

UNIVERSIDADE FEDERAL DE PELOTAS
Centro de Engenharias
Programa de Pós-Graduação em Ciências Ambientais
Mestrado em Ciências Ambientais

Dissertação



**Uso de Inteligência Artificial para escolha de sementes visando a
Sustentabilidade na Agricultura**

Ruan Bernardy

Pelotas, 2023

Ruan Bernardy

**Uso de Inteligência Artificial para escolha de sementes visando a
Sustentabilidade na Agricultura**

Dissertação apresentada junto ao Programa de Pós-Graduação em Ciências Ambientais da Universidade Federal de Pelotas, como requisito parcial para a obtenção do título de Mestre em Ciências Ambientais.

Orientadora: Prof^a. Dra. Gizele Ingrid Gadotti

Pelotas, 2023

Universidade Federal de Pelotas / Sistema de Bibliotecas
Catalogação na Publicação

B518u Bernardy, Ruan

Uso de inteligência artificial para escolha de sementes visando a sustentabilidade na agricultura / Ruan Bernardy ; Gizele Ingrid Gadotti, orientadora. — Pelotas, 2023.

99 f.

Dissertação (Mestrado) — Programa de Pós-Graduação em Ciências Ambientais, Centro de Engenharias, Universidade Federal de Pelotas, 2023.

1. Produção sustentável. 2. Aprendizado de máquinas.
3. Glycine max. 4. Redução de impactos. 5. Agricultura 4.0.
I. Gadotti, Gizele Ingrid, orient. II. Título.

CDD : 363.7

Ruan Bernardy

Uso de Inteligência Artificial para escolha de sementes visando a Sustentabilidade
na Agricultura

Dissertação aprovada, como requisito para obtenção do grau de Mestre em Ciências Ambientais, Programa de Pós-Graduação em Ciências Ambientais, Centro de Engenharias, Universidade Federal de Pelotas.

Data da defesa: 23 de fevereiro de 2023

Banca Examinadora:

Orientadora: Profa. Dra. Gizele Ingrid Gadotti
Universidade Federal de Pelotas – Centro de Engenharias

Prof. Dr. Marcelo Lemos Rossi
Universidade Federal de Pelotas – Centro de Engenharias

Prof. Dr. Érico Kunde Corrêa
Universidade Federal de Pelotas – Centro de Engenharias

Dedico à minha família e amigos que sempre estiveram do meu lado, apoiando as minhas decisões. Aos professores que sempre me incentivaram. A Nossa Senhora, por interceder todos os dias por mim. A Deus por sempre me mostrar o caminho a seguir.

Agradecimentos

Agradeço a Deus, por permitir que fizesse esse trabalho!

Aos meus familiares e amigos que me incentivaram com palavras afetuosas e me ajudaram sempre que preciso. São poucos, mas estão lendo estas linhas.

Aos meus colegas de trabalho no laboratório, por toda a ajuda nos artigos desenvolvidos e congressos participados.

Aos meus professores, pelos ensinamentos recebidos, especialmente pela minha orientadora, que sempre está ao meu lado, ajudando e guiando todos os meus passos.

À UFPel, por toda a estrutura disponibilizada, em especial, o curso de Pós-Graduação em Ciências Ambientais, com seus professores e colegas que sempre auxiliaram nesta jornada.

E por fim, agradeço a Nossa Senhora, por interceder pela minha vida e sempre estar ao meu lado nesta caminhada, me dando pessoas que considero como família.

Salve Maria: Puríssima!

Resumo

BERNARDY, Ruan. **Uso de Inteligência Artificial para escolha de sementes visando a Sustentabilidade na Agricultura**. Orientadora: Gizele Ingrid Gadotti. 2023. 98 f. Dissertação. (Mestrado em Ciências Ambientais) - Programa de Pós-Graduação em Ciências Ambientais. Centro de Engenharias. Universidade Federal de Pelotas, Pelotas, 2023.

O Brasil possui uma agricultura de destaque nos últimos quarenta anos, com a Inteligência Artificial chamando atenção de pesquisadores em geral. Porém, existe uma preocupação mundial que valoriza mais o produto sustentável, descartando das opções de compra quando não se encontram nesses padrões. Assim, o objetivo desse estudo foi utilizar o aprendizado de máquinas para permitir a classificação de sementes de soja com maior precisão e eficácia, fornecendo ao produtor rural sementes com alta qualidade, reduzindo a dependência de agroquímicos para alcançar alta produtividade, a fim de tornar o processo mais seguro e sustentável do ponto de vista ambiental. Foram utilizados dados provenientes de testes realizados nos laboratórios de análises de sementes da Universidade Federal de Pelotas – RS e também de empresas sementeiras. Os dados foram divididos em dois conjuntos, o C1 com três formas de classificação, e o C2 com menor quantidade de variáveis. Os classificadores utilizados foram J48, Random Forest, CVR, IBk, MLP e NãiveBayes. Para certificar a precisão dos algoritmos, utilizaram-se as métricas de acurácia: precisão, recall, F-Measure e ROC Area. Os algoritmos MLP e Random Forest obtiveram as melhores respostas em 2 formas de classificação do C1 e no C2, com mais de 95% de assertividade. Todos os algoritmos conseguiram classificar os lotes com taxa de acertos superior a 75%, sendo alguns ficando em 99%. O MLP apresentou os melhores resultados, mesmo quando os dados foram classificados a partir de uma única variável, que foi posteriormente retirada para não viciar o classificador. Com os resultados obtidos foi possível realizar a avaliação e ranqueamento com precisão e acurácia, utilizando a inteligência artificial em lotes de sementes de soja. Os atributos físicos não foram utilizados como principais variáveis, quando a classificação foi realizada com base no segundo teste de viabilidade por TZ. Com o C2, o Vigor e a Germinação se apresentaram primordiais para a classificação com precisão. Porém, alguns cientistas ainda salientam que para uma boa caracterização e ranqueamento do lote é necessário a realização de outros testes. Os impactos ambientais causados pelo excessivo uso de agroquímicos na implantação da cultura podem ser reduzidos devido à elevação da qualidade fisiológica e, dessa forma, tornar a cadeia produtiva mais sustentável utilizando a tecnologia de informação como ferramenta, visando melhorar a qualidade de vida das pessoas e atentar para o meio ambiente, de modo que todo o ecossistema receba os devidos cuidados.

Palavras-Chave: Produção Sustentável. Aprendizado de Máquinas. *Glycine max*. Redução de Impactos. Agricultura 4.0.

Abstract

BERNARDY, Ruan. **Use of Artificial Intelligence for seed selection aiming at sustainability in agriculture.** Advisor: Gizele Ingrid Gadotti. 2023. 98 f. Dissertation (Master in Environmental Sciences) - Graduate Program in Environmental Sciences. Engineering Center. Federal University of Pelotas, Pelotas, 2023.

Brazil has an outstanding agriculture in the last forty years, with Artificial Intelligence drawing attention from researchers in general. However, there is a worldwide concern that values more the sustainable product, discarding from the purchase options when they do not meet these standards. Thus, the objective of this study was to use machine learning to enable the classification of soybean seeds with greater accuracy and efficiency, providing the rural producer with high quality seeds, reducing dependence on agrochemicals to achieve high productivity, in order to make the process safer and more environmentally sustainable. Data from tests performed in the seed analysis laboratories of the Federal University of Pelotas - RS and also from seed companies were used. The data were divided into two sets, C1 with three forms of classification, and C2 with fewer variables. The classifiers used were J48, Random Forest, CVR, IBk, MLP and NaïveBayes. To certify the accuracy of the algorithms, the following accuracy metrics were used: precision, recall, F-Measure, and ROC Area. The MLP and Random Forest algorithms obtained the best responses in the 2 classification forms of C1 and in C2, with more than 95% of assertiveness. All algorithms were able to classify the plots with hit rates higher than 75%, with some getting 99%. The MLP presented the best results, even when the data were classified from a single variable, which was later removed so as not to bias the classifier. With the results obtained it was possible to perform the evaluation and ranking with precision and accuracy, using artificial intelligence on soybean seed lots. The physical attributes were not used as main variables when the ranking was performed based on the second viability test by TZ. With C2, Vigor and Germination were shown to be paramount for accurate classification. However, some scientists still point out that for a good characterization and ranking of the batch it is necessary to perform other tests. The environmental impacts caused by excessive use of agrochemicals in the implementation of the crop can be reduced due to the elevation of physiological quality and, thus, make the production chain more sustainable by using information technology as a tool, aiming to improve the quality of life of people and pay attention to the environment, so that the entire ecosystem receives proper care.

Keywords: Sustainable Production. Machine Learning. *Glycine max*. Impact Reduction. Agriculture 4.0.

Lista de Figuras

Figura 1	Projeção da produção de soja em grão para 2029/30.....	21
Figura 2	Sequência de operação utilizando mineração de dados	40
Figura 3	Processamento dos dados a partir do tipo de aprendizagem	43
Figura 4	Árvore decisória do algoritmo J48 com o Cenário 1 e dados C1	56
Figura 5	Árvore decisória do algoritmo J48 com o Cenário 2 e dados C1	61
Figura 6	Árvore decisória do algoritmo J48 com o Cenário 3 e dados C1	67
Figura 7	Árvore decisória do algoritmo J48 com os dados C2	72

Lista de Tabelas

Tabela 1	Consumo de agrotóxicos e fertilizantes nas lavouras do Brasil.	24
Tabela 2	Descrição dos atributos analisados pela mineração de dados do conjunto C1.....	48
Tabela 3	Descrição dos atributos analisados pela mineração de dados do conjunto C2.....	49
Tabela 4	Acurácia dos algoritmos após a classificação no Cenário 1 com dados C1.....	51
Tabela 5	Matriz de confusão do algoritmo MLP para o cenário 1 e dados C1	52
Tabela 6	Matriz de confusão do algoritmo Random Forest para o cenário 1 e dados C1.....	52
Tabela 7	Acurácias dos diferentes algoritmos utilizados, sendo Recall (sensibilidade), Precisão, Curva ROC (Receiver Operating Characteristic) e F Measure para o cenário 1 e dados C1	53
Tabela 8	Acurácia dos algoritmos após a classificação no Cenário 2 com dados C1.....	57
Tabela 9	Matriz de confusão do algoritmo IBk para o Cenário 2 e dados C1	58
Tabela 10	Matriz de confusão do algoritmo Random Forest para o Cenário 2 e dados C1.....	58

Tabela 11	Acurácias dos diferentes algoritmos utilizados, sendo Recall (sensibilidade), Precisão, Curva ROC (Receiver Operating Characteristic) e F Measure para o cenário 2 e dados C1	59
Tabela 12	Acurácia dos algoritmos após a classificação no Cenário 3 com dados C1.....	62
Tabela 13	Matriz de confusão do algoritmo MLP para o Cenário 3 e dados C1 ...	63
Tabela 14	Matriz de confusão do algoritmo Random Forest para o Cenário 3 e dados C1.....	64
Tabela 15	Acurácias dos diferentes algoritmos utilizados, sendo Recall (sensibilidade), Precisão, Curva ROC (Receiver Operating Characteristic) e F Measure para o cenário 3 e dados C1	65
Tabela 16	Acurácia dos algoritmos após a classificação de sementes de soja do conjunto C2	68
Tabela 17	Matriz de confusão do algoritmo MLP para o C2.	69
Tabela 18	Matriz de confusão do algoritmo Random Forest para o C2.....	70
Tabela 19	Acurácias dos diferentes algoritmos utilizados, sendo Recall (sensibilidade), Precisão, Curva ROC (Receiver Operating Characteristic) e F Measure para o conjunto C2	71

Sumário

1 Introdução	12
2 Objetivo Geral	17
2.1 Objetivos Específicos	17
3 Hipóteses	18
4 Referencial Teórico	18
4.1 A Cultura da Soja	19
4.2 Uso de Agroquímicos e a Poluição Ambiental.....	22
4.2.1 Tratamento de Sementes	25
4.2.2 Uso de Fertilizantes.....	27
4.2.3 Pegada de Carbono	28
4.2.4 Pegada Hídrica.....	29
4.3 Exigências Legais na Análise de Sementes.....	32
4.4 Métodos para Análise de Sementes.....	33
4.5 Inteligência Artificial: Aplicação na Agricultura e no Setor Sementeiro	35
4.6 Técnicas para Mineração de Dados.....	38
4.6.1 Regras de Classificação.....	41
4.6.2 Regras de <i>Clustering</i>	41
4.6.3 Regras de Associação.....	42
4.6.4 Aprendizado Supervisionado.....	43
4.6.5 Aprendizado Não-Supervisionado.....	44
4.6.6 Software WEKA.....	45
5 Material e Métodos	46
6 Resultados e Discussão	50
7 Considerações finais	75
Referências	76

1 Introdução

No decorrer dos anos, os produtores rurais foram incentivados a cultivar sementes com genética pura e elevada germinação (GHAFFARI; NAJAFABADI, 2022). De acordo com esses autores, sementes com alta qualidade são primordiais para uma produção alimentar sustentável e lucros estáveis. Por isso, organizaram-se processos para certificação desse material genético, facilitando assim o processo (NAGHASHZADEH; AZADBAKHT, 2018).

O Brasil vem apresentando, nos últimos quarenta anos, uma agricultura de destaque e a soja vem despontando como um dos principais produtos agrícolas cultivados no país, tornando-se o maior produtor do mundo com 40,7 milhões de hectares plantados (EMBRAPA, CONAB, 2022) e o maior exportador do grão. No Brasil, a leguminosa está na maior parte das lavouras destinadas a agricultura, devido ao bom retorno financeiro dado ao produtor pelas condições climáticas favoráveis e investimentos em tecnologias, que contribuem para colheitas recordes.

Segundo dados da CEPEA (2022), do ponto de vista das suas exportações, o agronegócio brasileiro é um setor bastante integrado internacionalmente, conseguindo, com isso, representar cerca de 25% do seu PIB, sendo que a comercialização é praticamente metade do que se produz em geral. A China é o principal país importador, recebendo 37% do total produzido, e a União Europeia, com 15%, se destaca em segundo lugar (CEPEA, 2022).

O superávit do mês de fevereiro em 2022 foi o maior desde 2017 (US\$ 4,05 bilhões), com os bens agropecuários alavancando o maior crescimento das exportações no mês (114,2%), tendo como destaque a soja (SISCOMEX, 2022).

De acordo com Pignati (2018), o agronegócio brasileiro segue em crescimento, com números expressivos a cada ano, tendo os produtos direcionados tanto para consumo interno quanto para o externo. Porém, o elevado uso de agroquímicos para produzir toda essa quantidade de alimento é notório, ocasionando grandes impactos ao ambiente (PIGNATI, 2018). De acordo com Costa *et al.* (2004), os princípios ativos destes produtos químicos possuem componentes que poluem o ambiente com metais pesados, surfactantes, entre outros, podendo causar danos irreversíveis ao meio ambiente. A utilização desses produtos é viável pela facilidade e melhora visível no

rendimento das lavouras, aumentando, portanto, os lucros, mas deixando o ambiente e saúde pública de lado (VIEIRA, 2021).

Entretanto, existe uma tendência mundial que valoriza mais o produto sustentável, fazendo com que os produtos que não sigam esses padrões sejam descartados das opções de compra. Por exemplo, no ano de 2006 foi criado um movimento liderado pelo Greenpeace, que induziu o boicote da compra de soja produzida em áreas desmatadas (NEPSTAD *et al.*, 2014). Isso mostra a necessidade de mudanças nos hábitos para produzir alimentos saudáveis, não só às pessoas, mas também para o meio ambiente.

De acordo com a *Food and Agriculture Organization of the United Nations* (FAO) (2017), o mundo deve atingir 9 bilhões de habitantes até 2050. Segundo a Embrapa, que realizou um estudo em 2021, a agricultura brasileira deve responder pela alimentação de mais de 800 milhões de pessoas no mundo. Posto isso, o país terá de ser extremamente responsável na questão de segurança alimentar.

Entretanto, para conseguir alimentar essa quantidade de pessoas, a degradação ambiental, devido a utilização inadequada de recursos naturais, se torna cada vez mais evidente, sendo esse o fator que levou os especialistas a buscarem reuniões e eventos de nível global com a finalidade de discutir possíveis medidas e mudanças na forma de extração destes recursos, melhorando a conservação do meio ambiente (RIBEIRO, 2019).

Na Conferência das Nações Unidas sobre mudanças climáticas realizada em 2021 (COP-26), o governo brasileiro mostrou que a partir da criação do Programa Agricultura de Baixa Emissão de Carbono (Plano ABC), em 2010, o país criou seis tecnologias descarbonizantes em mais de 52 milhões de hectares, evitando, com isso, a emissão de mais de 170 milhões de toneladas de carbono. O objetivo do programa é reduzir a emissão de carbono até 2030 em cerca de 1,1 bilhão de toneladas, somente no setor agropecuário, alcançando 72 milhões de hectares com técnicas sustentáveis.

Alguns estudos realizados no Brasil na década de 1960, sobre inovação tecnológica na agricultura, refletiram profundas mudanças ocorridas no setor: da estagnação ao acréscimo considerável na produtividade (VIEIRA FILHO; SILVEIRA, 2012). Com o passar dos anos, o uso da tecnologia se fez cada vez mais necessário no setor agropecuário, com novas opções de máquinas e sistemas que melhoram a vida do produtor rural e maximizam seus lucros.

Nesse contexto avaliativo, Vieira e Oliveira (2014) salientam que uma grande quantidade de dados cresce de forma muito acelerada em várias esferas do conhecimento, dificultando a interpretação destes, pois o volume é maior que o poder para explicá-los. Dessa forma, surge a necessidade de descobrir ferramentas e/ou técnicas automatizadas para solucionar este problema, as quais possam auxiliar o analista a transformar dados em conhecimento (HAN *et al.*, 2011).

Segundo Vieira (2021), a informática foi introduzida como ferramenta para simulação e gerenciamento de processos agrícolas, sendo a principal característica da agricultura 4.0, além de impactar diretamente na redução de custos das lavouras. Para o autor, os programas e sistemas foram desenvolvidos com objetivo de permitir a redução de frotas (caminhões e máquinas agrícolas), maximizando o lucro por hectare e otimizando as operações envolvidas na cadeia produtiva.

Com isso, é nítida a preocupação dos produtores rurais em relação ao conhecimento demandado para operar essas novas tecnologias que estão surgindo rapidamente. Segundo Soares (2018), é notório que a competitividade e a globalização tornam necessária a contratação de profissionais extremamente qualificados, elevando os custos de produção. Além disso, muitos agricultores necessitam de pessoas que consigam usufruir ao máximo de tudo que a tecnologia disponibiliza e, com isso, os custos operacionais das propriedades rurais se tornam ainda mais elevados.

Seguindo essa linha, empresas do ramo de produção de sementes estão buscando melhorar seus equipamentos para atender as necessidades do setor e depender menos de pessoas para operá-los, a fim de otimizar os processos e aumentar a precisão com que esses são realizados. No setor sementeiro, a tomada de decisão rápida e precisa ainda encontra muitos desafios, pois trabalha com grande quantidade de lotes de sementes e uma carga substancial de informações que, dependendo da empresa, poderá ter milhares de dados para processar em apenas uma safra, tornando o tratamento desses, de forma puramente manual, inviável (PATRÍCIO; RIEDER, 2018).

A maior parte das etapas de avaliação para certificar a qualidade de sementes gera uma grande quantidade de dados difíceis de analisar por métodos tradicionais (GHAFFARI; NAJAFABADI, 2022). Pinheiro *et al.* (2021) também comentam que analisar todos os testes que determinam a qualidade de sementes gera um tanto de informações que se torna quase impossível, a curto prazo, à capacidade intelectual humana de uma análise rápida e eficaz dentro de um laboratório de controle de qualidade. Portanto, resultados errôneos podem implicar em prejuízos econômicos para as empresas do setor (GHAFFARI; NAJAFABADI, 2022).

Alguns estudos já abordam a utilização de ferramentas que auxiliem nessa jornada (SINGH *et al.*, 2016). Conseqüentemente, estudos estão sendo focados em sistemas que utilizam a inteligência artificial (IA) para a realização de tarefas que possam auxiliar os trabalhadores em uma jornada de trabalho mais tranquila, com o objetivo de facilitar e melhorar o entendimento de modelos que utilizam esse sistema, otimizando recursos em diversos setores da agricultura, especialmente na avaliação da qualidade das sementes comercializadas.

A IA possui uma grande aplicação na agricultura, desde o desenvolvimento de robôs, monitoramento de solos e culturas, além da predição que possibilita enviar ao produtor informações que auxiliem na maximização de conhecimento (JHA *et al.*, 2019).

O aprendizado de máquinas tem impactado todos os setores econômicos, por sua grande capacidade de analisar e interpretar dados em velocidade superior ao do ser humano, tornando os investimentos e as adesões cada vez maiores, especialmente para tornar os processos mais rápidos, desde fábricas, transportes, setor imobiliário e agronegócio (MORETI *et al.*, 2021).

Para isso, são utilizadas técnicas de mineração de dados que consistem em encontrar padrões dentro de um grande conjunto de valores, usufruindo de métodos computacionais que permitam gerar informações mais precisas, onde o conhecimento é automaticamente extraído dos conjuntos de dados (CARDOSO; MACHADO, 2008; REDDY, 2021).

Deste modo, essa técnica surge como uma ferramenta importante para predição de dados na qualidade fisiológica das sementes. Segundo Moraes (2020), o ranqueamento de lotes de sementes comerciais em uma empresa é fundamental para ter rapidez de despacho dos lotes aos produtores, sendo necessário possuir os percentuais de vigor e germinação, entre outras informações de qualidade destes. Nesse sentido, a demanda por métodos eficientes e seguros está em uma crescente, sendo a tecnologia da informação uma ferramenta para esse fim.

Usar sementes de boa qualidade é a principal solução para se conseguir bons rendimentos, mantendo a área livre de patógenos (RAVA *et al.*, 2006). Segundo os autores, um exemplo que pode ser citado é o do feijão, que usando apenas sementes saudáveis resulta em aumento de mais de 45% na produtividade, sem elevar o uso de produtos químicos, que afetam diretamente o meio ambiente, ocasionando grandes impactos ambientais.

Porém, o intuito de conseguir sementes saudáveis, unido a grande ascensão da cadeia produtiva do agronegócio, trouxe consigo poluições, enfermidades e patologias devido ao uso indiscriminado de fertilizantes, para conseguir maior produtividade, e agrotóxicos, para combater o ataque de pragas e doenças no início do plantio, ocasionando danos ao meio ambiente e à saúde humana (VIEIRA, 2021). Esse uso desenfreado, acaba por si contaminando as lavouras, o meio ambiente e, conseqüentemente, a vida humana (CARNEIRO *et al.*, 2015). Com isso, a IA está sendo uma ferramenta que poderá reduzir o uso destes químicos, diminuindo os impactos conseqüentes.

Por ser considerada uma área em ascensão, trabalhos de pesquisa utilizando IA para avaliação de qualidade de sementes ainda são poucos. Mesmo para grandes pesquisadores do setor de tecnologias esse tipo de atividade ainda é muito novo, carecendo, portanto, de estudos para avaliar seu alcance na agricultura, encontrar seu máximo rendimento para tornar os processos rápidos, seguros e precisos, retirando, assim, das pessoas aquele processo repetitivo, como se fosse uma “atividade robotizada”.

O processo manual de classificação, normalmente empregado nas empresas, está se tornando inviável, pois utiliza mão de obra muitas vezes não mais encontrada no mercado de trabalho, devido a carga de processos demandada na operação. Já a IA irá oportunizar a entrada de novos profissionais, que estão se especializando cada vez mais na área de tecnologia da informação.

Aliado a isso, o uso desenfreado de produtos químicos, especialmente agrotóxicos, para elevar a qualidade das sementes, acarreta em diversos impactos ambientais, na contaminação daqueles que trabalham diretamente com o material e os consumidores desses alimentos no final da cadeia produtiva. Além disso, de acordo com os estudos realizados por Deuner *et al.* (2014), Costa *et al.* (2018) e Bem Junior (2020), o uso de alguns produtos estão afetando negativamente a germinação das sementes, sendo este outro ponto a ser discutido.

É essencial que essas sementes passem por um rigoroso controle na fase de avaliação e classificação, reduzindo erros e melhorando a precisão dos resultados, onde o produtor rural poderá se utilizar da própria resistência natural do material genético frente a doenças e pragas, diminuindo o uso de produtos químicos agrícolas.

2 Objetivo Geral

Assim, o objetivo desse estudo foi utilizar o aprendizado de máquinas para permitir a classificação de sementes de soja com maior precisão e eficácia, fornecendo ao produtor rural, sementes com alta qualidade genética, física, fisiológica e sanitária e retirando a dependência de fertilizantes e agrotóxicos para alcançar alta produtividade, a fim de tornar o processo mais seguro e sustentável do ponto de vista ambiental.

2.1 Objetivos Específicos

Com base no trabalho realizado recentemente com sementes de soja por Gadotti *et al.* (2022a), milho por Moraes (2020) e Gadotti *et al.* (2022b), e sorgo por Rocha *et al.* (2023), os objetivos específicos são:

1. Aplicar a técnica de aprendizagem de máquina em lotes de sementes de soja.
2. Utilizar algoritmos de IA para classificar as sementes, de modo a manter um padrão de qualidade;

3. Elevar as opções de algoritmos que consigam se adequar a classificação de sementes de soja, permitindo mais possibilidades dentro dos laboratórios de análise;
4. Tornar a cadeia produtiva sustentável e segura no ponto de vista ambiental.

3 Hipóteses

1. A técnica de aprendizagem de máquina é uma ferramenta que pode ser utilizada para o conjunto de dados proposto, avaliando de forma satisfatória os lotes de sementes de soja.

2. Os algoritmos utilizados nessa metodologia conseguem avaliar os lotes de sementes de soja, tornando o processo de análise mais preciso e confiável.

3. Através da utilização de dois conjuntos de dados, com variáveis e forma de classificação diferente, os algoritmos conseguem se adequar, realizando a classificação com precisão e indicando quais testes devem constar no banco de dados de sementes.

4 Referencial Teórico

Nesse tópico será realizado um estudo sobre os temas que envolvem esta pesquisa, bem como os trabalhos já realizados, sendo estes citados no decorrer da escrita. A análise dos aspectos relacionados diretamente a pesquisa é importante para que os leitores consigam entender o real objetivo dessa dissertação e fazendo com que, após o término, todos consigam fazer suas considerações para que novos trabalhos sejam desenvolvidos.

4.1 A Cultura da Soja

A soja é uma cultura que possui grande importância mundial, sendo amplamente utilizada em elaboração de rações animais, produção de óleo e outros subprodutos, além de ser consumida *in natura*, algo que vem crescendo nos últimos anos (PICCOLI, 2018).

