

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



Dissertação

**Método para diagnóstico e identificação de pacientes em risco de desenvolver
Alzheimer utilizando técnicas de Inteligência Artificial**

Guilherme Povala

Pelotas, 2018

Guilherme Povala

**Método para diagnóstico e identificação de pacientes em risco de desenvolver
Alzheimer utilizando técnicas de Inteligência Artificial**

Dissertação apresentada ao Programa de
Pós-Graduação em Computação da Universi-
dade Federal de Pelotas, como requisito par-
cial à obtenção do título de Mestre em Com-
putação

Orientador: Prof. Dr. Bruno Zatt
Coorientador: Prof. Dr. Eduardo Rigon Zimmer

Pelotas, 2018

Universidade Federal de Pelotas / Sistema de Bibliotecas
Catalogação na Publicação

P739m Povala, Guilherme

Método para diagnóstico e identificação de pacientes em risco de desenvolver Alzheimer utilizando técnicas de inteligência artificial / Guilherme Povala ; Bruno Zatt, orientador ; Eduardo Rigon Zimmer, coorientador. — Pelotas, 2018.

97 f. : il.

Dissertação (Mestrado) — Programa de Pós-Graduação em Computação, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, 2018.

1. Doença de Alzheimer. 2. Diagnóstico. 3. Inteligência artificial. 4. Alzheimer pré-clínico. I. Zatt, Bruno, orient. II. Zimmer, Eduardo Rigon, coorient. III. Título.

CDD : 005

Guilherme Povala

Método para diagnóstico e identificação de pacientes em risco de desenvolver Alzheimer utilizando técnicas de Inteligência Artificial

Dissertação aprovada, como requisito parcial, para obtenção do grau de Mestre em Ciência da Computação, Programa de Pós-Graduação em Computação, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas.

Data da Defesa: 07 de Março de 2018

Banca examinadora:


.....
Prof. Dr. Bruno Zatt (Orientador)
Doutor em Microeletrônica pela Universidade Federal do Rio Grande do Sul


.....
Prof. Dr. Evandro Manica
Doutor em Matemática Aplicada pela University of Pittsburg


.....
Prof. Dr. Ricardo Matsumura Araujo
Doutor em Ciência da Computação pela Universidade Federal do Rio Grande do Sul


.....
Prof. Dr. Marlilton Sanchotene de Aguiar
Doutor em Ciência da Computação pela Universidade Federal do Rio Grande do Sul

RESUMO

POVALA, Guilherme. **Método para diagnóstico e identificação de pacientes em risco de desenvolver Alzheimer utilizando técnicas de Inteligência Artificial.** 2018. 97 f. Dissertação (Mestrado em Computação) – Programa de Pós-Graduação em Computação, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, Pelotas, 2018.

Ao longo dos anos, os avanços em pesquisas relacionadas a saúde permitiram descobrir curas e tratamentos para inúmeras doenças. No entanto, ainda existem várias doenças para as quais a cura é desconhecida e os tratamentos existentes são ineficazes. Dentre elas, está a Doença de Alzheimer (DA), que é uma doença neurodegenerativa crônica e atinge principalmente pessoas com mais de 65 anos de idade. A DA não tem cura e o diagnóstico precoce é essencial para novas estratégias terapêuticas. Com isso, a busca por tratamentos que possam retardar a progressão da patologia se tornam essenciais, principalmente se a doença for descoberta precocemente. Desta forma, faz-se necessário identificar pacientes que estão em risco de desenvolver a doença, de preferência de uma forma não invasiva, com um baixo custo para implementação e aplicação e acessível para grande parte da população. Portanto, este trabalho se propõe a estudar a DA, suas características e a base de dados da *Alzheimer's Disease Neuroimaging Initiative* (ADNI), que contém biomarcadores clínicos, de imagem, genéticos e de bioespécime, obtendo assim o conhecimento necessário para desenvolver um método para classificação de pacientes que tendem a progredir de um estágio pré-clínico ou de pré-demência para DA, a fim de identificar indivíduos em risco de desenvolver a doença. Desta forma, é proposto o *Two-Step Classifier* (TSC), um método de classificação de pacientes em risco. O TSC foi modelado na ferramenta RapidMiner Studio e é arquitetado em duas etapas. O *first-step* é responsável pela classificação de pacientes em risco, enquanto o *second-step* realiza a predição do estado cognitivo dos pacientes. O TSC atinge uma precisão de 89,3% na identificação de pacientes em risco de desenvolver Alzheimer.

Palavras-Chave: doença de alzheimer; diagnóstico; inteligência artificial; alzheimer pré-clínico

ABSTRACT

POVALA, Guilherme. **Method for diagnosis and identification of patients at risk of developing Alzheimer's disease using Artificial Intelligence techniques.** 2018. 97 f. Dissertação (Mestrado em Computação) – Programa de Pós-Graduação em Computação, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, Pelotas, 2018.

Over the years, advances in health research have led to discovering cures and treatments for numerous diseases. However, there are still several diseases in which the cure is unknown and the existing treatments are ineffective. Among them is Alzheimer's Disease (AD), which is a chronic neurodegenerative disease and affects mainly people over 65 years old. The AD has no cure and early diagnosis is essential for new therapeutic strategies. Thus, the search for treatments that may delay the progression of the disease becomes essential, especially if the disease is discovered early. This way, it is necessary to identify patients who are at risk of developing the disease, preferably in a non-invasive manner, with a low cost for implementation and application and accessible to a large part of the population. Therefore, this work aims to study Alzheimer's disease, its characteristics and the database of the Alzheimer's Disease Neuroimaging Initiative (ADNI), which contains clinical, imaging, genetic and biospecimen biomarkers, thus obtaining the necessary knowledge to develop a method for classifying patients who tend to progress from a preclinical stage or from pre-dementia to AD, in order to identify individuals at risk of developing the disease. In this way, the Two-Step Classifier (TSC), a method of classifying patients at risk, is proposed. The TSC was modeled in the RapidMiner Studio tool and is architected in two steps. The *first-step* is responsible for the classification of patients at risk, while the *second-step* performs the prediction of patients' cognitive status. TSC reaches a precision of 89.3% in identifying patients at risk of developing Alzheimer.

Key-words: alzheimer disease; diagnosis; artificial intelligence; preclinical alzheimer

LISTA DE FIGURAS

Figura 1	Número projetado de pessoas com 65 anos ou mais (total e por faixa etária) na população dos EUA com demência de Alzheimer (ASSOCIATION, 2017)	16
Figura 2	Causas de morte por doença entre os anos de 2000 e 2014 (ASSOCIATION, 2017)	16
Figura 3	Redução de placas de A β com a aplicação do anticorpo: exemplos de imagens no início e após um ano (SEVIGNY et al., 2016)	17
Figura 4	Proteína amilóide precursora (OSMOSIS, 2016)	22
Figura 5	Processo de proteólise através das enzimas α -secretase e γ -secretase (OSMOSIS, 2016)	22
Figura 6	Processo de proteólise através das enzimas β -secretase e γ -secretase (OSMOSIS, 2016)	23
Figura 7	Placas de A β no cérebro (OSMOSIS, 2016)	23
Figura 8	Microtúbulos estabilizados pela proteína tau (ADEAR, 2008)	24
Figura 9	Ativação de quinases e a alteração da proteína tau (OSMOSIS, 2016) .	25
Figura 10	Diagrama de como os microtúbulos se desintegram com a doença de Alzheimer (ADEAR, 2008)	25
Figura 11	Mudanças que ocorrem no cérebro com o avanço da doença de Alzheimer (OSMOSIS, 2016)	26
Figura 12	Biomarcadores de Alzheimer ao longo da doença (ADNI, 2016a) . .	27
Figura 13	Evolução dos sintomas conforme a DA avança (OSMOSIS, 2016) .	29
Figura 14	Exemplo de uma árvore de decisão	42
Figura 15	Comparação entre um classificador simples, e classificadores gerados pela técnica de <i>Bagging</i> e de <i>Boosting</i> (XRISTICA, 2016) . .	43
Figura 16	Exemplo de uma rede neural <i>Multilayer Perceptron</i>	45
Figura 17	Exemplo de regressão linear	46
Figura 18	Exemplo de regressão logística	47
Figura 19	Exemplo de uma SVM	48
Figura 20	Exemplo do funcionamento do algoritmo kNN	49
Figura 21	Modelo genérico do <i>Stacked Generalization</i>	49
Figura 22	Exemplo de curva ROC	52
Figura 23	Arquitetura genérica do <i>two-step classifier</i>	58
Figura 24	Diagrama de Venn das instâncias classificadas incorretamente para o experimento D-PET-CSF-MRI-G	66

Figura 25	Acurácia obtida para os modelos avaliados no <i>second-step</i>	74
Figura 26	Arquitetura do <i>two-step classifier</i>	76

LISTA DE TABELAS

Tabela 1	Número de participantes nas diferentes fases do ADNI	33
Tabela 2	Estágios acompanhadas em cada fase do ADNI (ADNI, 2016a)	33
Tabela 3	Tipos de dados coletados nas diferentes fases do ADNI	34
Tabela 4	Dados de estudo disponibilizados pelo ADNI	36
Tabela 5	Resumo dos trabalhos relacionados	40
Tabela 6	Exemplo de matriz de confusão	50
Tabela 7	Métricas derivadas da matriz de confusão	50
Tabela 8	Conjuntos de biomarcadores avaliados para a predição do diagnóstico da DA	56
Tabela 9	Conjuntos de biomarcadores explorados para o <i>first-step</i> do <i>two-step classifier</i>	59
Tabela 10	Resultados para o experimento D-PET-CSF-N	62
Tabela 11	Ganho de Informação para os biomarcadores utilizados neste trabalho	63
Tabela 12	Resultados para o experimento D-PET-CSF-N*	63
Tabela 13	Resultados para o experimento D-PET-CSF	64
Tabela 14	Resultados para o experimento D-PET-CSF-MRI-G	65
Tabela 15	Matriz de confusão da Regressão Logística para o experimento D-PET-CSF-MRI-G	65
Tabela 16	Resultados do para o experimento D-PET-CSF-MRI-G-SG	67
Tabela 17	Matriz de confusão da MLP para o experimento D-PET-CSF-MRI-G com <i>Stacked Generalization</i>	67
Tabela 18	Resumo dos resultados obtidos para a classificação da DA	68
Tabela 19	Resultados do para o experimento TSC-D-PET-CSF-MRI	69
Tabela 20	Resultados do para o <i>first-step</i> do experimento TSC-D-PET-CSF	70
Tabela 21	Resultados do para <i>first-step</i> do experimento TSC-D-CSF-MRI	70
Tabela 22	Resultados do para <i>first-step</i> do experimento TSC-D-MRI	71
Tabela 23	Matriz de confusão do GBT para o experimento TSC-D-MRI	71
Tabela 24	Matriz de confusão do conjunto de validação para o experimento TSC-D-MRI utilizando GBT	72
Tabela 25	Resumo dos resultados obtidos para o <i>first-step</i> do TSC	72
Tabela 26	Matriz de confusão para o <i>second-step</i> utilizando GBT	74
Tabela 27	Matriz de confusão para o <i>second-step</i> utilizando MLG	75
Tabela 28	Matriz de confusão do conjunto de validação aplicado no <i>second-step</i> utilizando GBT	75

Tabela 29	Matriz de confusão do <i>second-step</i> utilizando GBT sobre para as instâncias classificadas como NOT_NL no <i>first-step</i> do experimento TSC-D-MRI	76
Tabela 30	Matriz de confusão da classificação de pacientes em risco após o processamento do TSC	77
Tabela 31	Comparação com os trabalhos relacionados	79
Tabela 32	Taxa de identificação dos <i>converters</i> para os experimentos realizados	80

LISTA DE ABREVIATURAS E SIGLAS

A β	Beta-amiloide
ADNI	Iniciativa de neuroimagem da doença de Alzheimer
ML	Aprendizagem de Máquina
APP	Proteína precursora de amiloide
AUC	Área sob a curva ROC
AV45	<i>Florbetapir</i>
CSF	Fluido cefalorraquidiano
NL	Cognitivamente normal
CT	Tomografia computadorizada
DA	Doença de Alzheimer
DAAD	Doença de Alzheimer Autossômica Dominante
DT	Árvore de decisão
DTI	Imagem de tensor de difusão
EMCI	Comprometimento cognitivo leve precoce
FDG	<i>Flodesoxiglicose</i>
GBT	<i>Gradient Boosted Trees</i>
fMRI	Imagem de ressonância magnética funcional
FN	Falso Negativo
FP	Falso Positivo
IA	Inteligência Artificial
KNN	k-vizinhos mais próximos
LMCI	Comprometimento cognitivo leve tardio
MCI	Comprometimento cognitivo leve
MLG	Modelo Linear Generalizado
MLP	<i>Multilayer Perceptron</i>
MRI	Imagem de ressonância magnética

NFT	Emaranhados neurofibrilares
PET	Tomografia por emissão de positrões
PiB	<i>Pittsburgh compound B</i>
pMCI	MCI que progrediu para DA
PS-1	<i>Presenilina-1</i>
PS-2	<i>Presenilina-2</i>
RF	<i>Random Forest</i>
RLog	Rregressão Logística
ROI	Regiões de interesse
SMC	Preocupação significativa de memória
sMCI	MCI estável
SPECT	Tomografia computadorizada por emissão de fóton único
SVM	Máquina de vetor de suporte
TN	Verdadeiro Negativo
TP	Verdadeiro Positivo
TSC	<i>Two-Step Classifier</i>

SUMÁRIO

1 INTRODUÇÃO	14
1.1 Objetivos	18
1.2 Organização do Trabalho	19
2 A DOENÇA DE ALZHEIMER	20
2.1 Patofisiologia	21
2.1.1 A deposição de placas de A β	21
2.1.2 O acúmulo de emaranhados neurofibrilares e a proteína tau	23
2.2 Neuropatologia	24
2.3 Biomarcadores	27
2.4 Sintomas e estágios da doença de Alzheimer	28
2.4.1 Comprometimento cognitivo leve	29
2.4.2 Estágio inicial	29
2.4.3 Estágio intermediário	30
2.4.4 Estágio avançado	31
2.5 Fatores de risco	31
2.5.1 Idade	31
2.5.2 Genética	31
2.6 Alzheimer's Disease Neuroimaging Initiative	32
2.7 Trabalhos Correlatos	35
2.7.1 Estudos <i>cross-sectional</i>	36
2.7.2 Estudos <i>follow-up</i>	38
2.8 Considerações finais do capítulo	38
3 INTELIGÊNCIA ARTIFICIAL E A ÁREA MÉDICA	41
3.1 Árvores de decisão	42
3.1.1 <i>Random Forest</i>	44
3.1.2 <i>Gradient Boosted Trees</i>	44
3.2 Rede neural artificial	44
3.2.1 <i>Multilayer perceptron</i>	44
3.3 Regressão linear	45
3.4 Regressão logística	46
3.5 Modelo Linear Generalizado	46
3.6 Máquina de vetor de suporte	47
3.7 <i>k</i>-vizinhos mais próximos	47
3.8 <i>Stacked Generalization</i>	48
3.9 Métricas de avaliação	50

3.9.1	Matriz de confusão	50
3.9.2	Métricas derivadas da matriz de confusão	50
3.9.3	Acurácia	51
3.9.4	Precisão	51
3.9.5	Sensibilidade	51
3.9.6	<i>F1 Score</i>	51
3.9.7	Área sob a curva ROC	51
4	METODOLOGIA E TRABALHO PROPOSTO	53
4.1	Biomarcadores utilizados	55
4.2	Classificação da Doença de Alzheimer	56
4.2.1	Conjunto de Dados	56
4.2.2	Classificadores	57
4.2.3	Avaliação dos Resultados	57
4.3	<i>Two-step classifier</i>	58
4.3.1	Conjunto de Dados	58
4.3.2	Classificadores	59
4.3.3	Avaliação dos Resultados	60
5	RESULTADOS	61
5.1	Classificação da Doença de Alzheimer	61
5.2	<i>Two-step classifier</i>	67
5.2.1	<i>First-step</i>	69
5.2.2	<i>Second-step</i>	72
5.2.3	Identificação de Pacientes em Fase Pré-Clínica	77
6	CONCLUSÃO	81
REFERÊNCIAS		83
ANEXO A ITENS NO ARQUIVO DE DADOS ADNI		93

1 INTRODUÇÃO

Nas últimas décadas, as pesquisas médicas, as descobertas científicas e as melhorias na educação e nas condições de vida no cenário mundial causaram uma redução significativa nas chances de desenvolvimento de pandemias causadas por agentes patogênicos infecciosos (G7, 2017). Desta forma, o crescimento populacional nas últimas décadas veio também acompanhado de um aumento progressivo do número de idosos. Em países desenvolvidos, a expectativa de vida já atinge valores acima dos 80 anos. No Japão, por exemplo, a expectativa de vida era de 83,7 anos em 2015 (WHO, 2015).

Hoje, as principais causas de morte de pessoas idosas nos Estados Unidos e na Europa ainda são as doenças cardiovasculares e o câncer (CDC, 2017; EUROSTAT, 2017). No entanto, as causas de morte por essas doenças tem decrescido anualmente (ASSOCIATION, 2017). O avanço na diminuição da mortalidade relacionada com estas duas doenças contribuiu ainda mais para o crescimento no número de cidadãos idosos, o que provocou um aumento sem precedentes no número de pessoas afetadas por doenças neurodegenerativas (GOOCH; PRACHT; BORENSTEIN, 2017).

Doenças neurodegenerativas são classificadas como uma série de condições que afetam principalmente os neurônios, encontrados no sistema nervoso central e periférico. Além disso, essas doenças são condições incuráveis e debilitantes que resultam em degeneração progressiva, o que reflete em disfunção motora ou cognitiva progressiva (JPND, 2017).

Dentre elas, temos a Doença de Alzheimer (DA), que se apresenta fortemente relacionada com a idade avançada de seus portadores e está entre as dez principais doenças sem cura conhecida que resultam em morte (WHO, 2017). Além disso, a DA é uma doença neurodegenerativa crônica que geralmente tem um início lento e progride com o passar do tempo (BURNS; ILIFFE, 2009; WHO, 2015). Ela se caracteriza pela deposição de placas de beta-amiloide ($A\beta$, do inglês *Amyloid-beta*) e emaranhados neurofibrilares (NFT, do inglês *Neurofibrillary Tangles*) compostos de proteína tau hiperfosforilada no cérebro, acompanhados por disfunção sináptica e neurodegeneração – morte neural progressiva no cérebro – em regiões cerebrais relacionadas com

aprendizado e memória, como o córtex e o hipocampo (BURNS; ILIFFE, 2009; WHO, 2015; SEVIGNY et al., 2016). Com o avanço gradual da doença, aparecem sintomas como confusão mental, desorientação, irritabilidade, agressividade, alterações de humor, falhas na linguagem e perda de memória, até o aparecimento das manifestações de demência (BURNS; ILIFFE, 2009; WHO, 2015).

A demência é uma maneira de descrever um conjunto de sintomas como deterioração da memória e incapacidade de absorver novas informações e faz com que seja muito difícil viver de uma forma independente (WHO, 2015). Normalmente causada por algum tipo de disfunção cerebral, ela pode se originar de uma série de doenças. A DA é a forma mais comum de demência, correspondendo a cerca de 60% a 70% dos casos (BURNS; ILIFFE, 2009; WHO, 2015). Só no ano de 2010, a demência resultou em cerca de 486 mil mortes (LOZANO et al., 2012).

Apesar de ser difícil identificar o início da DA, existem sintomas comuns que podem ajudar no diagnóstico. O sintoma inicial mais frequente é a dificuldade em recordar acontecimentos recentes (perda de memória de curto prazo) (BURNS; ILIFFE, 2009). Não raro, os primeiros sintomas da doença são confundidos com os processos normais de envelhecimento ou estresse (WALDEMAR et al., 2007), como (REPORTS, 2012):

- esquecer-se ocasionalmente das coisas;
- perder ocasionalmente objetos;
- pequenas perdas de memória a curto prazo;
- esquecer-se que ocorreram lapsos de memória.

Em 2015, havia cerca de 46 milhões de pessoas em todo o mundo com Alzheimer (WHO, 2015), com previsão de que 1 em cada 85 habitantes seja afetado com a doença no ano de 2050 (BROOKMEYER et al., 2007). Geralmente, a DA começa a manifestar em pessoas com mais de 65 anos de idade, apesar de 4% a 5% dos casos serem de início precoce (MENDEZ, 2012). A doença afeta 1% dos idosos entre os 65 e 70 anos, mas a prevalência aumenta exponencialmente com a idade, sendo de 6% aos 70, 30% aos 80 anos e mais de 60% após os 90 anos (BERMEJO-PAREJA et al., 2008). Como podemos visualizar no gráfico da Figura 1, a projeção do número de idosos com 65 anos ou mais, nos Estados Unidos, que desenvolverão Alzheimer e chegarão ao estágio de demência tende a crescer de forma acelerada nos próximos trinta anos.

Não obstante, como pode ser visto na Figura 2, a DA é a doença que teve o maior aumento em causas de morte entre os anos de 2000 e 2014 nos Estados Unidos. Inclusive, outras doenças, como câncer de mama e de próstata, doenças do coração, acidente vascular encefálico e HIV tiveram uma diminuição no número de mortes

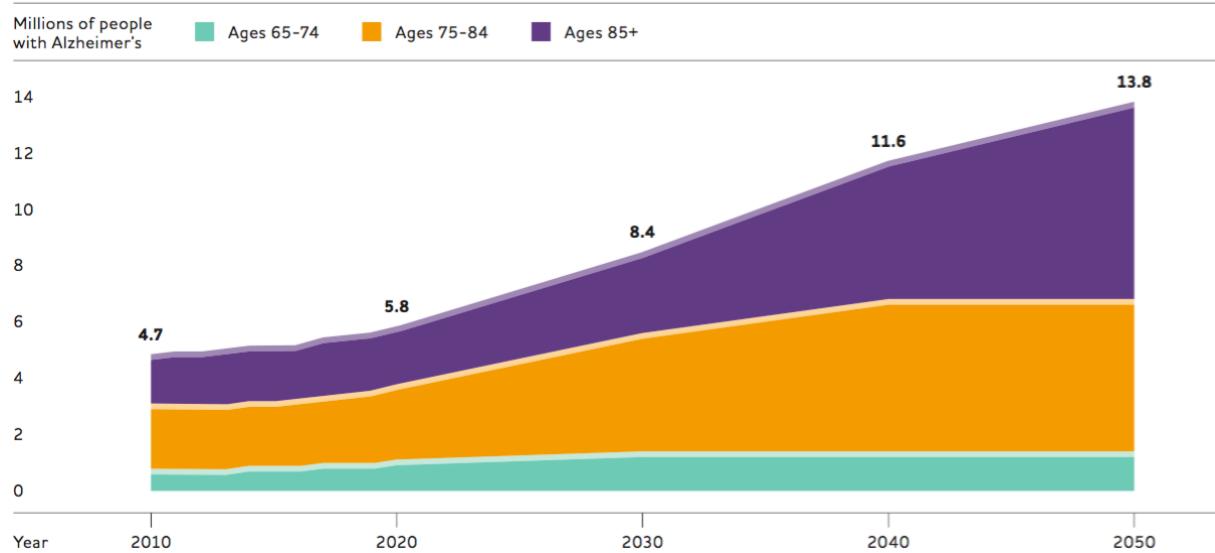


Figura 1: Número projetado de pessoas com 65 anos ou mais (total e por faixa etária) na população dos EUA com demência de Alzheimer (ASSOCIATION, 2017)

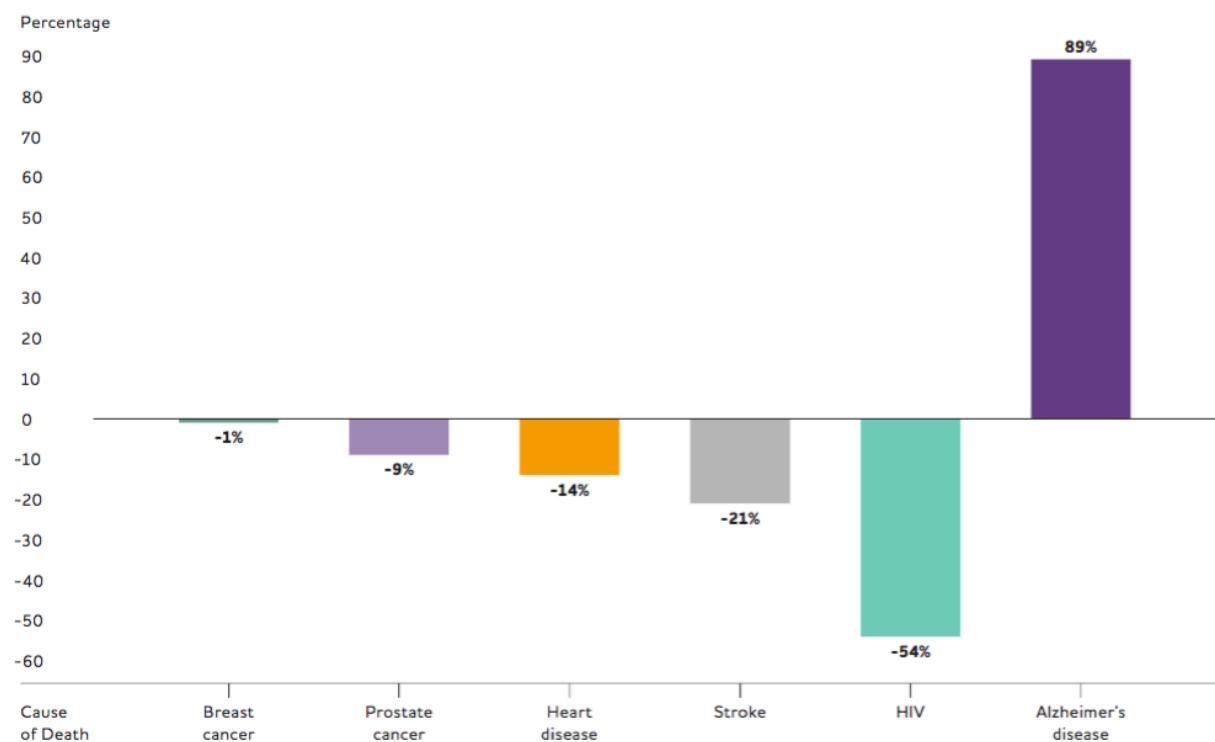


Figura 2: Causas de morte por doença entre os anos de 2000 e 2014 (ASSOCIATION, 2017)

nesse período, devido ao avanço e sucesso de estratégias de diagnóstico precoce e tratamento.

Embora a velocidade de progressão da doença possa variar, a esperança média de vida após o diagnóstico é de três a nove anos (QUERFURTH; LAFERLA, 2010; TODD et al., 2013). Mesmo que a DA reduza a expectativa de vida do paciente, é outra comorbidade ou doença que pode realmente levar o paciente a óbito. Esta outra

condição ou doença provavelmente listada como a causa no certificado de morte da pessoa é, comumente, a pneumonia, sendo ela a responsável por até dois terços das causas de morte das pessoas com demência (SOCIETY, 2012).

Apesar do grande esforço e pesquisas sobre a doença, não existe um tratamento efetivo para a DA. Todavia, acredita-se amplamente que a patogênese causadora da DA seja a disfunção do metabolismo da proteína precursora de amiloide (APP, do inglês *Amyloid Precursor Protein*), que leva a hiperprodução e deposição do peptídeo A β no cérebro em forma de placas de A β (MURPHY; LEVINE, 2010).

Em um artigo publicado recentemente por SEVIGNY et al. (2016), foi relatado que a imunoterapia com um anticorpo monoclonal, denominado *aducanumab*, remove agregados de A β seletivamente. Dependendo da dose utilizada, com um ano de tratamento as placas de A β foram drasticamente reduzidas. A Figura 3 mostra a redução de placas A β – medidas por exame de imagem por tomografia por emissão de pós-trons (PET, do inglês *Positron Emission Tomography*) – após um ano de tratamento com *aducanumab* em diferentes doses.

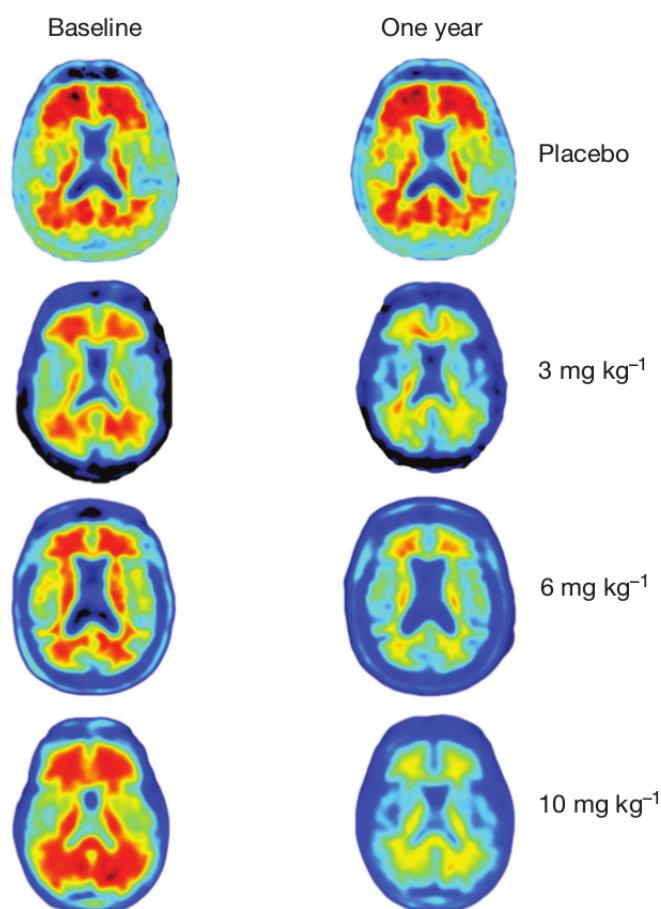


Figura 3: Redução de placas de A β com a aplicação do anticorpo: exemplos de imagens no início e após um ano (SEVIGNY et al., 2016)

Como pode ser visto na Figura 3, após um ano de tratamento, houve uma dimi-

nuição significativa nos grupos em que foram aplicadas doses de 3, 6 e 10 mg kg⁻¹, enquanto que a alteração para o grupo placebo foi mínima (SEVIGNY et al., 2016).

