo	bolsonaro urnas eletrônicas	Contagem	-0,06
Desinformativo	bolsonaro urnas eletrônicas	Data	-0,41

Desinformativo	bolsonaro urnas eletrônicas	Relevância	-0,69
Desinformativo	nova urna eletronica 2022	Classificação	-0,85
Desinformativo	nova urna eletronica 2022	Contagem	-0,84
Desinformativo	nova urna eletronica 2022	Data	-0,35
Desinformativo	nova urna eletronica 2022	Relevância	-1,00
Desinformativo	urna eletrônica	Classificação	-0,77
Desinformativo	urna eletrônica	Contagem	-0,48
Desinformativo	urna eletrônica	Data	-0,76
Desinformativo	urna eletrônica	Relevância	-1,00
Desinformativo	urnas eletronicas	Classificação	0,20
Desinformativo	urnas eletronicas	Contagem	-0,21
Desinformativo	urnas eletronicas	Data	-0,64
Desinformativo	urnas eletronicas	Relevância	-0,99
Informativo	bolsonaro embaixadores urnas	Classificação	-0,29
Informativo	bolsonaro embaixadores urnas	Contagem	-0,01
Informativo	bolsonaro embaixadores urnas	Data	-0,51
Informativo	bolsonaro embaixadores urnas	Relevância	-0,70
Informativo	bolsonaro urnas eletrônicas	Classificação	0,57
Informativo	bolsonaro urnas eletrônicas	Contagem	-0,15
Informativo	bolsonaro urnas eletrônicas	Data	-0,22
Informativo	bolsonaro urnas eletrônicas	Relevância	-0,68
Informativo	nova urna eletronica 2022	Classificação	-0,96
Informativo	nova urna eletronica 2022	Contagem	-0,83
Informativo	nova urna eletronica 2022	Data	-0,54
Informativo	nova urna eletronica 2022	Relevância	-1,00
Informativo	urna eletrônica	Classificação	-0,64
Informativo	urna eletrônica	Contagem	-0,58
Informativo	urna eletrônica	Data	-0,80
Informativo	urna eletrônica	Relevância	-0,85
Informativo	urnas eletronicas	Classificação	0,41
Informativo	urnas eletronicas	Contagem	-0,25
Informativo	urnas eletronicas	Data	-0,67
Informativo	urnas eletronicas	Relevância	-0,94
Misto	bolsonaro embaixadores urnas	Classificação	-0,18
Misto	bolsonaro embaixadores urnas	Contagem	-0,01
Misto	bolsonaro embaixadores urnas	Data	-0,12
Misto	bolsonaro embaixadores urnas	Relevância	-0,67
Misto	bolsonaro urnas eletrônicas	Classificação	0,34
Misto	bolsonaro urnas eletrônicas	Contagem	-0,16

Misto	bolsonaro urnas eletrônicas	Data	-0,48
Misto	bolsonaro urnas eletrônicas	Relevância	-0,74
Misto	nova urna eletronica 2022	Classificação	-0,92
Misto	nova urna eletronica 2022	Contagem	-0,83
Misto	nova urna eletronica 2022	Data	-0,49
Misto	nova urna eletronica 2022	Relevância	-1,00
Misto	urna eletrônica	Classificação	-0,80
Misto	urna eletrônica	Contagem	-0,54
Misto	urna eletrônica	Data	-0,83
Misto	urna eletrônica	Relevância	-1,00
Misto	urnas eletronicas	Classificação	0,29
Misto	urnas eletronicas	Contagem	-0,25
Misto	urnas eletronicas	Data	-0,79
Misto	urnas eletronicas	Relevância	-0,99
Neutro	bolsonaro embaixadores urnas	Classificação	-0,29
Neutro	bolsonaro embaixadores urnas	Contagem	-0,01
Neutro	bolsonaro embaixadores urnas	Data	-0,27
Neutro	bolsonaro embaixadores urnas	Relevância	-0,63
Neutro	bolsonaro urnas eletrônicas	Classificação	0,36
Neutro	bolsonaro urnas eletrônicas	Contagem	-0,20
Neutro	bolsonaro urnas eletrônicas	Data	-0,52
Neutro	bolsonaro urnas eletrônicas	Relevância	-0,76
Neutro	nova urna eletronica 2022	Classificação	-0,93
Neutro	nova urna eletronica 2022	Contagem	-0,83
Neutro	nova urna eletronica 2022	Data	-0,40
Neutro	nova urna eletronica 2022	Relevância	-1,00
Neutro	urna eletrônica	Classificação	-0,95
Neutro	urna eletrônica	Contagem	-0,53
Neutro	urna eletrônica	Data	-0,73
Neutro	urna eletrônica	Relevância	-1,00
Neutro	urnas eletronicas	Classificação	-0,18
Neutro	urnas eletronicas	Contagem	-0,37
Neutro	urnas eletronicas	Data	-0,85
Neutro	urnas eletronicas	Relevância	-0,98

APÊNDICE E – Gráficos do SERP-MS e distribuição percentual de vídeos
recomendados por Algoritmo de Ordenação e Termo de Pesquisa
para cada perfil da Etapa 3