Segundo a Embrapa (2023), a soja é oriunda da China e foi introduzida no país via EUA, chegando no Recôncavo Baiano pela primeira vez em 1882. O primeiro registro de cultivo comercial está datado em 1941, no município de Santa Rosa localizada no norte do RS, devido as semelhanças de clima encontrado entre essa região e a do Sul americano. Além disso, o país desenvolvia a produção de gado leiteiro em pequenas propriedades, demandando feno para alimentação desses animais e ocasionando uma necessidade estratégica da oleaginosa (EMBRAPA, 2019).

Na década seguinte ocorreu aumento acentuado no preço da leguminosa no mercado mundial, incentivando ainda mais os produtores e também o próprio governo brasileiro, detentor da vantagem competitiva para escoar a safra no decorrer da entressafra dos Estados Unidos (PEREIRA, 2021).

Porém, a partir de 1970, as condições climáticas favoráveis, unidas ao desenvolvimento econômico e genético, consolidou a soja como rotação de cultura ao trigo, fazendo com que a área plantada crescesse mais de 50 vezes entre as safras 1970/71 e 1979/80 (WESZ JÚNIOR, 2014).

Segundo Teixeira (2005), a produção de soja está diretamente vinculada à produção agrícola, que por sua vez, sofreu diversas mudanças no decorrer dos anos. Passando a Segunda Guerra Mundial, houve uma transformação chamada “modernização agrícola” que impactou o setor (HEDLUND; MARUJO; MELO, 2021). Teixeira (2005) ainda relata que o país sentiu os efeitos dessas mudanças após a década de 50, refletindo principalmente na região Sul e Sudeste. As transformações foram sentidas em diversos setores da agricultura, afetando o desenvolvimento tecnológico e também o mercado de trabalho (HEDLUND; MARUJO; MELO, 2021).

Atualmente, a soja é um dos principais produtos no agronegócio mundial e é utilizado como moeda de valor pelos agricultores, cerealistas e corretores, multiplicando ganhos de quem consiga entender como funciona o mercado do grão, que contribui diretamente para o aumento do PIB brasileiro (IBGE, 2021).

Na visão de Hirakuri e Lazzarotto (2014), a importância econômica está relacionada ao movimento de um grande número de agentes e organizações ligados aos mais diversos setores, como:

(...) empresas de pesquisa e desenvolvimento, fornecedores de insumos, indústrias de máquinas e equipamentos, produtores rurais, cooperativas agropecuárias, cooperativas agroindustriais, processadoras, produtores de óleo, fabricantes de ração e usinas de biodiesel, dentre outras. Em outros termos, o supracitado complexo é um vital gerador de riquezas, empregos e divisas, se transformando em um dos principais vetores de desenvolvimento regional do País (HIRAKURI; LAZZAROTTO, 2014, p. 56).

Já a modernização agrícola brasileira se deu pela cultura da soja, onde as principais conquistas foram:

1. Valor da produção primária: Na maioria dos anos, a soja remunerou os agricultores, acima do custo de oportunidade de outros produtos, fundamento de sua atratividade; 2. Cadeias produtivas dinâmicas, modernas e sofisticadas: Em função do avanço da cultura no país, implantaram-se fábricas ou centros de negócios no país, mormente de máquinas, fertilizantes e agrotóxicos a montante e de comercialização e industrialização de soja, à jusante; 3. Agregação de valor em carnes e outros alimentos; 4. Liderança nas exportações há quase 30 anos; 5. Receita de R\$285 bilhões na última década (2011 = R\$ 24 bilhões) nas exportações; 6. Responde por mais de 60% do saldo comercial do agronegócio (GAZZONI, 2020, p. 4).

Segundo Ávila e Albrecht (2010), a importância dessa leguminosa também vem sendo utilizada como alternativa na prevenção de doenças e na constituição da alimentação humana, sendo transformada em alimentos proteicos como farinha, leite, proteína texturizada e cremes.

Outro ponto importante da soja para a economia foi a consolidação de outras cadeias de produção, como milho e algodão cultivados em larga escala, bem como a criação intensiva de aves e de suínos com a utilização do farelo de soja (PICCOLI, 2018).

O Ministério da Agricultura, Pecuária e Abastecimento (MAPA) fez uma projeção do agronegócio da soja para os próximos 10 anos. De acordo com o estudo, a projeção de produção de soja em grão para 2029/30 é de 156,5 milhões de toneladas, o que representa um acréscimo de 30,1% em relação à produção de 2019/20, que foi de 120,33 milhões de toneladas (MAPA, 2020, p.37).

No que se refere as exportações da soja em grão no país, a projeção para 2029/30 é de 134,391 milhões de toneladas, um aumento de 23,1% em relação a quantidade exportada pelo Brasil em 2019/20 (MAPA, 2020, p. 39).

A Figura 1 mostra as perspectivas de produção, consumo e exportação da Soja (mil toneladas) para a safra 2029/30.

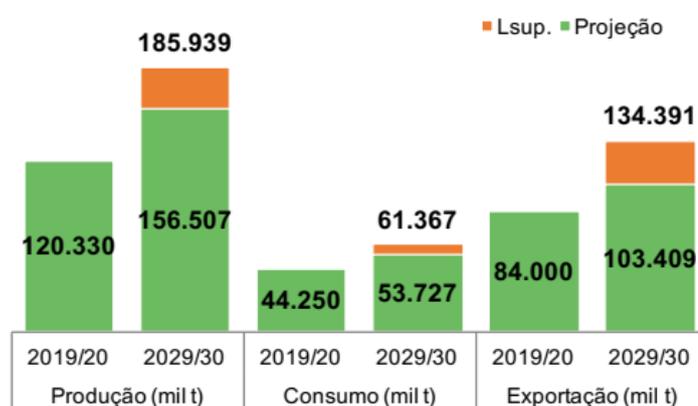


Figura 1 – Projeção da produção de soja em grão para 2029/30.
Fonte: CGEA/DCEE/SPA/Mapa e SGI/Embrapa, 2020.

Ainda segundo o MAPA (2020), a área plantada deve ter um acréscimo de 9,7 milhões de hectares até 2030, passando de 36,84 milhões para 46,6 milhões de ha, sendo a lavoura que mais deve sofrer expansão nos próximos anos, ocorrendo principalmente em regiões com maior disponibilidade de terras, seja virgens ou com pastagens. No caso específico do Brasil, a tendência é que a expansão ocorra principalmente sobre terras de pastagens naturais (IBGE, 2021), como é o caso da região sul do Rio Grande do Sul.

Nesse contexto, para conseguir uma produção ideal do grão, alguns fatores devem ser considerados, sendo: as exigências hídricas, a rotação de culturas, o manejo do solo, a correção de fertilidade, a cultivar, o manejo do cultivo, o espaçamento entre plântulas, o controle de ervas daninhas, o manejo de doenças e pragas, além da tecnologia utilizada no material genético e da colheita (EMBRAPA, 2021).

Assim, a instalação da lavoura sempre inicia pela escolha das sementes e tratamentos com fungicidas e inseticidas, que garantirão um bom estande de plantas, uma vez que a germinação destas depende da sua capacidade de suportar o clima desfavorável (ROMANI, 2017).

Logo, utilizar técnicas sofisticadas para analisar de forma adequada a qualidade das sementes é fundamental para a redução na aplicação desses fungicidas, ao ponto que o produtor tenha a intenção de salvar a sua lavoura do ataque de doenças, causando, com isso, problemas futuros no meio ambiente, nas pessoas envolvidas diretamente com a área cultivada e, também, no consumidor final a partir do uso desenfreado dessas substâncias químicas (LOPES; ALBUQUERQUE, 2018).

Em vista disso, é necessário analisar a utilização de produtos químicos, bem como a poluição causada por eles no meio ambiente e nas pessoas, para que o leitor entenda a importância de pesquisar alternativas para a redução drástica do seu uso, evitando, desse modo, impactos cada vez maiores.

4.2 Uso de Agroquímicos e a Poluição Ambiental

O crescimento exponencial da população mundial e, por consequência, a degradação do meio ambiente, é uma dura realidade que as pessoas estão vivendo nos últimos anos, o que exige rápidas mudanças para equilibrar esses efeitos impostos ao meio ambiente (MARACAJÁ *et al.*, 2014).

O ser humano sempre dependeu da utilização dos recursos naturais para sobreviver e os impactos negativos dessas extrações no meio ambiente são praticados há muitos anos (RIBEIRO, 2019). Porém, de acordo com o mesmo autor, a industrialização, urbanização e o crescimento rápido da população intensificaram ainda mais as ações explorativas e os danos ocasionados, formando discussões sobre os efeitos da poluição e diversos outros problemas, mas com estes sendo iniciados somente após essa industrialização.

Nesse contexto, de produção de alimentos em larga escala, é necessário a utilização de produtos altamente tóxicos, com a finalidade de controlar doenças, patógenos e aumentar a produtividade. Entretanto, os mesmos não foram inventados inicialmente para isto. Segundo Romani (2017), estes compostos foram elaborados durante a Primeira Guerra Mundial e utilizados também na Segunda Guerra Mundial, como armas químicas.

Os agroquímicos iniciaram sua caminhada no país a partir da Revolução Verde, sendo introduzidos, fortemente, através dos pacotes tecnológicos avançados (OPAS/OMS, 1996). O objetivo era a redução dos inimigos naturais, porém, acabou gerando sistemas cada vez mais resistentes devido ao seu uso inapropriado (OPAS/OMS, 1996).

A Organização das Nações Unidas (ONU), através do seu programa relacionado as questões da agricultura e alimentação *Food and Agriculture Organization* (FAO), definiu agroquímico, segundo o estudo realizado por Peres e Moreira (2003), como sendo:

(...) qualquer substância, ou mistura de substâncias, usadas para prevenir, destruir ou controlar qualquer praga – incluindo vetores de doenças humanas e animais, espécies indesejadas de plantas ou animais, causadoras de danos durante (ou interferindo na) a produção, processamento, estocagem, transporte ou distribuição de alimentos, produtos agrícolas, madeira e derivados – ou que deva ser administrada para o controle de insetos, aracnídeos e outras pestes que acometem os corpos de animais de criação (PERES; MOREIRA, 2003, p.24).

Segundo relatório da ABRASCO (CARNEIRO *et al.*, 2015), a partir de 2008 o Brasil se tornou o país com maior consumo de agrotóxicos do mundo. O relatório mostra que impactos gerados na saúde são enormes, atingindo grandes áreas e envolvendo diferentes grupos de pessoas, trabalhadores em vários setores, até moradores nos arredores de agroindústrias e fazendas, além dos consumidores destes produtos, pois o atual modelo de produção visa principalmente a produção de bens primários para exportação.

Mesmo existindo técnicas sustentáveis para produção que garantem uma qualidade de vida ao ser humano e ao ambiente, são escassas as políticas públicas que buscam reduzir o uso de agroquímicos no país (ALMEIDA, 2019).

De acordo com dados divulgados pela Agência Nacional de Vigilância Sanitária (Anvisa) e pelo Observatório da Indústria dos Agrotóxicos da Universidade Federal do Paraná, em abril de 2012, enquanto o mercado mundial de agrotóxicos cresceu 93% nos últimos 10 anos, o Brasil teve um acréscimo de 190% (CARNEIRO *et al.*, 2015). Este mercado movimentou, durante a safra de 2010/11, 936 mil toneladas de produtos, alcançando US\$8,5 bilhões, sendo as lavouras de soja, milho, algodão e cana-de-açúcar as principais consumidoras (ANVISA, 2016).

Com o avanço das áreas plantadas e a busca por maior produtividade nos cereais, surgem também riscos de epidemias de doenças foliares, muito em função da uniformidade genética de algumas culturas em relação a fungos, e a contaminação do grão quanto ao uso excessivo de fungicidas (DORNELLES, 2016). Segundo Vieira (2021), relacionar eficiências com baixo custo na produção de alimentos é um dos fatores de maior relevância para o conceito de produtividade.

A Tabela 1 mostra o crescimento do consumo de agroquímicos e fertilizantes na agricultura do país, sendo este proporcional ao aumento de monoculturas, dependentes desses produtos.

Tabela 1 – Consumo de agrotóxicos e fertilizantes nas lavouras do Brasil.

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Agrotóxicos (mi. de L)	599	643	693	706	687	686	673	725	827	852
Fertilizantes (mi. de kg)	4.91	5.38	6.21	6.55	6.17	6.07	6.24	6.47	6.49	6.74

Fonte: CARNEIRO *et al.*, 2015.

De acordo com relatório de comercialização de agrotóxicos apresentado pelo governo federal em dezembro de 2022, foram comercializadas 720,87 mil toneladas de ingredientes ativos em 2021, representando um aumento de 5,03% em relação ao ano anterior (BRASIL, 2022). Os principais ingredientes vendidos foram: glifosato, 2,4-D, Mancozebe, Atrazina, Acefato, Malationa, Cletodim, Enxofre e S-metolacoloro. Em 2020, os herbicidas ocuparam a primeira posição, com 58,45% do total, seguido pelos fungicidas (12,06%) e os inseticidas (10,1%), onde as demais categorias tiveram 19,39% do consumo nacional, com esses percentuais se mantendo estáveis nos últimos anos (RESENDE, 2020).

Esse uso desenfreado preocupa autoridades públicas e os sistemas de saúde. Um estudo realizado em todos os 26 estados brasileiros pelo Programa de Análise de Resíduos de Agrotóxicos da Anvisa no ano de 2011, coletando amostras de alimentos consumidos diariamente pelas pessoas, mostrou que um terço está contaminado com agrotóxicos (ANVISA, 2016).

O avanço tecnológico na agricultura foi importante para a produção de alimentos. O problema é a falta de conhecimento ou preparo por parte dos agricultores, que não conseguiram se aperfeiçoar a ponto de manejar estes produtos da forma adequada. O uso dos defensivos agrícolas de forma errada, em doses inadequadas, pode trazer prejuízos ao meio ambiente e a saúde dos seres vivos, além de aumentar o custo de produção e acarretando em perdas desnecessárias (LIGNANI; BRANDÃO, 2022).

4.2.1 Tratamento de Sementes

No âmbito da produção de sementes, a realização em ambientes inóspitos faz com que essas sejam portadoras das mais severas doenças, como por exemplo fungos de solo e mofo-branco em regiões produtoras de soja no Brasil, lembrando que o segundo requer medidas drásticas para ser controlado, como várias aplicações de fungicidas e/ou manejo diferenciado do solo (RAVA *et al.*, 2006). Soma-se a isso o tratamento de sementes já infectadas, onde mesmo usando mais de um produto, não elimina totalmente os patógenos presentes, caminhando a passos largos para a utilização cada vez maior de agroquímicos nas fases posteriores ao tratamento das sementes (RAVA *et al.*, 2006).

O tratamento de sementes pode ser realizado durante o beneficiamento desta para o comércio, ou pelo próprio produtor durante a pré-semeadura. Esse processo é realizado para que ocorra maior germinação e consiga suportar os problemas no campo (ROMANI, 2017). Para tanto, são utilizados vários produtos formulados para fungos, insetos, bem como reguladores de crescimento, antibióticos, corantes e inoculantes (CARVALHO, 2012).

Para Henning *et al.* (2010):

O volume de sementes tratadas com fungicidas, que na safra 1991/92 não atingia 5% da área semeada, hoje está em torno de 90-95% da área semeada com soja, no Brasil. Na década de 80, com a expansão da soja para o Brasil Central, os produtores utilizavam mais de dois sacos de 50 kg de sementes por hectare. Hoje, com o uso de tratamento de sementes é comum o agricultor utilizar 40-50 kg de semente/ha (HENNING *et al.*, 2010, p.3).

De acordo com Romani (2017), o tratamento químico utilizado em soja é uma técnica já bastante difundida pelos agricultores, pois devido aos altos custos para implantação de uma lavoura, acaba sendo mais rentável utilizar essa forma de proteção nas sementes, para não haver problemas no futuro, como perda de produtividade e replantio. Porém, na maioria das vezes o próprio agricultor realiza esse procedimento, sem os maiores cuidados fundamentais para a proteção do indivíduo (ROMANI, 2017).

O manejo utilizando agroquímicos é realizado amplamente na agricultura, porém, apesar de auxiliar diretamente no aumento de produtividade e controle de pragas e doenças, seu uso descomedido deixa a população exposta à intoxicação e doenças relacionadas, assim como gera impactos diretos no meio ambiente (PEREIRA, 2014).

De acordo com Santos (2018), doses altíssimas de agrotóxicos nas lavouras estão colocando em risco a saúde não só dos consumidores, mas também de agricultores, prejudicando o ecossistema como um todo. Koifman *et al.* (2002) analisaram o volume consumido de agroquímicos em 11 estados brasileiros e verificaram que existe uma correlação positiva na incidência de intoxicações crônicas provocadas pela poluição desses defensivos agrícolas, sendo observada a manifestação de enfermidades como neoplasias, malformações congênitas e desregulações endócrinas.

Conforme saliente Romani (2017), o tratamento e manuseio das sementes realizado pelo produtor é quase sempre realizado de forma errada, pois esse não possui a orientação adequada, estando em risco direto de contaminação, podendo gerar problemas futuros em sua saúde. Dessa forma, a falta de informação e a negligência sobre a forma correta de manejo do material tratado tem sido a causa de muitas doenças encontradas na população rural (PEREIRA, 2014).

Contudo, ao analisar o trabalho de Costa *et al.* (2018), os autores comentam que algumas pesquisas realizadas no âmbito sementeiro demonstraram que certos produtos químicos, quando aplicados em algumas culturas, podem ocasionar redução de germinação e sobrevivência de plantas.

4.2.2 Uso de Fertilizantes

De acordo com a Lei Federal nº 6.894, de 16 de dezembro de 1980, que “dispõe sobre a inspeção e a fiscalização da produção e do comércio de fertilizantes, corretivos, inoculantes, estimulantes ou biofertilizantes, remineralizadores e substratos para plantas, destinados à agricultura”, fertilizante é toda “substância mineral ou orgânica, natural ou sintética, fornecedora de um ou mais nutrientes vegetais” (BRASIL, 1980).

Estes constituintes químicos contêm nutrientes que podem ser elencados como macro e micronutrientes. Dentro dos macros existe o nitrogênio, fósforo e o potássio, já nos micros possuem boro, cloro, molibdênio, manganês e zinco, onde por sua vez são adicionados em dosagens muito menores ao solo quando comparados aos macronutrientes (COSTA; SILVA, 2012).

Segundo o Resende (2020), o país é quarto no ranking mundial em consumo de fertilizantes, ficando atrás apenas da China, Índia e EUA, sendo o fósforo o principal nutriente utilizado, enquanto nos outros países o nitrogênio detém essa posição.

Os produtos aplicados na agricultura devem ser calculados para que a qualidade da água se mantenha estável, dentro dos padrões considerados naturais (COSTA *et al.*, 2016). Por isso, com o aumento da sojicultura, é fundamental avaliar os impactos negativos que poderão surgir no ambiente, principalmente quanto ao uso dos produtos químicos nas lavouras (COSTA *et al.*, 2014).

Nesse contexto, a pegada ecológica é um dos grandes indicadores utilizados em nível mundial, introduzido ainda na década de 90 e difundido por Wackernagel e Rees (1996). O objetivo principal é analisar a pressão exercida pelo consumo das pessoas sobre os recursos naturais (WACKERNAGEL; REES, 1996; MARACAJÁ, 2016). Por isso, se faz importante analisar nos próximos tópicos a pegada de carbono e hídrica, de modo que o leitor entenda a importância destas nesse estudo.

4.2.3 Pegada de Carbono

As grandes mudanças climáticas provocadas pela emissão de gases de efeito estufa vêm ganhando mais espaço em debates e reuniões de cientistas no mundo todo como, por exemplo, o Painel Intergovernamental de Mudanças Climáticas (IPCC) (PANDEY *et al.*, 2011). O aumento na temperatura vem ao encontro da sensibilidade apresentada pelas atividades agrícolas, com essa elevação preocupando cada vez mais os empreendimentos que dependem dessas atividades, pois poderá ocasionar estresse hídrico e incidência maior de pragas e doenças (CAVERO, 2016).

Com isso, o cultivo de grãos é visto como um sistema que engloba a fabricação de materiais ricos em energia, por meio da atividade de fotossíntese e utilização de insumos demasiadamente dependente de combustíveis fósseis, fertilizantes e pesticidas (PIMENTEL; PIMENTEL, 2006). Na sustentabilidade, o uso de energia é o fator essencial para o desenvolvimento econômico, pois há uma dependência do combustível fóssil, sendo este causa primária de problemas ambientais como o aquecimento global (COWELL; PARKINSON, 2003).

Em consequência disso, surgiu a necessidade de determinar quais são as emissões totais desses gases, associados a produção agrícola, que contribuem para o aquecimento global (CARMO *et al.*, 2016). A pegada de carbono consegue fazer essa quantificação, tornando-se uma das métricas na avaliação do ciclo de vida no sistema de produção agrícola (ROTZ *et al.*, 2010). No país, é mais comum essa determinação em setores como o da construção civil (PASSUELLO *et al.*, 2014).

A Pegada de Carbono foi criada com o objetivo de controlar as emissões dos gases de efeito estufa, utilizada para medir os impactos causados por esses gases, e empregada em diversas pesquisas acadêmicas (PANDEY *et al.*, 2011). Esse indicador é muito importante para estimar os níveis de poluição de modo que se consiga diminuí-los, servindo também como parâmetro ambiental em vários setores, pois quantifica o impacto negativo das emissões de carbono no ambiente (HERTWICH; PETERS, 2009).

Para simplificar o contexto em que está inserido, a pegada de carbono indica as consequências das ações humanas sobre o ambiente, em especial, nas mudanças climáticas, consistindo na medida das emissões dos gases que provocam tal condição em toneladas de carbono, por atividades de ação direta das pessoas ou pela manufatura e ciclo de vida de produtos ou serviços (WIEDMANN; MINX, 2007; MARACAJÁ, 2016; GONÇALVES *et al.*, 2018).

É importante enfatizar que existem outras “pegadas”, entretanto, a pegada de carbono está sendo colocada como a mais importante nos últimos anos, conforme GALLI *et al.* (2012) comentam em seu trabalho. Montibeller Filho *et al.* (2012) salientam que essa pegada auxilia os tomadores de decisão a adquirirem um ponto de vista mais socioambiental.

A soja ganha cada vez mais relevância no agronegócio brasileiro, devendo em grande parte às exportações, elevada produtividade, baixo custo de produção e aumento desenfreado em áreas amazônicas (FAGUNDES *et al.*, 2013). O Brasil se tornou o maior produtor do grão a partir de 2011, onde o principal destino do grão exportado é para alimentação humana e animal (VASCONCELLOS; BELTRÃO; PONTES, 2016).

Assim, é de suma importância o levantamento da pegada de carbono advindo da produção sojícola. Em relatório apresentado pela Bayer, na safra 2021/22 os agricultores que participam do Programa Pro Carbono alcançaram uma pegada média de 783kg CO₂ eq por tonelada (et/t) de soja, significando cerca de 80% menor em comparação com outros países, como o EUA (1.200kg et/t) (BAYER, 2022). Isso demonstra a importância do emprego de programas e ações que visam a redução dos impactos ambientais sem comprometer a produção agrícola.

4.2.4 Pegada Hídrica

A Pegada Hídrica avalia o uso de água na manufatura de um produto ou realização de um serviço (RIBEIRO, 2019). Esta é definida como volume total de água usado para produzir bens e serviços consumidos por um indivíduo ou comunidade (HOEKSTRA; CHAPAGAIN, 2008; HOEKSTRA *et al.*, 2009; HOEKSTRA *et al.*, 2011).

Apesar do método de medição diferir da pegada de carbono, algumas características são comuns em ambas, como o uso de recursos naturais pelas pessoas e o estresse ambiental causado por isso (HOEKSTRA, 2009; HOEKSTRA *et al.*, 2011).

A pegada hídrica identifica o impacto do consumo humano sobre a quantidade global de água doce (BLENINGER; KOTSUKA, 2015), além de analisar a poluição na mesma, sendo expressa em volume anual para uma certa região demográfica (HOEKSTRA; CHAPAGAIN, 2007).

O conceito foi introduzido por Arjen Hoekstra, em 2002, durante uma conferência que reuniu peritos internacionais para abordar o comércio de água virtual, realizado na Holanda (RIBEIRO, 2019). O termo proposto foi correlacionar com a pegada ecológica já que esta é expressa em hectares e a pegada hídrica em volume de água doce (SILVA *et al.*, 2013).

Porém, segundo Hoekstra *et al.* (2011), o conceito de pegada hídrica é diferente do consumo usual de água, pois considera também o uso indireto. Mas, de acordo com Ribeiro (2019), esta não considera se a água retorna ao meio da qual foi retirada e, sim, contempla a água originária das chuvas, umidade de solo e água poluída.

Esse indicador possui uma visão ampla sobre a relação do consumidor com o produtor, não sendo considerada uma medida de impacto ambiental local relacionado a água, que, por sua vez, depende da vulnerabilidade do sistema hídrico da área, bem como a quantidade de consumidores e poluidores existentes (GUIMARAES *et al.*, 2020).

Por conseguinte, essa fornecerá dados que servirão para construção de discussões acerca do uso da água em um determinado local, servindo, também, como ferramenta para avaliar os impactos ambientais daquela área (HOEKSTRA *et al.*, 2011).

Segundo dados da Agência Nacional de Águas (ANA, 2021), o Brasil obteve cerca de $2.831,65 \text{ m}^3 \cdot \text{s}^{-1}$ de vazão retirada (89,36 trilhões de L.ano⁻¹), contabilizando a evaporação líquida, onde a maior contribuição foi para a irrigação no setor agrícola, respondendo por 50%. Supõe-se que haverá aumento de 42% na extração desse recurso até 2040, sendo mais 26 trilhões de litros retirados dos mananciais no ano (115,36 trilhões de L.ano⁻¹) (ANA, 2021).

Ao comparar esses dados com a vazão média das Cataratas do Iguaçu, por exemplo, os números são pouco superiores. Porém, a preocupação no uso desenfreado desse recurso natural se deve a prevenção e cuidado antes de ocorrer a escassez, pois essas previsões podem ser aceleradas devido a fatores econômicos favoráveis e alterações mais profundas nos setores. Além disso, mudanças climáticas podem ocasionar a redução de disponibilidade hídrica em muitos lugares. Segundo a ANA (2021), a irrigação pode ter acréscimo em 20% na demanda em relação aos números expostos acima, em função do cenário mais crítico de ocorrência de chuvas.