Além do mais, essa redução de placas de A β foi acompanhada pela diminuição do declínio clínico do paciente. Com estes resultados, o desenvolvimento de imunoterapias para o tratamento da DA fica cada vez mais justificado (SEVIGNY et al., 2016).

Dessa maneira, mesmo sabendo que ainda não há cura para a doença, estamos cada vez mais perto de um tratamento que diminua ou até possa impedir o declínio cognitivo do paciente. Portanto, busca-se, principalmente, retardar a progressão da doença, melhorar os sintomas, resolver os problemas comportamentais e melhorar a qualidade de vida. Assim, se a doença for diagnosticada em uma fase precoce, existem tratamentos que podem desacelerar temporariamente o desenvolvimento de sintomas de demência (WHO, 2015) ou até mesmo levar a cura, caso o trabalho de SEVIGNY et al. (2016) se confirme em pacientes portadores de Alzheimer.

Porém, para diagnosticar e definir em que estágio da DA se encontra determinado indivíduo, geralmente é necessária uma avaliação clínica composta por vários tipos de testes, como exames de imagem e exames clínicos, uma vez que ainda não existe um exame único que permita identificar a doença em qualquer paciente. Dentre os exames realizados, os principais são os testes neuropsicológicos, testes de imagem cerebral, como ressonância magnética (MRI, do inglês *Magnetic Resonance Imaging*) e PET, testes genéticos, testes de fluido espinal e outros.

Neste cenário, o custo econômico para o tratamento dos pacientes portadores da DA é altíssimo (BONIN-GUILLAUME et al., 2005; MEEK; MCKEITHAN; SCHUMOCK, 1998). Considerando apenas os testes de imagem nos EUA, o custo de um exame de MRI fica entre US\$1500 e US\$3000. Já o exame de PET custa entre US\$3000 e US\$6000 (SCOTT, 2016). Portanto, realizar todos os exames que normalmente são requisitados pelos médicos para o diagnóstico da DA torna-se caro – não escalando para número elevado de pessoas – e, ainda assim, não detecta a doença antes dela começar a causar danos cognitivos.

1.1 Objetivos

Levando em consideração a importância da saúde mental das pessoas e que o diagnóstico precoce da DA é de extrema importância para que haja tempo suficiente para que ações sejam tomadas para retardar, atenuar, ou até mesmo evitar o avanço da doença, este trabalho tem como objetivo principal propor um método de diagnóstico da DA que permita classificar pacientes em estado de risco, ou seja, que já possuem a DA ou estão em fase inicial de perda cognitiva.

Além disso, considerando os altos custos envolvidos nos exames médicos, foi definido um sub-conjunto de exames a serem utilizados que apresentam a melhor relação

entre precisão do diagnóstico da doença e acessibilidade para grande parte da população.

Desta forma, temos como objetivos específicos do trabalho:

- Estudar a doença de Alzheimer: patofisiologia, neuropatologia, biomarcadores, sintomas e estágios da doença e fatores de risco;
- Estudar a iniciativa ADNI e sua base de dados, possibilitando o entendimento dos biomarcadores ali presentes;
- Estudar técnicas de Inteligência Artificial que possam ser aplicadas para a solução do problema;
- Avaliar os exames que mais contribuem para uma melhor precisão de classificação e que sejam acessíveis a grande parte da população;
- Propor um método de diagnóstico da DA que permita classificar pacientes em estado de risco.

1.2 Organização do Trabalho

Este trabalho está dividido em seis capítulos. No capítulo 2 é apresentada a doença de alzheimer, suas principais características, a iniciativa ADNI e trabalhos correlatos. No capítulo 3 são discutidas as diferentes técnicas de Inteligência Artificial e métricas de avaliação utilizadas neste trabalho. No capítulo 4, a metodologia e o trabalho proposto são apresentados. No capítulo 5, são discutidos os resultados obtidos pelo trabalho. Por fim, o capítulo 6, conclui o trabalho e apresenta possibilidades de trabalhos futuros.

2 A DOENÇA DE ALZHEIMER

Os antigos filósofos e médicos gregos e romanos associavam a velhice com o aumento da demência (BERCHTOLD; COTMAN, 1998). Porém, em 1906/1907 o psiquiatra alemão Alois Alzheimer identificou o primeiro caso da patologia que ficou conhecida, posteriormente, como Doença de Alzheimer. Nesse caso, uma mulher de cinquenta anos de idade, chamada Auguste Deter, exibia declínio cognitivo progressivo, alucinações e incompetência psicossocial (STRASSNIG; GANGULI, 2005; ALZHEIMER, 1987; MAURER; MAURER, 2003). Ao realizar a autópsia do cérebro de Auguste Deter, foram encontradas deposições de algumas substâncias peculiares espalhadas pelo córtex cerebral (JUCKER, 2006).

Após a morte de Auguste, em 1906, onze casos semelhantes foram relatados na literatura médica nos cinco anos seguintes, sendo que alguns deles já usavam o termo DA (BERCHTOLD; COTMAN, 1998). Até o ano de 1977, a DA era caracterizada como uma doença pré-senil e seu diagnóstico era reservado para indivíduos com idades entre os 45 e os 65 anos que desenvolveram sintomas de demência. Nesse ano, em uma conferência sobre a DA, que concluiu-se que as manifestações clínicas e patológicas de demência pré-senil e senil eram quase idênticas, sem excluir, no entanto, a possibilidade de que elas tenham causas diferentes (KATZMAN; TERRY; BICK, 1978). Sendo assim, a idade deixou de ser um fator determinante para o diagnóstico da doença (BOLLER; FORBES, 1998). Desta forma, o termo Alzheimer começou a ser usado, eventualmente, para descrever indivíduos com qualquer idade que apresentem características padrões comuns da doença (AMADUCCI; ROCCA; SCHOENBERG, 1986).

Como mencionado anteriormente, ao ser feita a autópsia do cérebro de Auguste Deter, algumas substâncias foram encontradas no córtex cerebral. Hoje, sabe-se que essas substâncias são as principais características neuropatológicas que indicam a DA: a deposição de placas de A β e a formação de emaranhados neurofibrilares (NIS-BET et al., 2015; THAL et al., 2015). Essas características serão detalhadas na próxima seção.

2.1 Patofisiologia

A DA afeta os três processos que mantêm as células cerebrais saudáveis: a comunicação, o metabolismo e a reparação. Consequentemente, certas células neurais param de funcionar, perdem suas conexões com células adjacentes e finalmente morrem. A destruição e morte dessas células pode provocar a sintomatologia clássica da DA: falha de memória, mudanças de personalidade, problemas na realização de atividades diárias e outras características da doença (LAKHAN, 2016).

Embora os mecanismos que causam a DA não sejam completamente compreendidos, exceto para os casos em que diferenças genéticas são identificadas (ASSOCIATION, 2010a; REITZ; MAYEUX, 2014), há duas fortes hipóteses que são frequentemente citadas: (i) a deposição de placas de A β (HARDY; HIGGINS, 1992; HARDY; SELKOE, 2002; THAL et al., 2015) e o (ii) acúmulo de emaranhados neurofibrilares (NISBET et al., 2015).

2.1.1 A deposição de placas de A β

A DA é uma doença caracterizada por anomalias no enovelamento de proteínas¹, causada pelo acúmulo de proteínas A β e tau no cérebro (HASHIMOTO et al., 2003). A hipótese de amiloide postula que a deposição de placas de proteínas de A β é a principal causa da disfunção sináptica e subsequente neurodegeneração que está subjacente à progressão característica da DA (HARDY; HIGGINS, 1992; HARDY; SELKOE, 2002).

As placas são depósitos densos, principalmente insolúveis, de material proteico e celular que se encontram fora e ao redor dos neurônios. As placas são feitas de A β , um fragmento de proteína cortado a partir de uma proteína maior chamada APP. Estes fragmentos agrupam-se e são misturados com outras moléculas, neurônios e células não neuronais (LAKHAN, 2016). A Figura 4 mostra a localização da APP na membrana celular.

Como visto na Figura 4, temos a membrana celular de um neurônio no cérebro. Na membrana, está localizada a APP, uma proteína transmembrana encrustada na membrana celular. Uma das extremidades da APP está dentro da célula, e a outra extremidade está fora da célula. A APP auxilia o neurônio a crescer e a reparar-se após uma lesão. Uma vez que a APP é uma proteína, ela sofre proteólise² (OSMOSIS, 2016).

Normalmente, neste processo de clivagem proteolítica, duas enzimas - α -secretase e γ -secretase - que funcionam como uma tesoura molecular fazem com

¹DILL, K. A.; OZKAN, S. B.; SHELL, M. S.; WEIKL, T. R. The Protein Folding Problem. *Annual Review of Biophysics*, [S.I.], v.37, n.1, p.289–316, June 2008

²WIKIPEDIA. **Proteólise**, 2014. Page Version ID: 40134280. Disponível em: <<https://pt.wikipedia.org/w/index.php?oldid=40134280>>. Acesso em: 2016-11-21

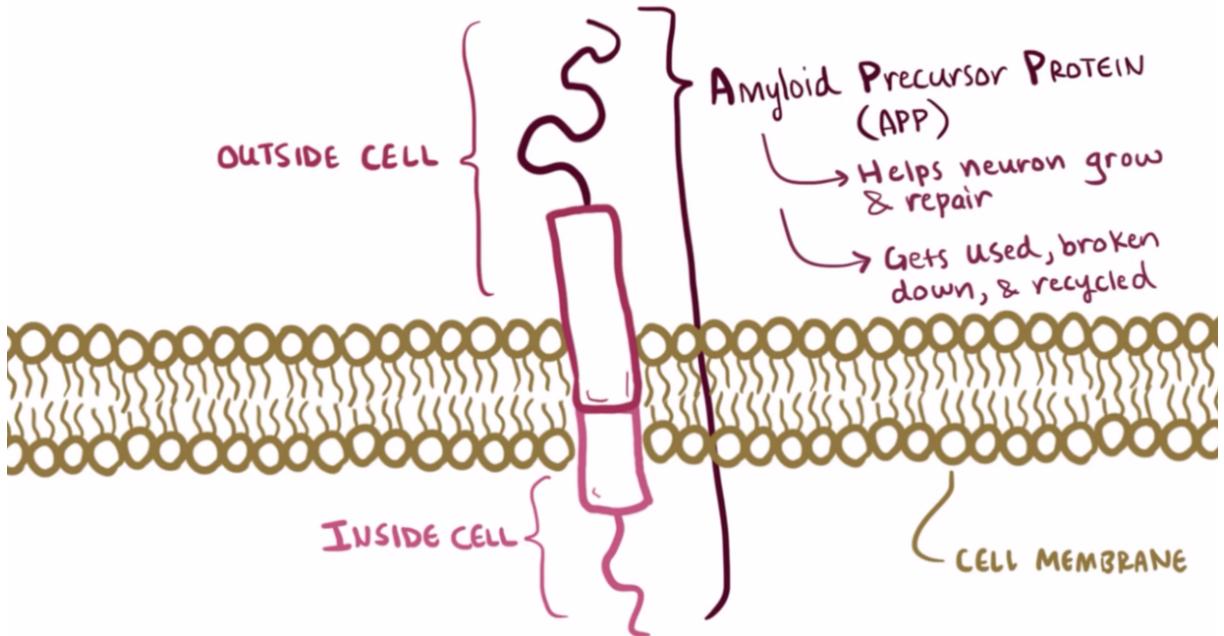


Figura 4: Proteína amilóide precursora (OSMOSIS, 2016)

que a APP seja dividida em fragmentos menores (HOOPER, 2005), que são solúveis e absorvidos. Esse processo pode ser visto na Figura 5.

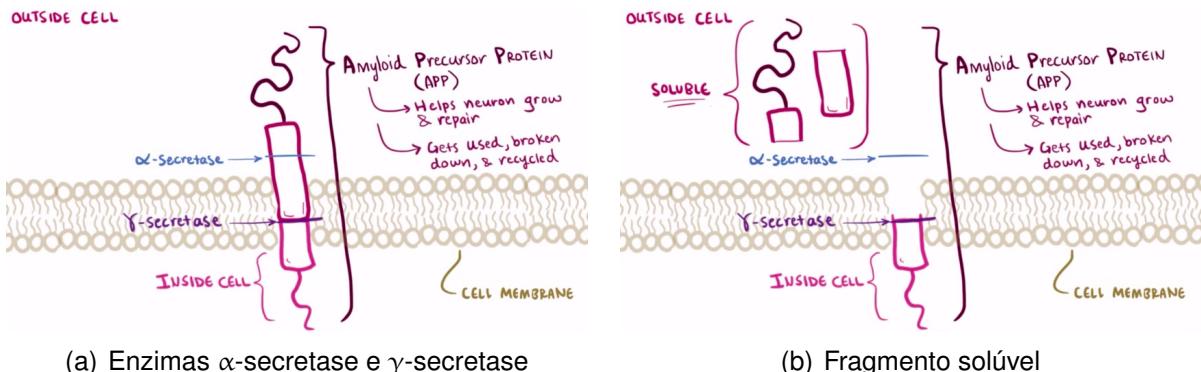


Figura 5: Processo de proteólise através das enzimas α -secretase e γ -secretase (OSMOSIS, 2016)

Porém, se outra enzima, a β -secretase, se aliar com a γ -secretase, como visto na Figura 6, é criado o monômero chamado A β . Esses monômeros tendem a ser mais quimicamente "pegajosos", e então, formam aglomerados insolúveis que se depositam no meio extracelular em formações densas conhecidas como placas de A β (OHNISHI; TAKANO, 2004; TIRABOSCHI et al., 2004; OSMOSIS, 2016).

Na DA, as placas se desenvolvem em áreas corticais e subcorticais do córtex cerebral que são usadas para codificar memórias, pensar e tomar decisões (LAKHAN, 2016). Estas placas, acabam interferindo nas sinapses que ocorrem entre os neurônios. Portanto, se as células cerebrais não puderem realizar sinapses e retransmitir

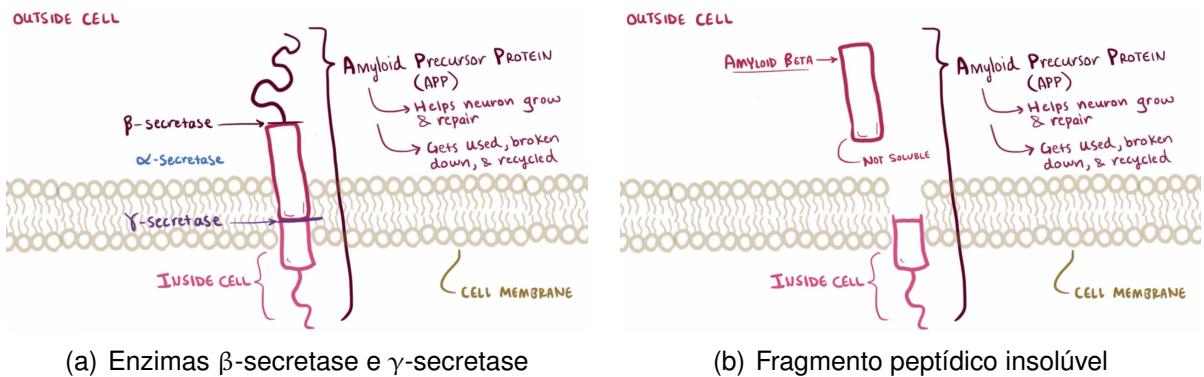


Figura 6: Processo de proteólise através das enzimas β -secretase e γ -secretase (OSMOSIS, 2016)

informações, as funções cerebrais como a memória podem ser gravemente prejudicadas (OSMOSIS, 2016). A Figura 7(a) mostra um exemplo das placas de A β interferindo nas sinapses.

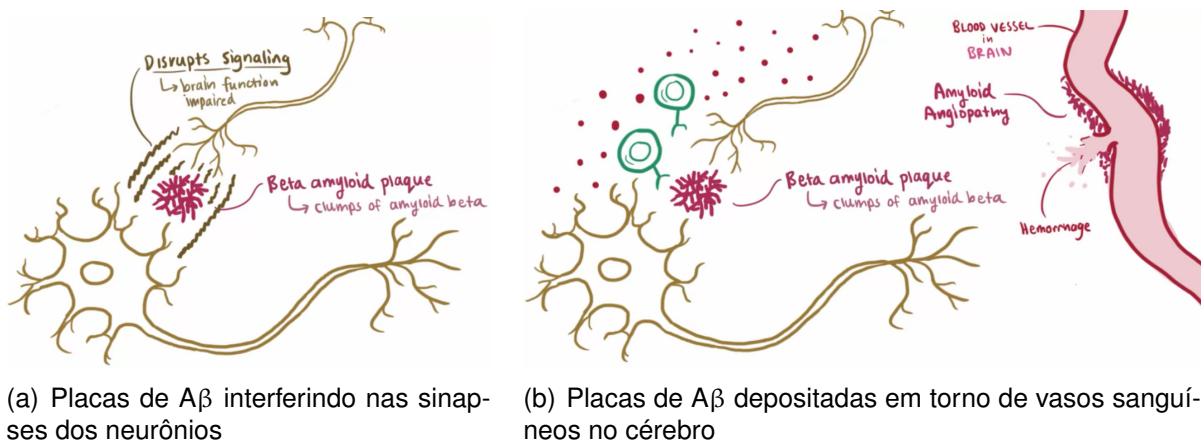


Figura 7: Placas de A β no cérebro (OSMOSIS, 2016)

Além do mais, essas placas podem iniciar uma resposta imune e causar inflamação que pode danificar os neurônios que se encontram ao redor. As placas de A β também podem se depositar em torno de vasos sanguíneos no cérebro, causando a chamada angiopatia amiloide, que enfraquece as paredes dos vasos sanguíneos e aumenta o risco de hemorragia ou ruptura e perda de sangue, como pode ser visto na Figura 7(b) (OSMOSIS, 2016).

2.1.2 O acúmulo de emaranhados neurofibrilares e a proteína tau

Como resultado direto da deposição de placas de A β , postula-se que é gerado o acúmulo de NFTs, seguido de morte celular, dano vascular e demência (HARDY; HIGGINS, 1992; HARDY; SELKOE, 2002).

Os NFTs, que são responsáveis por outra importante parte da DA, ao contrário das placas de A β , se encontram dentro da célula. Assim como outras células, os neurô-

nios são mantidos unidos por uma estrutura de suporte interna que é parcialmente feita de microtúbulos. Estes microtúbulos agem como trilhas, guiando nutrientes e moléculas do corpo da célula até as extremidades do axônio e de volta. Um tipo especial de proteína chamada tau que garante que essas trilhas não se separem. Essa proteína, ao ligar-se ao microtúbulos, estabiliza-os, funcionando como um laço ferroviário (LAKHAN, 2016). A Figura 8 mostra os microtúbulos estabilizados pela proteína tau dentro do neurônio.

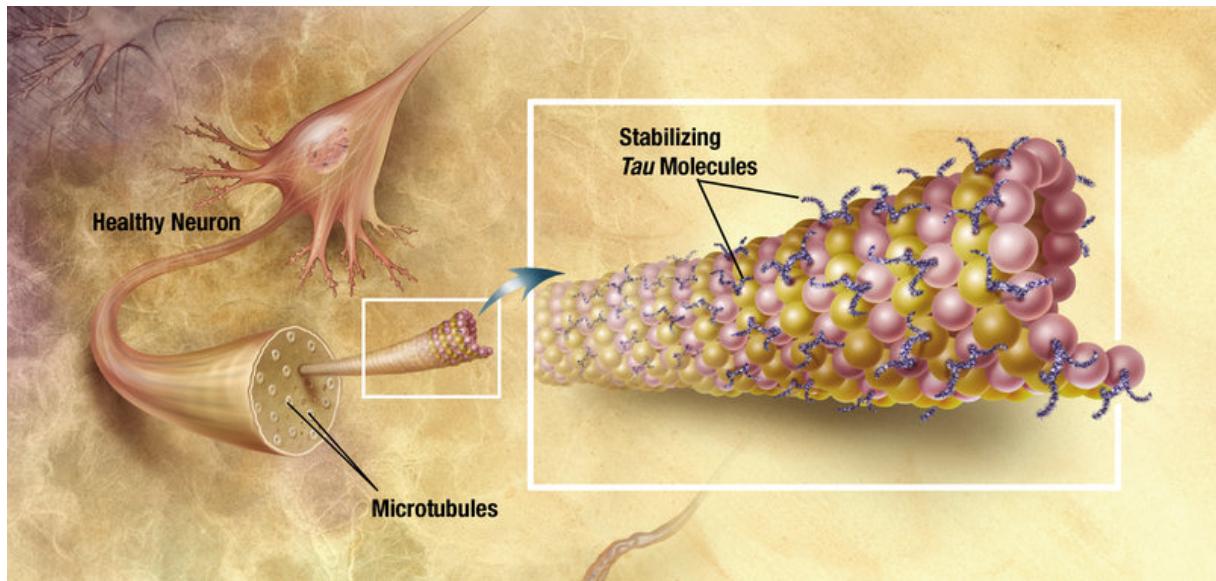


Figura 8: Microtúbulos estabilizados pela proteína tau (ADEAR, 2008)

Embora não completamente compreendido, supõe-se que as placas de A β levam à ativação de quinases, enzimas que transferem grupos fosfato para a proteína tau (OSMOSIS, 2016), como pode ser visto na Figura 9(a). Com isso, a proteína tau é quimicamente alterada, mudando de forma e parando de dar suporte aos microtúbulos, como na Figura 9(b). Quando isso acontece, os microtúbulos se desintegram, desmoronando o sistema de transporte do neurônio (IQBAL et al., 2005). Assim, as proteínas tau hiperfosforiladas passam a se aglomerar dentro do corpo das células nervosas, ficando enredadas, levando ao outro achado característico da DA, os NFTs (GOEDERT; SPILLANTINI; CROWTHER, 1991), como pode ser visto na Figura 10. A formação destes NFTs resultam inicialmente em problemas na comunicação entre neurônios, que mais tarde, acabam sofrendo morte celular (CHUN; JOHNSON, 2007; LAKHAN, 2016).

2.2 Neuropatologia

A perda de neurônios e sinapses geradas pela DA leva a diferenças claramente visíveis no tecido cerebral. Os NFTs são, no início, mais densamente distribuídos na

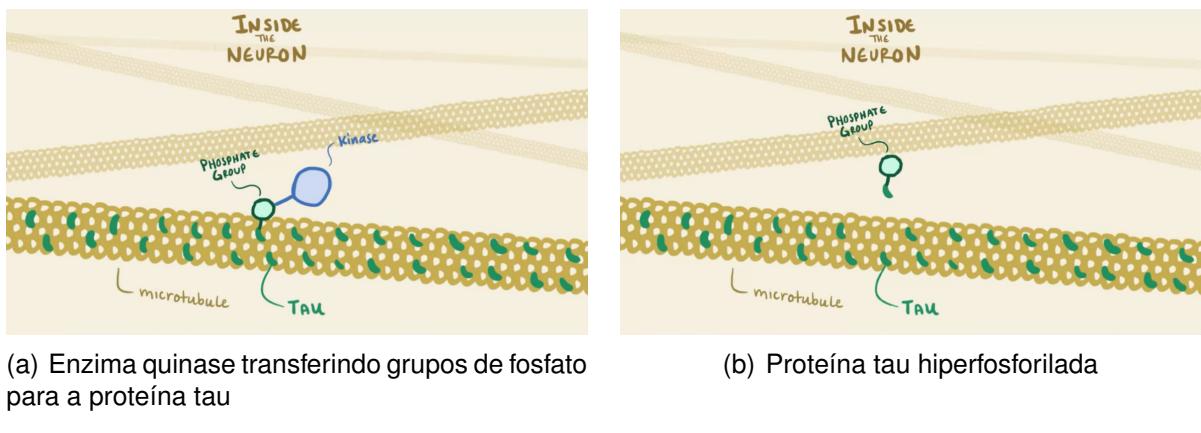


Figura 9: Ativação de quinases e a alteração da proteína tau (OSMOSIS, 2016)

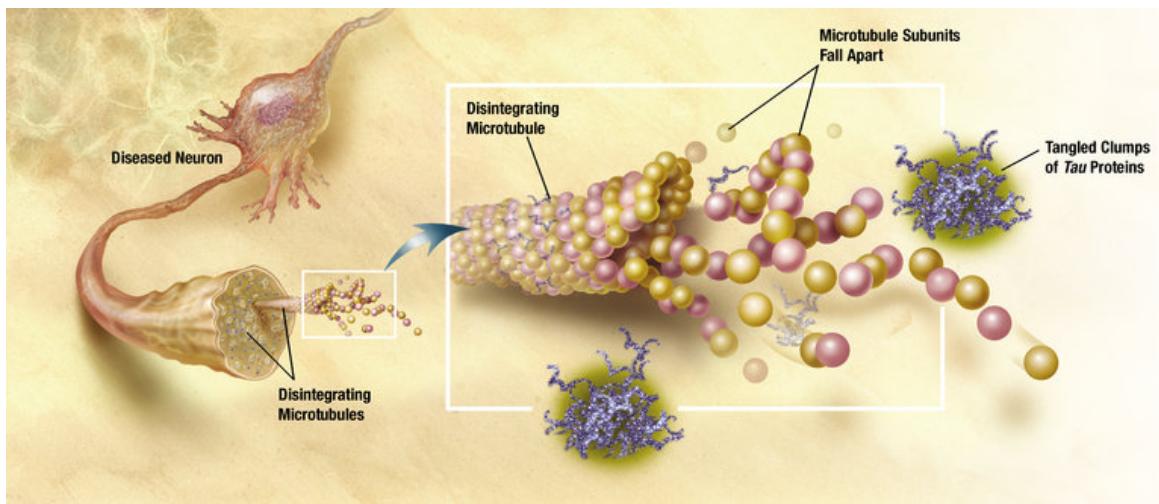


Figura 10: Diagrama de como os microtúbulos se desintegram com a doença de Alzheimer (ADEAR, 2008)

face medial e no polo do lobo temporal. Eles afetam o córtex entorrinal e o hipocampo de uma forma mais severa. À medida que a DA progride, os NFTs acumulam-se em outras regiões corticais, atingindo áreas de associação de alta ordem e com menor frequência, regiões motoras e sensoriais primárias (LAKHAN, 2016).

Assim como os NFTs, as placas de A_β também se acumulam primariamente nas áreas de associação do cérebro. Embora as placas de A_β e os NFTs sejam características da DA, eles não são patognomônicos. Os NFTs são encontrados em vários outros distúrbios neurodegenerativos. Já as placas de A_β podem ocorrer no processo normal de envelhecimento (LAKHAN, 2016).

Portanto, a mera presença dessas características não é suficiente para apoiar o diagnóstico de DA. As placas de A_β e os NFTs devem estar presentes em número suficiente e em uma distribuição topográfica característica para preencher os critérios histopatológicos atuais para a DA. Porém, há consenso de que a presença de números ainda baixos de NFTs no neocôrte cerebral com placas de A_β concomitantes é uma

característica da DA (LAKHAN, 2016).

À medida que os neurônios morrem, grandes mudanças começam a ocorrer no cérebro. O cérebro atrofia, e os giros - saliências do córtex cerebral - se tornam mais estreitos. À medida que estes ficam mais estreitos, os sulcos se tornam mais largos. Com a atrofia, os ventrículos - cavidades cheias de líquido no cérebro - ficam maiores. Essas mudanças são exemplificadas na Figura 11.

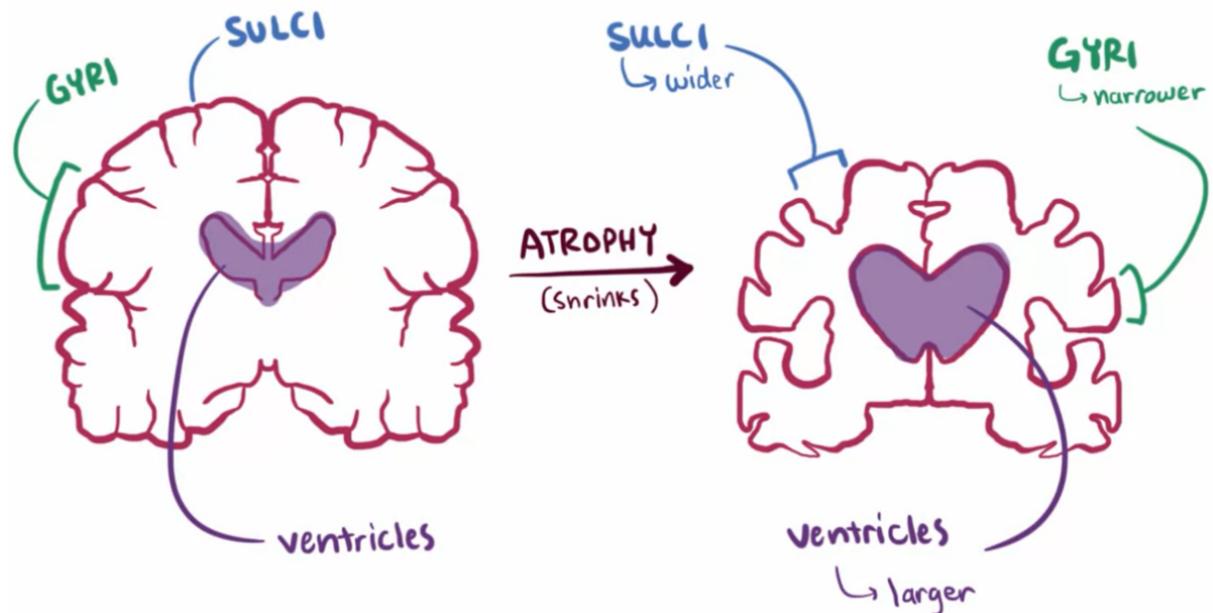


Figura 11: Mudanças que ocorrem no cérebro com o avanço da doença de Alzheimer (OSMOSIS, 2016)

Porém, alterações microscópicas no cérebro começam muito antes dos primeiros sinais de perda de memória, sendo que a presença de lesões médio-temporais começam a aparecer cerca de 5 a 6 anos antes do diagnóstico clínico de Alzheimer. Além disso, estudos mostram que a atrofia nos lobos frontais só ocorre mais perto do momento em que o diagnóstico da doença é feito (ARVESEN, 2015 apud BERNARD et al., 2014).