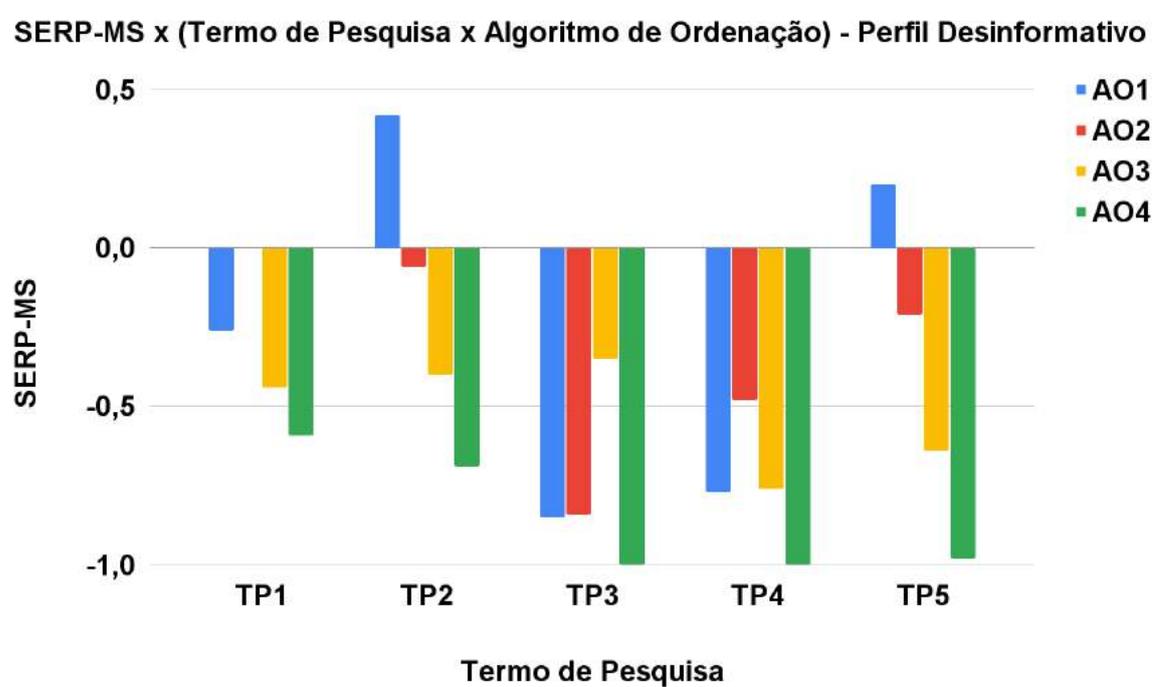
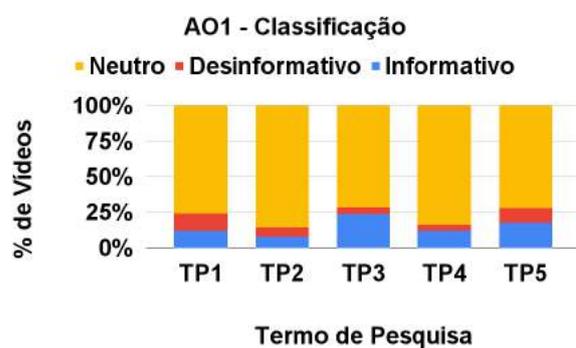
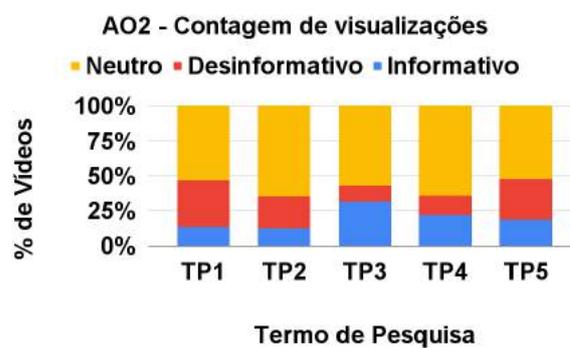


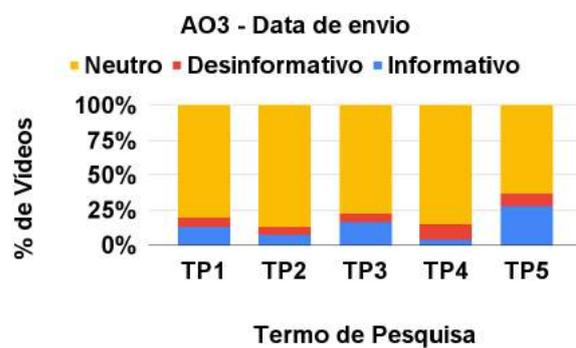
Figura 41 – SERP-MS x Pares (Termo de Pesquisa x Algoritmo de Ordenação) dos resultados da Etapa 3 – Perfil Desinformativo



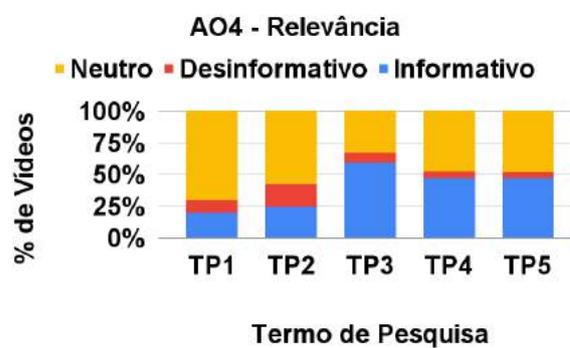
(a) Distribuição de vídeos classificados com AO1



(b) Distribuição de vídeos classificados com AO2



(c) Distribuição de vídeos classificados com AO3



(d) Distribuição de vídeos classificados com AO4

Figura 42 – Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 4 Algoritmos de Ordenação – Etapa 3 – Vídeos Recomendados – Perfil Desinformativo