Os dados reforçam a urgente necessidade de criar ações para planejar melhor o uso da água e, conseqüentemente, a segurança hídrica, evitando o surgimento de crises nesse setor, proporcionando o uso múltiplo da água, pois a utilização errada acarreta mudanças climáticas devido às alterações no ciclo da água (ANA, 2021).

Nessa linha, a pegada hídrica sempre considera três componentes no cálculo: azul, verde e cinza (GUIMARAES *et al.*, 2020). A Azul está relacionada ao consumo superficial e subterrâneo da água (RIBEIRO, 2019). A pegada verde representa a parcela oriunda da precipitação de chuvas (HOEKSTRA, 2011). Por fim, a PH cinza indica a poluição da água utilizada em determinada atividade, sendo o uso indireto do recurso (SILVA *et al.*, 2015).

Hoekstra *et al.* (2011) e Furlan e Palhares (2017) definem pegada hídrica cinza como sendo o volume essencial para converter toda a carga de poluentes às condições normais e/ou para padrões legais. Assim, no contexto agrícola, isso está relacionado com a aplicação de qualquer produto químico no solo, pois estes serão absorvidos em parte pelas plantas, ficando a outra parcela nos processos de lixiviação até o lençol freático ou escoamento para fontes de água.

De acordo com Ribeiro (2019), os poluentes químicos considerados na agricultura geralmente são compostos por fertilizantes macro e micronutrientes (já exposto acima) e agrotóxicos. Contudo, a pegada hídrica cinza oriunda das atividades agrícolas, pode ser amplamente reduzida adotando processos orgânicos ou sustentáveis, limitando o uso de substâncias químicas nocivas ao meio ambiente (HOEKSTRA *et al.*, 2011). “Portanto, a metodologia da pegada hídrica surge como um importante indicador e planejamento do uso eficiente da água na cadeia produtiva da soja” (COSTA *et al.*, 2016, p.3).

4.3 Exigências Legais na Análise de Sementes

Segundo Lopes e Nascimento (2009), a análise de sementes passou por um processo de evolução, iniciando entre 1900 e 1920, com o desenvolvimento de técnicas e procedimentos para testar a germinação destas. Em 1967, o Ministério da Agricultura oficializou as “Regras para Análise de Sementes – RAS”, baseado em regras internacionais, com a finalidade de melhor atender as possibilidades dos Laboratórios de Análise de Sementes brasileiros, especificando métodos padrões para obtenção de amostras e execução de testes de pureza física, espécies nocivas, germinação, tetrazólio, grau de umidade, sanidade de sementes e outros.

A RAS especifica métodos padrões e definições para o comércio internacional, por isso é extremamente necessário a precisão e acurácia nos testes. A produção de sementes passa por uma análise com dois objetivos: atender às exigências para a comercialização das sementes e controle de qualidade da produção, por isso que a RAS indica quais são os procedimentos para a obtenção de amostras, bem como para a execução dos testes (LOPES; NASCIMENTO, 2009).

A Instrução Normativa nº 45 de 17 de setembro de 2013, do Ministério da Agricultura, Pecuária e Abastecimento, descreve as exigências legais para comercialização de um lote de sementes de soja no Brasil. A pureza física mínima deve ser de 99%, germinação mínima de 75%, existência de outras espécies no lote não deve ser maior que 0,1%, número de sementes silvestres e nocivas deve ser zero (BRASIL, 2013).

4.4 Métodos para Análise de Sementes

Conforme exposto por Lopes e Nascimento (2009), a amostragem tem o objetivo de obter uma parcela da quantidade total do lote para ser analisado, sendo a primeira e a mais importante etapa dentro do processo de análise, já que se não for realizada da forma correta, compromete todas as seguintes. Os autores ainda comentam que a amostra é estabelecida em função da quantidade de sementes do lote e do tipo de embalagem, onde os resultados obtidos pela análise só terão validade se realizada em amostra representativa do lote.

A germinação é um fenômeno biológico que os botânicos consideram como a retomada do crescimento do embrião com o subsequente rompimento do tegumento pela radícula. Porém, de acordo com os tecnologistas em sementes, esse processo é definido como a emergência e o desenvolvimento das estruturas essenciais do embrião, manifestando sua capacidade de gerar uma nova plântula normal, sob condições favoráveis (NASSIF *et al.*, 1998).

O teste de germinação é o procedimento oficial para avaliação da capacidade das sementes produzirem plântulas normais em condições ideais, mas nem sempre indica diferenças de desempenho entre lotes de sementes durante o armazenamento ou em campo (CARVALHO; NAKAGAWA, 2000).

Já o teste de pureza física tem o objetivo de determinar a composição da amostra em exame e, conseqüentemente, a composição do lote de sementes, identificando as espécies e as partículas inertes que constituem a amostra (LOPES; NASCIMENTO, 2009).

Segundo Baalbaki (2009), vigor é a propriedade da semente que determina o potencial para uma emergência rápida e uniforme, desenvolvendo plântulas normais sob ampla diversidade de condições de ambiente. É o vigor da semente que permite a expressão do seu total potencial de produzir uma planta de alto desempenho agrônômico em condições de estresse, como profundidade excessiva de semeadura, compactação superficial do solo, excesso de chuvas ou estiagem prolongada após a semeadura ou ocorrência de baixas temperaturas (KRZYZANOWSKI *et al.*, 2018).

Alguns testes de vigor simulam as condições de campo, como o teste de frio, de velocidade e do percentual de emergência. Os resultados do teste indicam se o lote possui potencial de apresentar bom desempenho a campo se as condições ambientais não forem favoráveis (LOPES; NASCIMENTO, 2009).

Além de possibilitarem uma melhor germinação, expressando o máximo potencial produtivo, as sementes com alto vigor garantem também rápido desenvolvimento das plantas, mesmo em condições adversas relacionadas ao clima ou ataque de pragas. Além disso, elas reduzem os problemas na lavoura com ervas daninhas, pois as plantas na linha de semeadura preenchem o espaço mais cedo e suprimem o crescimento das mesmas, acarretando uma diminuição no uso de agroquímicos que possam vir a contaminar ainda mais o meio ambiente (BRASMAX, 2018).

Diversos testes podem ser utilizados para determinação do vigor na soja. Os mais utilizados são de tetrazólio e o envelhecimento acelerado. O teste de tetrazólio, conforme a metodologia proposta por Salton (2012), é o único método que estabelece uma classificação de vigor para os lotes de sementes, informação essa que ainda não é disponibilizada para os demais testes.

Assim, avaliação da qualidade fisiológica de sementes é frequentemente utilizada para determinar o desempenho dos lotes que serão comercializados, sendo a mensuração deste conjunto de parâmetros avaliada pela associação entre viabilidade e vigor (DELARMELO, 2012).

É de suma importância a obtenção de sementes com alta qualidade, pois assim será alcançado o máximo potencial de rendimento na cultura implantada. Segundo Menten e Moraes (2006), essa alta sanidade é fundamental para garantir boa qualidade na produção e manejo sustentável dos cultivos, onde a ocorrência de doenças pode reduzir em até 15% a produção de grãos.

Dois pontos relacionados às sementes e o manejo de doenças devem ser salientados: o primeiro é a possibilidade do transporte de patógenos pelas sementes, espalhando para lavouras vizinhas, debilitando plantas e, até mesmo, causando epidemias. O segundo é no que diz respeito ao tratamento de sementes, processo realizado para preservar ou aprimorar o seu desempenho. Sementes de alta qualidade reduzem a utilização de produtos químicos e, também, as atividades a serem realizadas no objetivo de conseguir uma lavoura de alto padrão com plantas saudáveis, que muitas vezes geram alto custo de investimento e baixo retorno financeiro (MENTEN; MORAES, 2006).

Na condução de estudos com testes de vigor de sementes é necessário que os lotes utilizados tenham qualidade fisiológica semelhante e que sejam comercialmente aceitos, visto que o principal objetivo destes testes é complementar os resultados de germinação, identificando diferenças de qualidade entre lotes com o potencial de germinação similar (MARCOS FILHO, 2005).

Em consequência disto, se faz necessário a implantação de novas ferramentas que auxiliem empresas e produtores a conseguir padronizar esses lotes. Nos próximos tópicos será explanado sobre a entrada da Inteligência Artificial nesse setor.

4.5 Inteligência Artificial: Aplicação na Agricultura e no Setor Sementeiro

O aprendizado de máquina (em inglês, *machine learning*) é um método de análise de dados que automatiza a construção de modelos analíticos. É um ramo da inteligência artificial baseado na ideia de que sistemas podem aprender com dados, identificar padrões e tomar decisões com o mínimo de intervenção humana (HUI LI, 2021; VAN DIJK *et al.*, 2021; YOOSEFZADEH-NAJAFABADI *et al.*, 2021a; YOOSEFZADEH-NAJAFABADI *et al.*, 2021b).

O objetivo dessa técnica de aprendizagem é desenvolver ferramentas e sistemas computacionais para adquirir conhecimento de forma automática, tomando decisões baseadas nas experiências anteriores, acumuladas através das soluções bem-sucedidas (MONARD; BARANAUSKAS, 2003).

O aprendizado de máquina passou por diversas transformações ao longo do tempo, nascendo do reconhecimento de padrões e da teoria de que computadores podem aprender, sem a intervenção humana em todo o tempo de trabalho, para realizar tarefas específicas. Pesquisadores interessados em IA queriam saber se as máquinas poderiam simular o comportamento do cérebro humano, expondo novos dados e repetindo o processo várias vezes, para saber se os computadores eram capazes de se adaptar independentemente, aprendendo com o histórico de processos para produzir decisões e resultados mais confiáveis e precisos, passíveis de repetição (HUI LI, 2021).

Enquanto a IA é a capacidade de uma máquina avaliar uma situação e em seguida tomar uma decisão através de uma busca no seu banco de dados ou séries históricas (FINLAY, 2020), a aprendizagem de máquina é uma vertente específica da IA, que treina máquinas para aprender com dados (HUI LI, 2021). Segundo Soam e Raghuphati (2018), a IA é bastante utilizada na área de pós-colheita para minimizar perdas e simplificar o transporte.

De acordo com Jha *et al.* (2019), o único propósito da aplicação dessa técnica é alimentar um sistema de dados com histórico criado por experiências e dados estatísticos para que ela possa executar sua tarefa atribuída e assim resolver o problema em questão. Isso nada mais é que uma abordagem matemática para construir máquinas inteligentes (PINHEIRO *et al.*, 2021).

De acordo com Pinheiro *et al.* (2021), essa técnica aplicada em sementes e a IA na agricultura possuem trabalhos empregados desde 1999, utilizando técnicas de modelagem e pesquisando soluções para os mais variados problemas.

A qualidade das sementes é determinada por testes de germinação e vigor (OSROUSH, 2010). Este processo é moroso e baseado em experiências e conhecimentos de peritos (NAGHASHZADEH; AZADBAKHT, 2018).

A FAO em 2017 mostrou que a população mundial precisa superar a produção de alimentos em mais de 70% até 2050, fazendo com que produtores agrícolas tenham que aumentar as terras ou o rendimento das lavouras com o auxílio de agricultura inteligente e de precisão (ELFERINK; SCHIERHORN, 2016).

Um estudo encomendado pela Microsoft, sobre as projeções dos impactos econômicos gerados pela IA, estima que em 2030 o uso dessa ferramenta em ações ambientais poderá contribuir com US\$ 5,2 trilhões para a economia global, criar 38,2 milhões de novos empregos e reduzir as emissões mundiais de gases de efeito estufa em 4%, uma quantidade equivalente, por exemplo, às emissões totais de Austrália, Canadá e Japão no mesmo ano (PWC BRASIL, 2019).

As tecnologias aplicadas atualmente no setor agrícola estão utilizando algoritmos de aprendizagem de máquina para maximizar o rendimento da colheita com diminuição nos custos da lavoura (REHMAN *et al.*, 2019), além de permitir que o agricultor aprimore os processos que envolvem a cadeia produtiva, como semeadura, previsão de safra, rendimento, aplicação de agroquímicos, irrigação e preço de venda (KAUR, 2016).

Conforme Talaviya *et al.* (2020), a IA é uma tecnologia emergente no campo da agricultura, com equipamentos e máquinas elevando o sistema agrícola de hoje a um nível diferente. Esta tecnologia melhorou o desenvolvimento agrícola e o monitoramento de colheita, agilizando o processamento e comercialização em tempo real (YANG *et al.*, 2007).

Atualmente, a utilização de máquinas e automação para processamento de tarefas está em evolução, pois o aprendizado de máquina é um dos métodos que simplifica os problemas do dia a dia. Nesse caso, o equipamento irá compreender seus resultados e assimilar conhecimentos a partir dos modelos anteriores, para que possa melhorar por conta própria, sem que necessite de orientações regulares para atualizar seu sistema (POOJA *et al.*, 2018).

Os algoritmos mais utilizados na inteligência artificial no setor agrícola são os que utilizam redes neurais e árvores de decisão. Pinheiro *et al.* (2021) comentam que a utilização de redes neurais chama mais a atenção dos pesquisadores, pois obtém respostas por meio de classificação utilizando multicamadas, trabalhando os dados brutos, agrupando-os e classificando-os, simulando assim, as conexões do cérebro humano.

Conforme Pinheiro *et al.* (2021), as soluções tecnológicas baseadas em IA, permitem aos agricultores aumentar a produção reduzindo a quantidade de insumos, melhorando a qualidade da produção e garantindo rápida colocação da safra no mercado. Estimou-se que em 2020, os agricultores estariam usando 75 milhões de dispositivos conectados, onde esse número poderá gerar cerca de 4,1 milhões de dados todos os dias (TALAVIYA *et al.*, 2020).

As recentes automatizações que utilizam robôs, máquinas agrícolas e drones contribuíram fortemente para o setor agroindustrial (PINHEIRO *et al.*, 2021). Diversos sistemas que se baseiam em computadores com alta tecnologia, são projetados para resolver vários problemas diariamente, detectando plantas daninhas a campo, rendimento de colheita e qualidade dos grãos (LIAKOS *et al.*, 2018), garantindo eficiência em todos os setores dentro da agricultura e ajudando a gerenciar os desafios enfrentados a campo e nas indústrias.

Assim, a IA entra como uma ferramenta essencial para realização de diversos testes em sementes, com alta precisão de resultados. Segundo Kumar (2018), se faz necessário a utilização da inteligência artificial para reinventar a agricultura, principalmente no setor de qualidade, ajustando os modelos para avaliar melhor os lotes de sementes.

4.6 Técnicas para Mineração de Dados

A extração de conhecimento através de um banco de dados passou a ser utilizado no processo decisório, sendo requisito importante que seja descoberto e compreensível pelos humanos, além de útil e interessante para os usuários do processo, fornecendo suporte para a tomada de decisão (FAYYAD *et al.*, 1996).

A área de mineração de dados, ou *data mining*, é multidisciplinar, incorporando técnicas utilizadas em banco de dados, IA e estatística, sendo considerada o núcleo do processo de descoberta de conhecimento (PIVATO, 2006). Ela consiste no processo de analisar grande quantidade de dados sob diferentes formas, a fim de encontrar informações úteis que normalmente não estão visíveis ou que dificilmente são descobertas (VASCONCELOS; CARVALHO, 2018).

Segundo Santos (2001, p.12), “mineração de dados consiste na aplicação de análise de dados e algoritmos de descobrimento que produzem uma enumeração de padrões (ou modelos) particulares sobre os dados”.

A sequência se inicia com a seleção de um conjunto de dados ou amostra com os quais o processo de descoberta será realizado. Estes poderão passar por uma etapa de pré-processamento, onde serão tratados problemas de desbalanceamento dos dados, sendo o passo seguinte, a tarefa de transformação dos mesmos, podendo-se reduzir o número de variáveis sob consideração ou encontrar representações que não variam (VASCONCELOS; CARVALHO, 2018).

É comum trabalhar com dados desbalanceados, onde esses podem ser oriundos do próprio processo de geração dos dados, da forma como foram adquiridos (em trabalhos a campo, por exemplo), no momento da transformação ou até mesmo de classes rotuladas de forma errônea (MONARD; BARANAUSKAS, 2003).

O processo de descoberta de conhecimento em banco de dados compreende três principais etapas: pré-processamento, mineração de dados e pós-processamento (TAN *et al.*, 2005). No primeiro, os dados são coletados e passam por processo de tratamento, retirando valores ruidosos, garantindo a qualidade destes (SILVA, 2018).

Na sequência, ocorre a mineração dos dados, aplicando processos exploratórios na busca por padrões, previsões, erros, correlações, entre outros (AMARAL, 2016). Por fim, o pós-processamento consiste em interpretar os resultados obtidos, visualizando possibilidades e analisando as informações extraídas em relação ao objetivo da mineração (SFERRA; CORRÊA, 2003; SILVA, 2018).

Passando essas etapas, o conjunto está apto para ser processado pela principal tarefa dentro de todo o processo: a mineração. A partir disso são empregados algoritmos, muitas vezes de forma repetitiva, que procuram por padrões e regras nos dados. Por fim, as informações descobertas são interpretadas e avaliadas, muitas vezes na forma de gráficos ou relatórios, selecionando os conhecimentos úteis (VASCONCELOS; CARVALHO, 2018). Esse processo pode ser compreendido pela Figura 2.

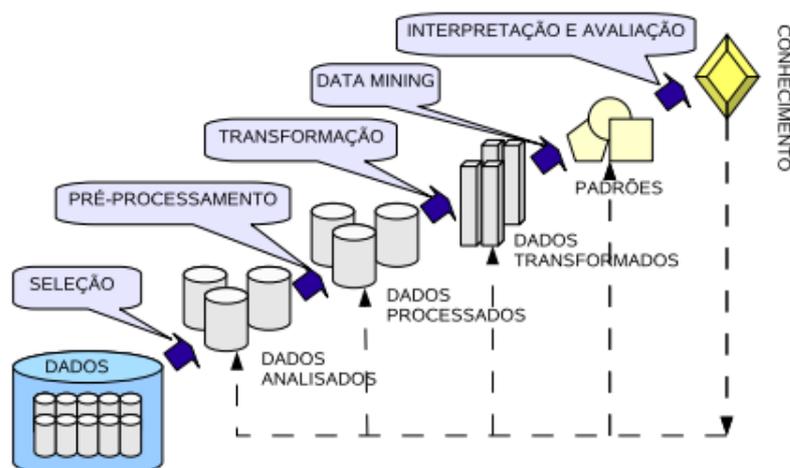


Figura 2 – Sequência de operação utilizando mineração de dados.
 Fonte: VASCONCELOS; CARVALHO, 2018.

Segundo Moraes (2020), a técnica de mineração de dados pode ser dividida em dois grupos: classificação e clusterização. As técnicas de classificação são projetadas para classificar amostras desconhecidas usando as informações fornecidas por um conjunto de amostras classificadas, onde o conjunto geralmente é chamado de conjunto de treinamento, pois é usado para treinar o algoritmo para realizar a classificação. Por exemplo, redes neurais e máquinas de vetores de suporte exploram conjuntos de treinamento para ajustar seus parâmetros, a fim de resolver um problema de classificação específico (GADOTTI *et al.*, 2022b).

Assim, mineração de dados tem se mostrado com forte potencial para avaliação e caracterização de problemas e soluções agrícolas, uma vez que permitem transformar as informações contidas em grandes volumes de dados e, portanto, são de grande aplicabilidade em estudos que utilizam séries temporais (WERNER *et al.*, 2019). As principais técnicas para mineração de dados incluem regras de classificação, *clustering* e associação (MUCHERINO *et al.*, 2009).

4.6.1 Regras de Classificação

Os filtros de classificação trabalham de forma a examinar as características de um conjunto e enquadrá-los em classes, objetivando a generalização e especialização dos dados que servem para distinguir essas classes, de modo a prever os dados ou registros não classificados automaticamente (VASCONCELOS; CARVALHO, 2018). Segundo Beniwal e Arora (2012), alguns algoritmos que utilizam esse conceito trabalham com árvores de decisão, redes neurais ou bayesianas, e vizinho mais próximo. Os mais pesquisados em trabalhos realizados recentemente são *Multilayer Perceptron* (redes neurais), *J48* e *Random Forest* (árvores de decisão).

4.6.2 Regras de *Clustering*

A tarefa de clusterização consiste em agrupar os dados em subconjuntos de maneira que as semelhantes sejam reunidas, descobrindo um novo conjunto de categorias ou grupos de interesse, sendo avaliados intrinsecamente (XU; WUNSCH, 2005).

Esse agrupamento é uma tarefa que procura classificar dados heterogêneos em subgrupos ou segmentos homogêneos (VASCONCELOS; CARVALHO, 2018). Os métodos utilizados nesse processo são os hierárquicos, de particionamento, baseados em densidade, baseados em modelo, entre outros (FAYYAD *et al.*, 1996).

4.6.3 Regras de Associação

De acordo com Patel e Kathiriya (2017), essa técnica é uma das mais eficientes na mineração de dados para pesquisar padrões escondidos em um grande conjunto. O objetivo da associação é encontrar relacionamentos entre os mais diversos itens em um banco de dados, descobrindo elementos que ocorrem repetidamente.

Por exemplo, a partir de um conjunto de dados em que são registradas as compras dos clientes, a estratégia para mineração usando esse método seria utilizar a seguinte regra: (cinto, bolsa) → (sapato), indicando que a pessoa que compra cinto e bolsa também compra sapato, com um certo grau de certeza, sendo esse grau definido por dois índices: o fator de suporte e o de confiança (VASCONCELOS; CARVALHO, 2018).

Segundo Vasconcelos e Carvalho (2018), o código de barras permitiu às empresas coletarem e armazenarem grande quantidade de dados referente às compras de seus clientes, conseguindo através disso, dirigirem estudos e estratégias de marketing para promover um layout e catálogos mais adaptados para alcançarem maiores resultados nas vendas.

Para Patel e Kathiriya (2017), as regras de classificação pode seguir a abordagem de aprendizado supervisionado, enquanto na tarefa de cluster, o aprendizado não-supervisionado é mais rotineiramente utilizado.

A partir da Figura 3, é possível entender quais são as etapas a serem seguidas após a inserção dos dados no software de aprendizagem.

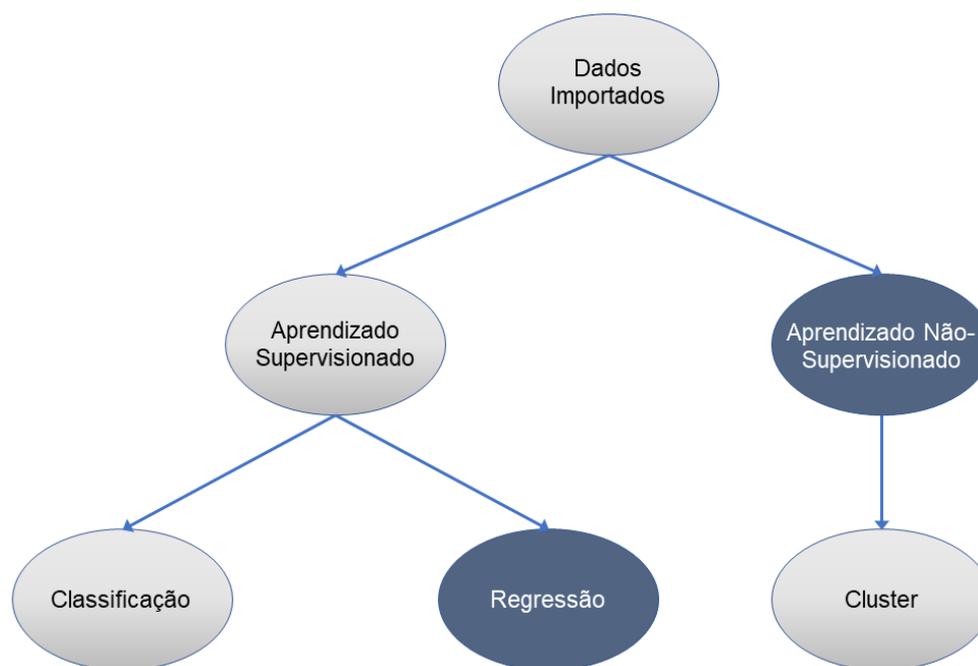


Figura 3 – Processamento dos dados a partir do tipo de aprendizagem.
Fonte: Adaptado de MONARD; BARANAUSKAS, 2003.

4.6.4 Aprendizado Supervisionado

De acordo com Monard e Baranauskas (2003), essa forma de aprendizagem permite ao algoritmo treinar utilizando um conjunto de dados onde o atributo de classe é conhecido, objetivando construir um classificador que consiga determinar corretamente a classe de outros dados não rotulados, predizendo com base nas características aprendidas na fase de treinamento qual classe cada conjunto se enquadra. Para rótulos de classe discretos, esse problema é conhecido como classificação e para valores contínuos como regressão.

Durante a fase de treinamento é fundamental que seja informado, previamente, a resposta esperada, para que o sistema, ao final dos testes, possa comparar os resultados obtidos, emitindo a precisão e eficiência do classificador escolhido em relação ao conjunto de dados em que foi submetido (SOUZA, 2011).

Após a realização dos testes, o classificador deve ser avaliado e poderá ser ajustado através de seus hiper parâmetros, como: número de repetições; atributos a serem utilizados ou excluídos; uso de filtros auxiliares que norteiam o classificador; entre outras configurações. Isso é feito para que consiga se adaptar da melhor forma aos dados (MONARD; BARANAUSKAS, 2003).

Alguns modelos nesse campo de aprendizagem utilizam redes neurais. Os algoritmos que utilizam esse método são amplamente empregados nos setores sementeiros (CHAUGULE, 2021). A estrutura do classificador desse gênero é estimulada pela capacidade de processar da mesma forma que o cérebro humano faz (GHAFARI; NAJAFABADI, 2022).

Um dos algoritmos mais utilizados na ciência de dados é o *Multilayer Perceptron* (MLP), desenvolvido por Pal e Mitra (1992). O MLP utiliza processo computacional não linear, sendo altamente eficiente para a classificação e regressão de dados complexos (CHEN; WANG, 2020; HESAMI *et al.*, 2020). O mesmo foi criado para resolução de problemas de classificação utilizando neurônios ocultos estruturados em camadas, onde estes por sua vez processam as informações obtidas das camadas anteriores e enviam o conhecimento gerado para as camadas da frente, para assim, chegar a uma resposta frente ao problema exposto (HECHT-NIELSEN, 1990).

4.6.5 Aprendizado Não-Supervisionado

Já no aprendizado não-supervisionado, o algoritmo analisa o padrão que os dados seguem, efetuando o agrupamento destes em clusters (CHEESEMAN; STUTZ, 1990). Após essa determinação é fundamental analisar se cada grupo formatado pelo filtro utilizado está correto, com base em uma possível classificação realizada por um especialista (MONARD; BARANAUSKAS, 2003).