Para analisar essas mudanças no cérebro de uma forma não invasiva, é utilizada a MRI. No entanto, para mostrar claramente os locais e forma em que as placas de A_β estão depositadas, é necessário utilizar a técnica chamada PET com um radiofármaco marcado com carbono-11 que tem afinidade pelas placas de A_β, o *Pittsburgh Compound B* (PiB). Esta técnica, mais invasiva, requer que um agente de contraste radioativo seja absorvido pelo cérebro do paciente. Além disso, essa é uma técnica mais recente, cara e não tão disponível como a MRI (ARVESEN, 2015 apud BERNARD et al., 2014).

2.3 Biomarcadores

Pesquisadores esperam descobrir uma maneira fácil e precisa de detectar a DA antes que esses sintomas comecem. Para tanto, acredita-se que biomarcadores - marcadores biológicos - oferecem um dos caminhos mais promissores. Um biomarcador é uma substância, medida ou indicador de um estado biológico. Os biomarcadores podem existir antes da ocorrência dos sintomas clínicos. Um exemplo de um biomarcador é o nível de glicose no sangue em jejum, que indica a presença ou não de diabetes, dependendo do valor encontrado para esse marcador (ASSOCIATION, 2010b).

Na DA, para indicar os estágios iniciais, vários biomarcadores estão sendo estudados, como níveis de A β e tau no fluido cefalorraquidiano (CSF, do inglês *Cerebrospinal Fluid*) e alterações cerebrais detectáveis por imagem (ASSOCIATION, 2010b).

Conforme a DA vai progredindo com o tempo, se hipotetiza que as magnitudes de alguns biomarcadores atingem níveis anormais em uma ordem previsível (ARVESEN, 2015). A Figura 12 mostra cinco biomarcadores como indicadores de demência. As curvas indicam mudanças causadas por esses biomarcadores.

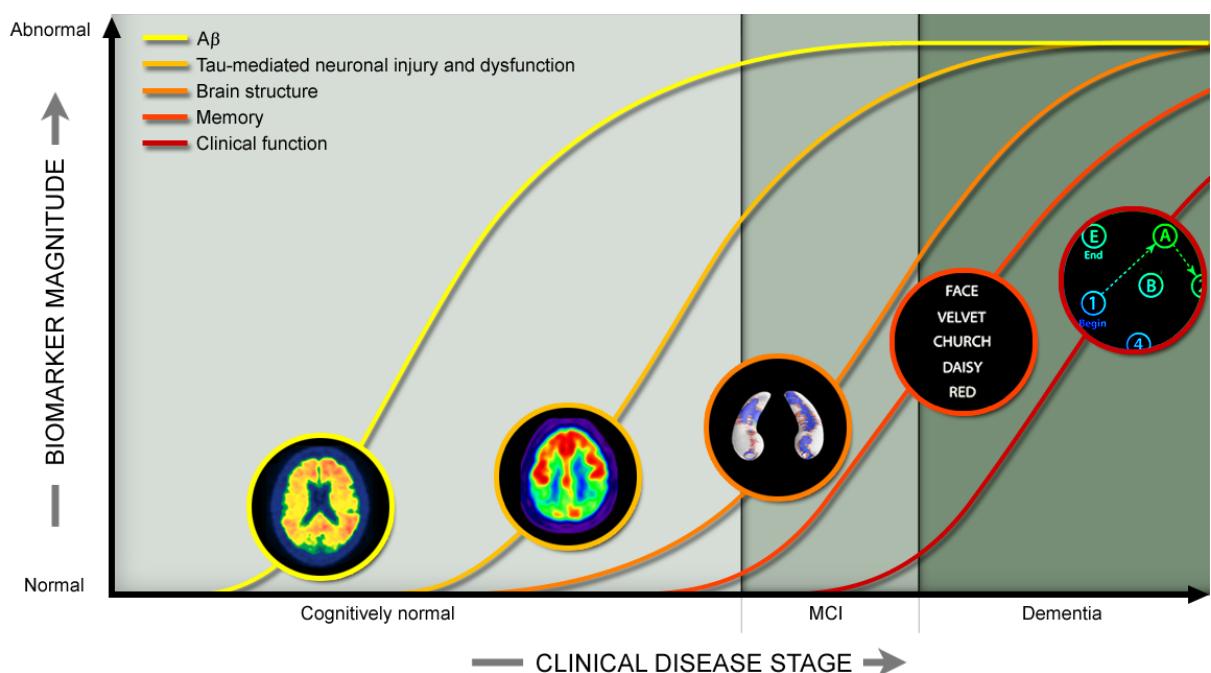


Figura 12: Biomarcadores de Alzheimer ao longo da doença (ADNI, 2016a)

Na Figura 12, são mostrados os seguintes biomarcadores:

1. Imagem de placas de A β detectadas por CSF e PET;
2. Neurodegeneração detectada pelo aumento do **CSF tau** e disfunção sináptica, medida através de PET tendo como marcador radiofármaco *fludesoxiglicose* (FDG);

3. **Atrofia do cérebro** medida com ressonância magnética;
4. **Perda de memória** medida por avaliação cognitiva;
5. **Declínio cognitivo geral** medido pela avaliação cognitiva.

As alterações de 1 a 3 estão indicadas por biomarcadores que podem ser observados antes do diagnóstico de demência. Já os itens 4 e 5 são os indicadores clássicos do diagnóstico de demência (ADNI, 2016a).

Além disso, como pode ser visto na Figura 12, conforme a doença vai avançando, a magnitude dos marcadores é alterada. Com isso, é importante identificar quais marcadores em potencial podem auxiliar no diagnóstico da DA anos antes dela começar apresentar seus sintomas característicos.

2.4 Sintomas e estágios da doença de Alzheimer

Os sintomas da DA pioram conforme as placas de A β e os NFTs vão se acumulando, e os danos neuronais aumentam. No início, os sintomas podem não ser detectáveis, mas à medida que a doença progride, há a perda da memória de curto prazo, a perda de habilidades motoras, a linguagem e comunicação são prejudicadas. Eventualmente, a memória de longo prazo é perdida e progressivamente, há a desorientação. Por fim, os pacientes ficam acamados, e acabam morrendo, sendo que a causa mais comum de morte é por uma infecção, como pneumonia (OSMOSIS, 2016). A Figura 13 mostra um gráfico da evolução dos sintomas conforme a DA avança.

Porém, os primeiros sintomas da doença são frequentemente confundidos com os processos normais de envelhecimento ou estresse (WALDEMAR et al., 2007). Por isso, uma série de testes devem ser feitos a fim de confirmar o diagnóstico, que é apoiado pela presença de características neurológicas e neuropsicológicas e pela exclusão de outras hipóteses (MENDEZ, 2006; KLAJKI et al., 2006). Para a exclusão de outras hipóteses são utilizadas técnicas de imagiologia médica, como tomografia computadorizada (CT, do inglês *Computerized Tomography*), MRI, tomografia computadorizada por emissão de fóton único (SPECT, do inglês *Single-Photon Emission Computed Tomography*) ou PET (NCCMH, 2007). Além disso, essas técnicas também auxiliam a prever a evolução dos estágios que possam indicar o início da DA (SCHROETER et al., 2009).

Esses estágios da DA, que podem ser divididos em quatro, apresentam um padrão progressivo de deficiência cognitiva e funcional.

ALZHEIMER DISEASE

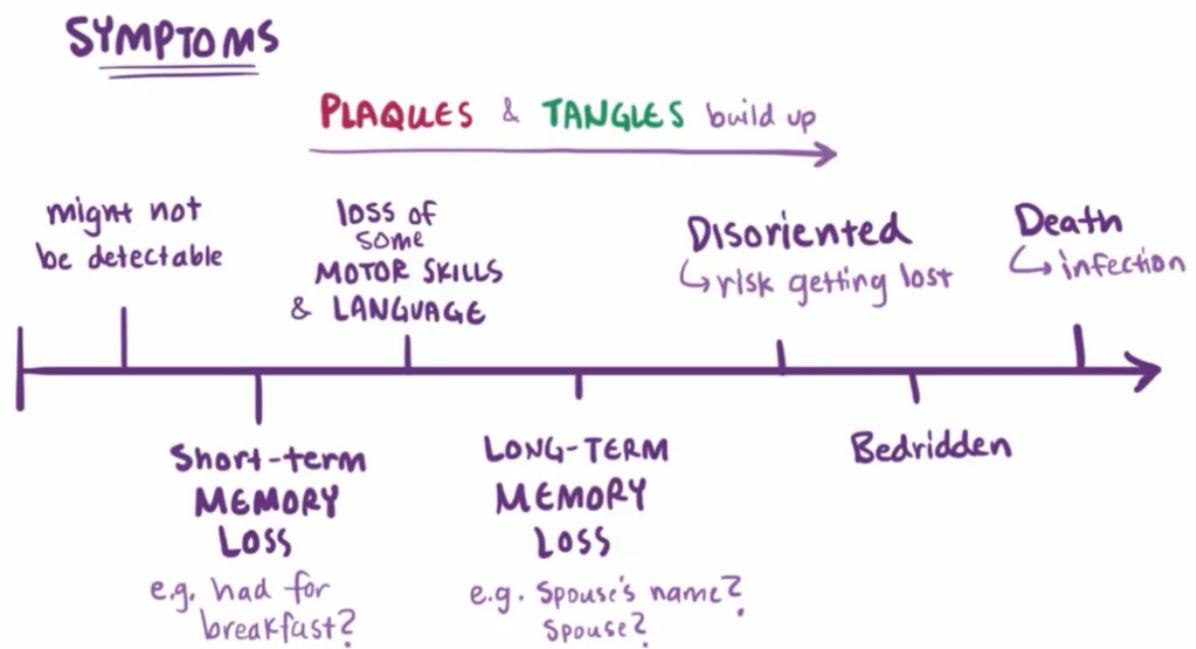


Figura 13: Evolução dos sintomas conforme a DA avança (OSMOSIS, 2016)

2.4.1 Comprometimento cognitivo leve

O primeiro estágio da doença é o comprometimento cognitivo leve (MCI, do inglês *Mild Cognitive Impairment*), que pode afetar algumas das mais complexas atividades do cotidiano, e apresenta a perda de memória a curto prazo como sintoma mais comum, dificultando a recordação de fatos recentes ou a memorização de novas informações (BÄCKMAN et al., 2004; ARNÁIZ; ALMKVIST, 2003).

Além da perda de memória, problemas com atenção, planejamento, flexibilidade, pensamento abstrato, memória semântica, sintomas depressivos, irritabilidade, e diminuição da percepção das próprias dificuldades de memória também podem ser sintomas nos estágios iniciais da doença (BÄCKMAN et al., 2004; MURRAY; BUTTNER; PRICE, 2012). Também nesta fase, é comum a observação de apatia, um sintoma neuropsiquiátrico que persiste durante todas as fases da doença (LANDES et al., 2001).

2.4.2 Estágio inicial

O estágio inicial da doença, devido ao prejuízo crescente tanto em memória quanto em aprendizagem, permite, eventualmente, dar um diagnóstico provável de Alzheimer. O que acontece é que as capacidades de memória não são afetadas igualmente. Enquanto memórias mais antigas, fatos já aprendidos e a memória implí-

cita (memória que guarda como lidar com as atividades do cotidiano, como andar, escovar os dentes, ...) são afetadas em menor grau, novos fatos e memórias tem um impacto maior (CARLESIMO; OSCAR-BERMAN, 1992; JELICIC; BONEBAKKER; BONKE, 1995).

Também, em alguns poucos casos, os problemas de memória não são os sintomas mais comuns, sendo mais proeminente as dificuldades de linguagem, funções executivas, percepção ou execução de movimentos (FÖRSTL; KURZ, 1999).

Nesta fase, é comum que a pessoa com a doença seja capaz de comunicar ideias básicas (FÖRSTL; KURZ, 1999; TALER; PHILLIPS, 2008; FRANK, 1994). Porém, com os problemas de linguagem, o vocabulário fica encolhido e há a diminuição na fluência das palavras, fazendo com que haja um empobrecimento na linguagem oral e escrita (FÖRSTL; KURZ, 1999; TALER; PHILLIPS, 2008).

Além disso, dificuldades na execução de tarefas motoras finas (escrever, desenhar, vestir, ...), na coordenação de movimento e de planejamento podem ocorrer, geralmente, de forma despercebida (FÖRSTL; KURZ, 1999). Com a progressão da DA, muitas vezes é possível continuar a desempenhar muitas tarefas de forma independente, porém, pode ser necessária assistência ou supervisão com as atividades cognitivamente mais exigentes (FÖRSTL; KURZ, 1999).

2.4.3 Estágio intermediário

Com o avanço da DA para o estágio intermediário, a independência fica comprometida, impossibilitando a realização de atividades mais comuns da vida diária (FÖRSTL; KURZ, 1999). O declínio na habilidade de fala, leitura e escrita também ocorre, dificultando a recordação do vocabulário, levando a frequentes substituições incorretas de palavras (FÖRSTL; KURZ, 1999; FRANK, 1994). A memória piora, levando ao esquecimento, inclusive, de parentes próximos. A memória de longo prazo, que até então estava intacta, começa ser prejudicada (FÖRSTL; KURZ, 1999). As capacidades motoras mais complexas se tornam menos coordenadas, podendo levar a quedas (FÖRSTL; KURZ, 1999).

Não obstante, alterações comportamentais e neuropsiquiátricas começam a ser mais predominantes. As manifestações comuns são: vagar pela casa, irritabilidade e incontinência emocional, levando o paciente a chorar, ter explosões de agressão não premeditadas, ou resistência a cuidados especiais (FÖRSTL; KURZ, 1999). Também pode ocorrer casos de síndrome do pôr-do-sol (VOLICER et al., 2001), que é um fenômeno clínico manifestado pela emergência ou incremento de sintomas neuropsiquiátricos noturnos (KHACHIYANTS et al., 2011).

Ademais, nesse estágio, os pacientes podem apresentar incontinência urinária, perda da percepção do processo e das limitações que a doença causa e sintomas delirantes (FÖRSTL; KURZ, 1999). Com a manifestação desses sintomas, um am-

biente de estresse é criado para parentes e cuidadores, o que pode ser reduzido, movendo a pessoa de cuidados domiciliários para clínicas de cuidados de longa duração (FÖRSTL; KURZ, 1999; GOLD et al., 1995).

2.4.4 Estágio avançado

Durante o estágio final da doença, o paciente fica completamente dependente de cuidadores (FÖRSTL; KURZ, 1999). Ocorre também a perda de habilidades de linguagem verbal - que é reduzida a frases simples ou apenas palavras isoladas - podendo levar à completa perda da fala (FÖRSTL; KURZ, 1999; FRANK, 1994). Porém, muitas vezes, as pessoas podem entender e retornar sinais emocionais, embora a agressividade ainda possa estar presente.

Outros sintomas comuns são a extrema apatia e a exaustão, sendo que em última análise, nem as tarefas mais simples podem ser feitas de forma independente. Também, a massa muscular e a mobilidade deterioram-se até o ponto em que o paciente é acamado e fica incapaz de se alimentar. Por fim, a causa da morte é geralmente um fator externo, como infecções ou pneumonia, e não a própria doença em si (FÖRSTL; KURZ, 1999).

2.5 Fatores de risco

A DA envolve a falha progressiva das células cerebrais, no entanto, a razão pela qual as células falham ainda não é clara. Como outras condições crônicas, os especialistas acreditam que a DA se desenvolve como um resultado complexo de múltiplos fatores ao invés de qualquer causa primordial. Os maiores fatores de risco para a DA são idade avançada, histórico familiar de Alzheimer e/ou portar o gene APOE-e4 (ASSOCIATION, 2010a).

2.5.1 Idade

Embora a DA não seja uma parte normal do envelhecimento, o maior fator de risco para a doença é o aumento da idade. Após a idade de 65 anos, o risco de desenvolver a doença dobra a cada cinco anos, e está em quase 50% ao atingir a idade de 85 anos (ASSOCIATION, 2010a).

2.5.2 Genética

Existem duas categorias de genes que influenciam se uma pessoa pode desenvolver uma doença: (i) genes de risco e (ii) genes determinísticos. Os genes de risco aumentam a probabilidade de desenvolver uma doença, mas não garantem que isso vá acontecer. Para a DA, foram descobertos vários genes que aumentam o risco desta acontecer, sendo o APOE-e4 o primeiro gene de risco identificado, e que continua a

ser o gene com maior impacto sobre o risco. APOE-e4 é uma das três formas comuns do gene APOE; Os outros são APOE-e2 e APOE-e3 (ASSOCIATION, 2010a).

Aqueles que herdam uma cópia de APOE-e4 têm um risco aumentado de desenvolver a DA. Aqueles que herdam duas cópias têm um risco ainda maior, mas não é uma certeza. Além de aumentar o risco, APOE-e4 fazer os sintomas aparecerem em uma idade mais jovem do que o habitual. Os cientistas estimam que APOE-e4 está implicado em cerca de 20% a 25% dos casos de Alzheimer (ASSOCIATION, 2010a).

Os genes determinísticos, por sua vez, causam uma doença diretamente, garantindo que quem herda um desses genes irá desenvolver um distúrbio. Na DA, foram descobertas variações que causam diretamente a DA nos genes que codificam três proteínas: APP, *presenilina-1* (PS-1) e *presenilina-2* (PS-2) (ASSOCIATION, 2007). Quando a DA é causada por estas variações determinísticas, é chamada de "DA Autossômica Dominante (DAAD)" ou "DA familiar", e muitos membros da família em várias gerações são afetados. Quando esses genes estão presentes, ocorre a manifestação precoce da DA, com o aparecimento dos sintomas entre os 40 e 50 anos de idade. Porém, esses genes estão presentes em menos de 5% dos casos de Alzheimer (ASSOCIATION, 2010a).

Embora os genes que causam a "DA familiar" sejam raros, a sua descoberta tem fornecido pistas importantes que ajudam a compreensão da DA como um todo. Todos estes genes afetam o processamento ou a produção de A β , que é um suspeito principal do declínio e da morte de células de cérebro (ASSOCIATION, 2010a). Atualmente, várias drogas em desenvolvimento visam dissolver os fragmentos de A β como uma estratégia para parar ou retardar significativamente a DA, como apresentado no trabalho (SEVIGNY et al., 2016).

Hoje, estão disponíveis testes genéticos tanto para APOE-e4 como para os genes raros que causam diretamente a DA. Porém, os profissionais de saúde atualmente não recomendam testes genéticos de rotina para a DA. Testes para APOE-e4 é por vezes incluído como parte de estudos de pesquisa (ASSOCIATION, 2010a).

2.6 *Alzheimer's Disease Neuroimaging Initiative*

A iniciativa de neuroimagem da DA (ADNI, do inglês *Alzheimer's Disease Neuroimaging Initiative*) é um esforço de pesquisa global que apoia a investigação e o desenvolvimento de tratamentos que retardem ou parem a progressão da DA. Este estudo avalia biomarcadores clínicos, de imagem, genéticos e de bioespécime, através do processo de envelhecimento normal até o comprometimento cognitivo leve precoce (EMCI, do inglês *Early Mild Cognitive Impairment*), o comprometimento cognitivo leve tardio (LMCI, do inglês *Late Mild Cognitive Impairment*), a demência ou DA. Com métodos padronizados para criação de imagens e coleta e análise de biomarcadores, o

ADNI facilita o caminho para quem deseja conduzir pesquisas e compartilhar dados compatíveis com pesquisadores de todo o mundo (ADNI, 2016a).

O estudo ADNI é dividido em três fases: ADNI 1, ADNI GO e ADNI 2. Novos participantes foram recrutados em toda a América do Norte durante cada fase do estudo e concordaram em realizar uma série de avaliações de imagem e clínicas. Os participantes são acompanhados e fazem reavaliações ao longo do tempo para que a patologia da doença seja acompanhada à medida que ela avança. Os resultados são então compartilhados pelo ADNI através do *Laboratory of Neuro Imaging's Image Data Archive* da Universidade da Carolina do Sul (USC), EUA, pelo site: <http://adni.loni.usc.edu/> (ADNI, 2016a).

A Tabela 1 resume o número de participantes alvo do estudo ADNI norte-americano, informado, para cada fase do ADNI, a quantidade de novos pacientes que foram inseridos no estudo em cada estágio da doença. O símbolo ↓ significa que os pacientes da fase anterior foram mantidos nesta fase.

Tabela 1: Número de participantes nas diferentes fases do ADNI

Fase	NL	SMC	EMCI	MCI	LMCI	DA
ADNI 1	200	-	-	400	-	200
ADNI GO	↓	-	200	↓	-	-
ADNI 2	150	100	150	↓	150	200

O objetivo do estudo ADNI é acompanhar a progressão da doença longitudinalmente usando biomarcadores para avaliar a estrutura e a função do cérebro ao longo de seis estágios de doença: cognitivamente normal (NL), preocupação significativa de memória (SMC, do inglês *Significant Memory Concern*), EMCI, MCI, LMCI e DA. A Tabela 2 resume os estágios da doença citados e em que fases do ADNI eles são avaliados.

Tabela 2: Estágios acompanhadas em cada fase do ADNI (ADNI, 2016a)

Sigla	Estágio	Fases
NL	Cognitivamente normal	ADNI 1/GO/2
SMC	Preocupação significativa de memória	ADNI 2
EMCI	Comprometimento cognitivo leve precoce	ADNI GO/2
MCI	Comprometimento cognitivo leve	ADNI 1
LMCI	Comprometimento cognitivo leve tardio	ADNI GO/2
DA	Doença de Alzheimer	ADNI 1/GO/2

Na fase ADNI 2, foi adicionado um novo estágio, o SMC. As preocupações com

a memória subjetiva mostraram-se correlacionadas com uma maior probabilidade de progressão da DA, minimizando assim a estratificação do risco entre os controles normais, além de abordar a diferença entre controles idosos saudáveis com quem está no estágio MCI. O principal critério de inclusão que distingue quem está no estágio SMC é o auto-relato de preocupação significativa com a memória (ADNI, 2016a).

No estágio NL, estão os indivíduos de controle no estudo ADNI. Eles não mostram sinais de depressão, nem comprometimento cognitivo leve ou demência. No estágio SMC, estão classificados os participantes que apresentam uma pontuação dentro da faixa normal de cognição, mas indicam que eles têm uma preocupação, e apresentam leve esquecimento. O informante não considera isso como comprometimento progressivo da memória nem considera isso como um esquecimento consistente. Já, nos estágios EMCI, MCI e LMCI estão os participantes que relataram uma preocupação de memória subjetiva de forma autônoma, através de um informante ou de um clínico. No entanto, não há níveis significativos de comprometimento em outros domínios cognitivos, nem em atividades essencialmente preservadas da vida diária, além de não apresentar sinais de demência. Por fim, no estágio DA, estão os indivíduos que foram avaliados e satisfazem critérios para provável DA (ADNI, 2016a).

Além dos estágios descritos anteriormente, temos o Alzheimer pré-clínico, que é um estágio recém-definido e reflete a evidência atual de que o processo fisiopatológico da DA e mudanças mensuráveis em biomarcadores no cérebro podem ocorrer anos, senão décadas, antes da ocorrência dos primeiros sintomas clínicos. Avanços recentes em neuroimagem, testes do líquido cefalorraquidiano e outros biomarcadores agora fornecem a capacidade de detectar evidências de alterações patológicas da DA *in vivo*. Este estágio pré-clínico da DA tornou-se um foco de pesquisa importante, pois o campo postula que a intervenção precoce pode oferecer a melhor chance de sucesso terapêutico. Até o momento, muito poucas evidências são estabelecidas nesta fase "silenciosa" da doença (SPERLING et al., 2011; DUBOIS et al., 2016). Os pacientes que estão no estágio pré-clínico da doença, no entanto, não são explicitamente descritos na base de dados ADNI.

Na Tabela 3, são apresentados os tipos de dados coletados nas diferentes fases do ADNI.

Tabela 3: Tipos de dados coletados nas diferentes fases do ADNI

Fase	MRI	fMRI	DTI	FDG	AV45	PiB	Bioamostras
ADNI I	✓			✓		✓	✓
ADNI GO	✓	✓	✓	✓	✓		✓
ADNI 2	✓	✓	✓	✓	✓		✓

No contexto de diagnóstico por imagem, o MRI é um exame de neuroimagem tradicional que permite a medição precisa do volume tridimensional das estruturas cerebrais, especialmente o tamanho do hipocampo e regiões relacionadas (RAMACHANDRAN, 2016).

A ressonância magnética funcional (fMRI, do inglês *functional Magnetic Resonance Imaging*) tem uma resolução temporal e espacial relativamente elevada e pode ser adquirida na mesma sessão que a imagem de ressonância magnética. Mais importante ainda, a fMRI pode fornecer informações úteis sobre a integridade funcional de redes cerebrais que suportam a memória e outros domínios cognitivos (SPERLING, 2011).

A imagem de tensor de difusão (DTI, do inglês *Diffusion Tensor Imaging*) é uma das ferramentas mais eficazes de ressonância magnética para a investigação da anatomia do cérebro. Além da substância cinzenta, estudos histopatológicos indicam que a substância branca é também um bom alvo tanto para o diagnóstico precoce da DA quanto para monitorar a progressão da doença, o que motiva a usar DTI para estudar pacientes com DA *in vivo* (KENICHI et al., 2011).

O FDG é um marcador radiofármaco utilizado nos exames de imagem PET, e serve para medir as taxas metabólicas cerebrais de glicose (MOSCONI et al., 2010).

Assim como o FDG, o *florbetapir* (AV45) e o PiB são compostos radiofármacos de varredura de PET utilizados para dar contraste nas placas de A β presentes no tecido neuronal (KLUNK et al., 2004; CAMUS et al., 2012). Já as bioamostras, são amostras biológicas derivadas de seres humanos, tais como tecido, liquor e sangue.

Além disso, os dados de estudo disponibilizados pelo ADNI podem ser divididos em diversas categorias. A Tabela 4 mostra essas principais categorias, bem como suas subcategorias. Dentre todos os dados de estudo, temos avaliações neuropsicológicas, dados genéticos, exames de imagiologia (MR e PET), histórico médico, características do paciente e outros. Para mais informações sobre os itens disponíveis no arquivo de dados ADNI, ver Anexo A.

2.7 Trabalhos Correlatos

Na literatura, são encontrados diversos trabalhos que abordam o diagnóstico da DA, que podemos separar em dois grandes grupos: estudo *cross-sectional* (WIKIPEDIA, 2018a) e estudo *follow-up* (POWELL; JANSSEN, 2013). O estudo *cross-sectional*, é um tipo de estudo observacional que analisa os dados coletados de uma população ou um subconjunto representativo em um ponto específico no tempo, isto é, dados transversais. Ou seja, no escopo da DA, é a classificação do estágio da doença em que o paciente se encontra. Já no estudo *follow-up*, um grupo de indivíduos é acompanhado ao longo do tempo para determinar o desenvolvimento do critério que está sendo observado. Na DA, este estudo é feito na tentativa de prever e classificar

Tabela 4: Dados de estudo disponibilizados pelo ADNI

Categoria Principal	Subcategoria
Avaliações	Diagnóstico Neuropsicológica
Bioespécime	Procedimentos de coleta de laboratório Resultados de laboratório
Genético	Informações sobre dados genéticos Resultados do genótipo
Imagen	MRI PET
Histórico médico	Eventos adversos Medicamentos Histórico médico Físico / Neurológico
Estudo	Dados e banco de dados Protocolos de estudo
Características do paciente	Histórico familiar Demografia

a progressão de um estado cognitivo para outro.

Nas próximas sub-seções, serão apresentados os principais trabalhos encontrados na literatura para estudos *cross-sectional* e *follow-up*.

2.7.1 Estudos *cross-sectional*

No artigo de ZHANG et al. (2011), foram combinadas três modalidades de biomarcadores (MRI, FDG-PET e biomarcadores CSF) para discriminar entre DA (ou MCI) e controles saudáveis. Para que a combinação dos três biomarcadores diferentes fosse eficaz, foram gerados três diferentes *kernels* que posteriormente foram combinados para uso em uma SVM com validação cruzada com 10-fold. Os dados dos pacientes foram extraídos do ADNI, sendo 52 pacientes NL, 99 MCI e 51 DA. Na comparação entre NL e DA, foi obtida uma precisão de 93,2% com uma sensitividade de 93% e especificidade de 93,3%. Já para a classificação de MCI e DA, foi atingida uma precisão de 76,4% com uma sensitividade de 81,8% e especificidade de 66%

Em (ORTIZ et al., 2016), é explorada a construção de um método de classificação baseado em arquiteturas de *deep learning* aplicadas em regiões cerebrais. As imagens da matéria cinzenta de cada área do cérebro foram divididas em *patches* 3D que são então usados para treinar diferentes *deep belief networks*. Um conjunto de *deep belief networks* é então composto, onde a previsão final é determinada por um

esquema de votação. No trabalho, são implementadas e comparadas duas estruturas baseadas em aprendizagem profunda e quatro esquemas de votação diferentes. O método resultante foi avaliado usando *cross-validation* com *10-fold* aplicado sobre a base de dados ADNI, sendo 68 pacientes NL, 111 com MCI, 26 com LMCI e 70 com DA. A arquitetura de classificação atinge uma precisão de até 90% e AUC de 0,95 para classificação de NL e DA. Já na comparação entre sMCI (MCI estável) e DA, a precisão é de 84% e AUC de 0,91. Para a classificação de NL e pMCI (MCI que progrediu para DA), a precisão é de 83% com AUC de 0,95.

No artigo proposto por LONG et al. (2017), é proposto um método de aprendizagem de máquina que utiliza uma SVM para distinguir pacientes com DA ou MCI de idosos saudáveis. A classificação é feita a partir do cálculo e análise das diferenças morfológicas regionais do cérebro entre os diferentes grupos através do uso de imagens de ressonância magnética de alta resolução. O estudo foi feito sobre a base de dados ADNI, sendo 135 indivíduos NL, 132 sMCI, 96 pMCI e 64 DA. A validação cruzada com *10-fold* foi adotada para estimar o desempenho da classificação, que atingiu uma acurácia de 96,5% e AUC de 0,995 na diferenciação de NL e DA.