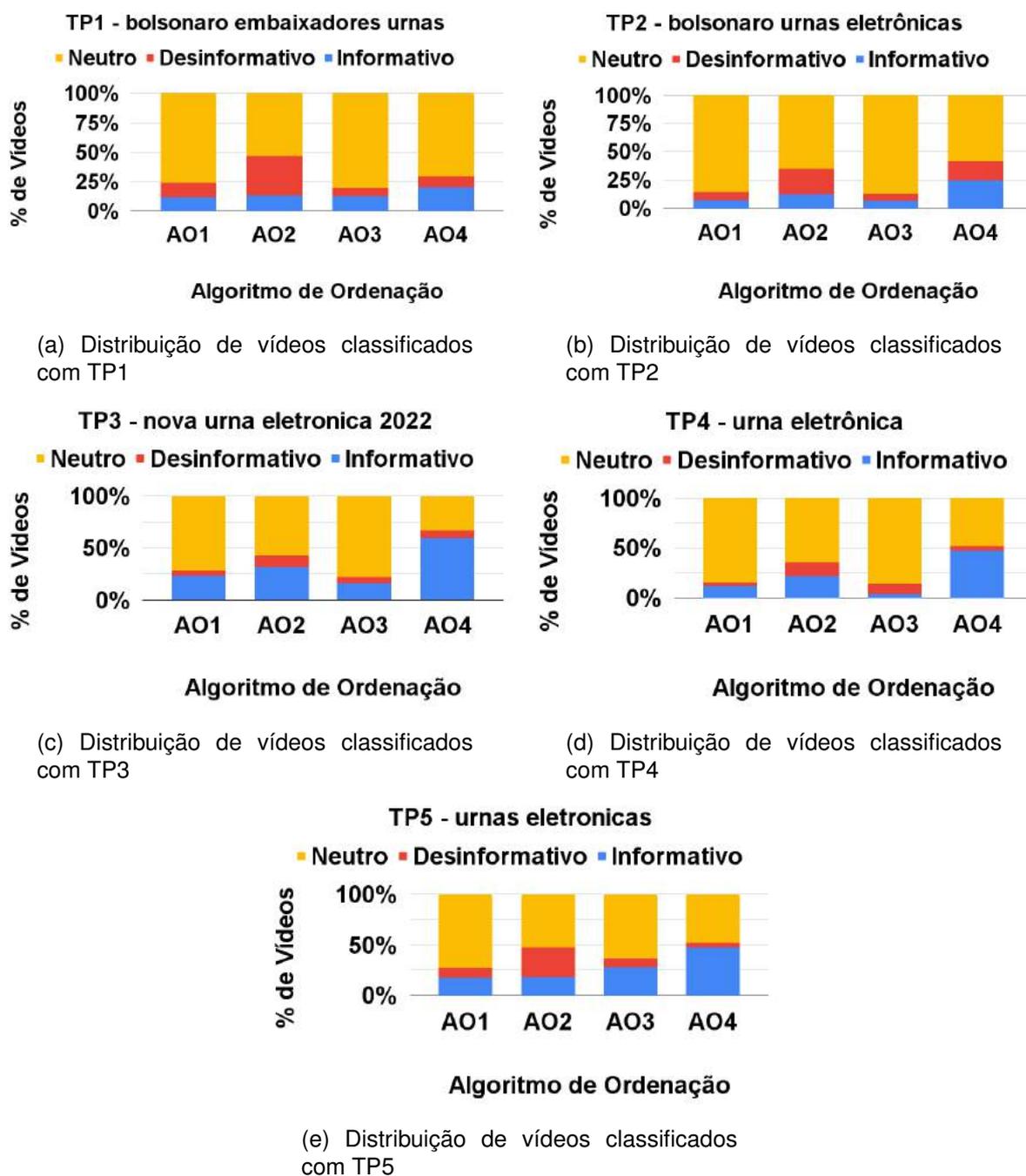


Figura 43 – Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 5 Termos de Pesquisa – Etapa 3 – Vídeos Recomendados – Perfil Desinformativo

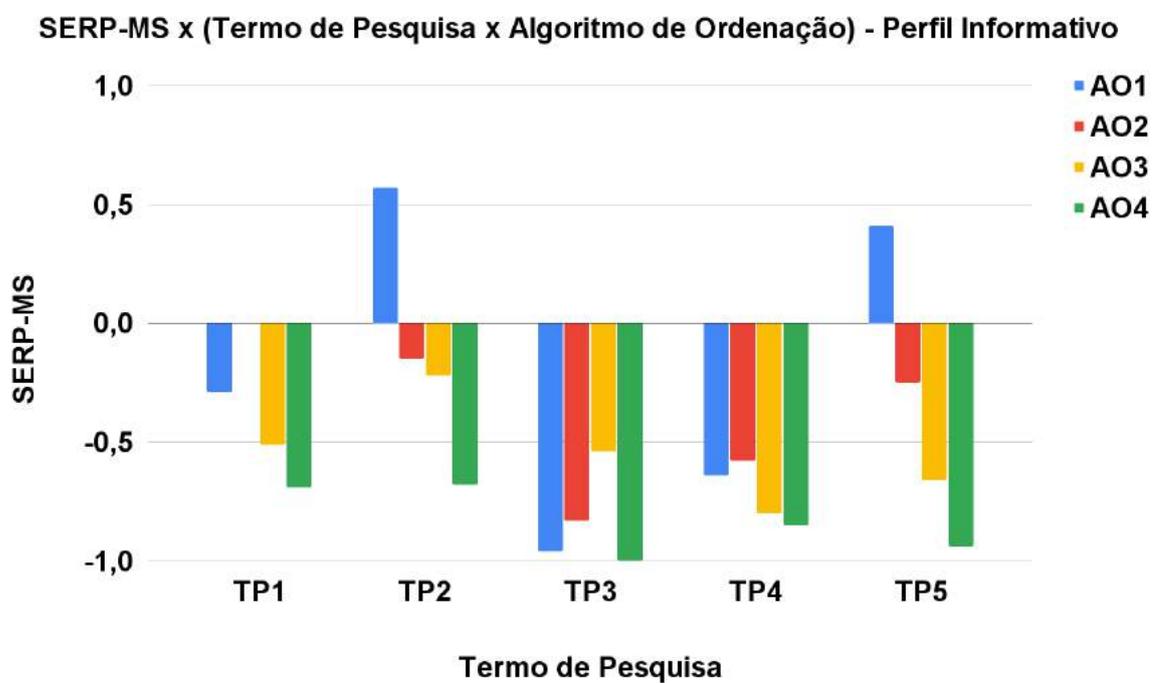


Figura 44 – SERP-MS x Pares (Termo de Pesquisa x Algoritmo de Ordenação) dos resultados da Etapa 3 – Perfil Informativo