4.6.6 Software WEKA

O software *Waikato Environment for Knowledge Analysis* (WEKA) é um ambiente para mineração de dados desenvolvido pela Universidade de Waikato, na Nova Zelândia (EIBE *et al.*, 2020). Este tem o objetivo de utilizar técnicas de aprendizagem de máquina a partir de algoritmos oriundo de diversas abordagens já estabelecidas, sendo o maior benefício de utilizar essa ferramenta a gama de algoritmos disponíveis (SILVA, 2018).

O Weka utiliza linguagem Java, que possui como maior vantagem a possibilidade de portabilidade, podendo ser implementada em diferentes sistemas operacionais, além de ser um software livre e de código aberto (SILVA, 2018).

Alguns modelos implementados no WEKA:

- **Regras de classificação** - Árvore de decisão, regras de aprendizagem, Naive Bayes, aprendizado baseado em instância, regressão lógica, *Multilayer Perceptron*, SVM.

- **Regras de *Clustering*** - Regressão linear, geradores de árvores modelo, regressão local de pesos, aprendizado baseado em instância, tabela de decisão, *Multilayer Perceptron*.

- **Regras de Agrupamento** - EM, Cobweb, SimpleKMeans, DBScan.

- **Regras de Associação** - Apriori, FPGrowth, PredictiveApriori, Tertius.

Além do mais, o software conta com uma variada quantidade de algoritmos que possibilitam a realização do pré-processamento dos dados, organizando os mesmos, que podem ser importados para a plataforma em formato de planilhas ou arquivos de texto (SILVA, 2018).

5 Material e Métodos

Foram utilizados dados provenientes de testes realizados nos laboratórios de análises de sementes da Universidade Federal de Pelotas – RS e também de empresas produtoras de sementes.

Para o processamento e predição dos lotes, inicialmente foi necessário a realização de pré-processamento dos dados, de modo a preparar o conjunto, para que a ferramenta realizasse a correta leitura e análise. Nessa etapa, os dados foram recebidos em formato .xls (Excel), sendo necessário colocar todos os atributos em uma única linha e, cada valor, em colunas, abaixo do seu respectivo atributo. Posteriormente, foi convertido o arquivo para formato .csv, com isso o conjunto de dados foi executado com auxílio do software Bloco de Notas do Microsoft Windows, necessitando a substituição das “vírgulas”, quando o valor atribuído for decimal (número com vírgulas), para “pontos” e dos “ponto e vírgula”, que realizam a divisão das colunas dos atributos, para “vírgulas”. Além disso, linhas com valores faltantes ou dados considerados equivocados foram excluídos durante esse processo prévio.

Os classificadores aplicados foram J48, Random Forest, CVR, IBk, MLP, NãiveBayes. Foi utilizada a validação cruzada para o aprendizado dos algoritmos, dividindo o conjunto de dados, treinamento e teste, em 10 subconjuntos (10 *folds*), onde os dados são subdivididos em dez partes, utilizando nove para treino e uma para teste, repetindo esse processo por dez vezes (*folds*), sempre alterando as partes utilizadas para treino e teste. Essa técnica reduz a probabilidade de que coincidências subavaliem ou sobreavaliem o desempenho, para uma determinada configuração. A média dessas precisões correspondeu ao desempenho do algoritmo sobre o conjunto de dados fornecido.

Para certificar a precisão dos algoritmos foram utilizadas as seguintes métricas de avaliação: Acurácia, Precisão, Recall, F-measure e Área ROC, de acordo com Lever *et al.* (2016).

Através da matriz de confusão, foram extraídos os valores de verdadeiros positivos (TP), falsos positivos (FP), verdadeiros negativos (TN) e falsos negativos (FN), com a finalidade de calcular as métricas Recall e Precisão, através das Equações 1 e 2, como proposto por Medeiros *et al.* (2020). Com os resultados obtidos determinou-se a melhor técnica de aprendizagem.

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FP}) \quad (1)$$

Onde:

TP = Verdadeiros Positivos

FP = Falsos Positivos

$$\text{Precisão} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (2)$$

Onde:

TP = Verdadeiros Positivos

TN = Verdadeiros Negativos

FP = Falsos Positivos

FN = Falsos Negativos

Para a tarefa de mineração de dados, utilizando os métodos de aprendizagem de máquina, foi utilizado o software Weka, versão 3.8.5.

Os dados de análises de sementes recebidos são desbalanceados por natureza, principalmente os oriundos das empresas que primam por alta qualidade de lotes. Com intuito de resolver esse problema e não tendenciar o algoritmo, utilizou-se o filtro Resample, um filtro de instância não supervisionada que mantém a distribuição de classe na subamostra, onde, alternativamente, pode ser configurado para enviar a distribuição de classes para uma distribuição uniforme (GADOTTI *et al.*, 2022a,b). A amostragem pode ser realizada com (padrão) ou sem reposição (WITTEN *et al.*, 2011).

A aprendizagem desequilibrada é um problema de classificação em que o número de observações de uma classe ultrapassa de longe o número de observações de outra classe, o que acontece em praticamente todos os bancos de dados trabalhados nas pesquisas com sementes. A técnica de sub-amostragem é a melhor entre as abordagens convencionais (SARADA; DEVI, 2019).

Testes prévios

Foram utilizados dois conjuntos de dados com diferentes parâmetros avaliados pelos laboratórios de análises. O primeiro é do Laboratório Didático de Análise de Sementes (LDAS) "Flávio Farias Rocha", da Universidade Federal de Pelotas, aqui nominado como C1. Esse conjunto possui informações de peneira, peso de mil sementes (PMS), peso hectolitro, umidade, dano mecânico, germinação, vigor e viabilidade de tetrazólio, e vigor em areia. Os testes de vigor foram conduzidos em dois momentos, imediatamente e com alguns dias de armazenamento, sendo diferenciado pelos algarismos arábicos 1 e 2 (TZ 1 e TZ 2, por exemplo). O conjunto possui 1062 lotes de sementes de soja, como segue na Tabela 2.

Tabela 2 – Descrição dos atributos analisados pela mineração de dados do conjunto C1.

Atributo	Descrição	Valor
Peneira	Peneira	{5,75 - 6,00 - 6,50 - 6,75}
Peso de Mil Sementes	PMS	Indeterminado
Peso Hectolitro	PH	{0-100}
Umidade	Umidade	{0-100}
Germinação	Germinação	{0-100}
Tetrazólio (sem armazenamento do material - 1)	Dano Mecânico	{0-100}
	Vigor TZ 1	{0-100}
	Viabilidade TZ 1	{0-100}
Areia	Vigor em Areia 1	{0-100}
Tetrazólio (com armazenamento do material - 2)	Vigor TZ 2	{0-100}
	Viabilidade TZ 2	{0-100}
Areia	Vigor em Areia 2	{0-100}
Classificação do lote	Decisão tomada	{aceito, rejeitado}

Fonte: Elaborado pelo autor, 2022.

Já o segundo conjunto é oriundo de uma junção de vários dados enviados por empresas do setor, colocados aqui como banco de dados e nominado como C2. Como as planilhas recebidas continham variáveis diferentes, foi necessário retirar algumas, analisando somente os percentuais de umidade, germinação, viabilidade e vigor de tetrazólio. Nesse conjunto, após a classificação do expert, têm 1287 lotes de sementes de soja, como segue a descrição na Tabela 3.

Tabela 3 – Descrição dos atributos analisados pela mineração de dados do conjunto C2.

Atributo	Descrição	Valor
Umidade	Umidade	{0-100}
Germinação	Germinação	{0-100}
Tetrazólio	Vigor	{0-100}
	Viabilidade	{0-100}
Classificação do lote	Decisão tomada	{aceito, rejeitado}

Fonte: Elaborado pelo autor, 2022.

Todas as informações contidas nos dados recebidos seguiram as Regras para Análises de Sementes (BRASIL, 2009). A classificação dos lotes de soja foi realizada durante essa pesquisa, pois as empresas não forneceram tais classes. Por isso, foi utilizado um *expert* na área com conhecimento científico para aprovar e rejeitar os lotes de acordo com três cenários de limitações propostas nessa pesquisa:

1. Cenário 1 → média de todas as variáveis de qualidade apresentadas pelo laboratório (germinação, vigor, viabilidade e danos), onde lotes que apresentavam resultados acima de 90% foram aceitos, abaixo disso foram colocados como rejeitados. Totalizando 708 lotes aceitos e 354 rejeitados.
2. Cenário 2 → classificação realizada a partir da variável germinação, sendo os lotes que apresentavam mais de 90% classificados como aceitos e abaixo disso rejeitados. Assim, o conjunto ficou com 751 lotes aceitos e 311 rejeitados.
3. Cenário 3 → classificação imposta a partir da viabilidade por TZ 2, onde os lotes que apresentavam 90% foram classificados como aceitos e, abaixo desse percentual, foram rejeitados. Neste conjunto, somou 789 lotes aceitos e 273 rejeitados.

Todos os cenários propostos foram utilizados no conjunto C1, certificando que os algoritmos fossem testados em uma grande quantidade de variáveis. O C2 foi utilizado para indicar quais seriam os parâmetros que devem constar em um banco de dados de sementes e quais podem ser considerados apenas como critério para elevar o campo amostral.

6 Resultados e Discussão

O objetivo do ranqueamento de lotes de sementes é fornecer informações precisas sobre os aspectos de qualidade fisiológica destas ao produtor rural, que irá buscar aqueles que estiverem com germinação e vigor próximo de 100%, o que significa que a semente é de boa qualidade, além de garantir o menor uso de agroquímicos na implantação da lavoura.

Moraes (2020) e Gadotti *et al.* (2022b) comentam que em mineração de dados, há diversas formas e modelos para predição na classificação desses lotes, utilizando diferentes algoritmos de aprendizagem de máquina. Esses modelos são criados a partir de listas para os quais os valores de classe são conhecidos, e essas listas são por sua vez obtidas a partir de sistemas reais que realizam a tarefa para a qual se deseja um modelo (GADOTTI *et al.*, 2022b).

De acordo com Pascuali (2012) e Costa *et al.* (2018), a qualidade do material genético a ser utilizado é fundamental no desenvolvimento da cultura, gerando plantas mais saudáveis, com elevada uniformidade na população, não havendo ataque de patógenos e maior capacidade de competição com outras plantas invasoras. Esses fatores são determinantes para a redução de fertilizantes e agrotóxicos a serem aplicados na lavoura, durante todo o desenvolvimento da cultura, diminuindo impactos negativos no ambiente.

Impondo o Cenário 1 (média das variáveis, com mínimo de 90% para lotes aceitos) de classificação, o Random Forest (98,8%) e o MLP (99,8%) obtiveram as melhores respostas, conseguindo prever a classificação dos lotes com ótima precisão. Na Tabela 4 pode ser visto o percentual de acerto de cada algoritmo para classificação das sementes de soja.

Tabela 4 – Acurácia dos algoritmos após a classificação no Cenário 1 com dados C1.

Algoritmo	Classificação Correta dos Atributos (%)
IBk	96,9
J48	95,5
CVR	97,9
MLP	99,8
NaiveBayes	94,5
Random Forest	98,8

Fonte: Elaborado pelo autor, 2022

Porém, é importante salientar que, em aprendizado de máquinas, percentuais de acerto próximos de 100% são reavaliados, pois pode se tratar de um super ajustamento do modelo frente o conjunto de dados proposto, mesmo utilizando técnicas de aprendizagem que reduzem esse risco.

Desta forma, a Matriz de Confusão se torna importante para avaliar melhor os erros e acertos dos classificadores empregados, para, a partir disso, conseguir escolher uma técnica que possua classificação mais realista. Essa matriz expressa em função das classes utilizadas, exibindo a distribuição dos dados de acordo com as classes reais e preditas pelo algoritmo, para comparar se o dado de uma determinada categoria foi classificado corretamente ou entendido como sendo de outra, pelo modelo computacional proposto (IBM, 2021).

Analisando as matrizes de confusão nas Tabelas 5 e 6 é possível visualizar que o MLP obteve alto desempenho, sendo que em trabalhos como de Gadotti *et al.*, (2022a) e Gadotti *et al.*, (2022b) esse mesmo classificador não foi escolhido pelos autores, pois ficou abaixo dos outros utilizados. A diferença aqui fica por conta da quantidade de lotes analisados, onde o algoritmo de Redes Neurais (RNs) se adaptou melhor com uma grande quantidade, como citado por Silva *et al.* (2012), onde os autores comentam como sendo a principal vantagem das RNs. A Tabela 5 mostra a matriz de confusão do algoritmo MLP.

Tabela 5 – Matriz de confusão do algoritmo MLP para o cenário 1 e dados C1.

		Predição	
		Aceito	Rejeitado
Classe Real	Aceito	706	2
	Rejeitado	0	354

Fonte: Elaborado pelo autor, 2022.

A partir da matriz acima, é possível visualizar que para a classe de Rejeitados, o classificador se mostrou extremamente assertivo. Como dito acima, pode ter acontecido um super ajustamento do algoritmo nesse processo.

Entretanto, de acordo com Alves (2020), o MLP ganha espaço no campo de dados supervisionados, sendo muito utilizado, por exemplo, no reconhecimento e classificação de idiomas em arquivos de áudio, além de estudos preditivos em séries temporais sobre poluição ambiental e evoluções de preços.

A Tabela 6 mostra a matriz de confusão do modelo Random Forest, onde esse utiliza uma série de árvores decisórias para classificar o lote em questão, montando uma floresta, para ao final efetuar uma espécie de eleição entre os resultados que mais aconteceram, mostrando a resposta que mais ocorreu quando se realizou a predição desse lote.

Tabela 6 – Matriz de confusão do algoritmo Random Forest para o cenário 1 e dados C1.

		Predição	
		Aceito	Rejeitado
Classe Real	Aceito	703	5
	Rejeitado	8	346

Fonte: Elaborado pelo autor, 2022.

Aqui, o classificador apresentou mais erros se comparado com o MLP, resultando numa classificação mais realista, do ponto de vista técnico. O Random Forest conseguiu separar os lotes com boa precisão e acurácia, com alguns erros que são considerados normais, diminuindo o risco de subavaliação ou vício para esse banco de dados.

Gadotti *et al.* (2022b) mostraram alta precisão na predição de classificação em lotes de sementes de milho, onde encontram-se três modelos computacionais com 100% de acertos, ao utilizar duas classes (aceitos e rejeitados), revelando ser mais fácil para o aprendizado de máquina, predizer os lotes com apenas duas classes, utilizando o limite mínimo nos percentuais de qualidade imposto pela legislação.

Em relação as métricas de avaliação, quando analisado a classe de aceitos, o classificador MLP obteve 99,7% de recall, 99,9% de F-Measure e 100% de ROC Area. O algoritmo Random Forest chegou a 99,3% de recall, 99,1% de F-Measure e 99,9% de ROC Area.

A Tabela 7 mostra os parâmetros de acurácia de cada algoritmo, de modo a embasar a melhor decisão de qual algoritmo se adapta no momento de auxiliar a classificação de sementes de soja.

Tabela 7 – Acurácias dos diferentes algoritmos utilizados, sendo Recall (sensibilidade), Precisão, Curva ROC (*Receiver Operating Characteristic*) e *F Measure* para o cenário 1 e dados C1.

Classificadores	Acurácia				
	Precision	Recall	F-Measure	ROC Area	Classe
IBk	0,972	0,982	0,977	0,970	Aceito
	0,963	0,944	0,953	0,970	Rejeitado
J48	0,971	0,960	0,966	0,954	Aceito
	0,923	0,944	0,933	0,954	Rejeitado
CVR	0,983	0,986	0,984	0,996	Aceito
	0,972	0,966	0,953	0,996	Rejeitado
<u>MLP</u>	1,000	0,997	0,999	1,000	Aceito
	0,994	1,000	0,997	1,000	Rejeitado
NaiveBayes	0,949	0,970	0,959	0,996	Aceito
	0,938	0,895	0,916	0,983	Rejeitado
<u>Random Forest</u>	0,989	0,993	0,991	0,999	Aceito
	0,986	0,977	0,982	0,999	Rejeitado

Fonte: Elaborado pelo autor, 2022.

Através da Tabela 7, é possível visualizar valores em 1,000 (100%) em algumas métricas. Isso confirma o exposto acima, onde pode ter ocorrido o super ajustamento dos modelos testados. Nesse caso, é fundamental efetuar novamente o teste em outro banco de dados, com variáveis e quantidades de dados diferentes, para validar as métricas de acurácia.

Soares *et al.* (2018) destacam que a vantagem principal da mineração de dados é a capacidade de oferecer soluções para questões complexas para qualquer área de conhecimento, pois se assemelha ao raciocínio lógico, ajudando no processo de classificação.

Em trabalho realizado por Ajaz e Hussain (2015), os autores encontraram 93,8% de recall, 93,8% de F-Measure e 98,9% de ROC Area quando utilizado o classificador CVR, com 10 folds para validação, e 95,2% de recall, 95,2% de F-Measure e 99,6% de ROC Area, utilizando o MLP como classificador, também utilizando 10 folds para validação, destacando o bom desempenho dos resultados encontrados na Tabela 7.

Gadotti *et al.* (2022a) ainda cita que a F-Measure demonstra valores médios de recall e precisão, sendo esse o motivo para as classes de maiores valores serem os de aceito e rejeitado. O ROC Area (*Receiver Operator Characteristic*) apresenta a relação entre a sensibilidade e a especificidade do classificador, ou seja, quanto maior o valor mais ajustado é a curva.

A técnica utilizada nessa pesquisa é adequada, pois as precisões encontradas nas diferentes classes de lotes aceitos e rejeitados, variou de 89% a 100% entre os modelos propostos.

No entanto, ainda é indispensável o esforço humano, pois os testes que resultam nos dados analisados, ainda são realizados de forma manual, requisito para atender as normativas (BOELT *et al.*, 2018). Para uma completa automatização, esse ponto deverá ser revisto pelos órgãos regulamentadores, e técnicas de visão computacional e engenharia. Em trabalho realizado por Ghaffari e Najafabadi (2022) já é possível verificar vários trabalhos sendo conduzidos na área de *machine learning*, para que essas ferramentas também possam ser utilizadas na certificação de sementes, com perspectiva a curto prazo para que sejam adicionadas na legislação vigente.

No trabalho realizado por Gadotti *et al.* (2022a), onde os autores trabalharam com soja, os algoritmos que apresentaram melhores resultados foram Random Forest e CVR. Já GADOTTI *et al.* (2022b), com trabalho realizado em sementes de milho, os dois mais indicados foram J48 e CVR, apontando que, alterando o conjunto de dados (com outra caracterização dos lotes), a metodologia deverá ser adequada novamente.

Para corroborar com a mudança de material testado, a pesquisa realizada por Gadotti *et al.* (2022a), onde os autores trabalharam com lotes de sementes de soja auxiliado pelo software Weka, o algoritmo que conseguiu classificar com boa acurácia foi o Random Forest, que chegou em 92,6% de recall e 90,9% de precisão para a classe “aceito”, 92,6% de recall e 73,6% de precisão para “rejeitado”.

Entretanto, é necessário avaliar a árvore de decisão construída pelo classificador J48, para entender quais foram as variáveis com maior peso na decisão para prever a classificação dos lotes de soja. A Figura 4 possibilita essa visualização.

A partir da Figura 4, é possível analisar que a viabilidade por TZ 2 foi o parâmetro escolhido com maior peso para iniciar a construção do primeiro nó. De acordo com Deminics *et al.* (2009), esse teste determina se as sementes estão vivas ou mortas. Por ser oficializado na RAS (Brasil, 2009), o teste de TZ é amplamente utilizado por ser rápido na obtenção de resultados (FERREIRA *et al.*, 2007; MEDEIROS, 2018). Analisando a continuação da árvore, os outros nós da sequência são construídos com base em testes de vigor, mostrando que os modelos computacionais apresentados aqui podem ser utilizados para prever lotes de sementes de soja, pois Rossi *et al.* (2017), Rau e Zanzi (2021) e Castro (2021) citam ser um teste fundamental para certificar a qualidade de sementes em geral.

Impondo o Cenário 2 (através do percentual de germinação, com mínimo de 90% para lotes aceitos), os algoritmos não obtiveram a mesma assertividade que o visualizado no Cenário 1. Nesse caso, o IBk (88,8%) e o Random Forest (93,0%) obtiveram as melhores respostas. Na Tabela 8 pode ser visto o percentual de acerto de cada algoritmo para classificação das sementes de soja.

Tabela 8 – Acurácia dos algoritmos após a classificação no Cenário 2 com dados C1.

Algoritmo	Classificação Correta dos Atributos (%)
IBk	88,8
J48	88,6
CVR	86,0
MLP	82,7
NaiveBayes	74,7
Random Forest	93,0

Fonte: Elaborado pelo autor, 2022

Contudo, através do exposto no cenário anterior, os percentuais mostrados na Tabela 8 demonstram uma melhor classificação, pois os modelos obtiveram precisão acima de 88%, mas ficando mais distante do super ajuste de 100%.

Além disso, os classificadores mostram a importância do atributo germinação na predição dos lotes, pois ao efetuar o ranqueamento das classes no pré-processamento, a variável foi retirada. Assim, os algoritmos reduziram sua assertividade. Na Tabela 9 é possível perceber que houve muitos lotes agrupados erroneamente quando utilizado o IBk, sendo a classe de rejeitados que mais obteve erros.

Tabela 9 – Matriz de confusão do algoritmo IBk para o Cenário 2 e dados C1.

		Predição	
		Aceito	Rejeitado
Classe Real	Aceito	702	49
	Rejeitado	70	241

Fonte: Elaborado pelo autor, 2022.

Na Tabela 10, o Random Forest demonstra respostas melhores, conseguindo apresentar menores erros quando comparado ao IBk, porém ainda, se confundindo muito na classe de Rejeitados. Ambos os algoritmos não conseguiram encontrar padrões como o encontrado no Cenário 1, quando os lotes foram classificados pela média de todos os testes realizados pelo laboratório. Como nesse caso o teste de germinação foi abstraído para não viciar os mesmos, isso mostra a importância do teste de germinação, para predição com precisão, pela aprendizagem de máquina.

Tabela 10 – Matriz de confusão do algoritmo Random Forest para o Cenário 2 e dados C1.

		Predição	
		Aceito	Rejeitado
Classe Real	Aceito	726	25
	Rejeitado	49	262

Fonte: Elaborado pelo autor, 2022.

Ambas as matrizes mostram diversos lotes classificados de forma errada. Isso se deve ao processo de retirada da variável responsável pelo ranqueamento do conjunto de dados, indisponibilizando o acesso pelos algoritmos. No cenário três, onde o procedimento de exclusão se repetiu, os resultados sofreram alterações.

Lobo Júnior *et al.* (2013) citam que diversos parâmetros devem ser avaliados, para chegar à conclusão, se o lote possui boa sanidade e potencial produtivo. Segundo Finch-Savage *et al.* (2016), para considerar uma semente de soja com alta qualidade, atributos como germinação e vigor, devem ser examinados para que a cultura possa desenvolver seu real potencial produtivo. Uma vez que não se tenha uma dessas informações, ou algum teste tenha sido realizado erroneamente, pode haver insucesso econômico (RAU e ZANZI, 2021).

Em relação as métricas de avaliação, quando analisado a classe de aceitos, o classificador IBk obteve 93,5% de recall, 92,2% de F-Measure e 86,3% de ROC Area. O algoritmo Randon Forest chegou a 96,7% de recall, 95,2% de F-Measure e 96,4% de ROC Area.

A Tabela 11 exibe os parâmetros de acurácia de cada algoritmo, para alicerçar melhor a decisão de qual algoritmo se adapta no momento de auxiliar a classificação de sementes de soja para o cenário proposto na metodologia.

Tabela 11 – Acurácias dos diferentes algoritmos utilizados, sendo Recall (sensibilidade), Precisão, Curva ROC (*Receiver Operating Characteristic*) e *F Measure* para o cenário 2 e dados C1.

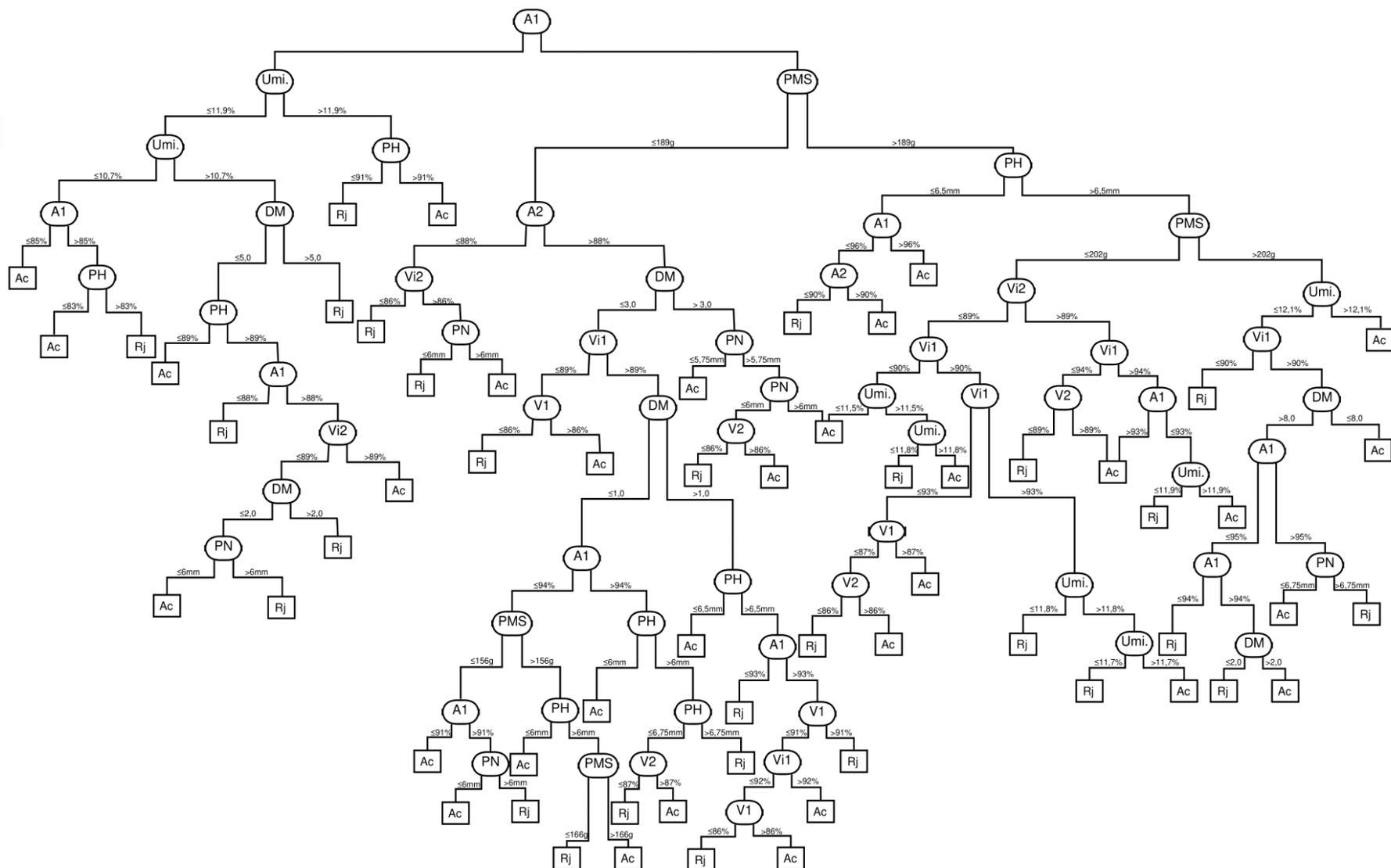
Classificadores	Acurácia				
	Precision	Recall	F-Measure	ROC Area	Classe
<u>IBk</u>	0,909	0,935	0,922	0,863	Aceito
	0,831	0,775	0,802	0,863	Rejeitado
J48	0,908	0,933	0,921	0,852	Aceito
	0,828	0,772	0,799	0,852	Rejeitado
CVR	0,884	0,924	0,904	0,899	Aceito
	0,794	0,707	0,748	0,899	Rejeitado
MLP	0,842	0,929	0,884	0,843	Aceito
	0,773	0,579	0,662	0,843	Rejeitado
NaiveBayes	0,788	0,877	0,830	0,790	Aceito
	0,593	0,431	0,499	0,789	Rejeitado
<u>Random Forest</u>	0,937	0,967	0,952	0,964	Aceito
	0,913	0,842	0,876	0,964	Rejeitado

Fonte: Elaborado pelo autor, 2022.