No trabalho de BEHESHTI; DEMIREL; MATSUDA (2017), foi desenvolvido um sistema de diagnóstico que utiliza um ranking de atributos e um algoritmo genético para analisar dados de imagem de MRI. O trabalho investiga a atrofia da matéria cinzenta global entre pacientes normais e com DA, gerando assim, regiões de interesse que são utilizados para todas as outras classes de diagnóstico. Esses valores entram em um vetor de atributos, onde é extraído um subconjunto ótimo através de um algoritmo genético. Finalmente, a classificação é feita utilizando uma SVM com validação cruzada de *10-fold*. O conjunto de dados foi extraído da base de dados ADNI, sendo 162 NL, 65 sMCI, 71 pMCI e 160 DA. A classificação de NL e DA tem uma precisão de 93% e AUC de 0,935.

Por fim, o trabalho de RODRIGUES et al. (2017) compara a capacidade de três métodos de seleção de atributos para obter combinações de biomarcadores que maximizem as taxas de classificação. Além disso, é utilizado o classificador kNN com pré-processamento de balanço de dados, subamostragem aleatória e sobreamostragem (SMOTE). Os dados utilizados no trabalho foram extraídos do ADNI, sendo 152 indivíduos NL, 414 MCI e 100 DA. Os resultados mostram que a abordagem proposta atinge 94,34% de precisão para reproduzir a definição das classes NL, MCI e DA.

Dentre os trabalhos na categoria *cross-sectional*, podemos observar a classificação de pacientes nos conjuntos {NL e MCI}, {NL e DA}, {MCI e DA} e {NL, MCI e DA}. Além disso, vemos também que diferentes técnicas são utilizadas para a classificação, tanto em termos de algoritmo utilizado para o classificador, quanto na questão de pré-processamento dos dados.

2.7.2 Estudos *follow-up*

O trabalho de MATHOTAARACHCHI et al. (2017) propõe um método probabilístico baseado em aprendizado de máquina chamado RUSRF, e que foi projetado para avaliar a progressão para demência em um período de 24 meses, com base apenas no biomarcador [¹⁸F]Florbetapir PET. No trabalho, foram avaliados pacientes retirados da base de dados ADNI, sendo 230 com sMCI e 43 com pMCI. O método utiliza o algoritmo RF com uma otimização que permite que as classes sejam balanceadas ao fazer o treinamento. O algoritmo utiliza um conjunto de teste independente e obteve uma precisão de 84% e AUC de 0,906.

No trabalho de LONG et al. (2017), além do estudo *cross-sectional*, um estudo *follow-up* de 36 meses também foi realizado. Os resultados mostram que, na diferenciação de sMCI e pMCI, uma acurácia de 89% é atingida com AUC de 0,932. Já ao classificar NL e pMCI, uma acurácia de 91,74% foi atingida com AUC de 0,971.

O sistema apresentado por BEHESHTI; DEMIREL; MATSUDA (2017), também é capaz de mostrar a previsão de conversão de MCI para AD entre 12 e 36 meses antes do diagnóstico clínico. Os resultados mostram uma precisão de 75% com AUC de 0,751.

Já no cenário dos trabalhos *follow-up*, os principais focos são a classificação da conversão de MCI para DA, a diferenciação entre pacientes MCI estáveis e MCI que converteram para DA e entre pacientes NL e com MCI progressivo. No entanto, nenhum dos trabalhos correlatos encontrados na literatura realiza a predição de pacientes em risco de desenvolver a DA, que é o foco deste trabalho.

A Tabela 5 mostra um resumo dos trabalhos apresentados nesta seção.

2.8 Considerações finais do capítulo

Neste capítulo, foi apresentada uma revisão sobre a DA juntamente com suas características, a iniciativa ADNI e trabalhos correlatos. Os principais tópicos abordados da doença foram: patofisiologia, neuropatologia, biomarcadores, sintomas e estágios da doença, além dos fatores de risco.

No escopo da patofisiologia, foram explanadas as duas mais forte hipóteses da causa da DA. São elas: (i) a deposição de placas de A_β e (ii) o acúmulo de emaranhados neurofibrilares. Com a perda de sinapses geradas pela DA, diversas mudanças começam a acontecer no cérebro e na vida do paciente. Essas mudanças foram vistas nas subseções neuropatologia e sintomas e estágios da DA.

Também, os principais fatores de risco da doença foram abordados, mostrando que não apenas a idade, mas a genética e o histórico familiar também podem contribuir para o aparecimento da DA. Além disso, foi feita uma breve explicação sobre biomarcadores, para que ele servem e como são úteis para detectar a doença antes

dos sintomas começarem.

Sabendo da importância dos biomarcadores para a detecção e rastreamento precoce da doença, um esforço global ADNI foi criado, com o propósito de desenvolver novos biomarcadores clínicos, de imagem, genéticos e bioquímicos.

Por fim, foram apresentados os principais trabalhos encontrados na literatura que abordam a classificação da DA por meio do uso de técnicas de IA. Os trabalhos foram divididos entre estudos *cross-sectional*, que se preocupam em classificar a doença em um ponto específico no tempo, e estudos *follow-up*, onde a variável tempo já entra como critério para a predição da doença em um período de tempo estabelecido.

Tabela 5: Resumo dos trabalhos relacionados

Trabalho	Tamanho da Amostra	Tipo da Validação	Classificação	Precisão (%)	AUC	Follow-up
ZHANG et al. (2011)	52 NL, 99 MCI, 51 DA	Cross-validation 10-fold	NL e DA MCI e DA	93,2 76,4	- -	-
ORTIZ et al. (2016)	68 NL, 111 MCI 26 LMCI, 70 DA	Cross-validation 10-fold	NL e DA sMCI e DA NL e pMCI	90 84 83	0,95 0,91 0,95	- - -
LONG et al. (2017)	135 NL, 132 sMCI 96 pMCI, 64 DA	Cross-validation 10-fold	NL e DA sMCI e pMCI NL e pMCI	96,5 89 91,74	0,995 0,932 0,971	- 36 36
BEHESHTI; DEMIREL; MATSUDA (2017)	162 NL, 65 sMCI 71 pMCI, 160 DA	Cross-validation 10-fold	NL e DA MCI e DA	93 75	0,935 0,751	- 12~36
RODRIGUES et al. (2017)	152 NL, 414 MCI, 100 DA	Cross-validation 10-fold	NL, MCI e DA	94,34	-	-
MATHOTAAARACHCHI et al. (2017)	230 sMCI, 43 pMCI	Conj. teste independente	sMCI e pMCI	84	0,906	24

3 INTELIGÊNCIA ARTIFICIAL E A ÁREA MÉDICA

Quando consideramos os avanços obtidos na medicina nas últimas décadas, é inegável que a computação foi uma grande contribuinte para que esses pudessem ser atingidos. Ademais, a computação se mostra uma grande aliada para ajudar no diagnóstico de doenças, uma vez que a Inteligência Artificial (IA) tem sido cada vez mais aplicada na área médica (HAMET; TREMBLAY, 2017).

Além disso, é crescente a preocupação na medicina para obter a maior quantidade possível de dados sobre a saúde ou doença dos pacientes e tomar decisões com base nisso. Antes, os médicos tinham que confiar em sua experiência, capacidade de julgamento e habilidade de resolução de problemas enquanto usavam ferramentas rudimentares e recursos limitados. No entanto, nos dias de hoje, as tecnologias disruptivas começaram a disponibilizar métodos avançados não só para profissionais da medicina, mas também para seus pacientes. Dentre estas tecnologias, está a genômica, biotecnologia, sensores portáteis e IA (MESKO, 2017).

Ao invés de desenvolver tratamentos para populações e tomar as mesmas decisões médicas com base em algumas características físicas semelhantes entre os pacientes, a medicina tem avançado para a prevenção, personalização e precisão, o que fará com que soluções médicas tradicionais *one-fits-all* migrem para tratamentos direcionados, terapias personalizadas e drogas exclusivas. Em outras palavras: medicina de precisão. Nesta mudança e transformação cultural, a IA é uma tecnologia chave que pode trazer esta oportunidade à prática diária (MESKO, 2017).

Neste aspecto, a aplicação de IA na medicina se divide em dois ramos principais: virtual e físico. O ramo físico envolve o uso de robôs. Este ramo não será explorado neste trabalho, uma vez que o escopo deste trabalho é o ramo virtual, representado pela Aprendizagem de Máquina (ML, do inglês *Machine Learning*), a qual utiliza algoritmos matemáticos que melhoram a aprendizagem através da experiência. Existem três tipos de algoritmos de ML: (i) não supervisionados (capacidade de encontrar padrões), (ii) supervisionados (algoritmos de classificação e predição baseados em exemplos anteriores) e (iii) aprendizagem de reforço (uso de sequências de recompensas e punições para formar uma estratégia para operação em um problema

específico) (HAMET; TREMBLAY, 2017).

Dentre as principais abordagens de ML podemos citar o aprendizado baseado em árvores de decisão, regressão logística, rede neural artificial, máquina de vetores de suporte, *clustering* e outros. Além disso, é importante salientar que dentro de cada uma dessas abordagens, são propostos na literatura uma gama de algoritmos diferentes que utilizam uma mesma abordagem conceitual. As próximas seções detalharão as principais abordagens descritas anteriormente juntamente com os algoritmos específicos de cada categoria que foram utilizados para o desenvolvimento deste trabalho.

3.1 Árvores de decisão

Árvores de decisão (DT, do inglês *Decision Trees*) são um método de aprendizado supervisionado não paramétrico usado para classificação e regressão. O objetivo é criar um modelo que preveja o valor de uma variável destino aprendendo regras de decisão simples inferidas a partir de dados recebidos como entrada. Após a geração de uma árvore de decisão, é possível extrair um conjunto de regras de decisão do tipo *if-then-else*. Quanto mais profunda a árvore, mais complexas são as regras de decisão e menos genérico é o modelo gerado. Dentre os algoritmos de árvore de decisão mais conhecidos estão o ID3 (QUINLAN, 1986) e o C4.5 (QUINLAN, 1993) (evolução do algoritmo ID3), que são amplamente utilizados nas mais diversas aplicações.

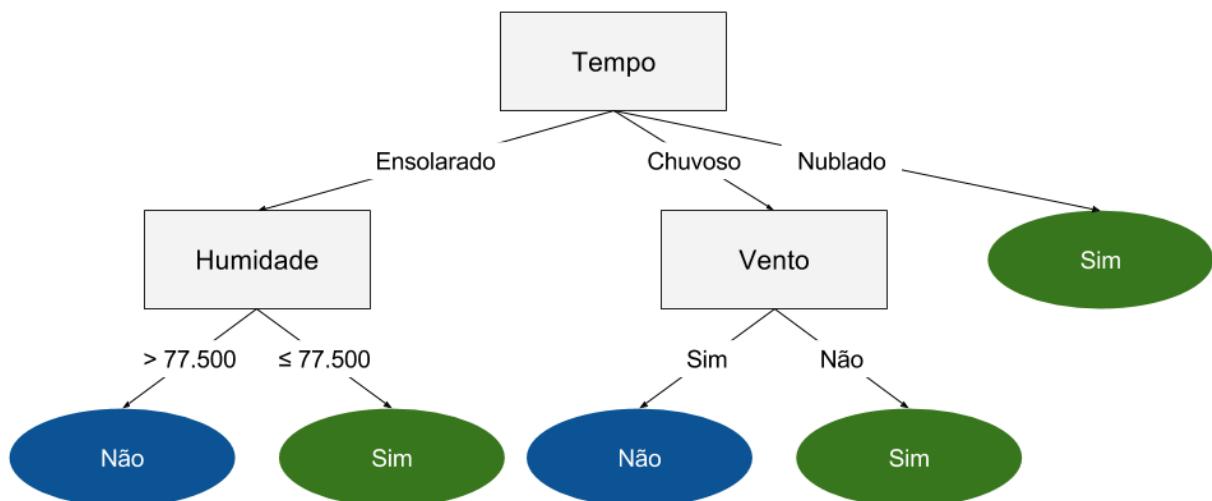


Figura 14: Exemplo de uma árvore de decisão

A Figura 14 mostra um exemplo de uma árvore de decisão que classifica se o dia é bom ou não para jogar golfe. As folhas representam o resultado da classificação, enquanto os nodos são os atributos as arestas são os possíveis valores dos atributos. Para saber o resultado da classificação, basta percorrer os nodos e arestas até encontrar uma folha. Por exemplo, caso *tempo = chuvoso* e *vento = não*, é um bom dia para jogar golfe. No entanto, caso *tempo = chuvoso* e *vento = sim*, não é um bom dia

para jogar golfe.

Quando tentamos prever a variável alvo usando qualquer técnica de aprendizado de máquina, temos três principais causas que geram a diferença entre o valor esperado e o valor predito. Essas causas são o ruído, a variância e a polarização. Para ajudar a reduzir esses fatores, pode ser utilizada a técnica de *ensemble* (GROVER, 2017).

Ensemble é definido como uma coleção de preditores que são utilizados em conjunto para gerar uma previsão final. A eficiência desta técnica se baseia na premissa de que utilizar diversos preditores diferentes para prever a mesma variável alvo é melhor do que utilizar um único preditor sozinho. As técnicas de *Ensembling* são ainda classificadas em *Bagging* e *Boosting* (GROVER, 2017).

Na técnica de *Bagging*, cada preditor é construído de forma independente. Um exemplo de *Bagging* são as *Random Forests*. No caso do *Boosting*, os preditores não são gerados independentemente, mas sim sequencialmente. Esta técnica emprega a lógica de que os preditores subsequentes aprendem com os erros dos preditores anteriores. Como os novos preditores aprendem com erros cometidos por preditores anteriores, menos iterações são necessárias para finalizar a execução do algoritmo. Um exemplo de *Boosting* são as *Gradient Boosted Trees* (GROVER, 2017).

A Figura 15 ilustra graficamente a diferença entre um classificador simples, um classificador gerado pela técnica de *Bagging* e um classificador gerado pela técnica de *Boosting*. Note que enquanto o classificador simples possui apenas um modelo, o *Bagging* gera vários modelos diferentes, criados independentemente. Por fim, podemos perceber a dependência dos modelos anteriores para a geração de um classificador *Boosting*.

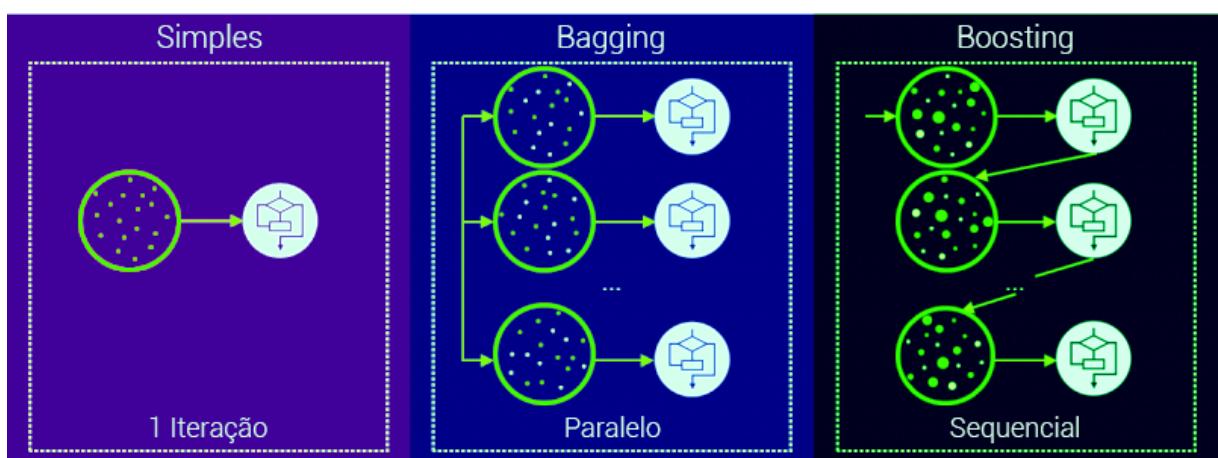


Figura 15: Comparação entre um classificador simples, e classificadores gerados pela técnica de *Bagging* e de *Boosting* (XRISTICA, 2016)

3.1.1 *Random Forest*

As *Random Forests* (RF) são definidas como uma combinação de preditores gerados por árvores de decisão, de modo que cada árvore é construída a partir de valores de um vetor aleatório amostrado independentemente e com a mesma distribuição para todas as árvores na floresta (BREIMAN, 2001). Desta forma, para gerar a classificação de um novo exemplo, o algoritmo escolhe a classificação de acordo com a maioria dos votos, que é calculado sobre todas as árvores na floresta.

3.1.2 *Gradient Boosted Trees*

As *Gradient Boosted Trees* (GBTs) são uma das técnicas para classificação e regressão. As GBTs produzem um modelo de predição robusto sob a forma de um conjunto de modelos de predição fracos, tipicamente árvores de decisão. Aqui, um modelo de predição fraco pode ser qualquer modelo que funcione um pouco melhor do que o acaso.

Como qualquer algoritmo de aprendizagem supervisionado, o objetivo das GBTs é definir uma função de perda e minimizá-la. Para isso, o algoritmo vai repetidamente encontrando um padrão nos erros do classificador gerado em cada iteração e utiliza esses erros para gerar um classificador ainda melhor. Estes erros significam entradas de dados que são difíceis de serem ajustadas por um modelo simples. Desta forma, o algoritmo se concentra nos dados que são difíceis de serem ajustados. Uma vez que chegamos a um estágio que os erros não possuem nenhum padrão que possa ser modelado, o algoritmo pode parar sua execução (GROVER, 2017).

3.2 Rede neural artificial

As redes neurais tem como principal objetivo simular células cerebrais interconectadas dentro de um computador, com o propósito de aprender novos conceitos, reconhecer padrões ou tomar decisões da forma mais parecida possível com que um humano faria. A maior vantagem de uma rede neural é que ela aprende sozinha, sem a necessidade de ser programada novamente a cada novo problema.

Uma rede neural típica tem de algumas dezenas a centenas, milhares ou até mesmo milhões de neurônios artificiais que são dispostos em uma série de camadas, cada uma das quais se conecta às camadas ao seu lado.

3.2.1 *Multilayer perceptron*

Uma das principais arquiteturas de redes neurais é a *Multilayer Perceptron* (MLP). As MLPs possuem uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. A camada de entrada é projetada para receber várias formas de informação do mundo exterior que a rede tentará aprender, reconhecer ou processar.

No lado oposto da rede, a camada de saída responde às informações que aprendeu. Entre essas duas camadas, as camadas ocultas formam a maioria do cérebro artificial. Além disso, os neurônios das MLPs estão totalmente conectados, o que significa que cada neurônio se conecta com todos os outros neurônios da camada anterior e da próxima (caso essas camadas existam). As conexões entre um neurônio e outro são representadas por um número chamado de peso, que pode ser positivo ou negativo. Quanto maior o peso, maior a influência que um neurônio tem sobre os outros. Além disso, para atualizar os pesos da rede, são propostos diferentes algoritmos na literatura. Esses algoritmos são chamados de *solvers*, sendo que o mais conhecido deles é a descida de gradiente. A Figura 16 mostra um exemplo de uma MLP com três neurônios na camada de entrada, três camadas ocultas com quatro, três e dois neurônios em cada camada, respectivamente, e uma camada de saída com apenas um neurônio.

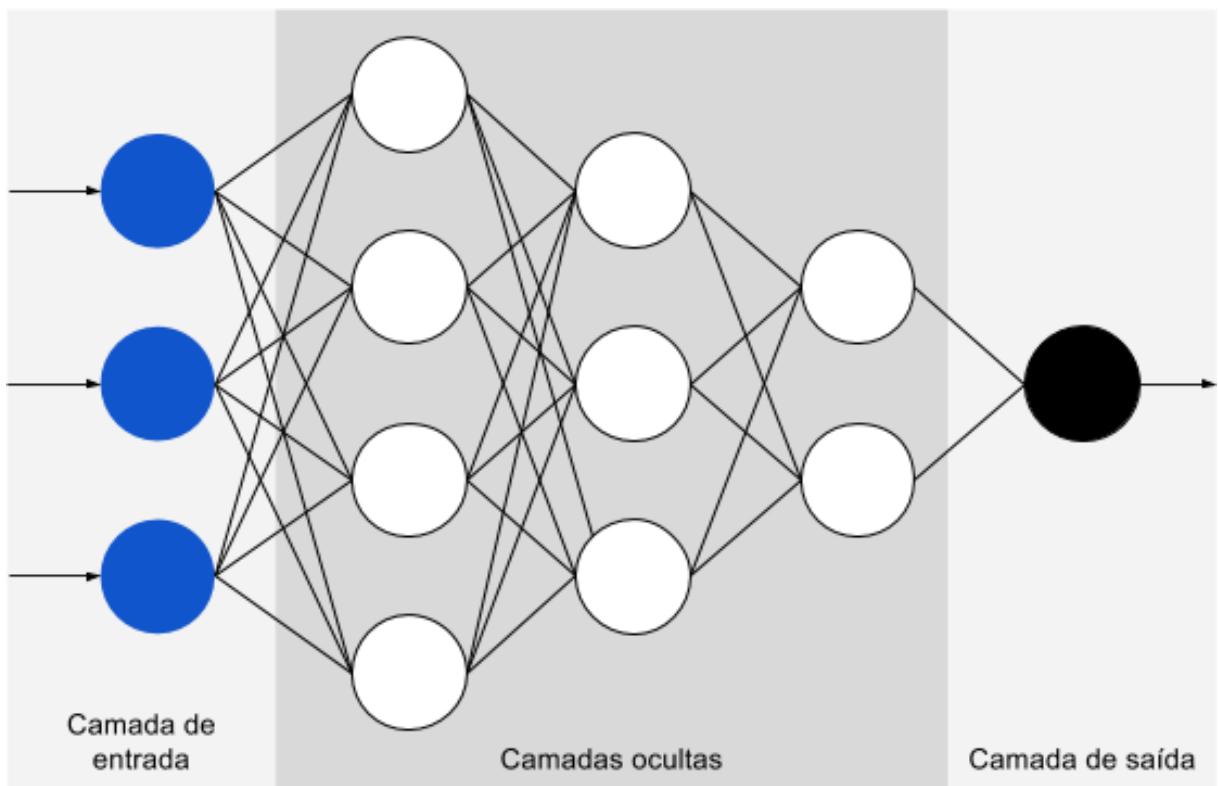


Figura 16: Exemplo de uma rede neural *Multilayer Perceptron*

3.3 Regressão linear

A regressão linear pode ser definida como o estudo de uma variável dependente em função de outras variáveis que podem auxiliar a entender o problema. Ou seja, estudar o que está disponível no momento para poder entender como reagir no futuro. Na regressão linear, o resultado (variável dependente) é contínuo, podendo assumir

qualquer valor de um número infinito de valores possíveis. Além disso, a regressão linear gera uma equação que é da forma $Y = mX + C$.

Matematicamente, podemos quantificar em números o quanto uma variável está relacionada a outra e como ela impacta em seu comportamento. Assim, com o uso da regressão, podemos obter qual a reta ou curva a ser traçada que melhor se ajusta nos dados disponibilizados. Na Figura 17, podemos observar um exemplo de uma reta gerada pela regressão linear para um conjunto de dados.

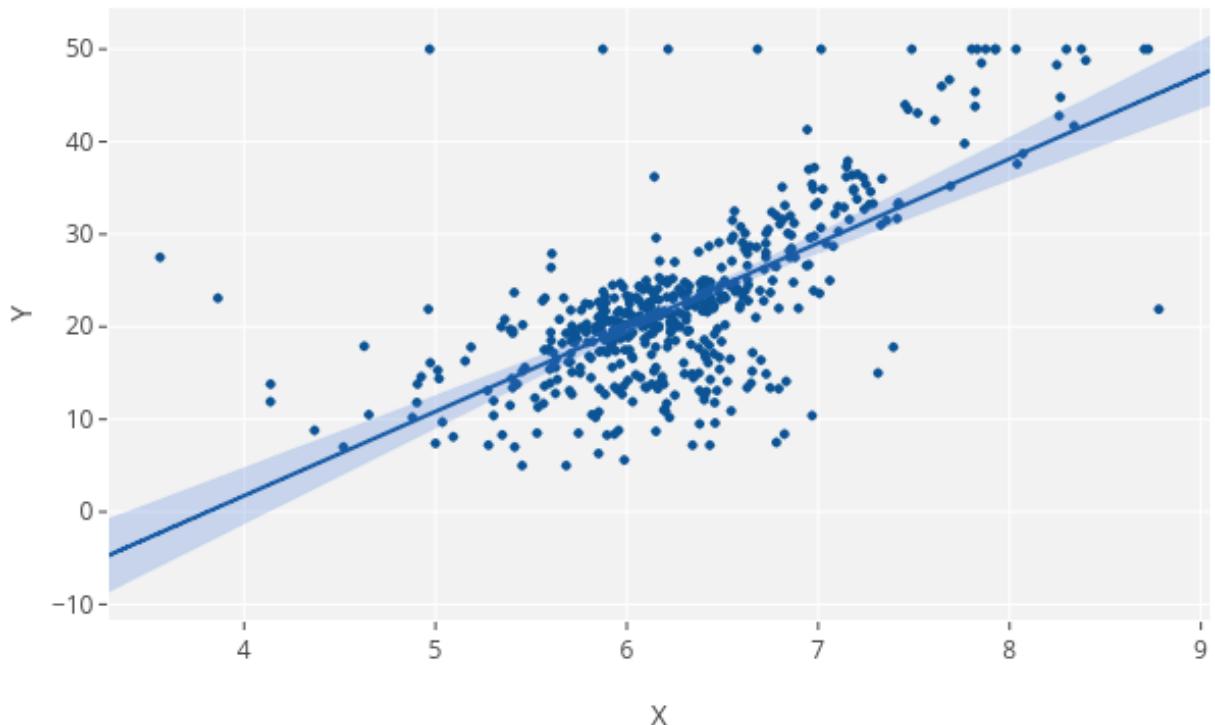


Figura 17: Exemplo de regressão linear

3.4 Regressão logística

Diferentemente da regressão linear, onde o resultado do classificador é definido por uma variável contínua, na Regressão Logística (RLog), o resultado pode assumir apenas um número limitado de valores possíveis. Desta forma, a RLog é usada quando a variável de resposta é de natureza categórica. Na RLog, uma equação da forma $Y = e^{X/1} + e^{-X}$ é gerada. Na Figura 18 temos um exemplo de RLog, e a comparação com o resultado de uma regressão linear.

3.5 Modelo Linear Generalizado

O Modelo Linear Generalizado (MLG), proposto por (NELDER; WEDDERBURN, 1972), é uma generalização flexível da regressão linear comum. Com ele, é possível

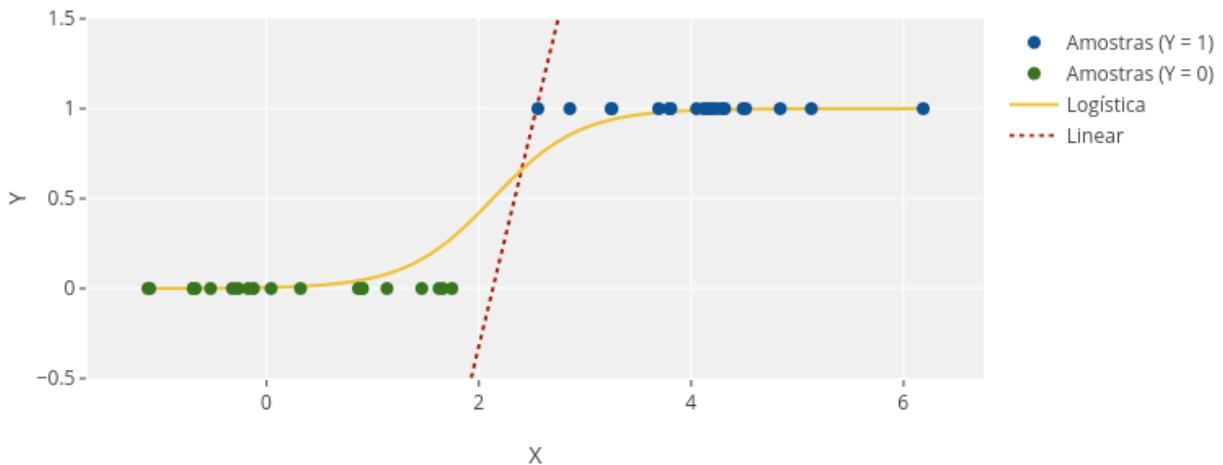


Figura 18: Exemplo de regressão logística

utilizar variáveis de saída que possuem modelos de distribuição de erros diferentes da distribuição normal. O MLG generaliza a regressão linear, permitindo que o modelo linear seja relacionado à variável de saída através de uma função de ligação e permite que a magnitude da variância de cada medida seja uma função do seu valor previsto. O MLG foi formulado como uma forma de unificar vários outros modelos estatísticos, incluindo regressão linear, regressão logística e regressão de Poisson.

3.6 Máquina de vetor de suporte

As máquinas de vetor de suporte (SVMs, do inglês *support vector machine*) são modelos de aprendizagem supervisionados que analisam os dados e reconhecem padrões. As SVMs são utilizadas para classificação e análise de regressão. Dado um conjunto de exemplos de treinamento, para cada entrada, o algoritmo define qual de duas possíveis classes a entrada faz parte. Isso faz com que as SVMs sejam um classificador linear binário não probabilístico. Além disso, um modelo SVM é representado por pontos distribuídos no espaço, mapeados para que os exemplos das categorias separadas sejam divididos por um espaço livre o mais amplo possível. Quando novos exemplos são recebidos, o algoritmo os mapeia para o mesmo espaço e define qual categoria ele pertencerá com base em qual lado da divisão ele está. A Figura 19 exemplifica uma SVM, onde podemos perceber que os exemplos foram divididos por uma reta em duas regiões distintas através de dois vetores de suporte.