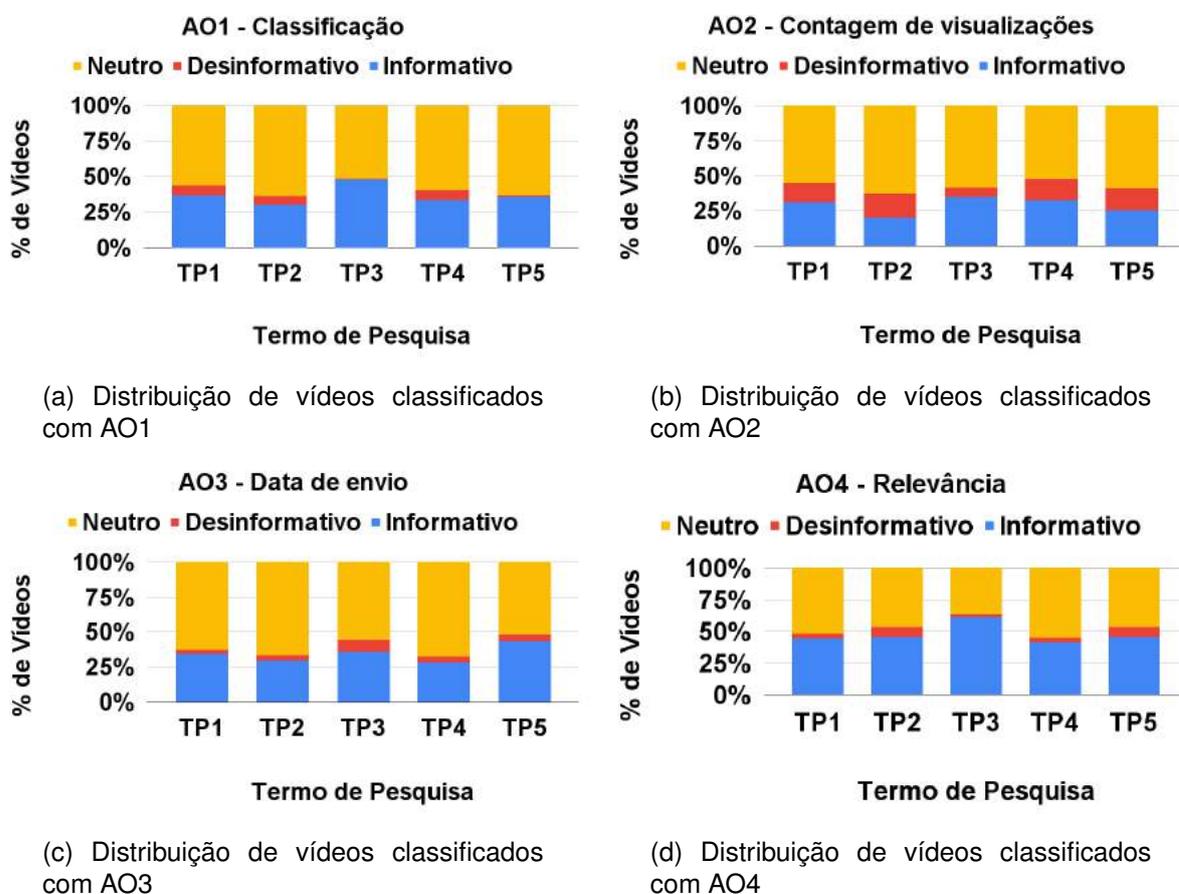


Figura 45 – Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 4 Algoritmos de Ordenação – Etapa 3 – Vídeos Recomendados – Perfil Informativo