Como visto já nas tabelas anteriores, alguns algoritmos conseguiram se adaptar melhor a essa metodologia proposta no cenário 2, como o Random Forest, que apresentou métricas bastante interessantes, mesmo alterando o método de classificação. Porém, como a variável germinação foi excluída dos dados no momento do pré-processamento desse cenário, para não viciar o algoritmo, isso demonstra que mesmo assim o Random Forest encontrou um padrão através das outras informações contidas nos lotes avaliados.

De acordo com Schiaffino (2020), esse modelo computacional consegue apresentar bons resultados em diversas avaliações realizadas pois, ao criar várias árvores, selecionando variáveis aleatórias para construção dos nós, este minimiza o erro, não ficando dependente de uma variável específica para classificar um conjunto de dados.

Novamente se faz importante avaliar a árvore de decisão, construída pelo classificador J48 para esse segundo cenário, para entender quais foram os parâmetros escolhidos na decisão para predizer a classificação quando utilizado a germinação em 90% como balizador. A Figura 5 possibilita essa visualização.



*Legenda: Umid. – Umidade; PMS – Peso de Mil Sementes; PH – Peso Hectolitro; A1 – Vigor Areia 1; A2 – Vigor Areia 2; Vi1 – Viabilidade 1; Vi2 – Viabilidade 2; DM – Dano Mecânico; V1 – Vigor 1; V2 – Vigor 2; PN – Peneira; Rj – Rejeitado; Ac – Aceito.

Figura 5 – Árvore decisória do algoritmo J48 com o Cenário 2 e dados C1.

Fonte: Elaborado pelo autor, 2022.

Embora muito utilizado na determinação de qualidade de sementes, o teste de germinação superestima o potencial da mesma, pois é guiado por condições favoráveis (HAESBAERT *et al.*, 2017). Isso faz os pesquisadores questionem o uso isolado desta avaliação para classificar o comportamento das sementes à campo, onde as condições nem sempre são favoráveis, sugerindo utilizar o teste de vigor para classificar o lote (TECNAL, 2022).

A partir disso, foi proposto o Cenário 3, onde a classificação foi realizada a partir do parâmetro de viabilidade TZ 2. Assim, a Tabela 12 mostra a acurácia dos algoritmos depois dos testes realizados na nova classificação. Dessa vez, os métodos que apresentaram melhores respostas foram o MLP (95%) e o Random Forest (96,9%), assim como no Cenário 1, onde ambos também expressaram melhores resultados em comparação com os outros. Nesse cenário, ao aceitar ou rejeitar o lote, o expert utilizou a informação de viabilidade TZ 2. Sendo assim, quando submetido às técnicas de aprendizagem de máquina, foi necessário abstrair esse atributo para não viciar os modelos.

Tabela 12 – Acurácia dos algoritmos após a classificação no Cenário 3 com dados C1.

Algoritmo	Classificação Correta dos Atributos (%)
IBk	94,3
J48	94,3
CVR	93,9
MLP	95,0
NaiveBayes	89,3
Random Forest	96,9

Fonte: Elaborado pelo autor, 2022.

Através da Tabela 12, é possível analisar que os percentuais de acerto foram maiores, quando comparado com o cenário 2. O procedimento de exclusão se repetiu, porém, os algoritmos mostraram maior dependência na variável germinação do que a viabilidade por TZ.

Em trabalho realizado por Pinheiro *et al.* (2021), o autor comenta que o teste de germinação é fundamental para entender qual a sanidade da semente. Porém, Paiva *et al.* (2008), Hampton e TeKrony (1995) e AOSA (2002) citam que o teste de vigor deve sempre complementar o de germinação, para detectar as diferenças na qualidade fisiológica, distinguindo potencial de desempenho no campo e armazenamento.

Assim, a Tabela 12 apresenta, através da assertividade dos algoritmos, que é necessário a informação de germinação para classificar o lote corretamente, ou seja, corrobora quais são os atributos para prever adequadamente os lotes. Entretanto, como a germinação só possuía um atributo no banco de dados, no cenário anterior, ao retirar a informação para evitar o vício, os classificadores não encontraram outro atributo que pudesse substituir essa variável, corroborando a importância do mesmo. No Cenário 3, como o atributo excluído foi a viabilidade TZ 2, outros como vigor em areia e TZ, além da própria viabilidade TZ 1, foram utilizadas pela mineração para decidir sobre o lote.

A Tabela 13 exibe a matriz de confusão do MLP, mostrando que ainda assim se confunde com alguns lotes, tanto para aceitos quanto para rejeitados, demonstrando a importância nesse conjunto de dados do atributo viabilidade TZ 2.

Tabela 13 – Matriz de confusão do algoritmo MLP para o Cenário 3 e dados C1.

		Predição	
		Aceito	Rejeitado
Classe Real	Aceito	763	26
	Rejeitado	27	246

Fonte: Elaborado pelo autor, 2022.

A matriz de confusão do Random Forest (Tabela 14) mostra erros menores quando comparado com o anterior, demonstrando novamente ser um classificador mais confiável justamente por não depender apenas de uma variável. Mesmo assim, como nesse cenário o atributo suprimido possuía outros para “substituir”, este conseguiu expressar resultados melhores, quando comparado com o cenário anterior.

Tabela 14 – Matriz de confusão do algoritmo Random Forest para o Cenário 3 e dados C1.

		Predição	
		Aceito	Rejeitado
Classe Real	Aceito	775	14
	Rejeitado	18	255

Fonte: Elaborado pelo autor, 2022.

Assim, as matrizes do cenário 3 possuem menores erros em relação ao segundo cenário, sendo superior ao cenário 1. Dessa forma, pela quantidade de erros apresentadas aqui, os modelos mostraram resultados mais realista tecnicamente, não sofrendo uma sub ou sobre avaliação dos dados.

Através das matrizes de confusão apresentadas em todos os cenários, é possível indicar que nesse conjunto C1 todos os parâmetros são igualmente importantes para o ranqueamento dos lotes com precisão e boa acurácia, sendo PMS, PH, Germinação, Dano Mecânico, Vigor e Viabilidade em TZ, Vigor em Areia, com as avaliações 1 e 2 (sem e com armazenamento, respectivamente).

Segundo Costa *et al.*, (2008), o PH é utilizado como medida comercial em vários países. Este demonstra a qualidade física do grão, como uniformidade, forma, densidade e tamanho, além de ser afetado pelo teor de materiais estranhos e sementes quebradas, indicando a qualidade do lote (MIRANDA; MORI; LORINI, 2008; COSTA, 2013). Já o PMS, sendo também uma característica física, está relacionado com o tamanho e a densidade do grão, podendo expressar a qualidade das sementes (COSTA, 2013). Ambos estão relacionados com outras características das sementes, afetando seu desempenho à campo, qualidade de grãos e produtividade, diminuindo consideravelmente o lucro do produtor.

Para o teste de tetrazólio, esse já demonstra ser um dos mais eficientes e completos para avaliar a qualidade fisiológica (SOUZA *et al.*, 2017). O teste tem como base a atividade enzimática das desidrogenases envolvidas na respiração.

De acordo com Peske *et al.* (2012) e Limberger *et al.* (2015), análises de vigor, danos e germinação são extremamente necessários para conseguir exprimir todo o conhecimento sobre o potencial fisiológico das sementes. Na Tabela 15 pode ser visualizado os parâmetros que permitem o detalhamento da acuracidade.

Tabela 15 – Acurácias dos diferentes algoritmos utilizados, sendo Recall (sensibilidade), Precisão, Curva ROC (*Receiver Operating Characteristic*) e *F Measure* para o cenário 3 e dados C1.

Classificadores	Acurácia				
	Precision	Recall	F-Measure	ROC Area	Classe
IBk	0,958	0,966	0,962	0,906	Aceito
	0,899	0,879	0,889	0,906	Rejeitado
J48	0,958	0,965	0,961	0,947	Aceito
	0,896	0,879	0,887	0,947	Rejeitado
CVR	0,951	0,967	0,959	0,985	Aceito
	0,900	0,857	0,878	0,985	Rejeitado
<u>MLP</u>	0,966	0,967	0,966	0,982	Aceito
	0,904	0,901	0,903	0,982	Rejeitado
NaiveBayes	0,911	0,948	0,929	0,944	Aceito
	0,830	0,733	0,778	0,944	Rejeitado
<u>Random Forest</u>	0,977	0,982	0,980	0,996	Aceito
	0,948	0,934	0,941	0,996	Rejeitado

Fonte: Elaborado pelo autor, 2022.

Limberger *et al.* (2015) comentam que avaliação física pode ser utilizada como forma de avaliar a sanidade das sementes e seu potencial para o surgimento de novas plantas. Porém, como dito por Vilani (2016), a qualidade da semente é o somatório dos atributos genéticos, fisiológicos, físicos e sanitários, não podendo afirmar o desempenho da semente em originar uma nova planta apenas por um destes atributos.

Os atributos genéticos são relacionados a fatores intrínsecos que definem os parâmetros fisiológicos e bioquímicos, interagindo com os fatores externos (ambiente físico e biótico) (MEDEIROS, 2022). Vários testes podem ser realizados para conseguir descrever, com detalhes, o material, como por exemplo as avaliações de resistência à pragas e doenças, ciclo da cultura e produtividade à campo.

Já os fisiológicos estão ligados aos atributos de metabolismo da semente, englobando todos os processos que ocorrem dentro destas (SOUZA, 2018). Os testes mais conhecidos são capacidade de emergência à campo, viabilidade e vigor. Essa qualidade pode ser afetada pelo ambiente de formação da semente e no momento do manuseio desta, principalmente durante a colheita, beneficiamento e armazenamento (FRANÇA NETO *et al.*, 2016).

Entretanto, é importante ter uma boa pureza física no lote de sementes, sendo os atributos físicos fundamentais para caracterizar a estrutura, composição, formato, peso, tamanho, integridade e aspectos destas (MACEDO *et al.*, 2023). Através de amostras, podem ser avaliados percentual de sementes puras, existência de outras espécies e materiais inertes, como terra, partes de plantas e outras impurezas (CRUZ *et al.*, 2022).

Por fim, é fundamental analisar os aspectos sanitários do lote, ou seja, a existência de agentes patogênicos que podem causar doenças, durante o armazenamento, principalmente, e diminuir a qualidade destas (GRIGOLETO *et al.*, 2019). Outros problemas podem ser a transmissão desses patógenos para outros lotes. Assim, efetuar a detecção de microrganismos antes da semeadura é indispensável, haja vista que o produtor terá que efetuar uma ressemeadura caso exista doenças que afetem o estabelecimento de um bom estante de plantas (MEDEIROS, 2022).

Todavia, Lorenzetti (2020), em trabalho realizado com sementes de soja avaliando a influência do diâmetro após a semeadura, detectou que sementes menores foram mais vigorosas e apresentaram melhor desenvolvimento inicial, igualmente às maiores, existindo pouca relação entre diâmetro e vigor das mesmas. Esse ponto ainda merece mais discussão em pesquisas futuras.

A Figura 6 expressa essa divergência, pois aqui, quando utilizado o mesmo conjunto de dados (C1) e alternado a classificação da germinação, onde o algoritmo utilizou atributos físicos para predição dos lotes, para o teste de viabilidade TZ 2, o classificador “ignorou” em um primeiro momento essas informações físicas, como PMS e PH, utilizando vigor, germinação e dano mecânico para construir os primeiros nós. Em um teste sem a presença desses parâmetros, os algoritmos expressariam resultados semelhantes, pois não são essenciais para a boa classificação, através do aprendizado de máquina.

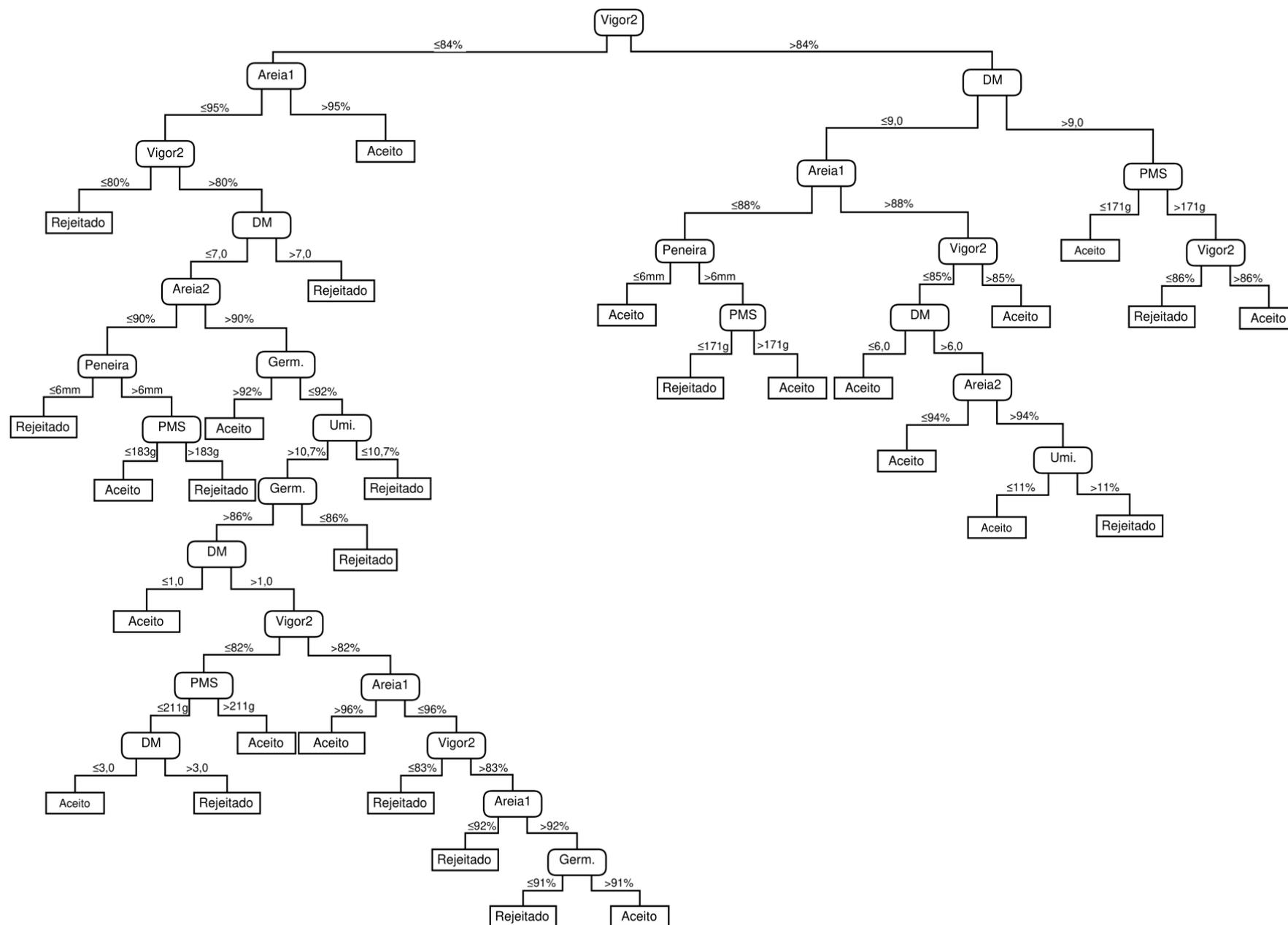


Figura 6 – Árvore decisória do algoritmo J48 com o Cenário 3 e dados C1.
 Fonte: Elaborado pelo autor, 2022.

Mais pesquisas devem ser realizadas para entender o que de fato exprime o máximo potencial da semente. Todavia, o vigor é a propriedade da semente que determina o potencial desta em originar uma nova planta sob diferentes condições de campo (MARCOS FILHO; NOVENBRE, 2009; ALMEIDA *et al.*, 2020). Os testes realizados para certificar os lotes comerciais devem sempre seguir as RAS (BRASIL, 2009) e os manuais de vigor da Associação Brasileira de Tecnologia de Sementes (LIMA JUNIOR, 2010; NORONHA *et al.*, 2018; KRZYZANOWSKI *et al.*, 2021).

Entretanto, em todos esses trabalhos, os autores conduziram os estudos com diferentes variáveis, não indicando quais são realmente necessárias para o bom desempenho dos algoritmos e quais podem ou não constar no banco de dados. Por isso, foram realizados testes com um segundo conjunto de dados, nominado acima de C2. A partir disso será possível apontar quais serão os testes fundamentais para o ranqueamentos dos lotes com maior precisão e rapidez. A Tabela 16 apresenta o percentual de acerto de cada algoritmo para classificação desse novo conjunto.

Tabela 16 – Acurácia dos algoritmos após a classificação de sementes de soja do conjunto C2.

Algoritmo	Classificação Correta dos Atributos (%)
IBk	97,8
J48	98,5
CVR	98,6
MLP	98,8
NaiveBayes	88,4
Random Forest	99,2

Fonte: Elaborado pelo autor, 2022.

A partir da Tabela 16, é possível analisar que todos os algoritmos tiveram um acréscimo de assertividade nos testes realizados com o C2, pois este possui menor quantidade de variáveis e maior número de linhas de dados, sendo mais fácil a tomada de decisão do modelo. Entretanto, dentre os classificadores utilizados, o NaiveBayes foi o que sofreu maior redução, apresentando 88,4% de classes corretamente agrupadas. Isso pode ser explicado através de Lira *et al.* (2019), onde o autor explica que esse modelo matemático funciona bem com variáveis categóricas, além de trabalhar com dados textuais.

Neste caso, os dados são desbalanceados, onde mesmo com a utilização de um filtro, alguns algoritmos não conseguem expressar um desempenho satisfatório. Casali (2021) também cita que para melhorar o desempenho do teorema de Bayes, outro modelo chamado Gaussian NaiveBayes deve ser utilizado.

A partir da análise das matrizes de confusão é possível visualizar que para esses dados o MLP apresentou maiores erros em ambas as classes. Como Manzan (2016) comenta, as amostras para treinamento são fundamentais para que o algoritmo adquira o máximo de informações sobre o problema e consiga resolvê-lo. Como no C2 existe menor número de variáveis, o erro na coleta de amostras é minimizado. Além disso, Manzan (2016) ainda cita que banco de dados supervisionados devem apresentar as classes de forma aleatória, assim aumentando o desempenho das RNs. Um ponto que aqui poderia ser considerado, seria a alteração na organização dos neurônios, possivelmente alterando a resposta final do classificador, como citam Cireşan *et al.* (2012). As matrizes estão nas Tabelas 17 e 18.

Tabela 17 – Matriz de confusão do algoritmo MLP para o C2.

		Predição	
		Aceito	Rejeitado
Classe Real	Aceito	860	10
	Rejeitado	5	412

Fonte: Elaborado pelo autor, 2022.

Por outro lado, o Random Forest apresentou melhora na classificação quando houve a alteração de conjunto, já que esse classificador conseguiu calcular com boa acurácia quando houve menor quantidade de dados e variáveis. Via de regra, a menor caracterização em um banco de dados se torna ruim para a modelagem, porém, durante o presente trabalho, um dos objetivos avaliou a precisão dos algoritmos caso houvesse, por decisão da empresa, quantidade menor de variáveis, validando a metodologia também para esses casos. Isso é possível perceber nos trabalhos de Gadotti *et al.* (2022a) e Gadotti *et al.* (2022b), onde os autores trabalham com conjuntos menores e o mesmo modelo computacional apresentou bons resultados.

Tabela 18 – Matriz de confusão do algoritmo Random Forest para o C2.

		Predição	
		Aceito	Rejeitado
Classe Real	Aceito	865	5
	Rejeitado	5	412

Fonte: Elaborado pelo autor, 2022.

Nesse método, para a construção das árvores decisórias, o algoritmo irá buscar duas ou mais variáveis de forma aleatória de modo a realizar os cálculos com base nessas amostras selecionadas, para assim definir qual irá ocupar o primeiro nó da árvore (MARINS *et al.*, 2021). Para o próximo nó, novamente será realizada uma escolha aleatória como ocorreu anteriormente, excluindo as já buscadas no primeiro nó. Refazendo esse processo até que se consiga solucionar o problema em questão (classificação dos lotes) (VANDERPLAS, 2016).

De acordo com Santos (2022), a desvantagem nesse processo de classificação é a escolha das variáveis, pois o modelo poderá selecionar as piores variáveis e prever de forma errada a classificação das sementes. Porém, como o C2 apresenta poucos parâmetros, o algoritmo teve esse risco minimizado, apresentando maior assertividade do que a predição no C1. Contudo, árvores de decisão possuem interpretações facilitadas, permitindo com que qualquer profissional possa entender o resultado da predição realizada, além dessas conseguirem trabalhar com atributos numéricos e nominais e conseguir trabalhar com base de dados com erros (SANTOS, 2022).

Para conseguir entender essa evolução de desempenho citada, a Tabela 19 apresenta os parâmetros de acurácia de cada algoritmo para esses dados.

Tabela 19 – Acurácias dos diferentes algoritmos utilizados, sendo Recall (sensibilidade), Precisão, Curva ROC (*Receiver Operating Characteristic*) e *F Measure* para o conjunto C2.

Classificadores	Acurácia				
	Precision	Recall	F-Measure	ROC Area	Classe
IBk	0,985	0,985	0,985	0,984	Aceito
	0,969	0,969	0,969	0,984	Rejeitado
J48	0,986	0,992	0,989	0,984	Aceito
	0,983	0,971	0,977	0,984	Rejeitado
CVR	0,985	0,994	0,990	0,997	Aceito
	0,988	0,969	0,978	0,997	Rejeitado
<u>MLP</u>	0,994	0,989	0,991	1,000	Aceito
	0,997	0,988	0,982	1,000	Rejeitado
NaiveBayes	0,880	0,960	0,918	0,969	Aceito
	0,896	0,727	0,803	0,969	Rejeitado
<u>Random Forest</u>	0,994	0,994	0,994	1,000	Aceito
	0,988	0,988	0,988	1,000	Rejeitado

Fonte: Elaborado pelo autor, 2022.

A partir do exposto na Tabela 19, comprova-se a piora do algoritmo NaiveBayes. Este é um classificador probabilístico que se baseia no “Teorema de Bayes”, criado por Thomas Bayes (1701 - 1761) para que esse conseguisse provar a existência de Deus (HOOSHMAND, 2020). De acordo com Sombra *et al.* (2020), esse modelo calcula a probabilidade de o lote ser aceito ou rejeitado multiplicando a probabilidade a priori com a do lote ser aceito ou rejeitado, por exemplo. Para isso, quanto maior for a caracterização do banco de dados, ou *features*, mais preciso será a resposta do algoritmo, aqui neste caso as variáveis utilizadas.

Contudo, é importante avaliar a árvore decisória que o J48 organizou para classificar os lotes, sendo este com 98,5% de assertividade na predição dos lotes. A partir da Figura 7 é possível visualizar que os algoritmos utilizaram o Vigor TZ para diferenciar os lotes aceitos de rejeitados. De acordo com Miquelão *et al.* (2018), a maior limitação no teste de germinação é a incapacidade de detectar o potencial da semente em sobreviver a estresses, sendo avaliado isso nos testes de vigor.

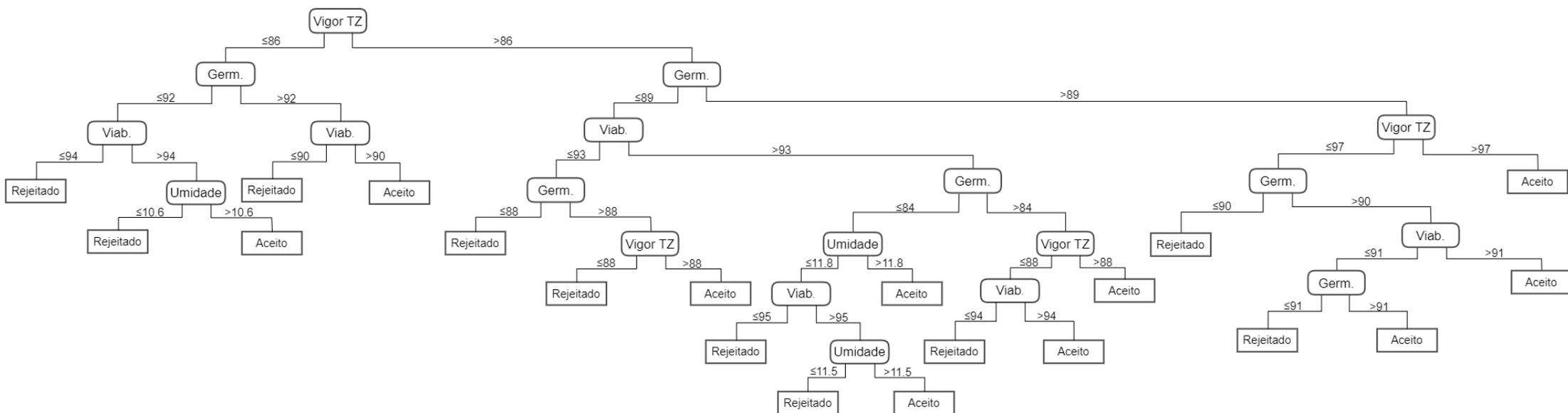


Figura 7 – Árvore decisória do algoritmo J48 com os dados C2.
 Fonte: Elaborado pelo autor, 2022.