3.7 k -vizinhos mais próximos

No reconhecimento de padrões, o algoritmo de k -vizinhos mais próximos (k NN) é um método não paramétrico usado para classificação e regressão, ou seja, o classificador não aprende nenhum parâmetro, pois não há processo de treinamento. Ao

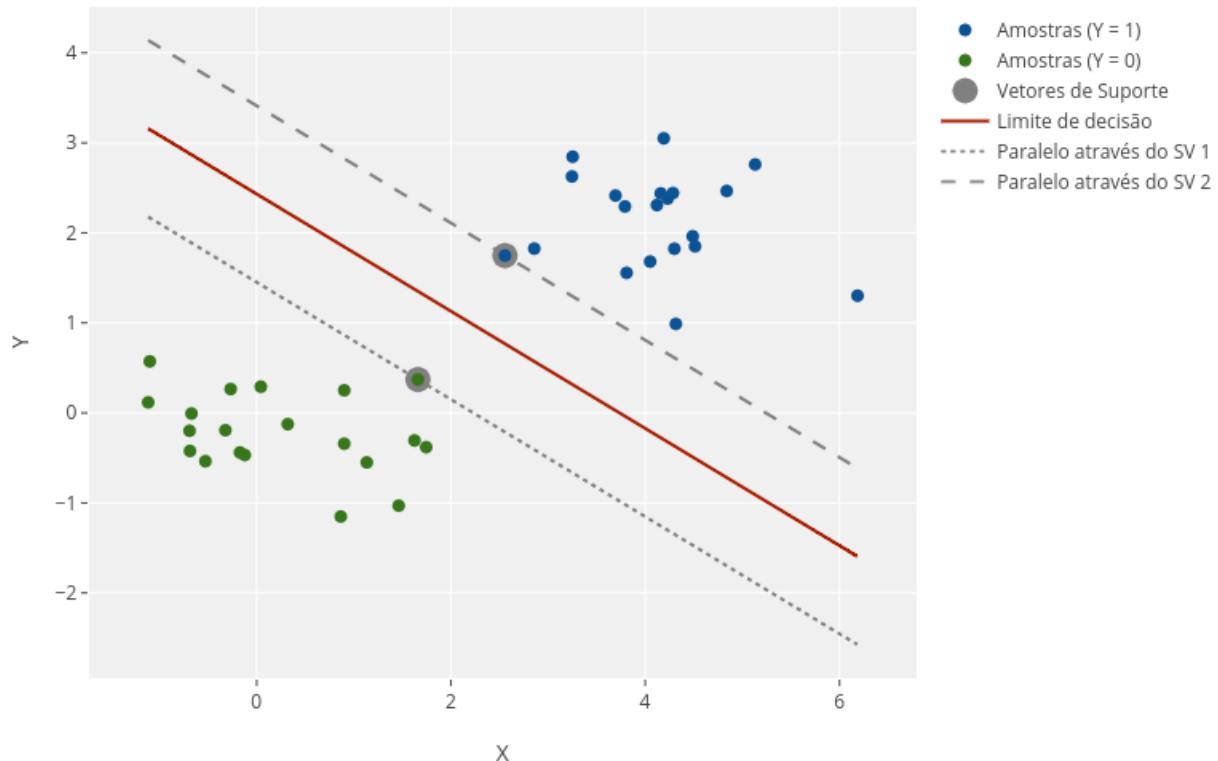


Figura 19: Exemplo de uma SVM

invés disso, a ideia é que sejam mantidas todas as amostras de treinamento, para que quando um novo exemplo seja recebido, o classificador possa medir a distância entre o novo ponto e todos os outros pontos que já foram processados.

Primeiramente, deve ser fornecido ao algoritmo os rótulos de cada exemplo de treinamento. Na sequência, para classificar um novo exemplo, o algoritmo encontrará os k pontos mais próximos nas amostras de treinamento e atribuirá ao exemplo o rótulo da classe mais comum entre esses vizinhos. A Figura 20 exemplifica o funcionamento do kNN.

Na Figura 20, para classificar o novo exemplo (representado por um círculo no centro da imagem), no caso de $k = 3$, a nova instância será classificada como sendo pertencente à Classe B, pois dos três pontos mais próximos, dois deles são da Classe B. No entanto, no caso de $k = 6$, a classificação será da Classe A, uma vez que dentre os 6 vizinhos mais próximos, quatro deles são da Classe A.

3.8 Stacked Generalization

O *Stacked Generalization* (WOLPERT, 1992) é um esquema proposto por Wolpert para minimizar a taxa de erro de classificadores. Para isso, ao contrário do *bagging* e do *ensemble*, ele combina modelos de diferentes tipos (i.e. árvores de decisão, redes neurais, regressão logística, ...), introduzindo o conceito de meta-aprendizagem. Inicialmente, todos os modelos são treinados da forma convencional, gerando seus

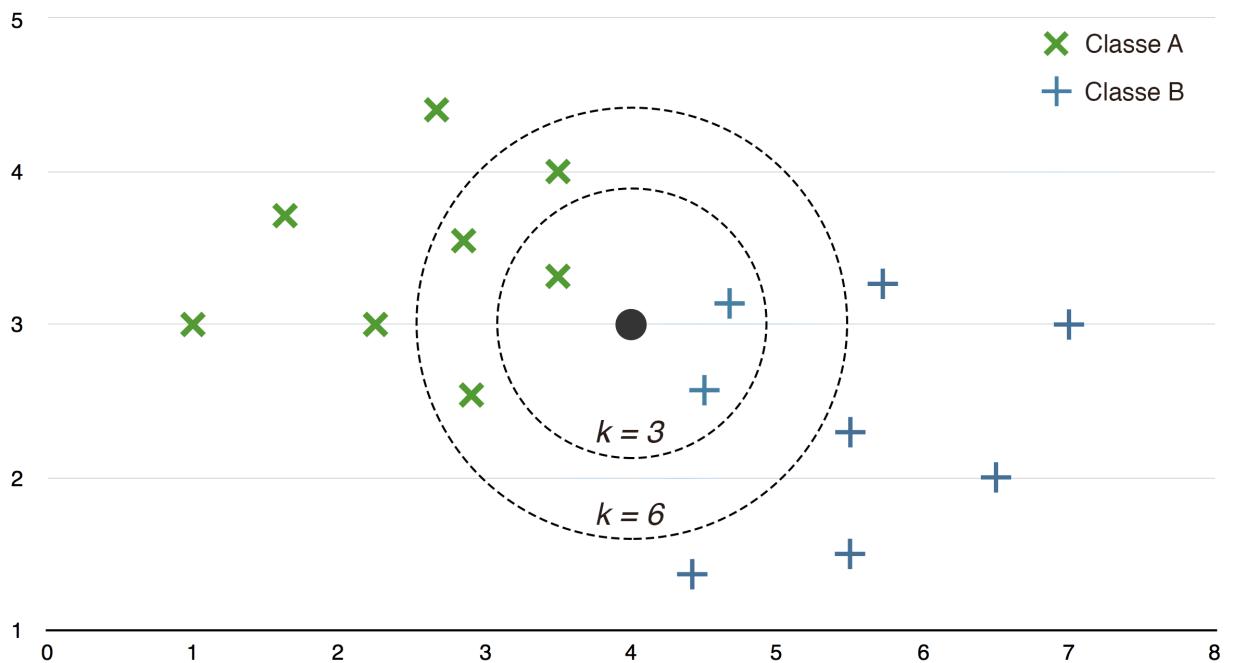


Figura 20: Exemplo do funcionamento do algoritmo kNN

modelos de predição. Usando as previsões geradas na primeira etapa como entrada e as respostas corretas como saída esperada, um classificador de nível superior é treinado. Desta forma, este classificador utilizará as previsões geradas nas saídas dos classificadores do nível inferior para gerar um modelo mais robusto. A Figura 21 exemplifica o *Stacked Generalization*.

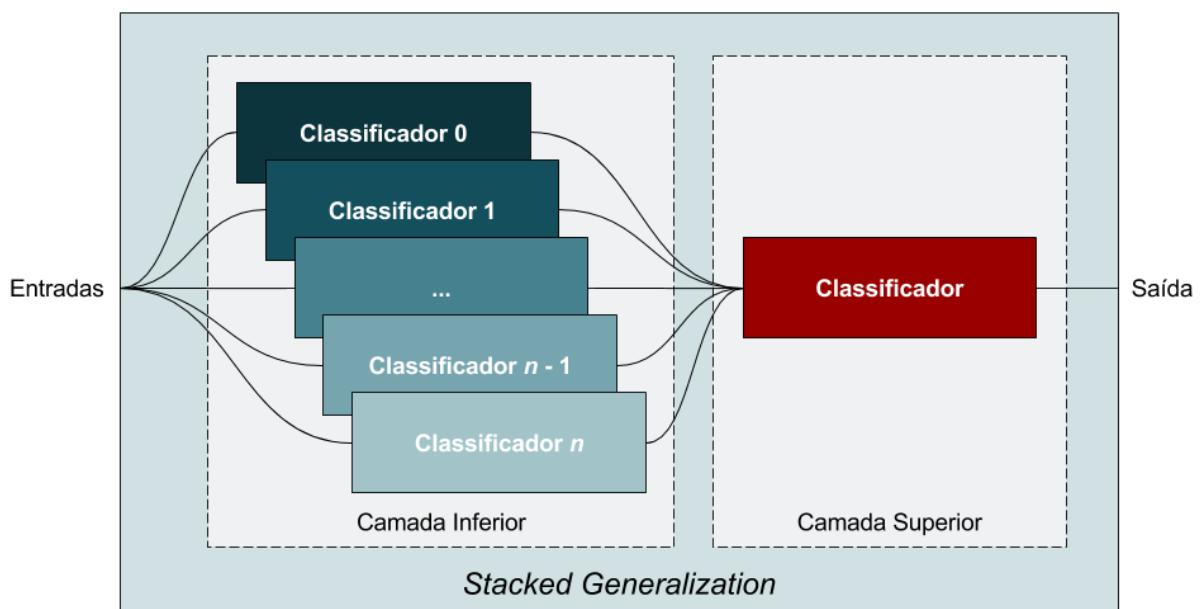


Figura 21: Modelo genérico do *Stacked Generalization*

3.9 Métricas de avaliação

Nas sub-seções seguintes serão apresentadas as principais métricas de avaliação dos algoritmos de aprendizado de máquina.

3.9.1 Matriz de confusão

Uma matriz de confusão é uma ferramenta de visualização tipicamente utilizada na aprendizagem supervisionada. Cada coluna da matriz representa as instâncias em uma classe predita, enquanto cada linha representa as instâncias em uma classe real. Na matriz de confusão, a diagonal principal mostra as instâncias corretamente classificadas (WIKIPEDIA, 2018b). Portanto, a matriz de confusão é usada para avaliar o desempenho de um classificador, e conta como um classificador é exato em fazer previsões sobre classificação. A Tabela 6 mostra um exemplo de matriz de confusão. Nesta tabela, percebe-se que cinco gatos, três cachorros e onze coelhos foram corretamente classificados. No entanto, dois gatos foram erroneamente preditos como cachorros, três cachorros foram preditos como gatos, dois cachorros como coelhos e um coelho foi predito como cachorro.

Tabela 6: Exemplo de matriz de confusão

		Classe predita		
		Gato	Cachorro	Coelho
Classe real	Gato	5	2	0
	Cachorro	3	3	2
	Coelho	0	1	11

3.9.2 Métricas derivadas da matriz de confusão

A partir de matriz de confusão, é possível extrair um conjunto de métricas que permite uma análise mais detalhada do que uma mera proporção de classificações corretas (WIKIPEDIA, 2018b). A Tabela 7 resume as métricas derivadas da matriz de confusão.

Tabela 7: Métricas derivadas da matriz de confusão

		Classe predita	
		Positivo	Negativo
Classe real	Positivo	Verdadeiro Positivo (TP)	Falso Negativo (FN)
	Negativo	Falso Positivo (FP)	Verdadeiro Negativo (TN)

A proporção de instâncias positivas que foram classificadas como positivas é chamada de Verdadeiro Positivo (TP, do inglês *True Positive*). Os exemplos negativos

classificados como positivos são chamados de Falso Positivo (FP, do inglês *False Positive*). As instâncias positivas classificadas como negativas dão origem ao Falso Negativo (FN, do inglês *False Negative*). Por fim, as instâncias negativas classificadas como negativas são chamadas de Verdadeiro Negativo (TN, do inglês *True Negative*).

3.9.3 Acurácia

A acurácia é uma métrica que responde, no geral, o quanto frequente o classificador está correto. A acurácia é definida como:

$$\text{Acurácia} = \frac{TP + TN}{\text{Total}} \quad (1)$$

A acurácia não é uma métrica confiável para o desempenho real de um classificador, porque produzirá resultados enganadores se o conjunto de dados estiver desequilibrado (ou seja, quando o número de observações em diferentes classes variam muito). Portanto, para que esta métrica possa ser utilizada, o conjunto de dados deve ser balanceado.

3.9.4 Precisão

A precisão é a fração de instâncias recuperadas que são relevantes, ou seja, mostra quanto os resultados da pesquisa são úteis. A precisão é definida como:

$$\text{Precisão} = \frac{TP}{TP + FP} \quad (2)$$

3.9.5 Sensibilidade

A sensibilidade nos informa a frequência em que o classificador encontra os exemplos de uma classe. Ou seja, é a fração de instâncias relevantes que são recuperadas. Traduzindo para fórmula:

$$\text{Sensibilidade} = \frac{TP}{TP + FN} \quad (3)$$

3.9.6 F1 Score

O F1 Score combina precisão e sensibilidade de modo a trazer um número único que indique a qualidade geral do seu modelo. A fórmula que define o F1 é a seguinte:

$$F1 = 2 \times \frac{\text{precisão} \times \text{sensibilidade}}{\text{precisão} + \text{sensibilidade}} \quad (4)$$

3.9.7 Área sob a curva ROC

A área sob a curva ROC (AUC, do inglês *Area Under the Curve*) nos diz a probabilidade de um classificador ranquear uma instância positiva escolhida aleatoriamente

melhor do que uma instância negativa, também escolhida aleatoriamente. Quando o valor da área está abaixo da diagonal 0,5 ou 50%, o modelo não tem validade, pois, os acertos e erros entram na mesma proporção e são devidos ao acaso. Analogamente, um valor igual a 1,0 ou 100% não chega a ser alcançado, pois, sempre há superposição na distribuição das proporções dos grupos. A Figura 22 mostra um exemplo de curva ROC.

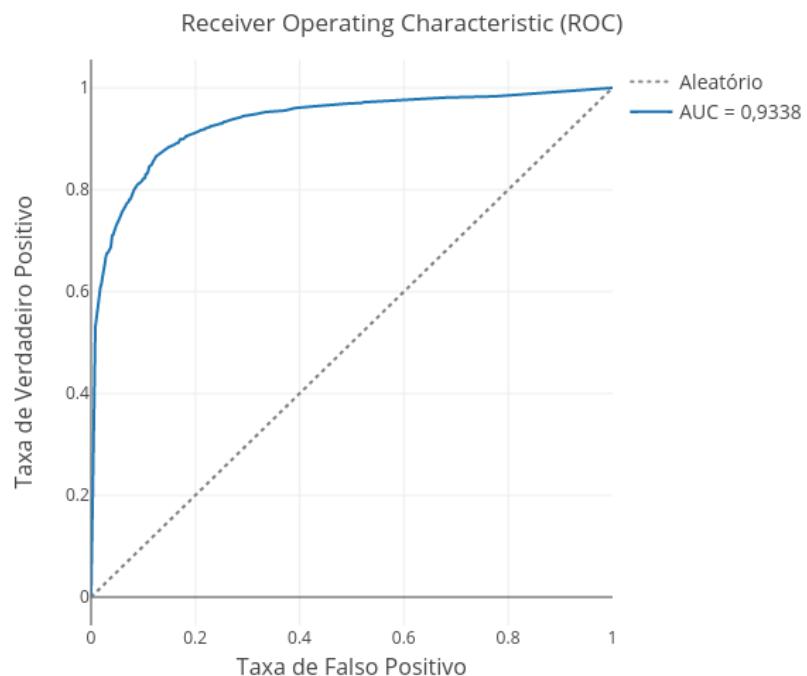


Figura 22: Exemplo de curva ROC

Quanto mais os resultados estão próximos do canto superior esquerdo do gráfico (valores próximos de 1), maior a sensibilidade e menor proporção de falsos positivos, tornando o modelo mais preciso.

4 METODOLOGIA E TRABALHO PROPOSTO

Inicialmente, foi realizado um estudo aprofundado da DA e suas características. Os principais tópicos estudados da DA foram sua patofisiologia, neuropatologia, biomarcadores, sintomas e estágios da doença e fatores de risco. Este estudo foi de extrema importância, uma vez que o assunto abordado não faz parte do conhecimento geral da área de computação e proporcionou um melhor entendimento da área.

Juntamente com o aprendizado da DA, foi feita uma exploração do ADNI, que é um esforço de pesquisa global que apoia a investigação e o desenvolvimento de tratamentos que retardem ou parem a progressão da DA. Como mencionado anteriormente, este estudo disponibiliza uma base de dados com biomarcadores clínicos, de imagem, genéticos e de bioespécime.

Desta forma, o problema foi estudado com foco nos biomarcadores disponíveis, uma vez que eles trazem resultados de diferentes tipos de exame para cada paciente. São eles: informações demográficas, exames neuropsicológicos, de imagem, de bioespécime, genéticos, histórico médico e outros. Além disso, os pacientes são monitorados ao longo do tempo, permitindo que a evolução da doença em cada um deles possa ser observada.

Ao se obter um entendimento mais profundo dos dados disponíveis na base de dados ADNI, pode-se definir o problema como sendo uma tarefa de classificação, uma vez que é desejado, a partir dos biomarcadores, determinar qual é o estado cognitivo do paciente. Portanto, propomos modelar o problema com o uso de técnicas de IA. Técnicas como SVM (CORTES; VAPNIK, 1995), árvores de decisão (e.g. RF (HO, 1995)), MLP e RLog (BISHOP, 2006) foram aplicadas em contextos semelhantes (e.g. BHATTACHARYA; BHATIA (2010), ZHANG et al. (2008), GIL; JOHNSSON (2009) e MUSLIMOVIC et al. (2005)) e foram avaliadas neste trabalho. Além das técnicas de classificação, foram aplicadas algumas técnicas de clusterização, como o kNN (ALTMAN, 1992). Essas técnicas foram comparadas utilizando as métricas de desempenho clássicas, como acurácia, precisão, sensibilidade, AUC e F1. Com essas métricas, foi possível escolher as melhores técnicas para serem aplicadas na solução do problema.

Além disso, foi realizada uma análise com os biomarcadores da base de dados

ADNI, a fim de selecionar os atributos mais relevantes para serem utilizados na construção dos modelos. Para isso, foi utilizada a taxa de ganho de informação. Não obstante, além do ganho de informação, tivemos a preocupação de reduzir ao máximo a quantidade de exames, e consequentemente, de biomarcadores, necessários para o desenvolvimento do trabalho. Isso foi feito pois, como dito anteriormente, desejamos que o método possa ser utilizado pela grande maioria da população, inclusive em zonas mais distantes dos grandes centros urbanos.

O estudo feito neste trabalho foi dividido em duas partes. A primeira, é um estudo em abrangência das técnicas de IA aplicadas para a classificação do estado cognitivo dos pacientes da base de dados ADNI. A segunda, é a proposta de um método para classificação de pacientes em estado de risco, ou seja, esse método se preocupa, principalmente, em informar se o paciente é definido como estável (baixa probabilidade de desenvolver a doença) ou em risco (alta probabilidade de desenvolver a doença).

Para o estudo em abrangência e a geração da taxa de ganho de informação dos biomarcadores, foi utilizada a ferramenta Orange3 (DEMŠAR et al., 2013). A Orange3 é uma ferramenta *open-source* que permite a modelagem rápida de técnicas de IA de uma forma interativa e visual.

No entanto, como a Orange3 é uma ferramenta direcionada para propósitos educacionais, ela carece de uma maior liberdade para ajustes finos nos algoritmos. Portanto, para a modelagem do método para classificação de pacientes em estado de risco, que é o objetivo principal deste trabalho, foi utilizada a ferramenta RapidMiner Studio (RAPIDMINER, 2017). O RapidMiner Studio é um ambiente de modelagem visual que possibilita a criação de fluxos de trabalho analíticos preditivos. Além disso, esta ferramenta possui centenas de algoritmos pré-definidos para pré-processamento de dados e aprendizagem de máquina.

Após a modelagem do método proposto, foram realizados testes com o intuito de validar o método implementado. Para isso, foi utilizado o método de *cross-validation* com *10-fold*, aplicado sobre um conjunto de dados balanceado obtidos da base de dados ADNI. Por fim, os resultados obtidos pelos diferentes algoritmos avaliados foram comparados para que pudesse ser escolhido aquele que entrega uma melhor precisão de predição.

Com a definição do algoritmo a ser utilizado para a resolução do problema, uma análise dos resultados de classificação foi realizada. Além de se observar as taxas de verdadeiro positivo e verdadeiro negativo (exemplos corretamente classificados), foram analisados os falsos negativos (indivíduos cognitivamente normais que foram classificados como pacientes em risco), a fim de observar se essas pessoas estão na fase pré-clínica da DA.

4.1 Biomarcadores utilizados

Os experimentos realizados neste trabalho utilizaram conjuntos com diferentes arranjos de biomarcadores para treinamento dos algoritmos de aprendizagem de máquina. Para uma melhor organização, estes biomarcadores serão descritos a seguir, separados por categoria.

- Dados Demográficos (D)
 - AGE: idade em anos
 - PTGENDER: gênero (*male / female*)
 - PTEDUCAT: educação em anos
- PET
 - FDG: metabolismo de glicose no cérebro
 - AV45: placas de A β no cérebro
- CSF - mede o fluido cerebrospinal (pg/mL)
 - ABETA: quantidade de proteína de A β no liquor
 - TAU: quantidade de tau no liquor
 - PTAU: quantidade de fosfo-tau no liquor
- Testes Neuropsicológicos (N)
 - ADAS11: teste que avalia a cognição e diferencia o funcionamento cognitivo normal da deterioração do funcionamento cognitivo
 - CDRSB: teste que avalia a memória, orientação, julgamento e resolução de problemas, assuntos da comunidade, lar e *hobbies* e cuidados pessoais
 - MMSE: teste que avalia a orientação, recuperação de palavras, atenção, cálculo, habilidades de linguagem e capacidade visuoespacial
 - RAVLT: teste que avalia a memória recente, aprendizagem, interferência, retenção e memória de reconhecimento
- MRI - mede o volume de diferentes regiões do cérebro (mm^3)
 - ICV: volume intracerebral
 - Entorhinal: córtex entorrinal
 - Fusiform: giro fusiforme, parte do lobo temporal e occipital
 - Hippocampus: hipocampo, estrutura localizada nos lobos temporais, considerada a principal sede da memória

- MidTemp: lobo médio temporal, região vital para a memória declarativa e de longo prazo
- Ventricles: ventrículos, uma rede comunicante de cavidades preenchidas com líquido cefalorraquidiano e localizadas no parênquima cerebral
- WholeBrain: volume de todo cérebro
- Genético (G)
 - APOE-e4: presença da apolipoproteína E4.

4.2 Classificação da Doença de Alzheimer

Com o propósito de diagnosticar o estágio da DA dos pacientes em sua primeira visita, foi realizado um estudo em abrangência com diferentes algoritmos de aprendizado de máquina com a ferramenta Orange3.

4.2.1 Conjunto de Dados

O atributo da base de dados a ser predito pelos classificadores é o *DX*, que indica o diagnóstico do paciente em uma determinada visita. Ele pode assumir os seguintes valores: *NL*, *MCI* ou *Dementia*.

Para este estudo, foram utilizados 900 instâncias de exames de pacientes em sua primeira visita (atributo *VISCODE* com valor *bl*) extraídas da base de dados ADNI, formando um conjunto de três classes balanceadas, sendo 300 *NL*, 300 *MCI* e 300 *Dementia*.

Para a predição do diagnóstico da DA, quatro conjuntos de biomarcadores foram avaliados. A Tabela 8 resume os conjuntos de biomarcadores que foram utilizados para a classificação da DA. O símbolo ✓ significa que todos os biomarcadores da categoria em questão foram utilizados.

Tabela 8: Conjuntos de biomarcadores avaliados para a predição do diagnóstico da DA

Categoria	D-PET-CSF-N	D-PET-CSF-N*	D-PET-CSF	D-PET-CSF-MRI-G
Dados Demográficos	✓	✓	✓	✓
PET	✓	✓	✓	✓
CSF	✓	✓	✓	✓
Testes Neuropsicológicos	✓	MMSE e RAVLT		
MRI				✓
Genéticos				✓

Nos quatro conjuntos, tanto os Dados Demográficos quanto os biomarcadores PET e CSF foram utilizados. O que diferencia os experimentos é a utilização de testes neuropsicológicos e o uso de MRI e biomarcadores genéticos. Enquanto no experimento D-PET-CSF-N todos os testes neuropsicológicos são utilizados, no D-PET-CSF-N* apenas os testes MMSE e RAVLT foram incluídos. O motivo pelo qual os testes ADAS11 e CDRSB foram removidos do experimento é que são determinantes clínicos para definir o diagnóstico. Ou seja, o diagnóstico do paciente é definido pelo valor obtido por estes dois exames.

Resumidamente, os testes neuropsicológicos avaliam o estado cognitivo do paciente. No entanto, a utilização destes testesN para o diagnóstico da DA tem três grandes problemas: (i) alto custo, (ii) a impossibilidade de escalar para uma grande quantidade de pessoas, uma vez que, para realizar o exame, são necessários profissionais treinados e (iii) estes exames só identificarão pacientes que já estão em estágios mais avançados de perda cognitiva. Portanto, o experimento D-PET-CSF não utiliza nem um teste neuropsicológico, mas somente dados demográficos e os biomarcadores PET e CSF.

No experimento D-PET-CSF-MRI-G, são adicionados ao teste anterior (D-PET-CSF) os biomarcadores extraídos de imagem de MRI e genéticos.

4.2.2 Classificadores

Os seguintes classificadores foram avaliados neste estudo: *Tree*, RF, MLP, kNN, RLog e SVM. Aqui, o algoritmo chamado *Tree* é uma implementação de árvore de decisão que foi desenvolvida pela própria Orange3.

4.2.3 Avaliação dos Resultados

Para treinamento e teste dos experimentos, foi utilizado o método de *cross-validation* com *10-fold*. No entanto, uma limitação do trabalho neste ponto é que não foi separada uma porção do conjunto de 900 exemplos para validação dos resultados.

Na tentativa de melhorar a taxa de predição do experimento D-PET-CSF-MRI-G, foi utilizado o método *Stacked Generalization*. Para o experimento D-PET-CSF-MRI-G, são utilizados no primeiro nível os algoritmos RLog, RF, MLP, SVM e *Tree*. Além disso, os biomarcadores que entram no sistema são repassados para todos os classificadores do primeiro nível. Para o classificador do segundo nível, as previsões geradas pelos modelos da camada inferior são utilizadas como entrada. A saída gerada por este classificador é a resposta final do sistema.

4.3 Two-step classifier

Após a definição dos experimentos realizados para o diagnóstico da DA, este estudo foi expandido para a identificação de pacientes em risco. Neste sentido, é proposto o *Two-Step Classifier* (TSC). Esse método tem como objetivo informar se o paciente é definido como um paciente estável (que provavelmente não desenvolverá déficit cognitivo) ou um paciente de risco (alta probabilidade de desenvolver a DA).

O TSC realiza a classificação em dois passos. No *first-step*, ocorre a identificação de pacientes em risco. Já no *second-step*, é realizado o diagnóstico do estado cognitivo dos indivíduos que foram classificados como em estado de risco pelo *first-step*. Mesmo que apenas pacientes em risco passem para o *second-step*, o classificador faz a classificação considerando os três valores possíveis: *NL*, *MCI* e *Dementia*. Como os classificadores sempre possuem um erro de predição, pacientes da classe *NL* podem ter sido incorretamente classificados como em risco. Com isso, no *second-step*, há a possibilidade destas classificações incorretas serem ajustadas, aumentando a taxa de acerto do sistema como um todo.

A Figura 23 apresenta uma arquitetura genérica para o TSC, assinalando as entradas e saídas do sistema e de cada etapa do classificador.

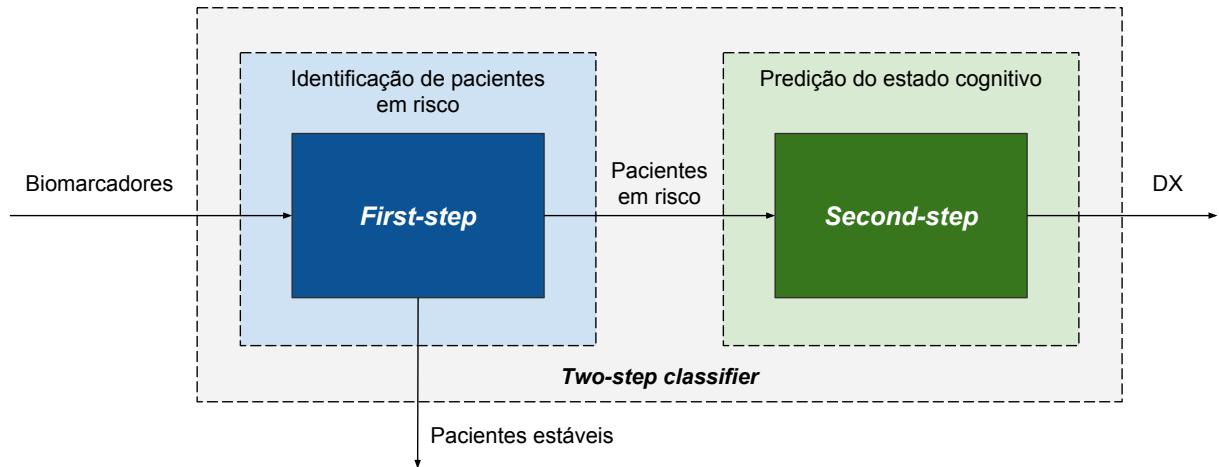


Figura 23: Arquitetura genérica do *two-step classifier*

4.3.1 Conjunto de Dados

No *first-step*, a classe *NL* é separada das classes *MCI* e *Dementia*. Ou seja, os pacientes em risco são aqueles pertencentes as classes *MCI* ou *Dementia*. Para isso, foi criado um novo atributo chamado *IS_NL*. O atributo *IS_NL* representa se o paciente é normal (*NL*), ou se ele possui algum déficit cognitivo (*NOT_NL*, pacientes das classes *MCI* ou *Dementia*).