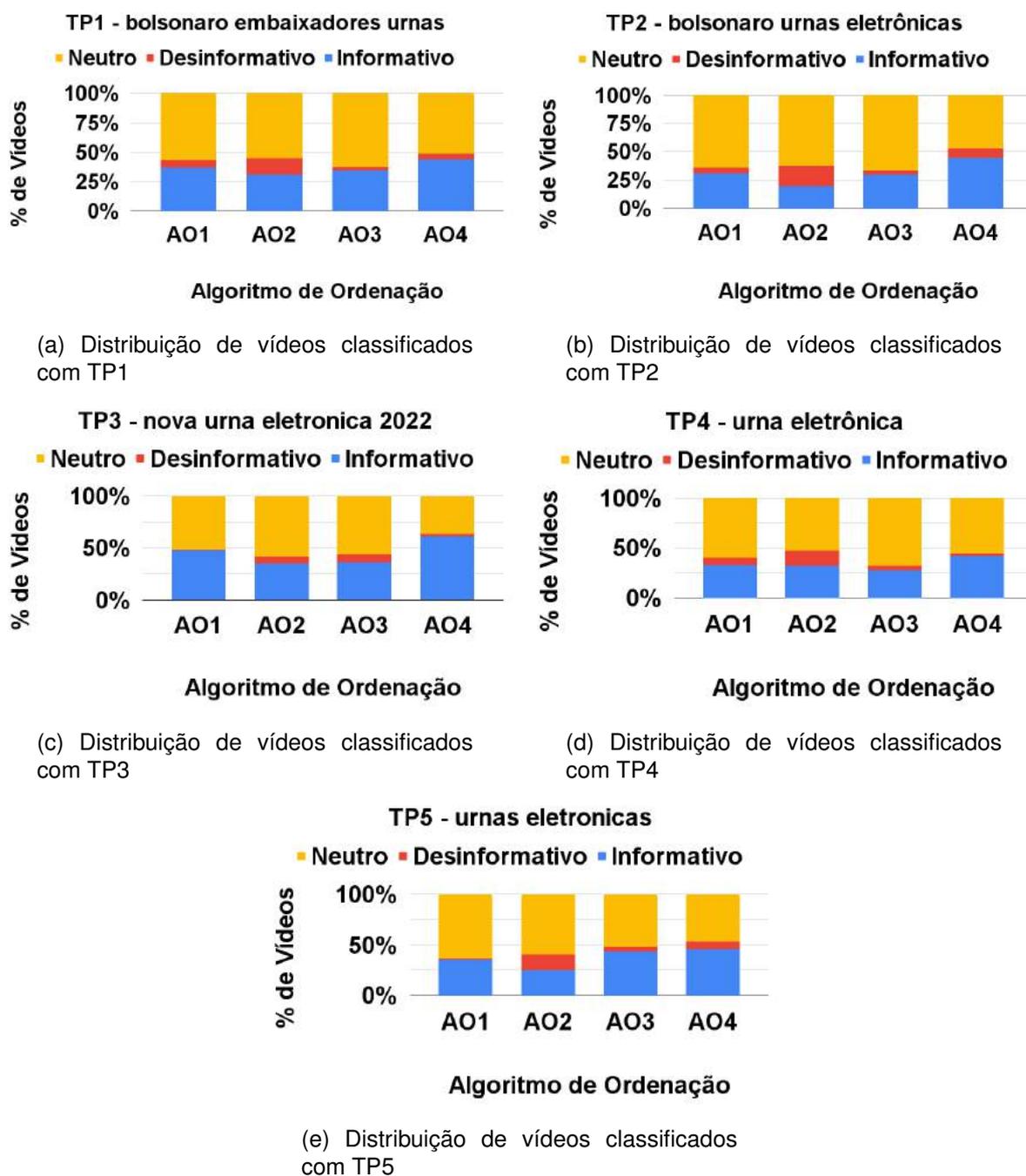


Figura 46 – Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 5 Termos de Pesquisa – Etapa 3 – Vídeos Recomendados – Perfil Informativo

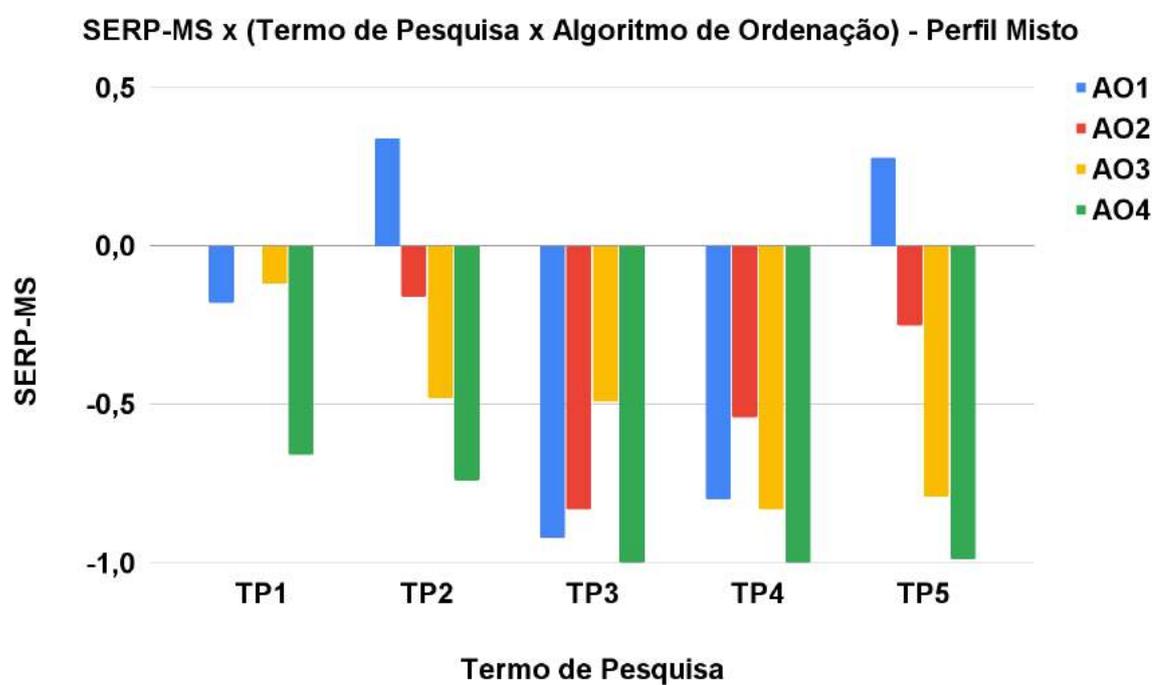


Figura 47 – SERP-MS x Pares (Termo de Pesquisa x Algoritmo de Ordenação) dos resultados da Etapa 3 – Perfil Misto