Entretanto, como visto nos percentuais de assertividade e parâmetros de acurácia na predição dos lotes do conjunto C2, as variáveis Vigor em TZ e Germinação são testes primordiais para a classificação com precisão através de algoritmos de aprendizagem de máquina, pois estes conseguiram expressar seu potencial de predição sem informações de danos, PMS, PH, entre outras variáveis disponíveis no C1. Porém, alguns pesquisadores ainda salientam que para uma boa caracterização e ranqueamento do lote, é fundamental realizar vários testes de vigor e germinação, como citado por Guedes (2012), Rubin e Travi (2020).

No entanto, para Miquelão *et al.* (2018), a germinação das sementes não é influenciada pela densidade ou peso destas, mas pelo vigor apresentado após os testes. Contudo, de acordo com Carvalho e Nakagawa (2000), sementes menores geram plantas menores, ao mesmo tempo que ocorre isso com sementes maiores, devido a diferença nas reservas das sementes.

De acordo com Pinheiro *et al.* (2021), conjuntos de dados trabalhados com técnicas de IA normalmente são grandes, sendo assim, a análise manual se torna demorada. O trabalho é lento, pois o analista precisa considerar a cultivar, o tratamento, a pureza e germinação, além de outros atributos de qualidade. Assim, pesquisar e testar modelos de aprendizagem de máquina é um trabalho essencial para encontrar um que corresponda um bom desempenho do algoritmo em relação ao produto analisado (PINHEIRO *et al.*, 2021), procurando também maximizar o desempenho destes classificadores, através do pré-processamento de variáveis realmente necessárias para este fim.

Conforme Jha *et al.* (2019), a finalidade principal do aprendizado de máquina é alimentar um determinado sistema com dados de testes anteriores e valores estatísticos, para assim executar a tarefa atribuída a este, resolvendo um problema específico. Essa técnica utiliza uma abordagem matemática para efetuar esse trabalho.

O MLP merece atenção especial, pois de acordo com Haykin (2001), Braga *et al.* (2007) e Silva *et al.* (2012), dentre as RNs com maior sucesso na classificação e reconhecimento de padrões de diversos bancos de dados, a MLP se destaca, sempre mostrando acurácia acima de 90% na solução de problemas.

Sobre a qualidade conferida as sementes, Peske *et al.* (2012) e Rodrigues *et al.* (2020) relatam que vigor facilita a melhor distribuição de lotes comercializados e assim melhor adequação à futuras condições climáticas adversas, sendo os lotes com maiores taxas enviados para regiões que apresentam limitações ambientais durante a semeadura. Para França-Neto e Krzyzanowski (2019, p.1) “a produção e utilização de sementes de alta qualidade são importantes e fundamentais para o sucesso da produção agrícola”.

Assim, é fundamental que o ranqueamento dos lotes seja realizado adequadamente, utilizando sistemas inteligentes que aferem maior agilidade e precisão aos resultados, para que ao comercializá-los, a região de destinação seja correta. Com isso, o agricultor receberá sementes que conseguem enfrentar melhor as adversidades do clima sem a necessidade de altas dosagens em fertilizantes e agrotóxicos, afetando positivamente toda a cadeia produtiva dos grãos, seja nas questões ambientais, como também custos de produção.

A sustentabilidade na agricultura se faz cada vez mais importante para a proteção do meio ambiente. A percepção que enfatiza o fato de que a ecologia sustenta a economia obriga as pessoas a operar em sintonia com os princípios da natureza (CAVALCANTI, 2001).

7 Considerações finais

É possível realizar a avaliação e ranqueamento, com precisão e acurácia, utilizando a inteligência artificial em lotes de sementes de soja, automatizando e agilizando esse processo em empresas produtoras de sementes.

Através dos testes realizados nos conjuntos C1 e C2, os algoritmos indicaram que os atributos vigor em TZ e germinação são fundamentais para a predição de lotes de sementes com precisão e boa acurácia.

Os atributos físicos não foram utilizados como principais variáveis, quando a classificação foi realizada com base teste de viabilidade por TZ no final da armazenagem, indicando que predizer a classificação de sementes por IA, parâmetros de vigor, viabilidade e germinação são essenciais para uma boa precisão de resultados.

É possível tornar a cadeia produtiva mais sustentável utilizando a tecnologia de informação como ferramenta, visando melhorar a qualidade de vida das pessoas e atentar para o meio ambiente, de modo que todo o ecossistema receba os devidos cuidados.

Com essas análises, espera-se que pegada de carbono e hídrica, citada durante o presente trabalho, sofra reduções frente a perspectiva de diminuição dos produtos químicos empregados pelos sojicultores, de modo que diversos outros impactos positivos poderão ser sentidos pela população em geral.

Deste modo, como consequência da certificação da qualidade de sementes utilizando aprendizado de máquina, espera-se que tenha uma diminuição na quantidade de produtos químicos aplicados, podendo haver redução nos custos de implantação da lavoura, fator importante para os agricultores, tornando a cadeia produtiva mais sustentável.

Por fim, recomenda-se o desenvolvimento de novas investigações, integrando a academia com o setor produtivo, construindo um banco de dados público, no qual pesquisadores de qualquer setor possam acessar de modo que outros trabalhos, nesse campo da tecnologia aplicada a agricultura, possam ser desenvolvidos.

Referências

AJAZ, Raja Haroon; HUSSAIN, Lal. Seed Classification using Machine Learning Techniques. **Journal Of Multidisciplinary Engineering Science and Technology**, Berlin, v. 2, n. 5, p. 1098-1102, maio 2015. Disponível em: <<https://www.jmest.org/wp-content/uploads/JMESTN42350753.pdf>>. Acesso em: 20 set. 2021.

ALMEIDA, Andréia da Silva; SILVA, Anderson Severo da; RODRIGUES, Helen Claudine Saliba; GONÇALVES, Vanessa Pinto; MELO, Alejandro Jose de; CARDOZO, Victoria da Silva; RODRIGUES, Daniele Brandstetter; TUNES, Lilian Vanussa Madruga de. Testes de vigor na avaliação da qualidade fisiológica de sementes de cenoura. **Brazilian Journal of Development**, [S.L.], v. 6, n. 6, p. 40985-40992, 2020. Brazilian Journal of Development. DOI: <http://dx.doi.org/10.34117/bjdv6n6-584>.

ALMEIDA, Edvilson de. **Da proibição do uso de agrotóxicos**. 2019. 59 f. TCC (Graduação) - Curso de Direito, Centro Universitário Toledo, Araçatuba, 2019. Disponível em: <<https://servicos.unitoledo.br/repositorio/handle/7574/2270>>. Acesso em: 26 out. 2021.

ALVES, Priscila Mello. **Inteligência Artificial e Redes Neurais**. 2020. Centro de Pesquisa em Ciência, Tecnologia e Sociedade. Disponível em: <https://www.ipea.gov.br/cts/pt/central-de-conteudo/artigos/artigos/106-inteligencia-artificial-e-redes-neurais#_ftn2>. Acesso em: 11 out. 2022.

AMARAL, F. **Aprenda Mineração de Dados: teoria e prática**. Rio de Janeiro: Alta Books, 2016. 220 p.

ANA. **Conjunturas de recursos hídricos no Brasil: relatório de 2021**. 2021. Agência Nacional de Águas e Saneamento Básico. Disponível em: <<https://relatorio-conjuntura-ana-2021.webflow.io>>. Acesso em: 07 set. 2022.

ANVISA - Agência Nacional de Vigilância Sanitária. **Seminário volta a discutir mercado de agrotóxicos em 2012**. 2016. Danilo Molina. Disponível em: <encurtador.com.br/NV369>. Acesso em: 02 set. 2022.

ASSOCIATION OF OFFICIAL SEED ANALYSTS. **Seed vigor testing handbook**. Lincoln: AOSA, 2002. 105p. (Contribution, 32).

ÁVILA, M. R.; ALBRECHT, L. P. Isoflavonas e a qualidade das sementes de soja. In: Workshop brasileiro sobre controle de qualidade de sementes, 3., 2010, Uberlândia. **Anais...** Brasília: Abrates, 2010.

BAALBAKI, R. Z. *et al.* (Ed.). **Seed vigor testing handbook**. Association of Official Seed Analysts, 2009. 341 p.

BAYER. **Agricultores do programa PRO Carbono apresentaram pegada de carbono até 80% menor do que padrões internacionais.** 2022.

JeffreyGroup. Disponível em: <<https://www.agro.bayer.com.br/nossa-bayer/release-pro-carbono>>. Acesso em: 02 set. 2022.

BENIWAL, Sunita; ARORA, Jitender. Classification and Feature Selection Techniques in Data Mining. **International Journal Of Engineering Research & Technology**, [S.L.], v. 1, n. 6, p. 1-6, ago. 2012. Disponível em: <https://www.researchgate.net/profile/Sunita-Beniwal/publication/263662705_Classification_and_feature_selection_techniques_in_data_mining/links/54296d700cf2e4ce940e6d76/Classification-and-feature-selection-techniques-in-data-mining.pdf>. Acesso em: 27 fev. 2023.

BEM JUNIOR, Luciano del; FERRARI, Jonas Leandro; DARIO, Gustavo; RAETANO, Carlos Gilberto. Impact of storage on the physiological quality of soybean seeds after treatment with fungicides and insecticides. **Journal Of Seed Science**, [S.L.], v. 42, p. e202042037, 2020. FapUNIFESP (SciELO). DOI: <http://dx.doi.org/10.1590/2317-1545v42236236>.

BLENINGER, T.; KOTSUKA, L. K. Conceitos de água virtual e pegada hídrica: estudo de caso da soja e óleo de soja no Brasil. **Revista de Recursos Hídricos**, v.36, n.1, p.15-24, 2015. Disponível em: <https://www.aprh.pt/rh/pdf/rh36_n1-2.pdf>. Acesso em: 07 set. 2022.

BOELT, Birte *et al.* Multispectral imaging – a new tool in seed quality assessment? **Seed Science Research**, [S.L.], v. 28, n. 3, p. 222-228, 27 jun. 2018. Cambridge University Press (CUP). DOI: <http://dx.doi.org/10.1017/s0960258518000235>.

BRAGA, A. P.; LUDENIR, T. B.; CARVALHO, A. C. P. L. F. **Redes Neurais Artificiais: teorias e aplicações.** Rio de Janeiro: Livros Técnicos e Científicos, 2007.

BRASIL. Constituição (1980). Lei nº 6894, de 16 de dezembro de 1980. Dispõe sobre a inspeção e a fiscalização da produção e do comércio de fertilizantes corretivos, inoculantes, estimulantes ou biofertilizantes, remineralizadores e substratos para plantas, destinados à agricultura, e dá outras providências. **Lei Nº 6.894, de 16 de Dezembro de 1980.** Brasília, DF: Presidência da República, 16 dez. 2013. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/1980-1988/l6894.htm. Acesso em: 02 set. 2022.

BRASIL. **Regras para análise de sementes.** Brasília: Ministério da Agricultura, Pecuária e Abastecimento, 2009. 398 p. Disponível em: https://www.gov.br/agricultura/pt-br/assuntos/insumos-agropecuarios/arquivos-publicacoes-insumos/2946_regras_analise__sementes.pdf. Acesso em: 27 fev. 2023.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Instrução Normativa nº45 de 17 de setembro de 2013**. Aprova os padrões para a produção e a comercialização de sementes. Brasília, Distrito Federal, 2013. Disponível em: <https://www.gov.br/agricultura/pt-br/assuntos/insumos-agropecuarios/insumos-agricolas/sementes-e-mudas/publicacoes-sementes-e-mudas/copy_of_INN45de17desetembrode2013.pdf>. Acesso em: 22 out. 2021.

BRASIL. **Relatórios de comercialização de agrotóxicos**. 12. ed. Brasília: Ibama - Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis, 2022. 1 p. Relatório online apresentado pelo Ibama. Disponível em: <https://www.gov.br/ibama/pt-br/assuntos/quimicos-e-biologicos/agrotoxicos/relatorios-de-comercializacao-de-agrotoxicos#boletinsanuais>. Acesso em: 07 fev. 2023.

BRASMAX. **A importância das sementes de alto vigor**. Paraná, 2018. Disponível em: <<https://www.brasmaxgenetica.com.br/blog/sementes-de-alto-vigor/>>. Acesso em: 17 jan. 2021.

CARDOSO, O.; MACHADO, R.; Gestão do conhecimento usando data mining: estudo de caso na Universidade Federal de Lavras. 2008. **Revista de Administração Pública**, v.42, p. 495-528.

CARMO, Heliton Fernandes do; MADARI, Beáta Emöke; WANDER, Alcido Elenor; MOREIRA, Flavia Rabelo Barbosa; GONZAGA, Augusto Cesar de Oliveira; SILVEIRA, Pedro Marques da; SILVA, Aluisio Goulart; SILVA, José Geraldo da; MACHADO, Pedro Luiz Oliveira de Almeida. Balanço energético e pegada de carbono nos sistemas de produção integrada e convencional de feijão-comum irrigado. **Pesquisa Agropecuária Brasileira**, [S.L.], v. 51, n. 9, p. 1069-1077, set. 2016. FapUNIFESP (SciELO). DOI: <<http://dx.doi.org/10.1590/s0100-204x2016000900006>>.

CARNEIRO, Fernando Ferreira; RIGOTTO, Raquel Maria; AUGUSTO, Lia Giraldo da Silva; FRIEDRICH, Karen; BÚRIGO, André Campos (org.). **Dossiê Abrasco**: um alerta sobre os impactos dos agrotóxicos na saúde. Rio de Janeiro/São Paulo: Expresso Popular, 2015. 628 p. Escola Politécnica de Saúde Joaquim Venâncio. Disponível em: <https://www.abrasco.org.br/dossieagrotoxicos/wp-content/uploads/2013/10/DossieAbrasco_2015_web.pdf>. Acesso em: 31 ago. 2022.

CARVALHO, N.M.; NAKAGAWA, J. **Sementes**: ciência, tecnologia e produção. 4.ed. Jaboticabal: FUNEP, 2000, 588p.

CARVALHO, N.M.; NAKAGAWA, J. **Sementes**: ciência, tecnologia e produção. 4.ed. Jaboticabal: FUNEP, 2000. 524p.

CARVALHO, Nelson Moreira de; NAKAGAWA, João. **Sementes**: ciência, tecnologia e produção. 5. ed. Jaboticabal: Funep, 2012. 590 p.

CASALI, Rudiney. **Naïve Bayes**: como funciona o algoritmo de machine learning. como funciona o algoritmo de machine learning. 2021. Digital House. Disponível em: <<https://www.digitalhouse.com/br/blog/naive-bayes/>>. Acesso em: 11 out. 2022.

CASTRO, Yasmin Regina Dias de. **A posição das sementes de soja no teste de crescimento de plântulas pode influenciar nos resultados fisiológicos?** 2021. 16 f. TCC (Graduação) - Curso de Agronomia, Instituto de Ciências Agrárias, Universidade Federal de Uberlândia, Uberlândia, 2021. Disponível em: <<https://repositorio.ufu.br/bitstream/123456789/33206/1/PosiçãoSementesSoja.pdf>>. Acesso em: 11 out. 2022.

CAVALCANTI, Clovis (org.). **Meio ambiente, desenvolvimento sustentável e políticas públicas**. Cortez, 2001.

CAVERO, Poholl Adan Sagratzki. **Impacto das mudanças climáticas na produtividade da cultura da soja (*Glycine max* (L.) Merr) na Amazônia: estudo de caso no município de santarém** - pa. 2016. 105 f. Tese (Doutorado) - Curso de Programa de Pós-Graduação em Clima e Ambiente, Instituto Nacional de Pesquisas da Amazônia, Universidade do Estado do Amazonas, Manaus, 2016. Disponível em: <<https://repositorio.inpa.gov.br/bitstream/1/12979/1/Poholl%20Adan.pdf>>. Acesso em: 07 set. 2022.

CEPEA. **Perspectivas para o agronegócio em 2022**. 2022. Elaborado por Geraldo Sant'Ana de Camargo Barros. Disponível em: <<https://www.cepea.esalq.usp.br/br/opiniao-cepea/perspectivas-para-o-agronegocio-em-2022.aspx#:~:text=O%20crescimento%20do%20volume%20de,%2C2%25%20para%20a%20pecuária>>. Acesso em: 16 mar. 2022.

CHAUGULE, Archana. Survey of Seed Classification techniques. **Turkish Journal of Computer and Mathematics Education**, [s. l.], v. 12, n. 13, p. 1236-1260, 2021. Disponível em: <<https://turcomat.org/index.php/turkbilmat/article/view/8678/6769>>. Acesso em: 12 set. 2022.

CHEESEMAN, P.; STUTZ, J. **Bayesian classification (Autoclass)**: Theory and results advances in knowledge discovery and data mining. 1990.

CHEN, Je-Chian; WANG, Yu-Min. Comparing Activation Functions in Modeling Shoreline Variation Using Multilayer Perceptron Neural Network. **Water**, [S.L.], v. 12, n. 5, p. 1281, 30 abr. 2020. MDPI AG. DOI: <<http://dx.doi.org/10.3390/w12051281>>.

CIREŞAN, D. C.; MEIER, U.; GAMBARDELLA, L. M.; SCHMIDHUBER, J. Deep big multilayer perceptrons for digit recognition. **Neural networks: Tricks of the trade**, p. 581–598. 2012. Springer. Disponível em: <<https://people.idsia.ch/~ciresan/data/NNtricks.pdf>>. Acesso em: 11 out. 2022.

CONAB. COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento de safra brasileiro – grãos**: sexto levantamento, março de 2022, safra 2021/2022. Brasília: Companhia Nacional de Abastecimento. 2022. Disponível em: <<https://www.conab.gov.br/info-agro/safras>>. Acesso em: 14 mar. 2022.

COSTA, C.N.; MEURER, E.J.; BISSANI, C.A. & SELBACH, P.A. Contaminantes e poluentes do solo e do ambiente. *In*: MEURER, E.J. **Fundamentos de química do solo**. 2.ed. Porto Alegre, Gênese, 2004. p.207-237.

COSTA, M. G; SILVA, E. L.; STAMFORD, T. L. M.; ANDRADE, S. A. C. Qualidade tecnológica de grãos e farinhas de trigo nacionais e importados. **Ciência e Tecnologia de Alimentos**, v. 28, n. 1, p. 220-225, 2008.

COSTA, Leticia Magalhães da; SILVA, Martim Francisco de Oliveira e. **A indústria química e o setor de fertilizantes**. Brasil: BNDES, 2012. 50 p. Disponível em: <https://web.bndes.gov.br/bib/jspui/bitstream/1408/2025/1/A%20indústria%20química%20e%20o%20setor%20de%20fertilizantes_P_A.pdf>. Acesso em: 07 set. 2022.

COSTA, Mariana Souza. **Avaliação da qualidade industrial de linhagens de trigo por meio de métodos físico-químicos, reológicos e de panificação**. 2013. 150 f. Dissertação (Mestrado) - Curso de Engenharia e Ciência de Alimentos, Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista "Júlio de Mesquita Filho", São José do Rio Preto, 2013. Disponível em: <<https://repositorio.unesp.br/bitstream/handle/11449/88423/000715110.pdf?sequence=1&isAllowed=y>>. Acesso em: 08 nov. 2022.

COSTA, D. C.; MARTORANO, L. G.; MARQUES, M. C.; LISBOA, L. S.; BARBOSA, P. S. P.; BARBOSA, A. M. S. Estimativa da Pegada Hídrica Cinza no controle da Lagarta-da-soja (*Anticarsia gemmatalis*), no polo de produção de grãos Paragominas, Pará. *In*: **Anais [...]**. Congresso Brasileiro de Meteorologia, 18, Recife, 2014.

COSTA, Douglas Cavalcante; MARTORANO, Lucieta Guerreiro; MARQUES, Marcelo Coelho; EL-HUSNY, Jamil Chaar; NACIF, Adelaide. PEGADA HÍDRICA COMO INDICADOR DE SUSTENTABILIDADE EM POLO DE GRÃOS NA AMAZÔNIA. **Enciclopédia Biosfera**, [S.L.], v. 13, n. 23, p. 920-929, 24 jun. 2016. Centro Científico Conhecer. DOI: <http://dx.doi.org/10.18677/enciclopedia_biosfera_2016_079>.

COSTA, Estevam Matheus; NUNES, Bruno de Moraes; VENTURA, Matheus Vinicius Abadia; ARANTES, Bruno Henrique Tondato; MENDES, Gabriel Ribeiro. Efeito Fisiológico De Inseticidas E Fungicida Sobre A Germinação E Vigor De Sementes De Soja (*Glycine max L.*). **Científic@ - Multidisciplinary Journal**, [S.L.], v. 5, n. 2, p. 77-84, 14 maio 2018. Associação Educativa Evangélica. DOI: <http://dx.doi.org/10.29247/2358-260x.2018v5i2.p77-84>.

COWELL, Sarah J; PARKINSON, Stuart. Localisation of UK food production: an analysis using land area and energy as indicators. **Agriculture, Ecosystems & Environment**, [S.L.], v. 94, n. 2, p. 221-236, fev. 2003. Elsevier BV. DOI: [http://dx.doi.org/10.1016/s0167-8809\(02\)00024-5](http://dx.doi.org/10.1016/s0167-8809(02)00024-5).

CRUZ, José de Oliveira; MARTINS, Cibele Chalita; SILVA, Patricia Cândido da Cruz; BARROS, Renato Teó de; SILVA, Givanildo Zildo da. Provenance, viability and physical quality in seeds of *Panicum maximum* 'Mombasa'. **Revista Ciência Agronômica**, Fortaleza, v. 53, p. e20207128, 2022. GN1 Sistemas e Publicacoes Ltd. DOI: <http://dx.doi.org/10.5935/1806-6690.20220042>.

DELARMELINO, L.M. **Composição química e qualidade fisiológica de sementes de soja**. 2012. 28 f. Dissertação (Mestrado em Ciência e Tecnologia de Sementes) - Faculdade de Agronomia Eliseu Maciel, Pelotas, 2012.

DEMINICIS, Bruno Borges; VIEIRA, Henrique Duarte; SILVA, Roberto Ferreira da. Teste de tetrazólio para avaliação da qualidade de sementes de *Clitorea ternatea* L. **Revista Brasileira de Sementes**, [s. l], v. 31, n. 2, p. 54-62, 2009. Disponível em: <https://www.scielo.br/j/rbs/a/qGWgrLDCKvVPbWsnnMTnZKp/?lang=pt&format=pdf>. Acesso em: 27 fev. 2023.

DEUNER, Cristiane; ROSA, Karla Crystina; MENEGHELLO, Géri Eduardo; BORGES, Carolina Terra; ALMEIDA, Andréia da Silva; BOHN, Alberto. Physiological performance during storage of corn seed treated with insecticides and fungicide. **Journal of Seed Science**, [S.L.], v. 36, n. 2, p. 204-212, jun. 2014. DOI: FapUNIFESP (SciELO). DOI: <http://dx.doi.org/10.1590/2317-1545v32n2928>.

DORNELLES, Eldair. Proposta de uma Solução para Avanços Científicos e Tecnológicos no Cultivo de Aveia: uso de modelagem matemática e inteligência artificial na categorização das cultivares de aveia quanto a resistência genética e uso de fungicida. *In*: Seminário de formação científica e tecnológica, 4., 2016, Santa Rosa. **IV SFCT Seminar 2016**. Santa Rosa: Grupo de Pesquisa em Computação Aplicada, 2016. p. 30-32. Disponível em: <https://docplayer.com.br/129034025-iv-seminario-de-formacao-cientifica-e-tecnologica.html>. Acesso em: 26 out. 2021.

EIBE, F.; MARK, A. H.; IAN, H. W. **The WEKA Workbench**. Online Appendix for Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, Fourth Edition, 2020.

ELFERINK, M.; SCHIERHORN, F. **Global demand for food is rising. Can we meet it?** Harvard Business Review. 4p. April 07. 2016. Disponível em: <https://hbr.org/2016/04/global-demand-for-food-is-rising-can-we-meet-it>. Acesso em: 22 out. 2021.

EMBRAPA. **VISÃO 2030**: o futuro da agricultura brasileira. Brasília: Secretaria de Inteligência e Relações Estratégicas, 2018. 213 p. Disponível em: <<https://www.embrapa.br/documents/10180/9543845/Visão+2030+-+o+futuro+da+agricultura+brasileira/2a9a0f27-0ead-991a-8cbf-af8e89d62829>>. Acesso em: 27 fev. 2023.

EMBRAPA. **Fertilização do solo**. 2021. Disponível em: <<https://www.embrapa.br/agencia-de-informacao-tecnologica/cultivos/arroz/producao/sistema-de-cultivo/arroz-de-terras-altas/correcao-do-solo-e-adubacao/fertilizacao-do-solo>>. Acesso em: 27 fev. 2023.

EMBRAPA. **História da soja**. 2023. Disponível em: <<https://www.embrapa.br/soja/cultivos/soja1/historia>>. Acesso em: 27 fev. 2023.

FAGUNDES, Mayra Batista Bitencourt; ARAÚJO, Caroline Giusti de; MORTARI, Valéria Silva; REBELO, Fábio Santos; RODRIGUES, Patricia. Análise dos efeitos da intervenção governamental na competitividade da produção de soja em grãos no Mato Grosso do Sul. **Gestão & Regionalidade**, [S.L.], v. 29, n. 87, p. 111-125, 18 dez. 2013. USCS Universidade Municipal de Sao Caetano do Sul. DOI: <<http://dx.doi.org/10.13037/gr.vol29n87.2038>>.

FAO. Food and Agriculture Organization of the United Nations, 2017. **The state of food and agriculture leveraging food systems for inclusive rural transformation**. 2017, p. 1-181.

FAYYAD, U., PIATETSKY-SHAPIRO G., SMYTH P. The KDD process for extracting useful knowledge from volumes of data. **Communication of the ACM**, v. 29, p. 27–34. 1996.

FERREIRA, R. A. *et al.* Comparação da viabilidade de sementes de *Schizolobium parahyba* (Vell.) Blake – *Leguminosae caesalpinioideae*, pelos testes de germinação e tetrazólio. **Revista Brasileira de Sementes**, Londrina, v. 29, n. 3, p. 83-89, 2007. DOI: <http://dx.doi.org/10.1590/S0101-31222007000300011>.