Os dados utilizados para o *first-step* foram extraídos da base de dados ADNI, sendo 300 *NL*, 150 *MCI* e 150 *Dementia*, totalizando 600 exemplos. Desta forma, temos

300 pacientes com $IS_NL = NL$ e 300 pacientes com $IS_NL = NOT_NL$. Já para o *second-step*, foram utilizados 900 exemplos da base de dados ADNI, com 300 instâncias para cada uma das classes (*NL*, *MCI* e *Dementia*). Além disso, os dados foram normalizados utilizando o método *Z-transformation*.

Assim como na predição do diagnóstico da DA, diferentes conjuntos de biomarcadores foram avaliados para o *first-step* do TSC. Estes conjuntos podem ser vistos na Tabela 9.

Tabela 9: Conjuntos de biomarcadores explorados para o *first-step* do *two-step classifier*

Categoría	TSC-D-PET-CSF-MRI	TSC-D-PET-CSF	TSC-D-CSF-MRI	TSC-D-MRI
Dados Demográficos	✓	✓	✓	✓
PET	✓	✓		
CSF	✓	✓	✓	
MRI	✓		✓	✓

No experimento TSC-D-PET-CSF-MRI, são avaliados os biomarcadores demográficos, PET, CSF e MRI. Já no TSC-D-PET-CSF, em relação ao experimento anterior, são removidos os biomarcadores de MRI. A terceira análise (TSC-D-CSF-MRI) mantém dados demográficos, CSF e MRI. Finalmente, o TSC-D-MRI utiliza apenas dados demográficos e MRI.

Dentre as categorias de biomarcadores, os dados demográficos são os mais fáceis de serem obtidos, uma vez que todos podem ser extraídos apenas das informações dos pacientes. Por outro lado, nos exames de PET e CSF são utilizados radiofármacos para obtenção dos biomarcadores, tornando estes exames invasivos. Já o MRI, um dos exames mais simples de serem feitos, necessita apenas de um equipamento de ressonância magnética, que pode ser encontrado próximo de praticamente todas as cidades do mundo. Assim quanto menos exames forem necessários e quanto mais simples estes forem, mais acessível se torna o método proposto.

No entanto, para o *second-step*, em busca de um melhor acerto na predição, e considerando que pacientes em risco passam para esta etapa, todos os biomarcadores de dados demográficos, PET, CSF e MRI foram utilizados. Além disso, regiões mais específicas dos biomarcadores AV45, FDG e MRI foram inseridas no sistema. Neste ponto, cabe lembrar que nenhum exame neuropsicológico está sendo utilizado no TSC.

4.3.2 Classificadores

O TSC foi modelado na ferramenta RapidMiner Studio e foi avaliado para os algoritmos GBT, RLog, MLG, KNN, RF, e MLP. Aqui é utilizado o módulo AutoMLP do

RapidMiner Studio, que é um algoritmo simples que ajusta a taxa de aprendizado e o tamanho das redes neurais durante o treinamento de forma automatizada. Para isso, o AutoMLP combina ideias de algoritmos genéticos e otimização estocástica.

4.3.3 Avaliação dos Resultados

Para treinamento e teste dos modelos, foi utilizado o método de *cross-validation* com 10-*fold* sobre 80% do conjunto de exemplos. Para teste, foram utilizados os 20% restantes do conjunto de dados.

5 RESULTADOS

Nesta seção serão apresentados os resultados obtidos para os dois estudos realizados neste trabalho, a classificação da DA e a identificação de pacientes em estado de risco. Os resultados serão separados por experimento realizado e discutidos individualmente.

Além disso, tanto na ferramenta Orange3 quanto no RapidMiner Studio, para cada algoritmo testado, os parâmetros foram ajustados para que uma melhor predição pudesse ser atingida. Para cada algoritmo de cada solução, o ajuste dos parâmetros (e.g. a profundidade das árvores de decisão, o valor de k para o algoritmo kNN e outros) foi realizado de forma manual, buscando encontrar o valor que entregasse melhores resultados. Portanto, ao descrever os resultados dos experimentos, serão informados os valores dos parâmetros que levaram às melhores taxas de acerto.

5.1 Classificação da Doença de Alzheimer

Para a classificação da DA (ADC - *Alzheimer's Disease Classification*), lembramos que a ferramenta Orange3 foi utilizada para modelagem dos experimentos. Além disso, foi utilizado o método de *cross-validation* de *10-fold* para teste e validação. Os dados de entrada foram extraído da base de dados ADNI, sendo 300 pacientes *NL*, 300 com *MCI* e 300 com *Dementia*, num total de 900 instâncias. Para cada experimento, são apresentadas as métricas AUC, acurácia, F1, precisão e sensibilidade.

Para os resultados do experimento D-PET-CSF-N, o algoritmo RF foi modelado com 15 árvores, com profundidade máxima de cinco níveis em cada uma delas e cinco atributos considerados para cada *split*. A SVM foi modelada com *kernel* linear, $C = 0,7$ e $\epsilon = 0,6$. A RLog utilizou a regularização Lasso (L1) com $C = 3$. A Tree foi definida para ter no mínimo cinco instâncias em cada folha e profundidade máxima de seis níveis. A MLP foi modelada com 25 neurônios na camada oculta, função de ativação identidade – função que dá como imagem de cada elemento o próprio elemento – e *solver* Adam. Por fim, o kNN utilizou a métrica Manhattan com $k = 21$.

A Tabela 10 apresenta os resultados para o experimento D-PET-CSF-N. Analis-

sando estes resultados, percebe-se que a acurácia atinge valores elevados, chegando a 92,4% para o método RF e AUC de 0,964. Com exceção do kNN, que teve uma acurácia de 0,666 e AUC de 0,705, todos os outros algoritmos apresentaram resultados satisfatórios, sendo que, além do RF, a SVM, RLog e Tree atingiram uma acurácia superior a 90% e AUC acima de 0,940.

No entanto, estes resultados são esperados, uma vez que os testes neuropsicológicos foram utilizados como atributos para classificação. Não obstante, os biomarcadores CDRSB e ADAS11 são determinantes clínicos. Ou seja, os valores destes dois biomarcadores são utilizados pelos profissionais da área para dar o diagnóstico da DA. Portanto, a precisão deste experimento só não chega a 100% devido a possíveis erros existentes na base de dados. Este fato pode ser constatado ao observar o ganho de informação de cada biomarcador, presente na Tabela 11.

Tabela 10: Resultados para o experimento D-PET-CSF-N

Método	AUC	Acurácia	F1	Precisão	Sensibilidade
RF	0,964	0,924	0,891	0,858	0,927
SVM	0,970	0,913	0,875	0,876	0,873
RLog	0,949	0,911	0,861	0,899	0,827
Tree	0,940	0,906	0,860	0,848	0,873
MLP	0,955	0,892	0,829	0,88	0,783
kNN	0,705	0,666	0,428	0,535	0,357

Como podemos ver na Tabela 11, os quatro biomarcadores que tem o maior ganho de informação são testes neuropsicológicos, sendo o CDRSB o primeiro e o ADAS 11 o terceiro. No entanto, não há nenhuma outra categoria de biomarcadores que se destaque entre as demais.

Neste ponto, foram eliminados os determinantes clínicos (CDRSB e ADAS11) para verificar qual o impacto na precisão do modelo quando estes não são utilizados. O experimento D-PET-CSF-N* faz esta avaliação.

Os parâmetros dos algoritmos para o experimento D-PET-CSF-N* são descritos a seguir. O algoritmo RF foi modelado com 14 árvores, com profundidade máxima de quatro níveis em cada uma delas e três atributos considerados para cada *split*. A SVM foi modelada com *kernel* linear, $C = 0,7$ e $\epsilon = 0,6$. A RLog utilizou a regularização Lasso (L1) com $C = 0,9$. A Tree foi definida para ter no mínimo três instâncias em cada folha e profundidade máxima de 10 níveis. A MLP foi modelada com 20 neurônios na camada oculta, função de ativação identidade e *solver* Adam. O kNN utilizou a métrica Manhattan com $k = 10$.

A Tabela 12 apresenta os resultados do experimento D-PET-CSF-N*. Como podemos observar, a acurácia do experimento teve uma diminuição significativa ao serem

Tabela 11: Ganho de Informação para os biomarcadores utilizados neste trabalho

Categoria	Biomarcador	Ganho de Informação
N	CDRSB	1,1587
N	MMSE	0,6567
N	ADAS11	0,5319
N	RAVLT	0,3450
MRI	Hippocampus	0,2003
PET	FDG	0,1930
MRI	Entorhinal	0,1811
CSF	ABETA	0,1185
CSF	TAU	0,1100
MRI	MidTemp	0,1072
PET	AV45	0,0934
G	APOE-e4	0,0795
MRI	Fusiform	0,0779
CSF	PTAU	0,0659
MRI	Ventricles	0,0588
MRI	WholeBrain	0,0489
D	PTEDUCAT	0,0222
D	AGE	0,0182
MRI	ICV	0,0098
D	PTGENDER	0,0078

removidos os determinantes clínicos, atingindo um valor máximo de 72,7% para o RF e AUC de 0,767. O kNN foi novamente o último colocado neste experimento. Os algoritmos SVN, MLP e RLog tiveram uma acurácia superior a 70% com AUC maior que 0,726.

Tabela 12: Resultados para o experimento D-PET-CSF-N*

Método	AUC	Acurácia	F1	Precisão	Sensibilidade
RF	0,767	0,727	0,500	0,658	0,403
SVM	0,773	0,712	0,542	0,588	0,503
MLP	0,764	0,707	0,512	0,589	0,453
RLog	0,726	0,704	0,463	0,601	0,377
Tree	0,666	0,673	0,480	0,524	0,443
kNN	0,625	0,62	0,307	0,463	0,23

Como já mencionado anteriormente, testes neuropsicológicos têm um ônus de serem utilizados. Portanto, o experimento D-PET-CSF avalia a não utilização destes

testes para a modelagem dos preditores.

Para o experimento D-PET-CSF, os seguintes parâmetros foram definidos. O algoritmo RF foi modelado com seis árvores, com profundidade máxima de cinco níveis em cada uma delas e cinco atributos considerados para cada *split*. A SVM foi modelada com *kernel* linear, $C = 0,7$ e $\epsilon = 0,6$. A RLog utilizou a regularização Lasso (L1) com $C = 0,9$. A Tree foi definida para ter no mínimo três instâncias em cada folha e profundidade máxima de 10 níveis. A MLP foi modelada com 20 neurônios na camada oculta, função de ativação identidade e *solver* Adam. O kNN utilizou a métrica Manhattan com $k = 23$.

Na Tabela 13, temos os resultados para o experimento D-PET-CSF. Como esperado, a acurácia atingida foi menor do que quando utilizados testes neuropsicológicos. Tanto a RLog quanto a SVM e RF tiveram uma acurácia de 54,7%. No entanto, a RLog e a SVM tiveram um AUC muito próximo (0,607 e 0,607, respectivamente). Já o RF, teve um AUC de 0,587.

Tabela 13: Resultados para o experimento D-PET-CSF

Método	AUC	Acurácia	F1	Precisão	Sensibilidade
RLog	0,607	0,547	0,355	0,429	0,303
SVM	0,606	0,547	0,383	0,423	0,350
RF	0,587	0,547	0,360	0,429	0,310
MLP	0,598	0,541	0,370	0,406	0,340
kNN	0,499	0,486	0,260	0,341	0,210
Tree	0,540	0,474	0,298	0,391	0,240

Para o experimento seguinte, biomarcadores de imagens MRI e genéticos foram adicionados aos dados demográficos, PET e CSF. A Tabela 14 apresenta os resultados para o experimento D-PET-CSF-MRI-G. Aqui, o algoritmo RF foi modelado com 12 árvores, com profundidade máxima de cinco níveis em cada uma delas e quatro atributos considerados para cada *split*. A SVM foi modelada com *kernel* linear, $C = 0,7$ e $\epsilon = 0,6$. A RLog utilizou a regularização Lasso (L1) com $C = 5$. A Tree foi definida para ter no mínimo três instâncias em cada folha e profundidade máxima de 10 níveis. A MLP foi modelada com 20 neurônios na camada oculta, função de ativação identidade e *solver* Adam. O kNN utilizou a métrica Manhattan com $k = 22$.

Ao adicionar estes dois conjuntos de biomarcadores, percebemos uma melhora na predição ao comparar com o experimento D-PET-CSF. De todos os métodos avaliados, a RLog, RF e MLP atingiram uma acurácia superior a 60%, com AUC acima de 0,623.

Para melhor analisar os resultados de classificação do experimento D-PET-CSF-MRI-G, foi extraída a matriz de confusão para o algoritmo RLog, uma vez que ele apresentou uma melhor combinação de Acurácia e AUC. A Tabela 15 mostra esta ma-

Tabela 14: Resultados para o experimento D-PET-CSF-MRI-G

Método	AUC	Acurácia	F1	Precisão	Sensibilidade
RLog	0,630	0,608	0,385	0,49	0,317
RF	0,626	0,601	0,375	0,474	0,310
MLP	0,641	0,600	0,418	0,465	0,380
SVM	0,644	0,597	0,386	0,459	0,333
Tree	0,541	0,519	0,316	0,438	0,247
kNN	0,584	0,468	0,369	0,407	0,337

triz.

Tabela 15: Matriz de confusão da Regressão Logística para o experimento D-PET-CSF-MRI-G

		Classe predita			
		NL	MCI	Dementia	Sensibilidade
Classe atual	NL	223	106	30	0,621
	MCI	58	95	41	0,490
	Dementia	19	99	229	0,660
Precisão		0,743	0,317	0,763	

Ao observar a matriz de confusão, vemos que a classe *NL* teve 223 instâncias preditas corretamente, o que equivale a 74,3% de acerto. Já para a classe *Dementia*, essa quantia foi de 229, correspondendo a 76,3%. No entanto, a classe *MCI* teve apenas 95 exemplos corretamente classificados, representando um acerto de apenas 31,67%. Este resultado, no entanto, não é surpreendente, uma vez que pacientes da classe *MCI* estão em uma zona nebulosa (não são nem cognitivamente normais, nem apresentam grandes perdas de funções cognitivas), e é exatamente neste ponto que há uma grande dificuldade para classificação.

Levando isto em consideração, foi realizada uma análise dos resultados de classificação dos preditores do experimento D-PET-CSF-MRI-G a fim de entender melhor os erros de classificação. Para tanto, foram selecionadas as instâncias incorretamente classificadas de cada preditor e estas foram dispostas em diagrama de Venn, que é apresentado na Figura 24. Como podemos observar no diagrama, de todas as instâncias, apenas 176 delas (19,5% de todo conjunto de dados) foram classificadas incorretamente por todos os algoritmos.

Utilizando desta constatação, foi aplicado o *Stacked Generalization* sobre os preditores gerados para o experimento D-PET-CSF-MRI-G, com exceção do kNN, que apresentou o pior desempenho de predição. Este experimento se chama D-PET-CSF-MRI-G-SG. Na camada inferior do D-PET-CSF-MRI-G-SG foram usados os algoritmos

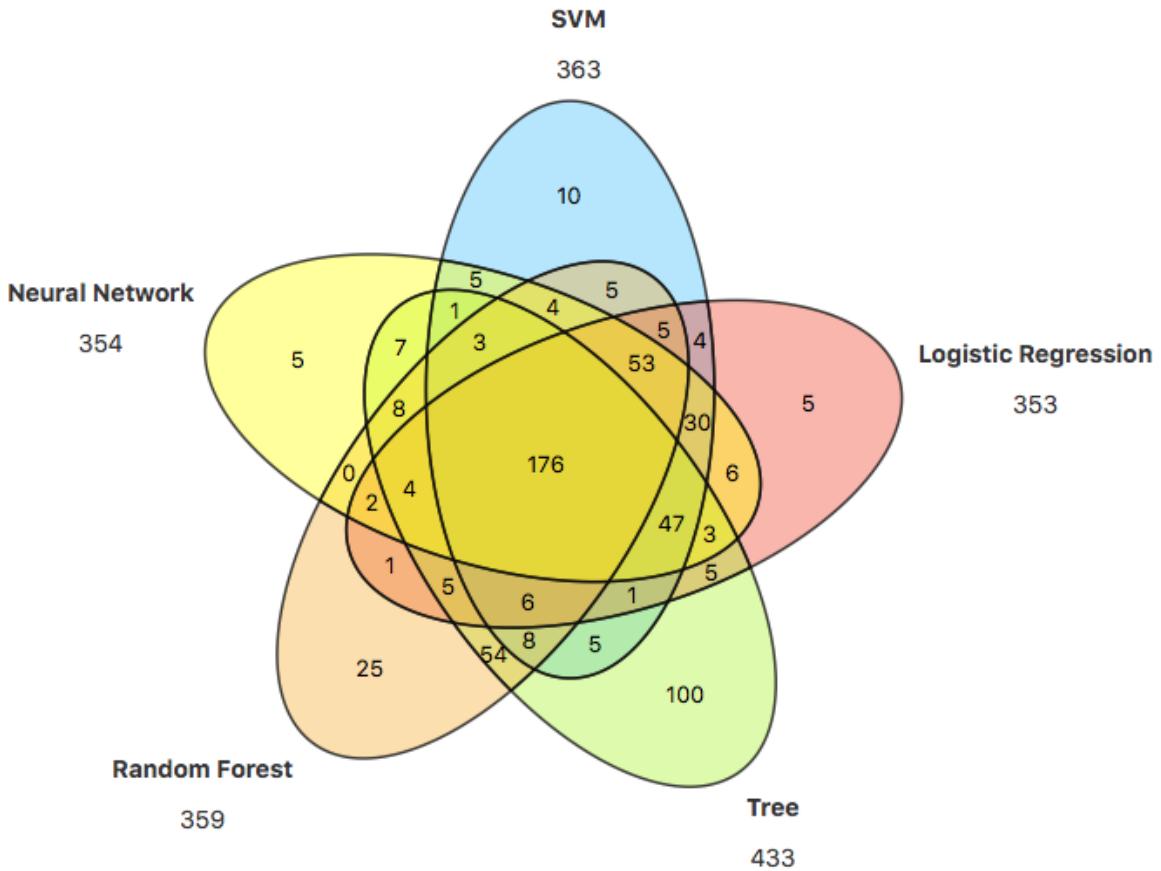


Figura 24: Diagrama de Venn das instâncias classificadas incorretamente para o experimento D-PET-CSF-MRI-G

RLog, RF, MLP, SVM e *Tree*. Para a camada superior, foram avaliados estes mesmos cinco algoritmos.

Os parâmetros dos algoritmos analisados na camada superior do D-PET-CSF-MRI-G-SG são descritos a seguir. O algoritmo RF foi modelado com 23 árvores, com profundidade máxima de cinco níveis em cada uma delas e quatro atributos considerados para cada *split*. A SVM foi modelada com *kernel* linear, $C = 0,7$ e $\epsilon = 0,6$. A RLog utilizou a regularização Lasso (L1) com $C = 5$. A *Tree* foi definida para ter no mínimo duas instâncias em cada folha e profundidade máxima de 10 níveis. A MLP foi modelada com 20 neurônios na camada oculta, função de ativação identidade e *solver* Adam.

Os resultados obtidos pelo D-PET-CSF-MRI-G-SG são mostrados na Tabela 16.

É notável, ao observar os resultados obtidos, que houve uma melhora na predição, ao comparar com o experimento D-PET-CSF-MRI-G. Enquanto no experimento D-PET-CSF-MRI-G, o melhor preditor atingia uma acurácia de 60,8% com AUC de 0,63, com o D-PET-CSF-MRI-G-SG foi possível atingir uma acurácia de 76,2% com AUC de 0,79 para a MLP. Além disso, todos os métodos avaliados tiveram resultados

Tabela 16: Resultados do para o experimento D-PET-CSF-MRI-G-SG

Método	AUC	Acurácia	F1	Precisão	Sensibilidade
MLP	0,790	0,762	0,683	0,742	0,633
RLog	0,776	0,759	0,679	0,735	0,63
RF	0,785	0,757	0,670	0,725	0,623
Tree	0,770	0,751	0,663	0,717	0,617
SVM	0,755	0,744	0,665	0,705	0,630

semelhantes, com acurácia entre 74,4% e 76,2% e AUC entre 0,755 e 0,79.

Desta forma, com o uso do *Stacked Generalization*, foi possível atingir resultados superiores do que o experimento D-PET-CSF-N*, onde os testes neuropsicológicos MMSE e RAVLT eram utilizados. Portanto, o uso do *Stacked Generalization* no conjunto de dados do experimento D-PET-CSF-MRI-G se mostra eficiente, uma vez que melhora significativamente a taxa de acerto da classe MCI, como podemos observar na matriz de confusão apresentada na Tabela 17.

Tabela 17: Matriz de confusão da MLP para o experimento D-PET-CSF-MRI-G com *Stacked Generalization*

		Classe predita			
		NL	MCI	Dementia	Sensibilidade
Classe atual	NL	238	45	16	0,796
	MCI	39	189	33	0,724
	Dementia	23	66	251	0,738
Precisão		0,793	0,630	0,837	

Além de mais instâncias serem classificadas corretamente para as classes *NL* e *Dementia*, a classe *MCI* teve um aumento de quase o dobro na taxa de acerto quando comparado ao não uso do *Stacked Generalization*. Neste caso, a taxa de acerto foi de 63%, comparados aos 31,7% do experimento D-PET-CSF-MRI-G. Já a classe *NL*, subiu de 74,3% para 79,3%. A classe *Dementia*, de 76,3%, foi para 83,7%.

Por fim, a Tabela 18 apresenta um resumo dos resultados dos experimentos D-PET-CSF-N (1), D-PET-CSF-N* (2), D-PET-CSF (3), D-PET-CSF-MRI-G (4) e D-PET-CSF-MRI-G-SG (5), mostrando os biomarcadores utilizados em cada estudo, bem como o algoritmo que gerou as melhores previsões, juntamente com a acurácia e AUC.

5.2 Two-step classifier

Nesta seção, serão apresentados os resultados obtidos pelo TSC. Para geração destes resultados, foi utilizada a ferramenta RapidMiner Studio. Além disso, foi utili-

Tabela 18: Resumo dos resultados obtidos para a classificação da DA

Biomarcadores	(1)	(2)	(3)	(4)	(5)
D	✓	✓	✓	✓	✓
PET	✓	✓	✓	✓	✓
CSF	✓	✓	✓	✓	✓
N	✓	MMSE e RAVLT			
MRI				✓	✓
G				✓	✓
Método	RF	RF	RLog	RLog	MLP
Acurácia	0,924	0,727	0,547	0,608	0,762
AUC	0,964	0,767	0,607	0,630	0,790

zado o método de *cross-validation* com 10-fold para treinamento e validação. A base de dados ADNI foi utilizada para extração dos dados para treinamento e teste dos algoritmos avaliados. Dentre estes dados, temos 300 pacientes *NL*, 150 com *MCI* e 150 com *Dementia*, num total de 600 indivíduos. Lembramos, que para o *first-step*, deseja-se prever a classe *IS_NL*, portanto, os pacientes *NL* recebem *IS_NL = NL* e pacientes *MCI* ou *Dementia* são classificados como *IS_NL = NOT_NL*. Para cada experimento do *first-step*, são apresentadas as métricas AUC, acurácia, F1, precisão e sensibilidade. Para as análises do *second-step*, é apresentada a matriz de confusão para cada experimento.

Neste ponto, é importante ressaltar que a nova implementação do RapidMiner Studio para os algoritmos GBT e RLog tiveram uma melhora significativa de desempenho. Este aumento no desempenho pode, em parte, ser atribuído ao fato de que esses algoritmos se ajustam automaticamente. Eles são projetados para encontrar as melhores configurações de parâmetros para otimizar a acurácia da predição. Isso não só proporciona uma melhor precisão; mas também reduz o esforço necessário realizar os ajustes desses algoritmos.

Além disso, o algoritmo aqui descrito como MLP é o operador chamado AutoMLP do RapidMiner Studio. O AutoMLP é uma implementação algorítmica simples para ajuste automático da taxa de aprendizagem e tamanho das redes neurais na fase de treinamento. O algoritmo combina ideias de algoritmos genéticos e otimização estocástica. Ele mantém um pequeno conjunto de redes que são treinadas em paralelo com diferentes taxas de aprendizado e diferentes configurações de camadas ocultas. Após um número pequeno e fixo de épocas, a taxa de erro é determinada em um conjunto de validação e as redes que tiveram as piores performances são substituídas por cópias das melhores redes, modificadas para ter diferentes taxas de aprendizado e configuração as camadas ocultas. As configurações das camadas ocultas e as ta-

xas de aprendizagem são definidas de acordo com as distribuições de probabilidade derivadas de taxas e configurações bem-sucedidas (BREUEL; SHAFAIT, 2010).

Para melhor organização e discussão dos resultados, esta seção será dividida em três sub-seções: *First-step*, *Second-step* e Identificação de Pacientes em Fase Pré-Clínica. Nessa última, será abordada a fase pré-clínica da DA.

5.2.1 *First-step*

Como mencionado anteriormente, para o *first-step* do método TSC, foram avaliados os algoritmos GBT, MLP, RLog e kNN. Os parâmetros destes algoritmos, para o experimento TSC-D-PET-CSF-MRI, são descritos a seguir. O GBT foi modelado com 50 árvores e profundidade máxima de cinco níveis. A RLog utilizou o solver IRLSM e o kNN foi configurado com $k = 25$.

A Tabela 19 apresenta os resultados para o experimento TSC-D-PET-CSF-MRI. Nesse experimento, são utilizados todos os exames descritos na seção 4.1, com exceção dos testes neuropsicológicos. Como comentado na seção anterior, testes neuropsicológicos não são desejados neste trabalho. Portanto, para o TSC, nenhum experimento utilizou-se desses exames.

Tabela 19: Resultados para o experimento TSC-D-PET-CSF-MRI

Método	AUC	CA	F1	Precisão	Sensibilidade
GBT	0,868	0,800	0,797	0,808	0,788
MLP	0,862	0,794	0,783	0,825	0,746
RLog	0,845	0,779	0,772	0,799	0,746
kNN	0,632	0,602	0,575	0,617	0,538

A partir da Tabela 19, observamos que o algoritmo GBT atingiu a melhor acurácia (80%) dentre os algoritmos avaliados, com AUC de 0,868. No entanto, a MLP obteve resultados próximos, com uma acurácia de 79,4% e AUC 0,862. O kNN apresentou os piores resultados, para todas as métricas avaliadas.

O TSC-D-PET-CSF analisa a qualidade na predição quando são removidos os exames de MRI do experimento TSC-D-PET-CSF-MRI. Os resultados obtidos pelo experimento TSC-D-PET-CSF são mostrados na Tabela 20.

Para a geração desses resultados, o GBT foi modelado com 40 árvores e profundidade máxima de quatro níveis. A RLog utilizou o solver IRLSM e o kNN foi configurado com $k = 15$.

Ao comparar com os resultados anteriores, percebemos que os resultados de predição tiveram uma queda de desempenho. No entanto, ao relembrar a Tabela 11, que apresenta o ganho de informação para cada biomarcador, se excluirmos a categoria de testes neuropsicológicos, dois dos três biomarcadores que trazem um melhor

Tabela 20: Resultados do para o *first-step* do experimento TSC-D-PET-CSF

Método	AUC	CA	F1	Precisão	Sensibilidade
GBT	0,816	0,731	0,731	0,732	0,729
RLog	0,781	0,710	0,702	0,722	0,683
MLP	0,791	0,704	0,674	0,750	0,613
kNN	0,708	0,671	0,650	0,693	0,613

ganho de informação são da categoria MRI (*Hippocampus* e *Entorhinal*).

No experimento TSC-D-PET-CSF, o GBT e o RLog apresentaram os melhores resultados, com acurácia de 73,1% e 71,0% e AUC de 0,816 e 0,781, respectivamente. O kNN, novamente, apresentou o pior desempenho, com acurácia de 67,1%.

O próximo experimento, TSC-D-CSF-MRI, pudemos avaliar a não utilização dos biomarcadores PET (FDG e AV45). Neste experimento, no entanto, são utilizados os biomarcadores extraídos de MRI. Os resultados obtidos são mostrados na Tabela 21.

Os parâmetros dos algoritmo para o experimento TSC-D-CSF-MRI são os seguintes. O GBT foi modelado com 30 árvores e profundidade máxima de seis níveis. A RLog utilizou o *solver* IRLSM e o kNN foi configurado com $k = 21$.

Tabela 21: Resultados do para *first-step* do experimento TSC-D-CSF-MRI

Método	AUC	CA	F1	Precisão	Sensibilidade
GBT	0,854	0,790	0,796	0,773	0,821
MLP	0,842	0,785	0,773	0,822	0,729
RLog	0,827	0,748	0,741	0,762	0,721
kNN	0,626	0,583	0,559	0,593	0,529

Como podemos notar, os resultados, apesar de mais baixos, quando comparados ao experimento TSC-D-PET-CSF-MRI, são similares. O algoritmo GBT teve uma acurácia de 79% com AUC de 0,854, contra os 80% do TSC-D-PET-CSF-MRI. Contudo, neste experimento, não são utilizados exames PET. O MLP teve uma acurácia de 78,5% com AUC de 0,842. Por último, temos o kNN, com acurácia de 58,3% e AUC 0,626.

Todavia, para os três experimentos anteriores, são utilizados os biomarcadores PET e/ou CSV. Para que esses biomarcadores possam ser obtidos, são utilizados radiofármacos (agentes radioativos injetados na corrente sanguínea do paciente), tornando estes exames invasivos. Além disso, poucas clínicas radiológicas realizam esses exames. Portanto, o experimento TSC-D-MRI remove essas duas classes de biomarcadores, deixando apenas dados demográficos e MRI. A Tabela 22 apresenta os resultados obtidos pelo TSC-D-MRI.