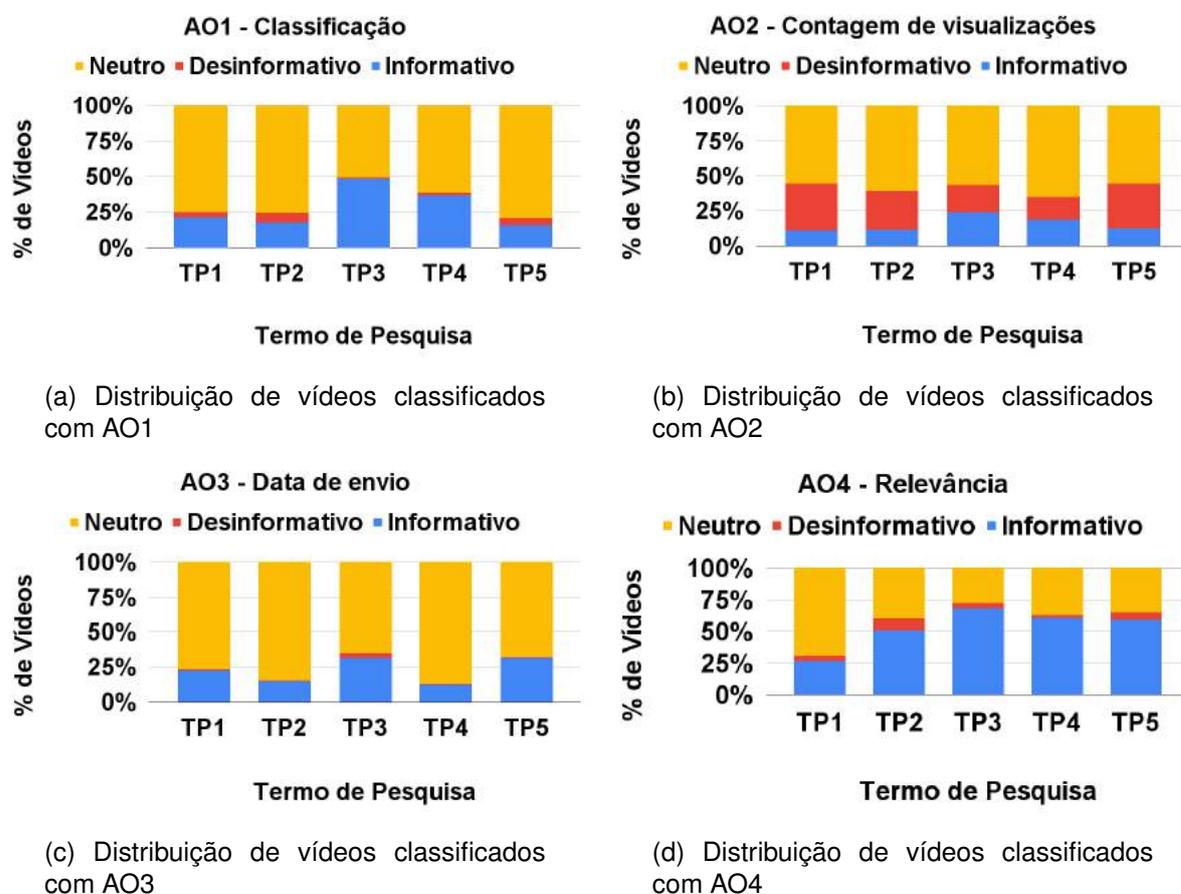


Figura 48 – Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 4 Algoritmos de Ordenação – Etapa 3 – Vídeos Recomendados – Perfil Misto