FINCH-SAVAGE, W. E.; BASSEL, G. W. Seed vigour and crop establishment: extending performance beyond adaptation. **Journal of Experimental Botany**, v.67, n.3, p. 567-591, 2016. DOI: <https://doi.org/10.1093/jxb/erv490>.

FINLAY, S. **Artificial intelligence for everyone**. Great Britain: Relativistic. 189p. 2020.

FRANÇA-NETO, J. B; KRZYZANOWSKI, F. C.; HENNING, A. A; PÁDUA, G. P; LORINI, I., HENNING, F. A. **Tecnologia da produção de semente de soja de alta qualidade**. Londrina: Embrapa Soja, 2016. 82p. (Documentos 380).

FRANÇA-NETO, José de Barros; KRZYZANOWSKI, Francisco Carlos. Tetrazolium: an important test for physiological seed quality evaluation. **Journal of Seed Science**, [S.L.], v. 41, n. 3, p. 359-366, jul. 2019. FapUNIFESP (SciELO). DOI: <http://dx.doi.org/10.1590/2317-1545v41n3223104>.

FURLAN, M; PALHARES, J. C. P. Pegada hídrica cinza de um sistema de produção de leite. In: simpósio internacional sobre gerenciamento de resíduos agropecuários e agroindustriais, 5., 2017, Foz do Iguaçu. **Anais [...]**. Concórdia: Sbera Embrapa Suínos e Aves, 2017. p. 491-494.

GADOTTI, Gizele I.; MORAES, Nicacia A. B.; SILVA, Joseano G. da; PINHEIRO, Romário de M.; MONTEIRO, Rita de C. M. PREDICTION OF RANKING OF LOTS OF CORN SEEDS BY ARTIFICIAL INTELLIGENCE. **Engenharia Agrícola**, Jaboticabal, v. 42, n. 4, p. 1-8, ago. 2022b. FapUNIFESP (SciELO). DOI: <http://dx.doi.org/10.1590/1809-4430-eng.agric.v42n4e20210005/2022>.

GADOTTI, Gizele Ingrid; ASCOLI, Carla Aparecida; BERNARDY, Ruan; MONTEIRO, Rita de Cassia Mota; PINHEIRO, Romário de Mesquita. Machine learning for soybean seeds lots classification. **Revista Engenharia Agrícola**, v. 42, special issue, 2022a. DOI: <http://dx.doi.org/10.1590/1809-4430-Eng.Agric.v42nepe20210101/2022>. Disponível em: <https://www.scielo.br/j/eagri/a/LtTLRpzgNQPWp5mw3qRMdtM/?format=pdf&lang=en>. Acesso em: 21 mar. 2022.

GALLI, Alessandro; WIEDMANN, Thomas; ERCIN, Ertug; KNOBLAUCH, Doris; EWING, Brad; GILJUM, Stefan. Integrating Ecological, Carbon and Water footprint into a “Footprint Family” of indicators: definition and role in tracking human pressure on the planet. **Ecological Indicators**, [S.L.], v. 16, p. 100-112, maio 2012. Elsevier BV. <http://dx.doi.org/10.1016/j.ecolind.2011.06.017>.

GAZZONI, Décio. **A sustentabilidade da soja no Brasil**. 2020. Disponível em: <https://www.noticiasagricolas.com.br/noticias/agronegocio/108656-asustentabilidade-dasoja-no-brasil>. Acesso em: 22 out. 2021.

GHAFFARI, Akram; NAJAFABADI, Mohsen Yoosefzadeh. Machine learning application in seeds certification process. **Experimental Results**. 2022. No prelo.

GONÇALVES, Carina de Farias; CORREA, Solange Rodrigues Santos; RODRIGUES, Luciano Brito; ALMEIDA NETO, José Adolfo. Pegada de carbono do ciclo de vida da celulose de eucalipto: estudo de caso numa empresa baiana. **Gaia Scientia**, [S.L.], v. 12, n. 4, p. 1-15, 28 dez. 2018. Portal de Periódicos UFPB. DOI: <http://dx.doi.org/10.22478/ufpb.1981-1268.2018v12n4.31816>.

GRIGOLETO, Marcela Regina Paganuchi; MEDINA, Priscila Fratin; PARISI, João José Dias; SANTOS, Franciele dos. Health aspects and ideal temperature for germination of peanut seeds. **Journal Of Seed Science**, [S.L.], v. 41, n. 2, p. 233-243, abr. 2019. FapUNIFESP (SciELO). DOI: <http://dx.doi.org/10.1590/2317-1545v41n2195483>.

GUEDES, Roberta Sales. **Testes de vigor para a avaliação da qualidade fisiológica de sementes de *Amburana cearensis* (Allemão) A.C. Smith.** 2012. 97 f. Tese (Doutorado) - Curso de Programa de Pós-Graduação em Agronomia, Universidade Federal da Paraíba, Areia, 2011. Disponível em: <<https://repositorio.ufpb.br/jspui/bitstream/tede/8023/2/arquivototal.pdf>>. Acesso em: 11 out. 2022.

GUIMARAES, Édson Avelar; NUNES, Elis Regina Nuffer; SILVA, Vinicius Vanderley Miguel da; CARVALHO, Angelica Oliveira; SANTOS, Giseli Gomes dos. Proposta de utilização da pegada hídrica como indicador socioeconômico e ambiental na gestão dos recursos hídricos no estuário do rio Macaé. **Estudos de Administração e Sociedade**, [S.L.], v. 5, n. 1, p. 44-51, 22 dez. 2020. Pro Reitoria de Pesquisa, Pós-Graduação e Inovação - UFF. DOI: <<http://dx.doi.org/10.22409/eas.v5i1.42173>>.

HAESBAERT, Fernando Machado; LOPES, Sidinei José; MERTZ, Liliane Marcia; LÚCIO, Alessandro Dal'col; HUTH, Caroline. Tamanho de amostra para determinação da condutividade elétrica individual de sementes de girassol. **Bragantia**, [S.L.], v. 76, n. 1, p. 54-61, 12 jan. 2017. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/1678-4499.389>.

HAMPTON, J. G.; TEKRONY, D. M. **Handbook of vigour test methods**. 3. ed. Zurich: ISTA, 1995. 117p.

HAN, Jiawei; KAMBER, Micheline; PEI, Jian. **Data Mining: concepts and techniques**. 3. ed. Massachusetts: Morgan Kaufmann Publishers, 2011. 703 p.

HAYKIN, S. **Redes Neurais: princípios e práticas**. 2. ed. São Paulo: BOOKMAN, 2001.

HECHT-NIELSEN, R. **Neurocomputing**. Addison (Wesley Co., Reading, MA). 1990.

HEDLUND, E. H.; MARUJO, A. V.; MELO, A. J de. Emprego formal e exportação: o caso da soja na região sul do Brasil. **Revista do Desenvolvimento Regional – COLÓQUIO**, Taquara-RS, v.18, n.4, 2021. DOI: <https://doi.org/10.26767/colóquio.v18i4.2221>.

HENNING, A. A.; FRANÇA NETO, J. B.; KRZYZANOWSKI, F. C.; LORINI, I. **Importância do tratamento de sementes de soja com fungicidas na safra 2010/2011, ano de “La Niña”**. Londrina: Embrapa Soja, 2010. 8p. (Embrapa Soja. Circular Técnica, 82). Disponível em: <<https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/866711/2/CT82.indd.pdf>>. Acesso em: 02 set. 2022.

HERTWICH, Edgar G.; PETERS, Glen P. Carbon Footprint of Nations: a global, trade-linked analysis. **Environmental Science & Technology**, [S.L.], v. 43, n. 16, p. 6414-6420, 15 jun. 2009. American Chemical Society (ACS). DOI: <<http://dx.doi.org/10.1021/es803496a>>.

HESAMI, Mohsen; NADERI, Roohangiz; TOHIDFAR, Masoud; YOOSEFZADEH-NAJAFABADI, Mohsen. Development of support vector machine-based model and comparative analysis with artificial neural network for modeling the plant tissue culture procedures: effect of plant growth regulators on somatic embryogenesis of chrysanthemum, as a case study. **Plant Methods**, [S.L.], v. 16, n. 1, p. 1-15, 13 ago. 2020. Springer Science and Business Media LLC. DOI: <http://dx.doi.org/10.1186/s13007-020-00655-9>.

HIRAKURI, M. H.; LAZZAROTTO, J. J. **O agronegócio da soja nos contextos mundial e brasileiro**. Londrina: Embrapa Soja, 2014. 37p. (Embrapa Soja. Circular Técnica, 349). Disponível em: <<https://ainfo.cnptia.embrapa.br/digital/bitstream/item/104753/1/O-agronegocio-da-soja-nos-contextos-mundial-e-brasileiro.pdf>>. Acesso em: 27 fev. 2023.

HOEKSTRA, A. Y.; CHAPAGAIN, A. K. Water footprints of nations: water use by people as a function of their consumption pattern. **Water Resources Management**, [S.L.], v. 21, n. 1, p. 35-48, 27 jun. 2007. Springer Science and Business Media LLC. DOI: <<http://dx.doi.org/10.1007/s11269-006-9039-x>>.

HOEKSTRA, A. Y.; CHAPAGAIN, A. K.; ALDAYA, M. M.; MEKONNEN, M. M. **Water Footprint Manual: State of the Art**. Ensched: Water Footprint Network, 2009. 131p.

HOEKSTRA, A.Y.; CHAPAGAIN, A. K. **Globalization of water: Sharing the planet's freshwater resources**. [S.I.]: Blackwell Publishing, 2008. 220p. ISBN: 978-1-4051-6335-4.

HOEKSTRA, Arjen Y.; CHAPAGAIN, Ashok K.; ALADAYA, Maite M.; MEKONNEN, Mesfin M. **Manual de Avaliação da Pegada Hídrica: estabelecendo o padrão global**. São Paulo: Earthscan, 2011. 216p.

HOOSMAND, Arash. Naive Bayesian Machine Learning to Diagnose Breast Cancer. **Research Square**, [S.L.], p. 1-12, 3 set. 2020. Research Square Platform LLC. DOI: <http://dx.doi.org/10.21203/rs.3.rs-60997/v1>.

HUI LI. **Machine Learning: o que é e qual sua importância?** Disponível em: <https://www.sas.com/pt_br/insights/analytics/machine-learning.html>. Acesso em: 21 out. 2021.

IBGE. Instituto Brasileiro de Geografia e Estatística, 2021. Disponível em: <<https://sidra.ibge.gov.br/acervo#/S/LA/A/47/T/Q>>. Acesso em: 22 out. 2021.

IBM. **Visualização da Matriz de Confusão**. 2021. Disponível em: <https://www.ibm.com/docs/pt-br/db2/10.5?topic=visualizer-confusion-matrix-view>. Acesso em: 05 set. 2022.

JHA, K.; DOSHI, A.; PATEL, P.; SHAH, M. A comprehensive review on automation in agriculture using artificial intelligence. **Artificial Intelligence in Agriculture**, v. 2, p. 1-12, 2019. DOI: 10.1016/j.aiia.2019.05.004.

KAUR, K. Machine learning: applications in Indian agriculture. **International Journal of Advanced Research in Computer and Communication Engineering**, v. 5, n. 4. 2016.

KOIFMAN, S.; KOIFMAN, R. J.; MEYER, A. Human reproductive system disturbances and pesticide exposure in Brazil. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 18, n. 2, p. 435-445, abr. 2002. DOI: <http://dx.doi.org/10.1590/s0102-311x2002000200008>.

KRZYZANOWSKI, Francisco Carlos; VIEIRA, Roberval Daiton; MARCOS FILHO, Julio; FRANÇA-NETO, José de Barros (ed.). **Vigor de sementes: conceitos e testes**. 2. ed. Londrina: ABRATES, 2021. 601 p.

KRZYZANOWSKI, F. C.; FRANÇA NETO, J. B.; HENNING, A. A. **A alta qualidade da semente de soja: fator importante para a produção da cultura**. Londrina: Embrapa Soja, 2018. 24p. (Embrapa Soja. Circular Técnica, 136). Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/177391/1/CT136-online.pdf>. Acesso em: 27 fev. 2023.

KUMAR, S. **Reinventing Agri-produce quality testing using AI**. National Workshop on Artificial Intelligence in Agriculture. Indian Council of Agricultural Research. New Delhi. p.44-45. 2018.

LEVER, Jake; KRZYWINSKI, Martin; ALTMAN, Naomi. Classification evaluation. **Nature Methods**, [S.L.], v. 13, n. 8, p. 603-604, 28 jul. 2016. Springer Science and Business Media LLC. DOI: <http://dx.doi.org/10.1038/nmeth.3945>.

LIAKOS, K.; BUSATO, P.; MOSHOU, D.; PEARSON, S.; BOCHTIS, D. Machine learning in agriculture: A review. **Sensors**, v. 18, n. 8, p. 2674, 2018. DOI: <http://dx.doi.org/10.3390/s18082674>.

LIGNANI, Leonardo de Bem; BRANDÃO, Júlia Lima Gorges. A ditadura dos agrotóxicos: o Programa Nacional de Defensivos Agrícolas e as mudanças na produção e no consumo de pesticidas no Brasil, 1975-1985. **História, Ciências, Saúde-Manguinhos**, Rio de Janeiro, v. 29, n. 2, p. 337-359, 19 ago. 2022. FapUNIFESP (SciELO). DOI: <http://dx.doi.org/10.1590/s0104-59702022000300023>.

LIMA JUNIOR, M. J V (ed.). **Manual de procedimentos de análise de sementes florestais**. 1. ed. Londrina: ABRATES, 2010. 146 p. Disponível em: l1nq.com/Jaw8x. Acesso em: 11 out. 2022. UFAM - Manaus-Amazonas, Brasil.

LIMBERGER, Lucas Martins; SCHUCH, Luis Osmar Braga; DÖRR, Caio Sippel; ALMEIDA, Andreia. Qualidade de Sementes de Soja: qualidade fisiológica de sementes de soja no município de santa rosa-rs. In: MENEGHELLO, Geri Eduardo; ALMEIDA, Andréia da Silva; VILLELA, Francisco Amaral; TUNES, Lilian Vanussa Madruga de (org.). **Produção técnico-científica em sementes**. Pelotas: Editora Santa Cruz, 2015. Cap. 9. p. 189-214. Disponível em: <https://wp.ufpel.edu.br/gem/files/2017/10/capitulo_09_-p_189_214.pdf>. Acesso em: 11 out. 2022.

LIRA, Larissa Duarte de Britto; MELO, Ingrid Rafaella dos Santos; SCHAFFER NETO, Herbert; RIBEIRO, Kátia Suely Queiroz Silva. Análise do algoritmo naive bayes na classificação de amostras do banco de dados hepatite. in: congresso nacional de pesquisa em ensino de ciências, 4., 2019, Paraíba. **Anais [...]**. Paraíba: CONAPESC, 2019. p. 1-12. Disponível em: <http://www.editorarealize.com.br/editora/anais/conapesc/2019/TRABALHO_EV126_MD1_SA10_ID2806_12082019105434.pdf>. Acesso em: 11 out. 2022.

LOBO JÚNIOR, M.; BRANDÃO, L. T. D.; MARTINS, B. É. de M.; **Testes para Avaliação da Qualidade de Sementes de Feijão Comum**. Goiás: Embrapa Arroz e Feijão, 2013. 4p. (Embrapa Arroz e Feijão. Circular Técnica, 90). Disponível em: <<https://www.infoteca.cnptia.embrapa.br/bitstream/doc/970251/1/circulartecnica90.pdf>>. Acesso em: 11 out. 2022.

LOPES, A. C. A.; NASCIMENTO, W. M. **Análise de Sementes de Hortaliças**. Brasília: Embrapa Hortaliças, 2009. 9p. (Embrapa Hortaliças. Circular Técnica, 83).

LOPES, Carla Vanessa Alves; ALBUQUERQUE, Guilherme Souza Cavalcanti de. Agrotóxicos e seus impactos na saúde humana e ambiental: uma revisão sistemática. **Saúde em Debate**, [S.L.], v. 42, n. 117, p. 518-534, jun. 2018. FapUNIFESP (SciELO). DOI: <<http://dx.doi.org/10.1590/0103-1104201811714>>.

LORENZETT, Geovani Tiago. **Germinação e vigor de sementes salvas de soja em função do diâmetro da semente e profundidade de semeadura**. 2020. 34 f. TCC (Graduação) - Curso de Agronomia, Universidade Federal da Fronteira Sul, Chapecó, 2020. Disponível em: <https://rd.uffs.edu.br/bitstream/prefix/4476/1/LORENZETT.pdf>. Acesso em: 11 out. 2022.

MACEDO, Jussara Aparecida Santos de; ARAËJO, Eduardo Fontes; ARAËJO, Roberto Fontes; GALVÃO, João Carlos Cardoso; CECON, Paulo Roberto; MACEDO, Leandro Roberto de; CAMPOS, Silvane de Almeida. Physical and physiological qualities and productivity of corn seeds fertilized with poultry waste. **Ciência Rural**, Santa Maria, v. 53, n. 1, p. e20210515, 2023. FapUNIFESP (SciELO). DOI: <http://dx.doi.org/10.1590/0103-8478cr20210515>.

MANZAN, José Ricardo Gonçalves. **Análise de desempenho de redes neurais artificiais do tipo *multilayer perceptron* por meio do distanciamento dos pontos do espaço de saída**. 2016. 131 f. Tese (Doutorado) - Curso de Programa de Pós-Graduação em Engenharia Elétrica, Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia, Uberlândia, 2016. Disponível em: <<https://repositorio.ufu.br/bitstream/123456789/17967/1/AnaliseDesempenhoRedes.pdf>>. Acesso em: 11 out. 2022.

MAPA. Ministério da Agricultura, Pecuária e Abastecimento. Secretaria de Política Agrícola. **Projeções do agronegócio: Brasil 2019/2020 a 2029/2030**. Brasília, 2020. Disponível em: <https://www.gov.br/agricultura/pt-br/assuntos/noticias/ao-completar-160-anos-ministerio-da-agricultura-preve-crescimento-de-27-na-producao-de-graos-do-pais-na-proxima-decada/ProjecoesdoAgronegocio2019_20202029_2030.pdf>. Acesso em: 22 out. 2021.

MARINS, Matheus A.; BARROS, Bettina D.; SANTOS, Ismael H.; BARRIONUEVO, Daniel C.; VARGAS, Ricardo E.V.; PREGO, Thiago de M.; LIMA, Amaro A. de; CAMPOS, Marcello L.R. de; SILVA, Eduardo A.B. da; NETTO, Sergio L. Fault detection and classification in oil wells and production/service lines using random forest. **Journal of Petroleum Science and Engineering**, [S.L.], v. 197, p. 107879, fev. 2021. Elsevier BV. DOI: <http://dx.doi.org/10.1016/j.petrol.2020.107879>.

MARACAJÁ, Kettrin Farias Bem. Modelo integrado das pegadas hídrica, ecológica e de carbono para o monitoramento da pressão humana sobre o planeta. **Espacios**, Campina Grande, v. 37, n. 17, p. 21, 2016. Disponível em: <https://www.revistaespacios.com/a16v37n17/16371721.html>. Acesso em: 07 set. 2022.

MARACAJÁ, Kettrin Farias Bem; ELOI, Lincoln; SILVA, Vicente de Paulo Rodrigues da. Regionalização da pegada hídrica do estado da Paraíba. **Reunir: Revista de Administração, Contabilidade e Sustentabilidade**, [S.L.], v. 4, n. 1, p. 105, 1 abr. 2014. Reunir Revista de Administração, Contabilidade e Sustentabilidade. DOI: <<http://dx.doi.org/10.18696/reunir.v4i1.206>>.

MARCOS-FILHO, J. **Fisiologia de sementes de plantas cultivadas**. Piracicaba: FEALQ, 2005. 495p.

MARCOS-FILHO, J.; NOVEMBRE, A.D.L.C. Avaliação do potencial fisiológico de sementes de hortaliças. In: NASCIMENTO, W.M. (Ed.). **Tecnologia de sementes de hortaliças**. Brasília: EMBRAPA HORTALIÇAS, 2009. p.185 – 243.

MEDEIROS, André Dantas de; CAPOBIANGO, Nayara Pereira; SILVA, José Maria da; SILVA, Laércio Junio da; SILVA, Clíssia Barboza da; DIAS, Denise Cunha Fernandes dos Santos. Interactive machine learning for soybean seed and seedling quality classification. **Scientific Reports**, [S.L.], v. 10, n. 1, p. e11267, 9 jul. 2020. Springer Science and Business Media LLC. DOI: <http://dx.doi.org/10.1038/s41598-020-68273-y>.

MEDEIROS, Gilberto. O que você precisa saber sobre qualidade de sementes. **Revista Cultivar**, Pelotas, mar. 2022. [Artigo Online]. Disponível em: <https://revistacultivar.com.br/artigos/o-que-voce-precisa-saber-sobre-qualidade-de-sementes>. Acesso em: 07 nov. 2022.

MEDEIROS, Josenilda Aprigio Dantas de. **Morfologia de plântulas de *Amburana cearensis* (alemão) a. c. smith (fabaceae) aplicada à análise de vigor de sementes**. 2018. 57 f. Dissertação (Mestrado) - Curso de Programa de Pós-Graduação em Ciências Florestais, Universidade Federal do Rio Grande do Norte, Macaíba, 2018. Disponível em: <https://repositorio.ufrn.br/bitstream/123456789/26581/1/MorfologiaplântulasAmburana_Medeiros_2018.pdf>. Acesso em: 11 out. 2022.

MENTEN, J. O. M.; MORAES, M. H. D. Sanidade das sementes previne doenças. **Revista Visão Agrícola**, São Paulo, v. 3, n. 6, p. 40-41, 2006.

MIQUELÃO, Luiz Felipe da Silva; OLIVEIRA, Daiane de Moura Costa; MACHADO, Maycon Rodrigues; BATISTA, Ricardo Augusto; SABUNDJIAN, Michelle Traete. Influência da classificação de sementes de soja (*Glycine max* (L.) Merrill) na germinação e vigor. **Revista Científica Eletrônica de Ciências Aplicadas da Fait**, [s. l.], v. 2, p. 1-6, nov. 2018. Disponível em: http://fait.revista.inf.br/imagens_arquivos/arquivos_destaque/ZJ3MWkBOljJbFm_y_2020-7-28-19-23-57.pdf. Acesso em: 11 out. 2022.

MIRANDA, M. Z.; MORI, C.; LORINI, I. **Qualidade Comercial do Trigo Brasileiro**: Safra, 2005. Embrapa, Passo Fundo, RS, 2008. (Documento 80).

MONARD, Maria Carolina, BARANAUSKAS, José Augusto. Conceitos Sobre Aprendizado de Máquina. In: MONARD, Maria Carolina, BARANAUSKAS, José Augusto. **Sistemas Inteligentes Fundamentos e Aplicações**. 1 ed. Barueri-SP: Manole Ltda, 2003. p. 89-114. ISBN 85-204-168.

MONTIBELLER FILHO, Gilberto; SOUZA, Glaucia Cardoso de; BÔLLA, Kelly Daiane Savariz. Economia Ecológica e Sustentabilidade Socioambiental. **Revista Brasileira de Ciências Ambientais**, [s. l.], n. 23, p. 23-25, 2012. Disponível em: https://rbciamb.com.br/Publicacoes_RBCIAMB/article/view/329. Acesso em: 07 set. 2022.

MORAES, Nicácia Andrade Borges. **Predição de ranqueamento de lotes de sementes de milho por inteligência artificial**. 2020. 30 f. Dissertação (Mestrado em Ciência e Tecnologia de Sementes) – Faculdade de Agronomia, Universidade Federal de Pelotas, Pelotas, 2020.

MORE: **Mecanismo online para referências, versão 2.0**. Florianópolis: UFSC Rexlab, 2013. Disponível em: <<http://www.more.ufsc.br/>>. Acesso em: 12 set. 2022.

MORETI, M. P.; OLIVEIRA, T.; SARTORI, R.; CAETANO, W. Inteligência artificial no agronegócio e os desafios para a proteção da propriedade intelectual. **Cadernos de Prospecção**, v. 14, p. 60, 2021. DOI: <http://dx.doi.org/10.9771/cp.v14i1.33098>

MUCHERINO, A.; PAPAJORGJI, P.; PARDALOS, P. M. A survey of data mining technique applied to agriculture. **Operational Res.**, v. 9, n. 2. p. 121-140, 2009.

NAGHASHZADEH, M. R.; AZADBAKHT, N. **Principles of seed production and certification**. Agricultural education and promotion institute. 2018.

NASSIF, S.M.L.; VIEIRA, I.G.; GOMES, I.; FERNANDES, G.D. **Fatores externos (ambientais) que influenciam a germinação de sementes**. Informativo de Sementes IPEF, 1998.

NEPSTAD, Daniel; MCGRATH, David; STICKLER, Claudia; ALENCAR, Ane; AZEVEDO, Andrea; SWETTE, Briana; BEZERRA, Tathiana; DIGIANO, Maria; SHIMADA, João; MOTTA, Ronaldo Seroa da. Slowing Amazon deforestation through public policy and interventions in beef and soy supply chains. **Science**, [S.L.], v. 344, n. 6188, p. 1118-1123, 6 jun. 2014. American Association for the Advancement of Science (AAAS). <http://dx.doi.org/10.1126/science.1248525>.

NORONHA, Bruno Gomes de; MEDEIROS, André Dantas de; PEREIRA, Márcio Dias. Avaliação da qualidade fisiológica de sementes de *Moringa oleifera* Lam. **Ciência Florestal**, [S.L.], v. 28, n. 1, p. 393-402, 2 abr. 2018. Universidad Federal de Santa Maria. DOI: <http://dx.doi.org/10.5902/1980509831615>.

OPAS/OMS - Organização Pan-americana de Saúde/Organização Mundial de Saúde. **Manual de vigilância da saúde de populações expostas a agrotóxicos**. Brasília: Representação do Brasil, 1996. Disponível em: <<https://bvsms.saude.gov.br/bvs/publicacoes/livro2.pdf>>. Acesso em: 26 out. 2021.

OSROUSH, S. **Cereal registration and certification**. Seed and plant registration and certification institute. 2010.

PAIVA, Auricleia Sarmiento de; RODRIGUES, Teresinha de Jesus Deléo; CANCIAN, Antonio João; LOPES, Magnólia de Mendonça; FERNANDES, Adriano Carlos. QUALIDADE FÍSICA E FISIOLÓGICA DE SEMENTES DA LEGUMINOSA FORRAGEIRA *Macrotyloma axillare* cv. Java. **Revista Brasileira de Sementes**, [S.I.], v. 30, n. 2, p. 130-136, 2008. Disponível em: <https://www.scielo.br/j/rbs/a/9cDJB3ZFpwvbXty5zgNBzH/?format=pdf&lang=p>. Acesso em: 11 out. 2022.