O GBT do TSC-D-MRI foi modelado com 30 árvores e profundidade máxima de cinco níveis. A RLog utilizou o solver IRLSM e o kNN foi configurado com $k = 25$.

Tabela 22: Resultados do para *first-step* do experimento TSC-D-MRI

Método	AUC	CA	F1	Precisão	Sensibilidade
GBT	0,847	0,781	0,780	0,785	0,775
MLP	0,813	0,748	0,731	0,785	0,683
RLog	0,816	0,731	0,723	0,747	0,700
kNN	0,633	0,602	0,575	0,617	0,538

Como podemos notar, a acurácia do TSC-D-MRI atingiu 78,1% para o algoritmo GBT, com AUC de 0,847. Para o algoritmo MLP, temos uma acurácia de 74,8% e AUC 0,813. O último classificado foi o kNN, com acurácia de 60,2% e AUC 0,633.

Dentre os experimentos realizados para o *first-step* do TSC, o TSC-D-MRI é o mais desejado, uma vez que utiliza apenas dados demográficos e biomarcadores extraídos de imagem MRI. Além disso, atinge resultados próximos do experimento que utiliza todas as classe de biomarcadores, o TSC-D-PET-CSF-MRI. Ao comparar, o TSC-D-MRI tem uma acurácia de 78,1%, comparada com os 80% do TSC-D-PET-CSF-MRI.

Para melhor entender a classificação do TSC-D-MRI para o *first-step* do TSC, foi gerada a matriz de confusão para o algoritmo GBT.

Tabela 23: Matriz de confusão do GBT para o experimento TSC-D-MRI

		Classe predita		Sensibilidade
		NL	NOT_NL	
Classe atual	NL	189	51	0,788
	NOT_NL	54	186	0,775
Precisão		0,778	0,785	

Ao analisar a matriz de confusão, percebemos que 51 de 240 pacientes foram classificados como *NOT_NL*, mas deveriam ter sido classificados como *NL*. Ou seja, esses pacientes foram sinalizados como em risco. Por outro lado, 54 indivíduos de 240 foram incorretamente classificados como *NL*. Neste ponto, os pacientes classificados como *NOT_NL* passam para o *second-step*, a fim de dar o diagnóstico final da progressão da DA. Fora isso, foi realizado um estudo sobre os falsos negativos (preditos como *NOT_NL* mas que são *NL*), com a finalidade de verificar se esses indivíduos estão na fase pré-clínica da DA. Essa análise é apresentada na seção 5.2.3.

Neste ponto, aplicamos a validação dos resultados sobre os 20% dos dados que foram separados do conjunto de treinamento. A Tabela 24 apresenta a matriz de confusão para o experimento TSC-D-MRI utilizando o algoritmo GBT.

Tabela 24: Matriz de confusão do conjunto de validação para o experimento TSC-D-MRI utilizando GBT

		Classe predita		Sensibilidade
		NL	NOT_NL	
Classe atual	NL	49	11	0,817
	NOT_NL	21	39	0,650
		Precisão	0,700	0,780

Para o conjunto de validação, foi possível alcançar uma acurácia de 73,3% para o modelo como um todo, com um AUC de 0,777, mostrando que o modelo avaliado atinge resultados significativos, mesmo utilizando apenas dados demográficos e imagem de MRI.

Por fim, a Tabela 25 resume os resultados obtidos para os experimentos TSC-D-PET-CSF-MRI (1), TSC-D-PET-CSF (2), TSC-D-CSF-MRI (3) e TSC-D-MRI (4), mostrando as classes de biomarcadores utilizadas em cada estudo, juntamente com o algoritmo que gerou as melhores previsões. É interessante notar, que em todos os experimentos, o GBT foi o que apresentou os melhores resultados de classificação.

Tabela 25: Resumo dos resultados obtidos para o *first-step* do TSC

Biomarcadores	(1)	(2)	(3)	(4)
D	✓	✓	✓	✓
PET	✓	✓		
CSF	✓	✓	✓	
MRI	✓		✓	✓
Método	GBT	GBT	GBT	GBT
Acurácia	0,800	0,731	0,790	0,781
AUC	0,868	0,816	0,854	0,847

5.2.2 Second-step

Nesta seção, serão apresentados os resultados para o *second-step* do TSC. Como o objetivo do *second-step* é dar o diagnóstico do estágio da DA, a saída esperada aqui é o atributo *DX*, que pode assumir os valores *NL*, *MCI* ou *Dementia*. Para isso, o *second-step* foi treinado com uma porção de 80% de um conjunto de dados balanceado de 900 exemplos (300 para cada classe). O treinamento e validação foi realizado com o método de *cross-validation* com 10-fold. Para teste, foram utilizados os 20% restantes do conjunto de dados.

Para o *second-step*, como os pacientes que passam para essa etapa estão em

risco de desenvolver DA. Com a finalidade de dar o melhor diagnóstico possível, são utilizados todas as classes de biomarcadores descritas na seção 2.3, com exceção dos testes neuropsicológicos.

Além disso, um conjunto estendido de biomarcadores, com várias regiões de interesse (ROI, do inglês *Region of Interest*) extraídas de imagem PET utilizando os radiofármacos FDG e AV45 também são utilizados.

No contexto do radiofármaco FDG, temos os ROIs *Left Angular Gyrus*, *Right Angular Gyrus*, *Bilateral Posterior Cingular*, *Left Inferior Temporal Gyrus* e *Right Inferior Temporal Gyrus*. Para cada ROI, é disponibilizado valor mínimo, máximo, média, mediana, moda e desvio padrão para o metabolismo da glicose no cérebro. Também, temos o número total de *voxels* e de *voxels* zero ou NaN para cada ROI (ADNI, 2018a).

Já para o radiofármaco AV45, são extraídas ROIs a partir de quatro grandes regiões: *Frontal*, *Anterior/Posterior Cingulate*, *Lateral Parietal* e *Lateral Temporal*. Para cada uma dessas regiões, são extraídos biomarcadores para diversas sub-regiões. Além disso, é disponibilizada a média para cinco outras regiões de referência: *Cerebellar Grey Matter*, *Whole Cerebellum*, *Brainstem/pons*, *Eroded Subcortical White Matter* e *Composite Reference Region* (ADNI, 2018b).

Deste ponto em diante, serão discutidos os resultados obtidos para o *second-step*, considerando o uso de todo esse conjunto de biomarcadores. Os algoritmos avaliados para o *second-step* foram MLG, RF, GBT e kNN. Para os algoritmos MLG, RF, GBT, foi utilizado o novo módulo do RapidMiner Studio chamado *Auto Model*, que faz a modelagem e otimização automática de parâmetros para diferentes modelos.

A Figura 25 mostra um gráfico com a acurácia para os modelos testados para o *second-step*. Aqui, o kNN foi configurado com $k = 100$, com a *flag weighted vote* ativa. A DT foi modelada com profundidade máxima de 25 níveis. A RF tem 100 árvores com profundidade máxima de sete níveis cada uma delas. Por fim, a GBT foi configurada com 40 árvores com profundidade máxima de 10 níveis. Os demais parâmetros dos algoritmos foram deixados com seu valor *default* no RapidMiner Studio.

Dentre os algoritmos testados, o kNN, DT e RF tiveram uma acurácia abaixo ou próxima de 50%. No entanto, o algoritmo MLG atingiu uma acurácia próxima a 58%. O melhor colocado neste experimento foi a GBT, com acurácia de 60,14%. Para melhor entender esses resultados, serão apresentadas as matrizes de confusão para o MLG e a GBT. A Tabela 26 apresenta a matriz de confusão para o algoritmo GBT.

Como podemos perceber na Tabela 26, a precisão de acerto da classe *MCI* é de 45,4%, com uma sensibilidade de 38,8%. Para a classe *NL*, dos 240 pacientes, 164 foram corretamente classificados. O complemento desse valor é a taxa de falsos negativos. Ou seja, 76 de 240 foram incorretamente classificados como *NL*, o que corresponde a 31,67% desta classe.

Sob outra perspectiva, juntando as classes *MCI* e *Dementia*, 83 indivíduos de

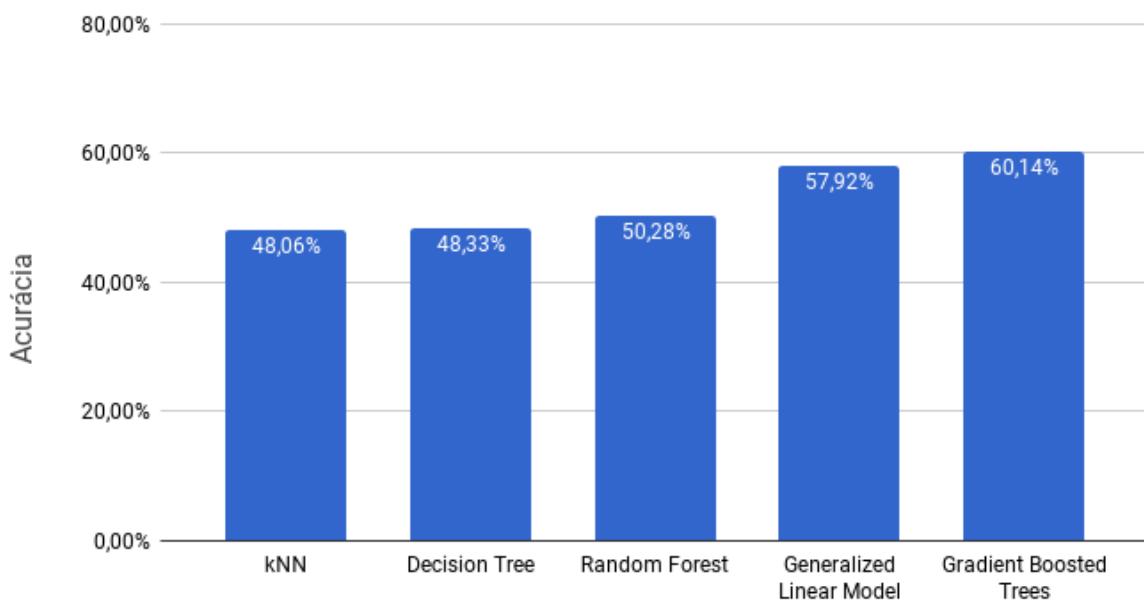


Figura 25: Acurácia obtida para os modelos avaliados no *second-step*

Tabela 26: Matriz de confusão para o *second-step* utilizando GBT

		Classe predita			
		NL	MCI	Dementia	Sensibilidade
Classe atual	NL	164	61	15	0,683
	MCI	70	93	77	0,388
	Dementia	13	41	176	0,733
Precisão		0,664	0,454	0,657	

480 foram classificados como *NL* com o uso da GBT, representando 17,29% deste conjunto. Ou seja, dos pacientes com déficit cognitivo, no *second-step*, 17,29% foram classificados como normais.

Aqui, queremos que esses dois parâmetros sejam os mais baixos possíveis, evitando que pacientes com déficit cognitivo sejam classificados como normais, e que pacientes normais sejam marcados como pertencentes às classes *MCI* ou *Dementia*.

A Tabela 27 apresenta a matriz de confusão para o algoritmo MLG.

Os resultados atingidos com o uso do MLG são próximos aos resultados da GBT. Com o uso do MLG, a sensibilidade da classe *NL* é de 70%. Ou seja, dos pacientes normais, 70% foram corretamente classificados como normais.

Essa classificação da classe *NL* no *second-step* faz com que o erro de predição da classe *NL* gerado no *first-step* seja diminuído. A taxa de falsos negativos para a classe *NL* com o MLG é de 30%. Ao unir as classes *MCI* e *Dementia*, 107 indivíduos de 480 foram classificados como *NL*, representando 22,29% de falsos negativos para

Tabela 27: Matriz de confusão para o *second-step* utilizando MLG

		Classe predita			
		NL	MCI	Dementia	Sensibilidade
Classe atual	NL	168	50	22	0,700
	MCI	91	75	74	0,313
	Dementia	16	50	174	0,725
	Precisão	0,612	0,429	0,644	

essas classes.

Portanto, entre a GBT e o MLG, ao analisar por essas duas óticas, o que traz um melhor *trade-off* entre essas duas características é a GBT.

No entanto, o *second-step* sofre do mesmo problema de classificação discutido na seção 5.1. Além disso, o maior problema está em indivíduos que são da classe *MCI* ou *Dementia* mas são classificados como *NL*. Desta forma, é altamente desejável diminuir os falsos positivos da classe *NL*. Por outro lado, deve-se levar em consideração que no *first-step*, todos esses indivíduos foram apontados pacientes em risco. Assim, o diagnóstico final do indivíduo, após o *second-step*, deve ser suportado por testes neuropsicológicos.

A Tabela 28 apresenta a matriz de confusão gerada para o algoritmo GBT sobre os 20% dos dados que foram separados para validação.

Tabela 28: Matriz de confusão do conjunto de validação aplicado no *second-step* utilizando GBT

		Classe predita			
		NL	MCI	Dementia	Sensibilidade
Classe atual	NL	44	15	1	0,733
	MCI	19	27	14	0,450
	Dementia	3	16	41	0,683
	Precisão	0,667	0,466	0,732	

Como podemos ver, a precisão e sensibilidade de cada classe ficaram muito próximas à atingida pelo conjunto de teste. Ademais, a taxa de falsos negativos para a classe *NL* é de 26,67%. E a taxa de falsos negativos para as classes *MCI* e *Dementia* em conjunto é de 18,33%.

Em adição a esses resultados, foi testado o conjunto de exemplos classificados como *NOT_NL* no *first-step* para o experimento TSC-D-MRI. Para tanto, esses 310 exemplos foram aplicados no modelo de GBT gerado para o *second-step*. A Tabela 29 mostra a matriz de confusão para esse conjunto de dados.

Aqui, notamos uma alta taxa de acerto do *second-step* utilizando GBT sobre o

Tabela 29: Matriz de confusão do *second-step* utilizando GBT sobre para as instâncias classificadas como NOT_NL no *first-step* do experimento TSC-D-MRI

		Classe predita			
		NL	MCI	Dementia	Sensibilidade
Classe atual	NL	52	26	1	0,658
	MCI	2	95	1	0,969
	Dementia	3	20	110	0,827
	Precisão	0,912	0,674	0,982	

conjunto de dados de pacientes classificados como em risco pelo *first-step*. Para o modelo como um todo, temos uma acurácia de 82,9%. Quando olhamos apenas para a classe *NL*, precisão é de 91,2%. Para a classe *MCI* esse valor é de 67,4% e para a classe *Dementia*, a precisão é de 98,2%.

Desta forma, a arquitetura final do modelo proposto para o TSC é composta pelo conjunto de dados TSC-D-MRI e o algoritmo GBT para o *first-step*. No *second-step*, temos um conjunto de biomarcadores estendido, com outra modelagem do algoritmo GBT. A Figura 26 mostra a modelagem final da arquitetura do TSC.

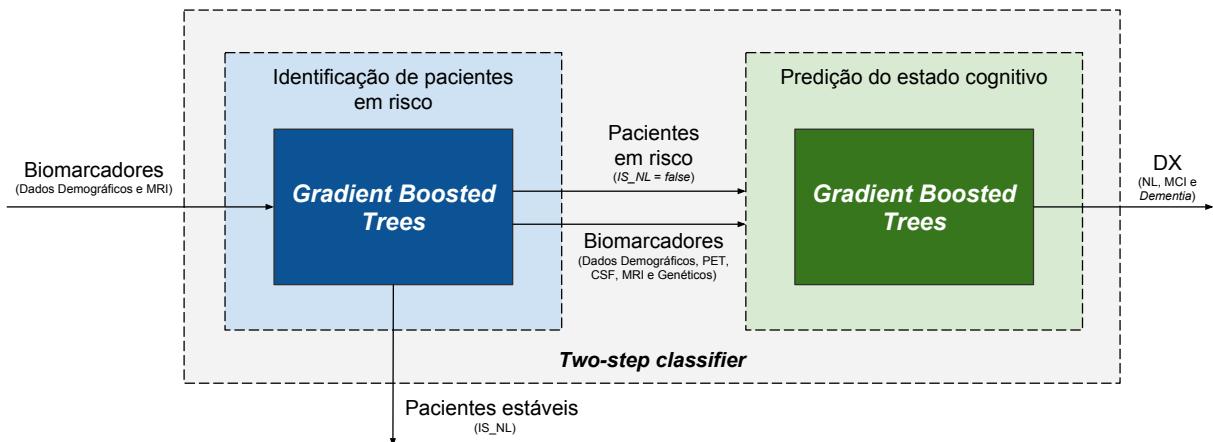


Figura 26: Arquitetura do *two-step classifier*

Após os dados passarem por todo o fluxo do TSC, podemos atualizar a matriz de confusão da Tabela 23, para que ela reflita a classificação final de pacientes em risco. A Tabela 30 apresenta a matriz de confusão atualizada.

Ao comparar a matriz de confusão da Tabela 23 com a apresentada na Tabela 30, percebemos um aumento na precisão tanto da classe *NL* quanto da classe *NOT_NL*. Já a sensibilidade da classe *NL*, teve um aumento considerável, passando de 73,7% para 91%. Contudo, a sensibilidade da classe *NOT_NL* teve uma pequena diminuição, uma vez que a taxa de falsos positivos aumentou de 69 para 74. Desta forma, o sistema como um todo atingiu uma precisão de 83,3%, ao classificar corretamente 499 pacientes de um total de 600.

Tabela 30: Matriz de confusão da classificação de pacientes em risco após o processamento do TSC

		Classe predita		
		NL	NOT_NL	Sensibilidade
Classe atual	NL	273	27	0,910
	NOT_NL	74	226	0,753
		Precisão	0,787	0,893

Ao comparar com trabalhos relacionados, lembramos que nenhum dos trabalhos encontrados na literatura fazem a classificação de pacientes em risco de desenvolver a DA da forma como foi explorada neste trabalho. Portanto, qualquer comparação seria injusta. De qualquer forma, na Tabela 31, acrescentamos os resultados dos dois métodos propostos por este trabalho (classificação da DA - ADC e o *Two-Step Classifier* - TSC).

5.2.3 Identificação de Pacientes em Fase Pré-Clínica

Esta seção tem como objetivo analisar os resultados obtidos pelo *first-step* do TSC, com o objetivo de identificar pacientes que estão na fase pré-clínica da DA. Para tanto, foram separados os pacientes classificados como *NOT_NL* mas que são, atualmente, da classe *NL*. Ou seja, os falsos negativos. Esse estudo foi realizado para os algoritmos que atingiram melhores resultados para experimentos TSC-D-CSF-PET-MRI, TSC-D-CSF-PET, TSC-D-CSF-MRI e TSC-D-MRI.

Como todos as instâncias de exames utilizadas neste trabalho são referentes a primeira visita (*VISCODE = bl*) dos pacientes, para cada um desses, foi selecionado o diagnóstico referente ao último conjunto de exames realizado. Desta forma, foi possível observar aqueles pacientes que converteram de *NL* para *MCI* ou *Dementia*. Esse valor contabilizado na variável chamada *Converters*. Ou seja, são classificados como *Converters*, aqueles pacientes que após certo tempo converteram de *NL* para *MCI* ou de *NL* para *Dementia*.

A Tabela 32 mostra a taxa de identificação de pacientes *Converters* para os experimentos mencionados acima.

Na coluna *Pred. NOT_NL* temos a quantidade de pacientes classificados como *NOT_NL* pelo método correspondente. A coluna *True NL* contabiliza a quantidade de indivíduos que são verdadeiros *NL*, isto é, que não converteram para *MCI* ou *Dementia* até o momento atual do estudo ADNI. A coluna *Converters* exibe a quantidade de pacientes que converteram para *MCI* ou *Dementia* e a coluna *Taxa de Converters* exibe a taxa de identificação dos indivíduos que converteram.

Como podemos observar, para o conjunto de dados TSC-D-CSF-PET-MRI e algoritmos GBT e RLog, a taxa de *Converters* foi de 33,33%. Ou seja, dos pacien-

tes classificados como *NOT_NL*, 33,33% converteram para os estágios de *MCI* ou *Dementia* com o passar do tempo. Já para o conjunto de dados TSC-D-MRI, a taxa de *Converters* foi de 24,66%. Para os outros conjuntos de dados e algoritmos, essa taxa ficou entre 25% e 30%.

Assim, o TSC se mostra eficiente em dois eixos: identificação de pacientes em risco e na fase pré-clínica da DA. Desta forma, o método proposto neste trabalho pode ser utilizado para auxiliar profissionais da área da saúde, colaborando com o diagnóstico e a identificação de pacientes em risco de desenvolverem a DA.

Tabela 31: Comparação com os trabalhos relacionados

Trabalho	Tamanho da Amostra	Tipo da Validação	Classificação	Precisão (%)	AUC	Follow-up
ADC	300 NL, 300 MCI, 300 DA	<i>Cross-validation 10-fold</i>	NL, MCI e DA	76,2	0,790	-
TSC	240 NL, 240 MCI, 240 DA	Conj. val. indep.	Pacientes em Risco	83,3	-	-
ZHANG et al. (2011)	52 NL, 99 MCI, 51 DA	<i>Cross-validation 10-fold</i>	NL e DA MCI e DA	93,2 76,4	- -	-
ORTIZ et al. (2016)	68 NL, 111 MCI 26 LMCI, 70 DA	<i>Cross-validation 10-fold</i>	NL e DA sMCI e DA NL e pMCI	90 84 83	0,95 0,91 0,95	-
LONG et al. (2017)	135 NL, 132 sMCI 96 pMCI, 64 DA	<i>Cross-validation 10-fold</i>	NL e DA sMCI e pMCI NL e pMCI	96,5 89 91,74	0,995 0,932 0,971	- 36 36
BEHESHTI; DEMIREL; MATSUDA (2017)	162 NL, 65 sMCI 71 pMCI, 160 DA	<i>Cross-validation 10-fold</i>	NL e DA MCI e DA	93 75	0,935 0,751	- 12~36
RODRIGUES et al. (2017)	152 NL, 414 MCI, 100 DA	<i>Cross-validation 10-fold</i>	NL, MCI e DA	94,34	-	-
MATHOTAAARACHCHI et al. (2017)	230 sMCI, 43 pMCI	Conj. teste independente	sMCI e pMCI	84	0,906	24

Tabela 32: Taxa de identificação dos *converters* para os experimentos realizados

Conjunto de Dados	Método	Pred. NOT_NL	True NL	Converters	Taxa de Converters
TSC-D-CSF-PET-MRI	GBT	72	48	24	33,33%
	RLog	54	36	18	33,33%
TSC-D-PET-CSF	GBT	77	54	23	29,87%
TSC-D-CSF-MRI	GBT	77	57	20	25,97%
	RLog	59	42	17	28,81%
TSC-D-MRI	GBT	79	60	19	24,05%

6 CONCLUSÃO

Este trabalho apresentou um estudo sobre os principais conceitos e características da DA. O estudo realizado abordou o histórico da doença, o entendimento de como a doença se comporta, juntamente com seus sintomas e estágios. As duas principais hipóteses que explicam os mecanismos que causam a doença foram explanadas. São elas: a deposição de placas de A β e o acúmulo de emaranhados neurofibrilares.

Ainda, foi feito um estudo acerca do esforço de pesquisa global ADNI, que apoia a investigação e desenvolvimento de tratamentos para a DA. Este estudo avalia biomarcadores clínicos, de imagem, genéticos e de bioespécime, através do processo de envelhecimento normal até o comprometimento cognitivo leve precoce, o comprometimento cognitivo leve tardio, a demência ou DA.

Com esse estudo, foi proposto um método para diagnóstico da DA, onde foram avaliados diferentes conjuntos de biomarcadores e algoritmos. Os algoritmos avaliados foram as *Random Forests*, *Multilayer Perceptron*, Regressão Logística, Árvores de Decisão e Máquina de Vetores de Suporte. Além disso, foi aplicado o método de *Stacked Generalization* (SG), que utiliza a saída de diversos preditores para melhorar o erro de predição. Com o uso do SG, foi possível obter uma acurácia de 76,2% com AUC de 0,790 para a classificação do estágio da DA sem a necessidade de utilizar testes neuropsicológicos.

Também, é proposto neste trabalho o *Two-Step Classifier* (TSC), método para diagnóstico da DA realizado em duas etapas. No *first-step*, é feita a identificação de pacientes em risco de desenvolver a doença. No *second-step*, é realizado o diagnóstico do estágio da DA. Além disso, com o TSC, foi feito um estudo sobre a fase pré-clínica da DA, em busca de identificar pacientes que ainda não apresentam déficit cognitivo, mas já são biomarcadores-positivos para a DA. Ou seja, já apresentam sinais que irão desenvolver a doença mesmo sem expor os sintomas iniciais. Com o TSC, foi possível identificar 33,33% de pacientes que estão na fase pré-clínica da DA. Ademais, foi possível identificar, utilizando apenas dados demográficos e imagem de MRI, com uma acurácia de 78,1% e AUC de 0,847, indivíduos que estão em risco de desenvolver a DA.

Como trabalhos futuros, o estudo será expandido para outras doenças neurodegenerativas, como a Doença de Parkinson e a Doença de Huntington. Também, serão estudados os diversos tipos de imagens de exames médicos, como MRI, PET e outros, para aplicação de técnicas de *Deep Learning* com o propósito de extrair padrões que auxiliem no diagnóstico atual e futuro de indivíduos com essas doenças.

REFERÊNCIAS

- ADEAR. **Diagram of how microtubules disintegrate with Alzheimer's disease**, 2008. Disponível em:
<http://www.nia.nih.gov/Alzheimers/Resources/MediaRoom.htm>. Acesso em: 2016-11-23.
- ADNI. **ADNI | Study Design**, 2016. Disponível em:
<http://adni.loni.usc.edu/study-design/>. Acesso em: 2016-11-30.
- ADNI. **ADNI | Data Inventory**, 2016. Disponível em:
<http://adni.loni.usc.edu/data-samples/adni-data-inventory/>. Acesso em: 2016-11-30.
- ADNI. **UC Berkeley - FDG analysis — ucberkeleyfdg**, 2018. Disponível em:
<https://adni.bitbucket.io/reference/ucberkeleyfdg.html>. Acesso em: 2018-02-06.
- ADNI. **UC Berkeley - AV45 analysis — ucberkeleyav45**, 2018. Disponível em:
<https://adni.bitbucket.io/reference/ucberkeleyav45.html>. Acesso em: 2018-02-06.
- ALTMAN, N. S. An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. **The American Statistician**, [S.I.], v.46, n.3, p.175–185, Aug. 1992.
- ALZHEIMER, A. About a peculiar disease of the cerebral cortex. **Alzheimer Disease and Associated Disorders**, [S.I.], v.1, n.1, p.3–8, 1987.
- AMADUCCI, L. A.; ROCCA, W. A.; SCHOENBERG, B. S. Origin of the distinction between Alzheimer's disease and senile dementia: how history can clarify nosology. **Neurology**, [S.I.], v.36, n.11, p.1497–1499, Nov. 1986.
- ARNÁIZ, E.; ALMKVIST, O. Neuropsychological features of mild cognitive impairment and preclinical Alzheimer's disease. **Acta Neurologica Scandinavica. Supplementum**, [S.I.], v.179, p.34–41, 2003.
- ARVESEN, E. **Automatic classification of alzheimer's disease from structural MRI**. 2015. Master's Thesis in Computer Science — Østfold University College.

ASSOCIATION, A. **Alzheimer's & Dementia Risk Factors**, 2007. Disponível em: <http://www.alz.org/alzheimers_disease_causes_risk_factors.asp>. Acesso em: 2017-04-17.

ASSOCIATION, A. **Alzheimer's & Dementia Causes, Risk Factors**, 2010.

Disponível em:

<http://www.alz.org/research/science/alzheimers_disease_causes.asp>. Acesso em: 2016-11-12.

ASSOCIATION, A. **Alzheimer's & Dementia Testing for Earlier Diagnosis**, 2010.

Disponível em:

<http://www.alz.org/research/science/earlier_alzheimers_diagnosis.asp>. Acesso em: 2016-11-24.

ASSOCIATION, A. 2017 Alzheimer's disease facts and figures. **Alzheimer's & Dementia: The Journal of the Alzheimer's Association**, [S.I.], v.13, n.4, p.325–373, Apr. 2017.

BÄCKMAN, L. et al. Multiple cognitive deficits during the transition to Alzheimer's disease. **Journal of Internal Medicine**, [S.I.], v.256, n.3, p.195–204, Sept. 2004.

BEHESHTI, I.; DEMIREL, H.; MATSUDA, H. Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm.

Computers in Biology and Medicine, [S.I.], v.83, p.109–119, Apr. 2017.

BERCHTOLD, N. C.; COTMAN, C. W. Evolution in the conceptualization of dementia and Alzheimer's disease: Greco-Roman period to the 1960s. **Neurobiology of Aging**, [S.I.], v.19, n.3, p.173–189, June 1998.

BERMEJO-PAREJA, F. et al. Incidence and subtypes of dementia in three elderly populations of central Spain. **Journal of the Neurological Sciences**, [S.I.], v.264, n.1-2, p.63–72, Jan. 2008.

BERNARD, C. et al. Time course of brain volume changes in the preclinical phase of Alzheimer's disease. **Alzheimer's & Dementia**, [S.I.], v.10, n.2, p.143–151.e1, Mar. 2014.

BHATTACHARYA, I.; BHATIA, M. P. S. SVM Classification to Distinguish Parkinson Disease Patients. In: Amrita ACM-W Celebration ON Women IN Computing IN India, 1., 2010, New York, NY, USA. **Proceedings...** ACM, 2010. p.14:1–14:6. (A2CWIC '10).

BISHOP, C. M. **Pattern recognition and machine learning**. [S.I.]: Springer, 2006.