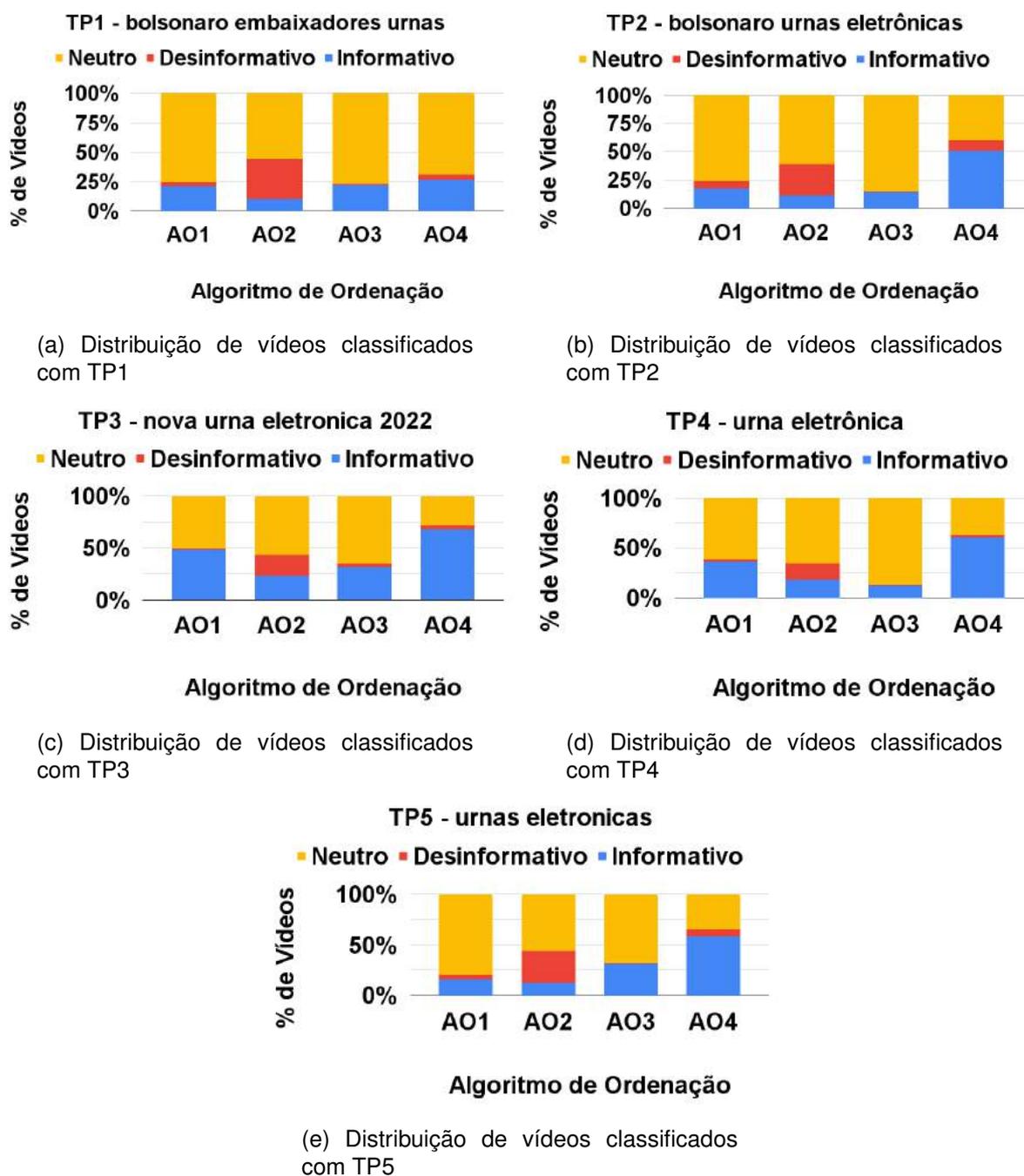


Figura 49 – Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 5 Termos de Pesquisa – Etapa 3 – Vídeos Recomendados – Perfil Misto

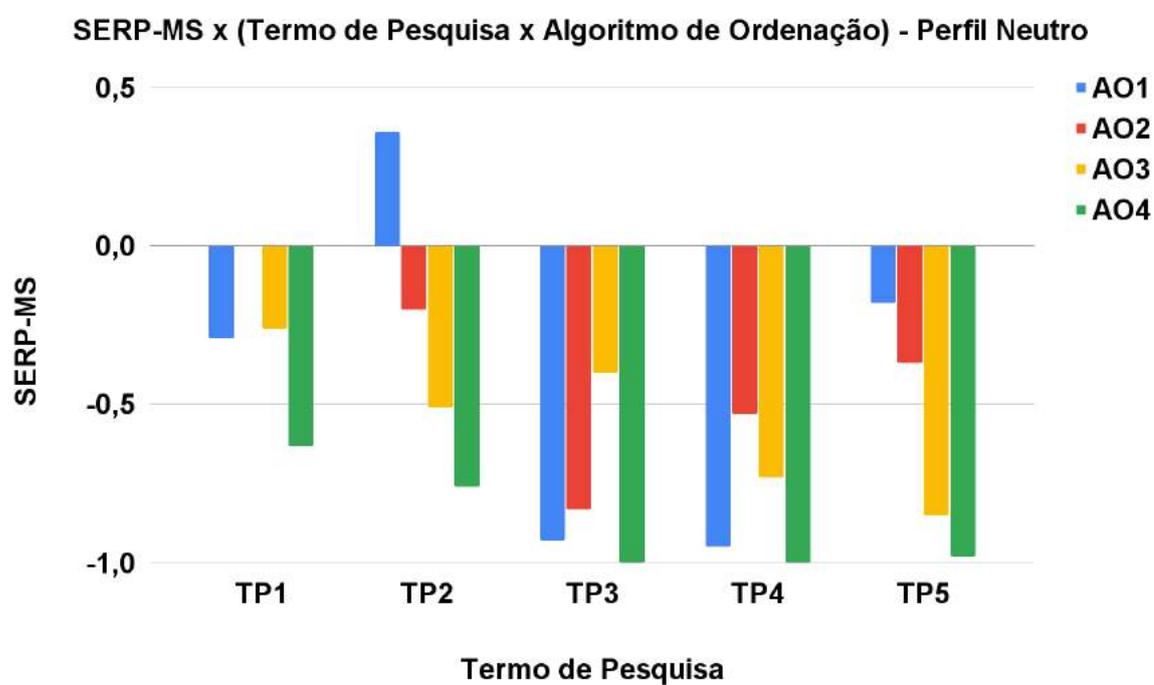


Figura 50 – SERP-MS x Pares (Termo de Pesquisa x Algoritmo de Ordenação) dos resultados da Etapa 3 – Perfil Neutro