- BOLLER, F.; FORBES, M. M. History of dementia and dementia in history: an overview. **Journal of the Neurological Sciences**, [S.I.], v.158, n.2, p.125–133, June 1998.
- BONIN-GUILLAUME, S. et al. Impact économique de la démence. **La Presse Médicale**, [S.I.], v.34, n.1, p.35–41, Jan. 2005.
- BREIMAN, L. Random Forests. **Machine Learning**, [S.I.], v.45, n.1, p.5–32, Oct 2001.
- BREUEL, T.; SHAFAIT, F. AutoMLP: Simple, Effective, Fully Automated Learning Rate and Size Adjustment. In: THE LEARNING WORKSHOP, 2010. **Anais...** Online, 2010. Extended Abstract.
- BROOKMEYER, R.; JOHNSON, E.; ZIEGLER-GRAHAM, K.; ARRIGHI, H. M. Forecasting the global burden of Alzheimer's disease. **Alzheimer's & Dementia: The Journal of the Alzheimer's Association**, [S.I.], v.3, n.3, p.186–191, July 2007.
- BURNS, A.; ILIFFE, S. Alzheimer's disease. **BMJ**, [S.I.], v.338, p.b158, Feb. 2009.
- CAMUS, V. et al. Using PET with 18F-AV-45 (florbetapir) to quantify brain amyloid load in a clinical environment. **European Journal of Nuclear Medicine and Molecular Imaging**, [S.I.], v.39, n.4, p.621–631, Apr. 2012.
- CARLESIMO, G. A.; OSCAR-BERMAN, M. Memory deficits in Alzheimer's patients: a comprehensive review. **Neuropsychology Review**, [S.I.], v.3, n.2, p.119–169, June 1992.
- CDC. **Older Persons' Health**, 2017. Disponível em:
<https://www.cdc.gov/nchs/fastats/older-american-health.htm>. Acesso em: 2017-12-12.
- CHUN, W.; JOHNSON, G. V. W. The role of tau phosphorylation and cleavage in neuronal cell death. **Frontiers in Bioscience: A Journal and Virtual Library**, [S.I.], v.12, p.733–756, Jan. 2007.
- CORTES, C.; VAPNIK, V. Support-vector networks. **Machine Learning**, [S.I.], v.20, n.3, p.273–297, Sept. 1995.
- DEMŠAR, J. et al. Orange: Data Mining Toolbox in Python. **J. Mach. Learn. Res.**, [S.I.], v.14, n.1, p.2349–2353, Jan. 2013.
- DILL, K. A.; OZKAN, S. B.; SHELL, M. S.; WEIKL, T. R. The Protein Folding Problem. **Annual Review of Biophysics**, [S.I.], v.37, n.1, p.289–316, June 2008.

DUBOIS, B. et al. Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria. **Alzheimer's & Dementia: The Journal of the Alzheimer's Association**, [S.I.], v.12, n.3, p.292–323, Mar. 2016.

EUROSTAT. **Causes of death statistics - people over 65 - Statistics Explained**, 2017. Disponível em: <http://ec.europa.eu/eurostat/statistics-explained/index.php/Causes_of_death_statistics_-_people_over_65>. Acesso em: 2017-12-16.

FÖRSTL, H.; KURZ, A. Clinical features of Alzheimer's disease. **European Archives of Psychiatry and Clinical Neuroscience**, [S.I.], v.249, n.6, p.288–290, 1999.

FRANK, E. M. Effect of Alzheimer's disease on communication function. **Journal of the South Carolina Medical Association (1975)**, [S.I.], v.90, n.9, p.417–423, Sept. 1994.

G7. The challenge of neurodegenerative diseases in an aging population.
[S.I.: s.n.], 2017.

GIL, D.; JOHNSSON, M. Diagnosing Parkinson by using artificial neural networks and support vector machines. **Global Journal of Computer Science and Technology**, [S.I.], v.9, n.4, p.63–71, 2009.

GOEDERT, M.; SPILLANTINI, M. G.; CROWTHER, R. A. Tau proteins and neurofibrillary degeneration. **Brain Pathology (Zurich, Switzerland)**, [S.I.], v.1, n.4, p.279–286, July 1991.

GOLD, D. P.; REIS, M. F.; MARKIEWICZ, D.; ANDRES, D. When home caregiving ends: a longitudinal study of outcomes for caregivers of relatives with dementia. **Journal of the American Geriatrics Society**, [S.I.], v.43, n.1, p.10–16, Jan. 1995.

GOOCH, C. L.; PRACHT, E.; BORENSTEIN, A. R. The burden of neurological disease in the United States: A summary report and call to action. **Annals of Neurology**, [S.I.], v.81, n.4, p.479–484, Apr. 2017.

GROVER, P. **Gradient Boosting from scratch**, 2017. Disponível em: <<https://medium.com/mlreview/gradient-boosting-from-scratch-1e317ae4587d>>. Acesso em: 2018-02-01.

HAMET, P.; TREMBLAY, J. Artificial intelligence in medicine. **Metabolism**, [S.I.], v.69, p.S36–S40, Apr. 2017.

HARDY, J. A.; HIGGINS, G. A. Alzheimer's disease: the amyloid cascade hypothesis. **Science (New York, N.Y.)**, [S.I.], v.256, n.5054, p.184–185, Apr. 1992.

- HARDY, J.; SELKOE, D. J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. **Science (New York, N.Y.)**, [S.I.], v.297, n.5580, p.353–356, July 2002.
- HASHIMOTO, M.; ROCKENSTEIN, E.; CREWS, L.; MASLIAH, E. Role of protein aggregation in mitochondrial dysfunction and neurodegeneration in Alzheimer's and Parkinson's diseases. **Neuromolecular Medicine**, [S.I.], v.4, n.1-2, p.21–36, 2003.
- HO, T. K. Random decision forests. In: International Conference ON Document Analysis AND Recognition, 3., 1995. **Proceedings...** [S.I.: s.n.], 1995. v.1, p.278–282 vol.1.
- HOOPER, N. M. Roles of proteolysis and lipid rafts in the processing of the amyloid precursor protein and prion protein. **Biochemical Society Transactions**, [S.I.], v.33, n.Pt 2, p.335–338, Apr. 2005.
- IQBAL, K. et al. Tau pathology in Alzheimer disease and other tauopathies. **Biochimica Et Biophysica Acta**, [S.I.], v.1739, n.2-3, p.198–210, Jan. 2005.
- JELICIC, M.; BONEBAKKER, A. E.; BONKE, B. Implicit memory performance of patients with Alzheimer's disease: a brief review. **International Psychogeriatrics**, [S.I.], v.7, n.3, p.385–392, 1995.
- JPND. **What is neurodegenerative disease?**, 2017. Disponível em: <<http://www.neurodegenerationresearch.eu/about/what/>>. Acesso em: 2017-12-16.
- JUCKER, M. **Alzheimer**: 100 years and beyond. Berlin; New York: Springer, 2006. OCLC: 262693411.
- KATZMAN, R.; TERRY, R. D.; BICK, K. L. (Ed.). **Alzheimer's Disease**: Senile Dementia and Related Disorders. New York: Raven Press, 1978.
- KENICHI, O. et al. DTI Analyses and Clinical Applications in Alzheimer's Disease. **Journal of Alzheimer's Disease**, [S.I.], n.s3, p.287–296, 2011.
- KHACHIYANTS, N.; TRINKLE, D.; SON, S. J.; KIM, K. Y. Sundown syndrome in persons with dementia: an update. **Psychiatry Investigation**, [S.I.], v.8, n.4, p.275–287, Dec. 2011.
- KLAFKI, H.-W.; STAUFENBIEL, M.; KORNHUBER, J.; WILTFANG, J. Therapeutic approaches to Alzheimer's disease. **Brain: A Journal of Neurology**, [S.I.], v.129, n.Pt 11, p.2840–2855, Nov. 2006.

- KLUNK, W. E. et al. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B: Imaging Amyloid in AD with PIB. **Annals of Neurology**, [S.I.], v.55, n.3, p.306–319, Mar. 2004.
- LAKHAN, S. E. **Alzheimer Disease**: Practice Essentials, Background, Anatomy, 2016. Disponível em: <<http://emedicine.medscape.com/article/1134817-overview>>. Acesso em: 2016-11-21.
- LANDES, A. M.; SPERRY, S. D.; STRAUSS, M. E.; GELDMACHER, D. S. Apathy in Alzheimer's disease. **Journal of the American Geriatrics Society**, [S.I.], v.49, n.12, p.1700–1707, Dec. 2001.
- LONG, X. et al. Prediction and classification of Alzheimer disease based on quantification of MRI deformation. **PLOS ONE**, [S.I.], v.12, n.3, p.1–19, 03 2017.
- LOZANO, R. et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. **The Lancet**, [S.I.], v.380, n.9859, p.2095–2128, Dec. 2012.
- MATHOTAARACHCHI, S. et al. Identifying incipient dementia individuals using machine learning and amyloid imaging. **Neurobiology of Aging**, [S.I.], v.59, p.80–90, Nov. 2017.
- MAURER, K.; MAURER, U. **Alzheimer**: the life of a physician and the career of a disease. New York: Columbia University Press, 2003.
- MEEK, P. D.; MCKEITHAN, K.; SCHUMOCK, G. T. Economic considerations in Alzheimer's disease. **Pharmacotherapy**, [S.I.], v.18, n.2 Pt 2, p.68–73; discussion 79–82, Apr. 1998.
- MENDEZ, M. F. The accurate diagnosis of early-onset dementia. **International Journal of Psychiatry in Medicine**, [S.I.], v.36, n.4, p.401–412, 2006.
- MENDEZ, M. F. Early-onset Alzheimer's disease: nonamnestic subtypes and type 2 AD. **Archives of Medical Research**, [S.I.], v.43, n.8, p.677–685, Nov. 2012.
- MESKO, B. The role of artificial intelligence in precision medicine. **Expert Review of Precision Medicine and Drug Development**, [S.I.], v.2, n.5, p.239–241, Sept. 2017.
- MOSCONI, L. et al. Pre-Clinical Detection of Alzheimer's Disease Using FDG-PET, with or without Amyloid Imaging. **Journal of Alzheimer's Disease**, [S.I.], v.20, n.3, p.843–854, May 2010.
- MURPHY, M. P.; LEVINE, H. Alzheimer's Disease and the β -Amyloid Peptide. **Journal of Alzheimer's disease : JAD**, [S.I.], v.19, n.1, p.311, Jan. 2010.

- MURRAY, E. D.; BUTTNER, E. A.; PRICE, B. H. Depression and Psychosis in Neurological Practice. In: **Bradley's Neurology in Clinical Practice, 2-Volume Set: Expert Consult - Online and Print, 6e.** 6 edition.ed. Philadelphia, PA: Saunders, 2012.
- MUSLIMOVIC, D.; POST, B.; SPEELMAN, J. D.; SCHMAND, B. Cognitive profile of patients with newly diagnosed Parkinson disease. **Neurology**, [S.I.], v.65, n.8, p.1239–1245, Oct. 2005.
- NCCMH. **Dementia: A NICE-SCIE Guideline on Supporting People With Dementia and Their Carers in Health and Social Care.** Leicester (UK): British Psychological Society, 2007. (National Institute for Health and Clinical Excellence: Guidance).
- NELDER, J. A.; WEDDERBURN, R. W. M. Generalized Linear Models. **Journal of the Royal Statistical Society. Series A (General)**, [S.I.], v.135, n.3, p.370, 1972.
- NISBET, R. M.; POLANCO, J.-C.; ITTNER, L. M.; GÖTZ, J. Tau aggregation and its interplay with amyloid- β . **Acta Neuropathologica**, [S.I.], v.129, p.207–220, 2015.
- OHNISHI, S.; TAKANO, K. Amyloid fibrils from the viewpoint of protein folding. **Cellular and molecular life sciences: CMLS**, [S.I.], v.61, n.5, p.511–524, Mar. 2004.
- ORTIZ, A.; MUNILLA, J.; GÓRRIZ, J. M.; RAMÍREZ, J. Ensembles of Deep Learning Architectures for the Early Diagnosis of the Alzheimer's Disease. **International Journal of Neural Systems**, [S.I.], v.26, n.07, p.1650025, Apr. 2016.
- OSMOSIS. **Alzheimer's disease - plaques, tangles, causes, symptoms & pathology**, 2016. Disponível em: https://www.youtube.com/watch?v=v5gdH_Hydes. Acesso em: 2016-11-21.
- POWELL, L. H.; JANSSEN, I. Follow-up Study. In: **Encyclopedia of Behavioral Medicine**. [S.I.]: Springer, New York, NY, 2013. p.806–808. DOI: 10.1007/978-1-4419-1005-9_314.
- QUERFURTH, H. W.; LAFERLA, F. M. Alzheimer's Disease. **New England Journal of Medicine**, [S.I.], v.362, n.4, p.329–344, Jan. 2010.
- QUINLAN, J. R. Induction of Decision Trees. **Mach. Learn.**, [S.I.], v.1, n.1, p.81–106, Mar. 1986.
- QUINLAN, J. R. **C4.5: Programs for Machine Learning**. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1993.
- RAMACHANDRAN, T. S. **Alzheimer Disease Imaging**: Overview, CT Scan, MRI, 2016. Disponível em: <http://emedicine.medscape.com/article/336281-overview>. Acesso em: 2017-04-19.

RAPIDMINER. **Data Science Platform**, 2017. Disponível em: <<https://rapidminer.com/>>. Acesso em: 25 de junho de 2017.

REITZ, C.; MAYEUX, R. Alzheimer disease: epidemiology, diagnostic criteria, risk factors and biomarkers. **Biochemical Pharmacology**, [S.I.], v.88, n.4, p.640–651, Apr. 2014.

REPORTS, C. **Treating Alzheimer's disease**: Surprising facts about the effectiveness and safety of these commonly used drugs, 2012. Disponível em: <<http://www.consumerreports.org/cro/2012/07/evaluating-drugs-to-treat-alzheimer-s-disease/index.htm>>. Acesso em:

2016-11-23.

RODRIGUES, Y. et al. Wrappers Feature Selection in Alzheimer's Biomarkers Using kNN and SMOTE Oversampling. **Trends in Applied and Computational Mathematics**, [S.I.], v.18, n.1, p.15, 2017.

SCHROETER, M. L.; STEIN, T.; MASLOWSKI, N.; NEUMANN, J. Neural correlates of Alzheimer's disease and mild cognitive impairment: a systematic and quantitative meta-analysis involving 1351 patients. **NeuroImage**, [S.I.], v.47, n.4, p.1196–1206, Oct. 2009.

SCOTT, P. S. **Alzheimer's Tests**, 2016. Disponível em: <<https://www.caring.com/articles/alzheimers-tests-pros-cons>>. Acesso em: 2017-04-10.

SEVIGNY, J. et al. The antibody aducanumab reduces A β plaques in Alzheimer's disease. **Nature**, [S.I.], v.537, n.7618, p.50–56, Aug. 2016.

SOCIETY, A. **Factsheet 417**: The later stages of dementia, 2012. Disponível em: <https://www.alzheimers.org.uk/site/scripts/documents_info.php?documentID=101>. Acesso em: 2016-11-13.

SPERLING, R. A. et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. **Alzheimer's & Dementia: The Journal of the Alzheimer's Association**, [S.I.], v.7, n.3, p.280–292, May 2011.

S PERLING, R. The potential of functional MRI as a biomarker in early Alzheimer's disease. **Neurobiology of Aging**, [S.I.], v.32, p.S37–S43, Dec. 2011.

STRASSNIG, M.; GANGULI, M. About a Peculiar Disease of the Cerebral Cortex. **Psychiatry (Edgmont)**, [S.I.], v.2, n.9, p.30–33, Sept. 2005.

- TALER, V.; PHILLIPS, N. A. Language performance in Alzheimer's disease and mild cognitive impairment: a comparative review. **Journal of Clinical and Experimental Neuropsychology**, [S.I.], v.30, n.5, p.501–556, July 2008.
- THAL, D. R.; WALTER, J.; SAIDO, T. C.; FÄNDRICH, M. Neuropathology and biochemistry of A β and its aggregates in Alzheimer's disease. **Acta Neuropathologica**, [S.I.], v.129, n.2, p.167–182, Feb. 2015.
- TIRABOSCHI, P.; HANSEN, L. A.; THAL, L. J.; COREY-BLOOM, J. The importance of neuritic plaques and tangles to the development and evolution of AD. **Neurology**, [S.I.], v.62, n.11, p.1984–1989, June 2004.
- TODD, S.; BARR, S.; ROBERTS, M.; PASSMORE, A. P. Survival in dementia and predictors of mortality: a review. **International Journal of Geriatric Psychiatry**, [S.I.], v.28, n.11, p.1109–1124, Nov. 2013.
- VOLICER, L. et al. Sundowning and circadian rhythms in Alzheimer's disease. **The American Journal of Psychiatry**, [S.I.], v.158, n.5, p.704–711, May 2001.
- WALDEMAR, G. et al. Recommendations for the diagnosis and management of Alzheimer's disease and other disorders associated with dementia: EFNS guideline. **European Journal of Neurology**, [S.I.], v.14, n.1, p.e1–e26, Jan. 2007.
- WHO. **Life expectancy at birth**, 2015. Disponível em:
http://gamapserver.who.int/gho/interactive_charts/mbd/life_expectancy/atlas.html. Acesso em: 2017-12-12.
- WHO. **Dementia**, 2015. Disponível em:
<http://www.who.int/mediacentre/factsheets/fs362/en/>. Acesso em: 2016-11-09.
- WHO. **The top 10 causes of death**, 2017. DOI:
<http://www.who.int/mediacentre/factsheets/fs310/en/index.html>. Disponível em:
<http://www.who.int/mediacentre/factsheets/fs310/en/>. Acesso em: 2017-12-12.
- WIKIPEDIA. **Proteólise**, 2014. Page Version ID: 40134280. Disponível em:
<https://pt.wikipedia.org/w/index.php?oldid=40134280>. Acesso em: 2016-11-21.
- WIKIPEDIA. **Cross-sectional study**, 2018. Page Version ID: 818659824. Disponível em: https://en.wikipedia.org/wiki/Cross-sectional_study. Acesso em: 2018-01-15.
- WIKIPEDIA. **Confusion matrix**, 2018. Page Version ID: 821226853. Disponível em: https://en.wikipedia.org/wiki/Confusion_matrix.
- WOLPERT, D. H. Stacked Generalization. **Neural Networks**, [S.I.], v.5, p.241–259, 1992.

XRISTICA. **What is the difference between Bagging and Boosting?**, 2016.

Disponível em:

<<https://quantdare.com/what-is-the-difference-between-bagging-and-boosting/>>.

Acesso em: 2018-02-01.

ZHANG, D. et al. Multimodal classification of Alzheimer's disease and mild cognitive impairment. **NeuroImage**, [S.I.], v.55, n.3, p.856–867, Apr. 2011.

ZHANG, J. et al. CSF Multianalyte Profile Distinguishes Alzheimer and Parkinson Diseases. **American Journal of Clinical Pathology**, [S.I.], v.129, n.4, p.526–529, Apr. 2008.

ANEXO A ITENS NO ARQUIVO DE DADOS ADNI

Tabela 33: Itens no Arquivo de Dados ADNI - Categoria Assessments (ADNI, 2016b)

Subcategory	Item
Diagnosis	Conversions Confirmed by Conversion Committee [ADNI1]
Diagnosis	Diagnosis and Symptoms Checklist [ADNI1,GO,2]
Diagnosis	Diagnostic Summary – Baseline Changes [ADNI1,GO,2]
Diagnosis	Diagnostic Summary [ADNI1,GO,2]
Neuropsychological	ADAS Sub-Scores and Total Scores [ADNI1]
Neuropsychological	Alzheimer's Disease Assessment Scale (ADAS) [ADNI1]
Neuropsychological	Alzheimer's Disease Assessment Scale (ADAS) [ADNIGO,2]
Neuropsychological	Clinical Dementia Rating Scale (CDR) [ADNI1,GO,2]
Neuropsychological	Everyday Cognition – Participant Self Report [ADNIGO,2]
Neuropsychological	Functional Activities Questionnaire (FAQ) [ADNI1,GO,2]
Neuropsychological	Geriatric Depression Scale (GDS) [ADNI1,GO,2]
Neuropsychological	Mini-Mental State Examination (MMSE) [ADNI1,GO,2]
Neuropsychological	Modified Hachinski Ischemia Scale [ADNI1,GO,2]
Neuropsychological	Montreal Cognitive Assessment (MoCA) [ADNIGO, 2]
Neuropsychological	Neuropsychiatric Inventory (NPI) [ADNI2]
Neuropsychological	Neuropsychiatric Inventory Questionnaire (NPI-Q) [ADNI1,GO,2]
Neuropsychological	Neuropsychological Battery [ADNI1,GO,2]

Tabela 34: Itens no Arquivo de Dados ADNI - Categoria Biospecimen (ADNI, 2016b)

Subcategory	Item
Lab Collection Procedures	ApoE Genotyping – Draw Data [ADNI1,GO]
Lab Collection Procedures	Clinical Laboratory Tests [ADNI1,GO,2]
Lab Collection Procedures	Laboratory Data [ADNI1,GO]
Lab Results	ADNI Biomarker Core Laboratory. Baseline Isoprostanes Data Dictionary [ADNI1]
Lab Results	ADNI Biomarker Core Laboratory. Baseline Isoprostanes Data [ADNI1]
Lab Results	ApoE – Results [ADNI1,GO]
Lab Results	Biomarker Samples [ADNI1,GO,2]
Lab Results	CSF – Local Lab Results [ADNI1,GO,2]
Lab Results	CSF Multiplex Proteomics
Lab Results	Homocysteine – Results [ADNI1,GO]
Lab Results	Rules Based Medicine Plasma Multiplex QC Data Dictionary [ADNI1]
Lab Results	Rules Based Medicine Plasma Multiplex QC Data Dictionary [ADNI1]
Lab Results	Rules Based Medicine Plasma Multiplex QC Data [ADNI1]
Lab Results	Rules Based Medicine Plasma Multiplex Raw Data [ADNI1]
Lab Results	UPENN – Biomarker Data [ADNI1]
Lab Results	UPENN – Longitudinal Biomarker Data (3 yr) Dictionary [ADNI1]
Lab Results	UPENN – Longitudinal Biomarker Data (3 yr) [ADNI1]
Lab Results	UPENN – Longitudinal Biomarker Data Dictionary [ADNI1]
Lab Results	UPENN – Longitudinal Biomarker Data [ADNI1]
Lab Results	UPENN Plasma Biomarker Data Dictionary [ADNI1]
Lab Results	UPENN Plasma Biomarker Data [ADNI1]

Tabela 35: Itens no Arquivo de Dados ADNI - Categoria Enrollment (ADNI, 2016b)

Subcategory	Item
Enrollment	Additional Comments [ADNI1,GO,2]
Enrollment	Arm [ADNI1,GO,2]
Enrollment	Early Discontinuation and Withdrawal [ADNI1,GO,2]
Enrollment	Exclusion Criteria [ADNI1]
Enrollment	Inclusion Criteria [ADNI1]
Enrollment	Registry [ADNI1,GO,2]
Enrollment	Roster [ADNI1,GO,2]
Enrollment	Visits [ADNI1,GO,2]

Tabela 36: Itens no Arquivo de Dados ADNI - Categoria Genetic (ADNI, 2016b)

Subcategory	Item
Genetic Data Info	About the Genetic Data
Genetic Data Info	About the TOMM40 PolyT Variant Data
Genetic Data Info	DNA Source Reference
Genotype Results	Deprecated-SNP genotype data – set 12 of 11 (CSV Format)
Genotype Results	Deprecated-SNP genotype data – set 12 of 11 (CSV Format)
Genotype Results	SNP genotype data – set 01 of 11 (CSV Format)
Genotype Results	SNP genotype data – set 02 of 11 (CSV Format)
Genotype Results	SNP genotype data – set 03 of 11 (CSV Format)
Genotype Results	SNP genotype data – set 04 of 11 (CSV Format)
Genotype Results	SNP genotype data – set 05 of 11 (CSV Format)
Genotype Results	SNP genotype data – set 06 of 11 (CSV Format)
Genotype Results	SNP genotype data – set 07 of 11 (CSV Format)
Genotype Results	SNP genotype data – set 08 of 11 (CSV Format)
Genotype Results	SNP genotype data – set 09 of 11 (CSV Format)
Genotype Results	SNP genotype data – set 10 of 11 (CSV Format)
Genotype Results	SNP genotype data – set 11 of 11 (CSV Format)
	TOMM40 PolyT Variant Data (CSV Format)

Tabela 37: Itens no Arquivo de Dados ADNI - Categoria Imaging (ADNI, 2016b)

Subcategory	Item
MR Image Acquisition	1.5T MRI Scan Information [ADNI1,GO,2]
MR Image Acquisition	3T MRI Scan Information [ADNI1,GO,2]
MR Image Acquisition	MRI B1 Calibration [ADNI1,GO,2]
MR Image Acquisition	MRI Clinical Read [ADNI1]
MR Image Acquisition	MRI Phantom [ADNI1,GO,2]
MR Image Acquisition	MRI Protocol [ADNI1,GO,2]
MR Image Acquisition	MRI Serial [ADNI1,GO,2]
MR Image Acquisition	MRI Subject Inclusion [ADNI1,GO,2]
MR Image Quality	MRI MPRAGE Process [ADNI1,GO,2]
MR Image Quality	MRI MPRAGE Ranking [ADNI1,GO,2]
MR Image Quality	MRI Quality [ADNI1,GO,2]
MR Imaging Analysis	Deprecated-Stroke Summary Dictionary [ADNI1]
MR Imaging Analysis	Deprecated-Stroke Summary [ADNI1]
MR Imaging Analysis	Stroke Summary Dictionary Version 2 [ADNI1]
MR Imaging Analysis	Stroke Summary Version 2 [ADNI1]
MR Imaging Analysis	UA – MRI SPM Voxel Based Morphometry (VBM) Analysis Dictionary [ADNI1]
MR Imaging Analysis	UA – MRI SPM Voxel Based Morphometry (VBM) Analysis [ADNI1]
MR Imaging Analysis	UCL – Boundary Shift Integral Summaries Dictionary [ADNI1]
MR Imaging Analysis	UCL – Boundary Shift Integral Summaries [ADNI1]
MR Imaging Analysis	UCLA – Average Jacobian Dictionary [ADNI1]
MR Imaging Analysis	UCLA – Average Jacobian [ADNI1]
MR Imaging Analysis	UCSD – Derived Volumes Dictionary [ADNI1]
MR Imaging Analysis	UCSD – Derived Volumes [ADNI1]
MR Imaging Analysis	UCSF – Cross-Sectional FreeSurfer Dictionary [ADNI1]
MR Imaging Analysis	UCSF – Cross-Sectional FreeSurfer [ADNI1]
MR Imaging Analysis	UCSF – Longitudinal FreeSurfer (FreeSurfer Version 4.4) Dictionary [ADNI1]
MR Imaging Analysis	UCSF – Longitudinal FreeSurfer (FreeSurfer Version 4.4) [ADNI1]
MR Imaging Analysis	UCSF – Longitudinal FreeSurfer Dictionary [ADNI1]
MR Imaging Analysis	UCSF – Longitudinal FreeSurfer [ADNI1]
MR Imaging Analysis	UCSF – Regional Atrophy Rates Dictionary [ADNI1]
MR Imaging Analysis	UCSF – Regional Atrophy Rates [ADNI1]
MR Imaging Analysis	UCSF – SNT Hippocampal Volumes Dictionary [ADNI1]
MR Imaging Analysis	UCSF – SNT Hippocampal Volumes [ADNI1]
MR Imaging Analysis	UPENN – SPARE-AD Dictionary [ADNI1]
MR Imaging Analysis	UPENN – SPARE-AD [ADNI1]
MR Imaging Analysis	UWO – Ventricular Volumes [ADNI1]
PET Image Acquisition	AV-45 PET Scan Information [ADNIGO,2]
PET Image Acquisition	FDG PET Scan Information [ADNI1]
PET Image Acquisition	FDG PET Scan Information [ADNIGO/ADNI2]
PET Image Acquisition	FDG/AV45 PET Scan Information
PET Image Acquisition	PIB Scan Information [ADNI1]
PET Image Quality	AV-45 PET QC Tracking [ADNIGO,2]
PET Image Quality	PET QC Tracking [ADNI1,GO,2]
PET Image Quality	PIB QC Tracking [ADNI1]
PET Imaging Analysis	Banner Alzheimer's Institute NMRC Summaries Dictionary [ADNI1]
PET Imaging Analysis	Banner Alzheimer's Institute NMRC Summaries [ADNI1]
PET Imaging Analysis	UCB – PET ROI Analysis Dictionary [ADNI1]
PET Imaging Analysis	UCB – PET ROI Analysis [ADNI1]
PET Imaging Analysis	UPitt – PIB PET Analysis Dictionary [ADNI1]
PET Imaging Analysis	UPitt – PIB PET Analysis [ADNI1]
PET Imaging Analysis	UU – PET Analysis Dictionary [ADNI1]
PET Imaging Analysis	UU – PET Analysis [ADNI1]

Tabela 38: Itens no Arquivo de Dados ADNI - Categorias Medical History, Study Info e Subject Characteristics (ADNI, 2016b)

Subcategory	Item
Adverse Events	Adverse Events/ Hospitalizations [ADNI1,GO,2]
Adverse Events	AV-45 24-48 Hour Follow-Up [ADNIGO,2]
Drugs	Concurrent Medications Log [ADNI1,GO,2]
Drugs	Key Background Medications [ADNIGO,2]
Medical History	Documentation of Baseline Symptoms Log [ADNI1,GO,2]
Medical History	Medical History [ADNI1,GO,2]
Medical History	Recent Medical History Details Log [ADNI1,GO,2]
Physical/Neurological	AV-45 Pre and Post Injection Vitals [ADNIGO,2]
Physical/Neurological	Baseline Symptoms Checklist [ADNI1,GO,2]
Physical/Neurological	NACC Autopsy Information Dictionary [ADNI1]
Physical/Neurological	NACC Autopsy Information [ADNI1]
Physical/Neurological	Neurological Exam [ADNI1,GO,2]
Physical/Neurological	Physical Exam [ADNI1,GO,2]
Physical/Neurological	Vital Signs [ADNI1,GO,2]
Data & Database	Data dictionary [ADNI1,GO,2]
Study Protocols & CRFs	ADNI Neuropsychological Assessment Matrix (PDF)
Family History	Family History Questionnaire Subtable [ADNI1,GO,2]
Family History	Family History Questionnaire [ADNI1,GO,2]
Subject Demographics	Subject Demographics [ADNI1,GO,2]