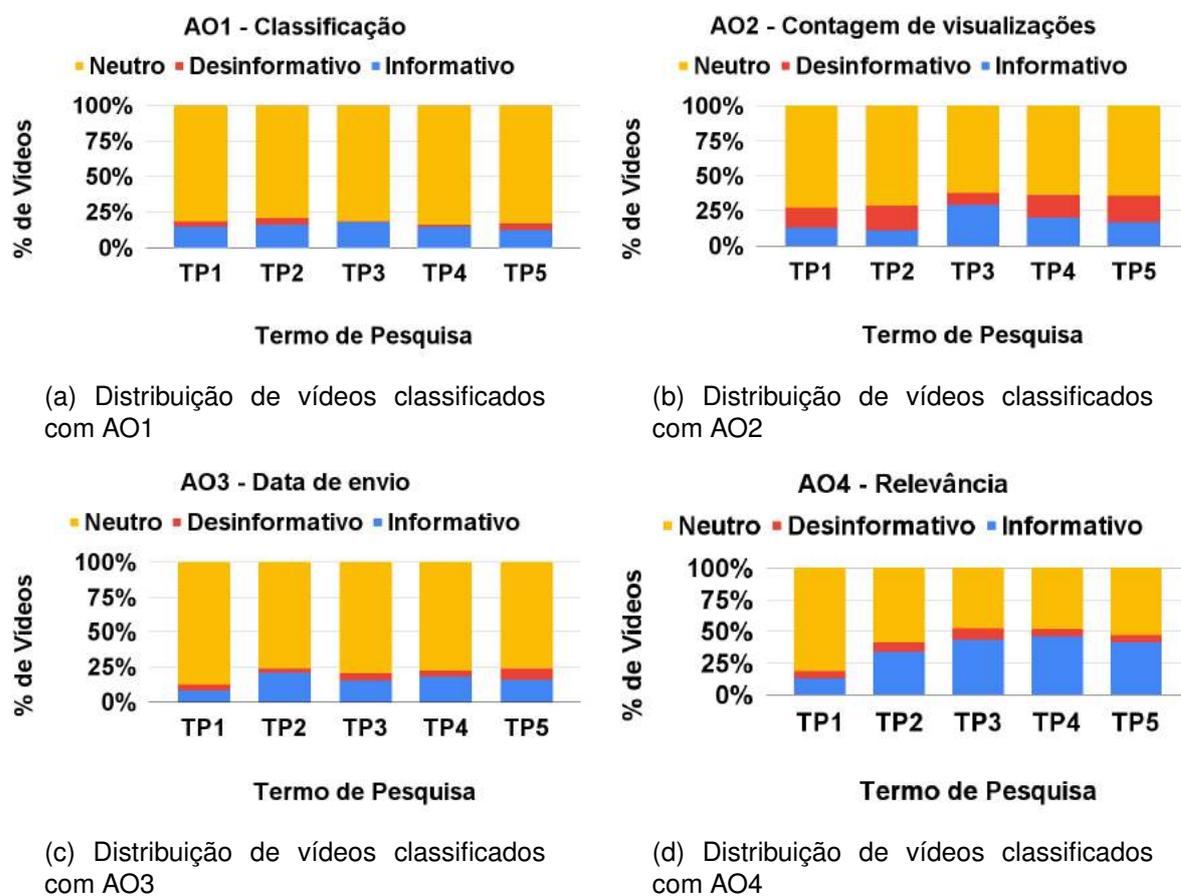


Figura 51 – Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 4 Algoritmos de Ordenação – Etapa 3 – Vídeos Recomendados – Perfil Neutro

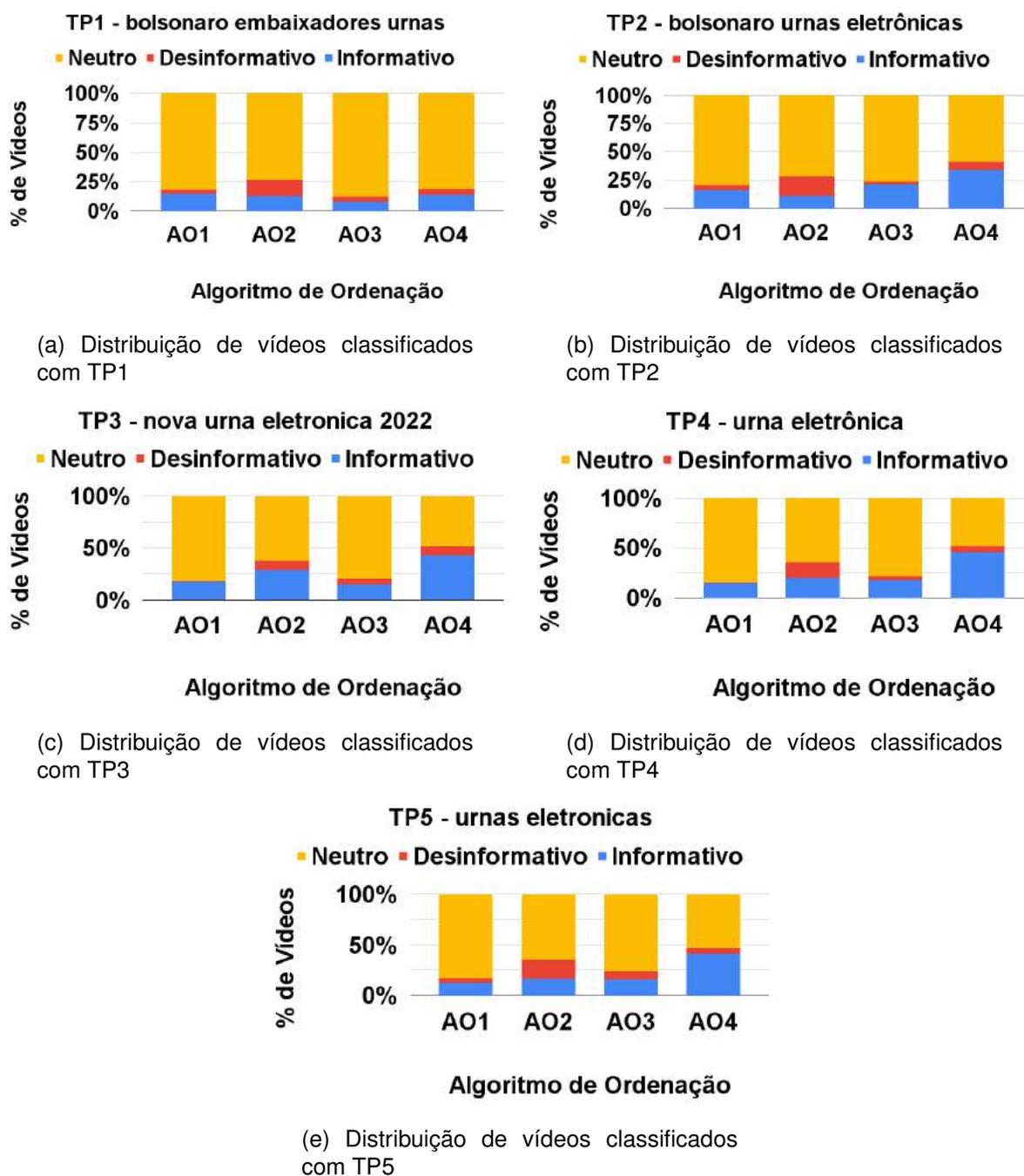


Figura 52 – Distribuições percentuais de cada tipo de vídeo classificado para cada um dos 5 Termos de Pesquisa – Etapa 3 – Vídeos Recomendados – Perfil Neutro