PANDEY, D.; AGRAWAL, M.; PANDEY, J. S. Carbon footprint: current methods of estimation. **Environ Monit Assess**, 178, 135–160, 2011. DOI: <<https://doi.org/10.1007/s10661-010-1678-y>>.

PASCUALI, L. C. **Estimativa do potencial de armazenamento de soja através do vigor das sementes**. 2012. 52 f. Tese (Doutorado) - Curso de Programa de Pós Graduação em Ciência e Tecnologia em Sementes, Faculdade de Agronomia Eliseu Maciel, Universidade Federal de Pelotas, Pelotas, 2012.

PASSUELLO, Ana Carolina Badalotti; OLIVEIRA, Alexandre Führ de; COSTA, Eugênio Bastos da; KIRCHHEIM, Ana Paula. Aplicação da Avaliação do Ciclo de Vida na análise de impactos ambientais de materiais de construção inovadores: estudo de caso da pegada de carbono de clínqueres alternativos. **Ambiente Construído**, [S.L.], v. 14, n. 4, p. 7-20, dez. 2014. FapUNIFESP (SciELO). DOI: <<http://dx.doi.org/10.1590/s1678-86212014000400002>>.

PATEL, A. A.; KATHIRIYA, D. R. Data mining trends in agriculture: a review. **AGRES – An International E. Journal**, v. 6, n. 4, p. 637-645, 2017.

PATRÍCIO, D.I.; RIEDER, R. Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. **Computers and Electronics in Agriculture**, v. 153, p. 69-81, 2018. DOI: <https://doi.org/10.1016/j.compag.2018.08.001>.

PEREIRA, Beatriz Montenegro. **Evolução técnico-econômica da produção de soja no brasil e na argentina**: uma análise histórico-comparada. 2021. 108 f. Dissertação. (Mestrado em Agronegócios) - Faculdade de Agronomia e Medicina Veterinária, Universidade de Brasília, Brasília, 2021.

PEREIRA, Leopoldo Mensch. **Análise de riscos ambientais na aplicação de defensivos agrícolas em propriedade rural**. 2014. 97 f. Monografia (Especialização) - Curso de Curso de Pós-Graduação Lato Sensu em Engenharia de Segurança do Trabalho, Departamento de Ciências Exatas e Engenharias, Universidade Regional do Noroeste do Estado do Rio Grande do Sul, Ijuí, 2014. Disponível em: <<https://bibliodigital.unijui.edu.br:8443/xmlui/bitstream/handle/123456789/2506/Monografia%20Versão%20Final%20pdf.pdf?sequence=1&isAllowed=y>>. Acesso em: 02 set. 2022.

PERES, F; MOREIRA, J. C. **É veneno ou é remédio?** - os desafios da comunicação rural sobre agrotóxicos. Rio de Janeiro: Fiocruz, 2003. 384 p.

PESKE, S. T.; BARROS, A. C. S. A.; SCHUCH, L. O. B. Produção de sementes. In: PESKE, S.T.; VILLELA, F.A.; MENEGHELLO, G.E. **Sementes: fundamentos científicos e tecnológicos**. 3ª edição. Pelotas: Ed. Universitária, p. 13-100, 2012.

PESKE, S. T; ROSENTHAL, M. D; ROTA, G. R. M. **Sementes: fundamentos científicos e tecnológicos**. 3. ed. Pelotas: Rua Pelotas, 2012. 573 p.

PICCOLI, Everton. **A importância da soja para o agronegócio**: uma análise sob o enfoque do aumento da produção de agricultores no município de Santa Cecília do Sul. 2018. 46 f. TCC (Graduação) - Curso de Administração, Faculdade e Escola, Tapejara, 2018. Disponível em: <<https://fatrs.com.br/faculdade/uploads/tcc/d464ec1e2f2c450aa33bb0e990b54878.pdf>>. Acesso em: 27 fev. 2023.

PIGNATI, Wanderlei. Uso de agrotóxicos no Brasil: perspectiva da saúde do trabalhador e ambiental. **Revista Brasileira de Medicina do Trabalho**, Mato Grosso, v. 16, n. 1, p. 37-37, nov. 2018. FRACTAL EDITORA LTDA. DOI: <http://dx.doi.org/10.5327/z16794435201816s1019>. Disponível em: <<https://cdn.publisher.gn1.link/rbmt.org.br/pdf/v16s1a19.pdf>>. Acesso em: 24 nov. 2021.

PIMENTEL, David; PIMENTEL, Marcia. Global environmental resources versus world population growth. **Ecological Economics**, [S.L.], v. 59, n. 2, p. 195-198, set. 2006. Elsevier BV. DOI: <http://dx.doi.org/10.1016/j.ecolecon.2005.11.034>>.

PINHEIRO, Rodrigo Malheiros. **Avaliação de diferentes substratos para teste de germinação de sementes de *Cunninghamia lanceolata***. 2021. 35 f. TCC (Graduação) - Curso de Engenharia Florestal, Centro de Ciências Rurais, Universidade Federal de Santa Catarina, Curitiba, 2021. Disponível em: <<https://repositorio.ufsc.br/bitstream/handle/123456789/223416/Rodrigo%20Malheiros%20Pinheiro.pdf?sequence=1&isAllowed=y>>. Acesso em: 11 out. 2022.

PINHEIRO, Romário de Mesquita; GADOTTI, Gizele Ingrid; MONTEIRO, Rita de Cassia Mota; BERNARDY, Ruan. Inteligência artificial na agricultura com aplicabilidade no setor sementeiro. 2021. **Diversitas Journal**, v. 6, p. 2984-2995. DOI: https://doi.org/10.48017/Diversitas_Journal-v6i3-1857.

PIVATO, M. A. **Mineração de regras de associação em dados georreferenciados**. 2006. 110 f. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, São Carlos, 2006.

POOJA, I; SHARMA, A.; SHARMA, A. Machine Learning: A Review of Techniques of Machine Learning. **JASC: Journal of Applied Science and Computations**. Volume 5, Issue 7, July /2018. p. 538-541, 2018.

PWC BRASIL. **Como a inteligência artificial pode promover um futuro sustentável?** São Paulo, 2019. Disponível em: <<https://www.pwc.com.br/pt/estudos/preocupacoes-ceos/mais-temas/2019/como-a-inteligencia-artificial-pode-promover-um-futuro-sustentavel.html>>. Acesso em: 17 jan. 2021.

RAU, Bruna Arteaga; ZANZI, José Vinícius dos Santos. **Influência do vigor de sementes na produção e na qualidade de sementes de soja**. 2021. 28 f. TCC (Graduação) - Curso de Agronomia, Faculdade de Ciências Agrárias, Universidade Federal da Grande Dourados, Dourados, 2021. Disponível em: <https://repositorio.ufgd.edu.br/jspui/bitstream/prefix/4668/1/BrunaArteagaRau_JoséViníciusdosSantosZanzi.pdf>. Acesso em: 11 out. 2022.

RAVA, C.; AIDAR, H.; COSTA, J. G. C. da; KLUTHCOUSKI, J. **Sementes Infectadas: sustentabilidade das lavouras questionada**. Brasília: Embrapa Arroz e Feijão, 2006. 3p. (Embrapa Arroz e Feijão. Informações Agronômicas, 115). Disponível em: <<https://ainfo.cnptia.embrapa.br/digital/bitstream/item/215745/1/CNPAF-2006-ia.pdf>>. Acesso em: 26 out. 2021.

REDDY, P.; V.; S.; Data mining and fuzzy data mining using map reduce algorithms. 2021. *In*: BIRANT, D. **Data Mining: methods, applications and systems**, 1. ed. Cap. 11, 2021. p.178-203. DOI: 10.5772/intechopen.92232. E-book.

REHMAN, T. U.; SULTAN MAHMUD, M. D.; CHANG, Y. K.; JIN, J.; SHIN, J. Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. **Computers and Electronics in Agriculture**, v. 156, p. 585-605. 2019. DOI: <https://doi.org/10.1016/j.compag.2018.12.006>.

RESENDE, Guilherme Mendes (ed.). **Cadernos do Cade: mercado de insumos agrícolas**. Brasília: Ministério da Justiça e Segurança Pública, 2020. 100 p. Disponível em: <<https://cdn.cade.gov.br/Portal/centrais-de-conteudo/publicacoes/estudos-economicos/cadernos-do-cade/mercado-de-insumos-agricolas-2020.pdf>>. Acesso em: 02 set. 2022.

RIBEIRO, Gabrielle de Araújo. **Análise da sustentabilidade ambiental na agroindústria canavieira utilizando a pegada hídrica**. 2019. 127 f. Tese (Doutorado) - Curso de Programa de Pós-Graduação em Recursos Naturais, Centro de Tecnologia e Recursos Naturais, Universidade Federal de Campina Grande, Campina Grande, 2019.

RODRIGUES, Marília Hortência Batista Silva; SANTOS, Adriana da Silva; MELO, Edinete Nunes de; SILVA, Joyce Naiara da; OLIVEIRA, Carlos Jardel Andrade. Vigor de sementes: métodos para análise e fatores que o influenciam. **Meio Ambiente (Brasil)**, [s. l], v. 2, n. 3, p. 43-52, 2020. Disponível em: <<https://www.meioambientebrasil.com.br/index.php/MABRA/article/view/53/47#>>. Acesso em: 07 set. 2022.

ROMANI, Samara. **Tratamento de sementes de soja com agrotóxicos: riscos e medidas de proteção**. 2017. 38 f. Monografia (Especialização) - Curso de Curso de Pós-Graduação "Lato Sensu" em Engenharia de Segurança do Trabalho, Universidade do Contestado, Concórdia, 2017.

ROSSI, R. F.; CAVARIANI, C.; FRANÇA-NETO, J. B. Vigor de sementes, população de plantas e desempenho agronômico de soja. **Revista de Ciências Agrárias**, v. 60, n. 3, p. 215- 222, 2017. DOI: <http://dx.doi.org/10.4322/rca.2239>.

ROTZ, C. A.; MONTES, F.; CHIANESE, D. S. The carbon footprint of dairy production systems through partial life cycle assessment. **Journal Of Dairy Science**, [S.L.], v. 93, n. 3, p. 1266-1282, mar. 2010. American Dairy Science Association. DOI: <<http://dx.doi.org/10.3168/jds.2009-2162>>.

RUBIN, Pedro Henrique; TRAVI, Magdalena Reschke Lajús. **Interferência na germinação e vigor de sementes de soja (*Glycine max* L.) submetidas a diferentes tratamentos de sementes**. 2020. 11 f. Monografia (Especialização) - Curso de Agronomia, UCEFF Faculdades, Chapecó, 2020.

SALTON, Júlio César. **Sistema Plantio Direto: o produtor pergunta, a embrapa responde**. 5. ed. Dourados: Embrapa-SPI, 2012. 254 p. (Coleção 500 Perguntas, 500 Respostas). Disponível em: <<https://www.infoteca.cnptia.embrapa.br/bitstream/doc/98258/1/500perguntassistemaplantiodireto.pdf>>. Acesso em: 27 fev. 2023.

SANTOS, Deyvyd Costa dos. **Proposta de ensino à distância utilizando metodologia ativa de aprendizagem: com uso do *random forest***. 2021. 79 f. Monografia (Especialização) - Curso de Licenciatura em Física, Instituto de Física, Universidade Federal Fluminense, Niterói, 2021. Disponível em: <<https://app.uff.br/riuff/bitstream/handle/1/23661/Deyvyd%20Costa%20dos%20Santos.pdf?sequence=1&isAllowed=y>>. Acesso em: 11 out. 2022.

SANTOS, Gislaíne Bergamo dos. **A percepção dos agricultores em relação ao uso de agrotóxicos e sementes transgênicas**. 2018. 73 f. Dissertação (Mestrado) - Curso de Programa de Pós-Graduação em Desenvolvimento Rural Sustentável, Centro de Ciências Agrárias da Unioeste, Universidade Estadual do Oeste do Paraná, Marechal Cândido Rondon, 2018. Disponível em: <[http://tede.unioeste.br/handle/tede/4034#:~:text=Conclui-se%20que%20os%20agrotóxicos,das%20pessoas%20e%20do%20ambiente](http://tede.unioeste.br/handle/tede/4034#:~:text=Conclui-se%20que%20os%20agrot%C3%B3xicos,das%20pessoas%20e%20do%20ambiente)>. Acesso em: 26 out. 2021.

SANTOS, Luiz Sérgio dos. Extração de conhecimento em bases de dados *knowledge discovery in database*. **Estudos**, Campinas, p. 11-19, jan. 2001. Revista da Faculdade de Ciências Humanas da Unimar. Disponível em: <http://ojs.unimar.br/index.php/estudos/article/viewFile/921/564>. Acesso em: 05 set. 2022.

SANTOS, Vinícius Cesar dos. **Uso de imagens térmicas para identificação da composição de alimentos**. 2022. 40 f. TCC (Graduação) - Curso de Bacharelado em Ciência da Computação, Universidade Estadual de Londrina, Londrina, 2022. Disponível em: https://sites.uel.br/dc/wp-content/uploads/2022/09/TCC_VINICIUS_CESAR_DOS_SANTOS.pdf. Acesso em: 11 out. 2022.

SARADA, C.; DEVI, M. S.; DEVI, M. Sathya. Imbalanced Big Data Classification using Feature Selection Under-Sampling. **Cvr Journal of Science & Technology**, [S.L.], v. 17, n. 1, p. 78-82, 1 dez. 2019. CVR College of Engineering. DOI: <http://dx.doi.org/10.32377/cvrjst1714>.

SCHIAFFINO, Matheus Costa. **Desenvolvimento de um método para classificação de comportamentos de ratos wistar utilizando o algoritmo de aprendizado supervisionado florestas aleatórias (*random forests*)**. 2020. 53 f. Dissertação (Mestrado) - Curso de Mestre em Neurociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2020.

SFERRA, Heloisa Helena; CORRÊA, Ângela M. C. Jorge. Conceitos e Aplicações de Data Mining. **Revista de Ciência & Tecnologia**, [S. L.], v. 11, n. 22, p. 19-34, 2003. Disponível em: <<https://www.sindipetroprsc.org.br/uploads/campanha/20091218130501.pdf>>. Acesso em: 07 set. 2022.

SILVA, Renato M.; ALMEIDA, Tiago A.; YAMAKAMI, Akebo. Análise de desempenho de redes neurais artificiais para classificação automática de web spam. **Revista Brasileira de Computação Aplicada**, Passo Fundo, v. 4, n. 2, p. 42-57, 28 nov. 2012. UPF Editora. <http://dx.doi.org/10.5335/rbca.2012.2195>. Disponível em: <http://seer.upf.br/index.php/rbca/article/view/2195/1880>. Acesso em: 11 out. 2022.

SILVA, Rodrigo. **Introdução ao WEKA**: software para mineração de dados. Software para mineração de dados. 2018. Disponível em: <<https://www.linkedin.com/pulse/introdução-ao-weka-software-para-mineração-de-dados-rosa-da-silva/?originalSubdomain=pt>>. Acesso em: 02 set. 2022.

SILVA, V. de P. R. da; ALEIXO, D. de O.; DANTAS NETO, J.; MARACAJÁ, K. F. B.; ARAÚJO, L. E. de. Uma medida de sustentabilidade ambiental: Pegada Hídrica. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.17, n.1, p.100-105, 2013.

SILVA, Vicente de Paulo Rodrigues da; ALEIXO, Danilo de Oliveira; ALMEIDA, Rafaela Silveira Rodrigues; CAMPOS, João Hugo Baracuy da Cunha; ARAÚJO, Lincoln Eloi de. An integrated approach based on water, ecological and carbon footprint for tracking human pressure on the planet. **Ambiência**, [S.L.], v. 11, n. 3, p. 639-649, dez. 2015. GN1 Genesis Network. DOI: <<http://dx.doi.org/10.5935/ambiencia.2015.03.09>>.

SINGH, Arti; GANAPATHYSUBRAMANIAN, Baskar; SINGH, Asheesh Kumar; SARKAR, Soumik. Machine Learning for High-Throughput Stress Phenotyping in Plants. **Trends In Plant Science**, [S.L.], v. 21, n. 2, p. 110-124, fev. 2016. Elsevier BV. DOI: <http://dx.doi.org/10.1016/j.tplants.2015.10.015>.

SISCOMEX - Ministério da Economia. **Superávit de fevereiro é o maior desde 2017 e comércio exterior bate novos recordes**. 2022. Disponível em: <<http://www.siscomex.gov.br/superavit-de-fevereiro-e-o-maior-desde-2017-e-comercio-exterior-bate-novos-recordes/>>. Acesso em: 16 mar. 2022.

SOAM, S.K.; RAGHUPHATI, B. **Artificial Intelligence in Agriculture: Global Status**. National Workshop on Artificial Intelligence in Agriculture. Indian Council of Agricultural Research. New Delhi. P.1-7. 2018.

SOARES, E. A. de M. G.; DAMASCENA, L. C. L. de; LIMA, L. M. M.; MORAES, R. M. de. Analysis of the Fuzzy Unordered Rule Induction Algorithm as a Method for Classification. In: CONGRESSO BRASILEIRO DE SISTEMAS FUZZY, 5., 2018, Fortaleza. **Recentes Avanços em Sistemas Fuzzy**. Fortaleza: Sbmec, 2018. p. 17-28.

SOARES, M. G. A **Quarta Revolução Industrial e seus possíveis efeitos no direito, economia e política**. Monografia. Universidade Autónoma de Lisboa, 2018. 33p. Disponível em: <https://www.migalhas.com.br/arquivos/2020/6/B86DDA9403078E_AQuartaRevolucaoIndustrialeseu.pdf>. Acesso em: 22 mar. 2022.

SOMBRA, Tobias Ribeiro; NUNES, Manuella Paula de Mesquita; SERRÃO, Gabriel Xavier; Araújo, Jonas Carneiro; SOUSA, Marco Antonio Paula; MORAIS, Emerson Cordeiro; DAHER, Luciara Celi Chaves; SILVA, André Guimarães Maciel. Utilização de redes bayesianas através do algoritmo naïve bayes para classificação de carcaças de ovinos. **Brazilian Journal of Development**, [S.L.], v. 6, n. 3, p. 10476-10498, 2020. Brazilian Journal of Development. DOI: <http://dx.doi.org/10.34117/bjdv6n3-068>.

SOUSA, Danielle Marie Macedo; BRUNO, Riselane de Lucena Alcântara; SILVA, Katiane da Rosa Gomes; TORRES, Salvador Barros; ANDRADE, Alberício Pereira. Viabilidade e vigor de sementes de *Poincianella pyramidalis* (Tul.) L. P. Queiroz pelo teste de tetrazólio. **Revista Ciência Agrônômica**, Fortaleza, v. 48, n. 2, p. 381-388, abr./maio 2017. Disponível em: <<https://www.scielo.br/j/rca/a/6vVSKhML5RlPxq4zGvhrbGw/?format=pdf&lang=pt>>. Acesso em: 10 nov. 2022.

SOUZA, Cibelle Noemia de Melo Vieira. **Qualidade de sementes de soja em função do armazenamento**. 2018. 52 f. Dissertação (Mestrado) - Curso de Agronomia, Instituto de Ciências Agrárias e Ambientais, Universidade Federal de Mato Grosso, Sinop, 2018. Disponível em: <https://ri.ufmt.br/bitstream/1/2939/1/DISS_2018_Cibelle%20Noemia%20de%20Melo%20Vieira%20Souza%20%281%29.pdf>. Acesso em: 10 nov. 2022.

SOUZA, Murilo Barreto. **Rede de aprendizado supervisionado como método de auxílio na detecção do ceratocône**. 2011. 111 f. Tese (Doutorado em Ciências) - Faculdade de Medicina da Universidade de São Paulo. São Paulo, 2011.

TALAVIYA, T.; SHAH, D.; PATEL, N.; YAGNIK, H.; SHAH, M. Implementation of artificial intelligence in agriculture for optimization of irrigation and application of pesticides and herbicides. **Artificial Intelligence in Agriculture**, v. 4, p. 58-73, 2020. DOI: <https://doi.org/10.1016/j.aiia.2020.04.002>.

TAN, P; STEINBACH, M; KUMAR, V. **Association analysis: basic concepts and algorithms**. Introduction to data mining. [s. l], p. 327-414, 2005. Disponível em: <<https://www-users.cse.umn.edu/~kumar001/dmbook/ch6.pdf>>. Acesso em: 07 set. 2022.

TECNAL. **Teste de vigor e germinação como parâmetros na qualidade de sementes**: sementes. sementes. 2022. Disponível em: <https://tecnal.com.br/pt-BR/blog/201_teste_de_vigor_e_germinacao_como_parametros_na_qualidade_de_sementes>. Acesso em: 11 out. 2022.

TEIXEIRA, J. C. Modernização da agricultura no Brasil: impactos econômicos, sociais ambientais. **Revista Eletrônica da Associação dos Geógrafos Brasileiros**, Três Lagoas-MS, v.1, n.2, p.21-42, 2005.

VAN DIJK, Aalt Dirk Jan; KOOTSTRA, Gert; KRUIJER, Willem; RIDDER, Dick de. Machine learning in plant science and plant breeding. **Iscience**, [S.L.], v. 24, n. 1, p. 101890, jan. 2021. Elsevier BV. DOI: <http://dx.doi.org/10.1016/j.isci.2020.101890>.

VANDERPLAS, J. **Python Data Science Handbook: Essential Tools for Working with Data**. 1. ed. Sebastopol: O'Reilly Media, 2016.

VASCONCELLOS, Renan Coelho de; BELTRÃO, Norma Ely Santos; PONTES, Altem Nascimento. Estimativa da pegada de carbono no setor de soja no município de Paragominas, Pará. **Revista Ibero-Americana de Ciências Ambientais**, [S.L.], v. 7, n. 1, p. 142-149, 9 jan. 2016. Companhia Brasileira de Produção Científica. DOI: <<http://dx.doi.org/10.6008/spc2179-6858.2016.001.0012>>.

VASCONCELOS, L. M. R. de; CARVALHO, C. L. de. Aplicação de regras de associação para mineração de dados na web. **Technical Report - RT-INF_004-04**, 2018.

VIEIRA FILHO, J. E. V.; SILVEIRA, J. M. F. J. da. Mudança tecnológica na agricultura: uma revisão crítica da literatura e o papel das economias de aprendizado. **Revista de Economia e Sociologia Rural**. 2012, v. 50, n. 4, p. 721-742. DOI: <https://doi.org/10.1590/S0103-20032012000400008>.

VIEIRA, Fábio Danilo; OLIVEIRA, Stanley Robson de Medeiros. Mineração de dados: conceitos e um estudo de caso sobre certificação racial de ovinos. In: MASSRUHÁ, S. M. F. S.; LEITE, M. A. de A.; LUCHIARI JUNIOR, A.; ROMANI, L. A. S. (ed.). **Tecnologias da informação e comunicação e suas relações com a agricultura**. Brasília: Embrapa, 2014. Cap. 14. p. 273-292. Disponível em: <<https://ainfo.cnptia.embrapa.br/digital/bitstream/item/126974/1/capitulo14-111-14.pdf>>. Acesso em: 02 set. 2022.

VIEIRA, Lucas Augusto. **Redução de uso de agrotóxicos por meio de inteligência artificial**. 2021. 95 p. Dissertação (Mestrado em Engenharia Ambiental) - Universidade Tecnológica Federal do Paraná. Apucarana, 2021.

VILANI, Ismael. **Avaliação da qualidade física e fisiológica de sementes de cultivares de trigo**. 2016. 40 f. TCC (Graduação) - Curso de Agronomia, Departamento de Estudos Agrários, Universidade Regional do Noroeste do Estado do Rio Grande do Sul – Unijuí, Ijuí, 2016. Disponível em: <<https://bibliodigital.unijui.edu.br:8443/xmlui/bitstream/handle/123456789/3744/Ismael%20Vilani.pdf?sequence=1&isAllowed=y>>. Acesso em: 11 out. 2022.

WACKERNAGEL, M.; REES, W. **Our Ecological Footprint: Reducing Human Impact on the Earth**. Philadelphia, PA: New Society Publishers, 1996. 160p.

WERNER, J. P. S.; ESQUERDO, J. C. D. M.; OLIVEIRA, S. R. de M. Uso da mineração de dados na classificação do algodão utilizando séries-temporais de imagens modis. Simpósio Brasileiro de Sensoriamento Remoto, 19, 2019, Santos, SP. **Anais [...]**. Santos, SP: INPE, 2019.

WESZ JÚNIOR, V. J. O mercado da soja no Brasil e na Argentina: semelhanças, diferenças e interconexões. **Século XXI: Revista de Ciências Sociais**, v. 4, n. 1, p. 114-161, 2014.

WIEDMANN, Thomas; MINX, Jan. A definition of 'carbon footprint'. **Ecological Economics Research Trends**, [S. L], p. 1-11, 2007. Disponível em: <encurtador.com.br/hwBGL>. Acesso em: 07 set. 2022.

WITTEN, Ian H.; FRANK, Eibe; HALL, Mak A. **Data Mining: practical machine learning tools and techniques**. 3. ed. Usa: Morgan Kaufmann Publishers, 2011. 665 p.

XU, R.; WUNSCH, D. Survey of clustering algorithms. **IEEE Transactions on Neural Networks**, v. 16, n. 3, p. 645-678, 2005.

YANG, H.; LIUSHENG, W.; HONGLI JUNMIN, X. Wireless Sensor Networks for Intensive Irrigated Agriculture. **Consumer Communications and Networking Conference**, 2007. CCNC 2007. 4th IEEE, p. 197-201. Las Vegas, Nevada. Jan. 2007.

YOOSEFZADEH-NAJAFABADI, Mohsen; TULPAN, Dan; ESKANDARI, Milad. Application of machine learning and genetic optimization algorithms for modeling and optimizing soybean yield using its component traits. **Plos One**, [S.L.], v. 16, n. 4, p. e0250665, 30 abr. 2021a. Public Library of Science (PLoS). DOI: <http://dx.doi.org/10.1371/journal.pone.0250665>.

YOOSEFZADEH-NAJAFABADI, Mohsen; TULPAN, Dan; ESKANDARI, Milad. Using Hybrid Artificial Intelligence and Evolutionary Optimization Algorithms for Estimating Soybean Yield and Fresh Biomass Using Hyperspectral Vegetation Indices. **Remote Sensing**, [S.L.], v. 13, n. 13, p. 2555, 30 jun. 2021b. MDPI AG. DOI: <http://dx.doi.org/10.3390/rs13132555>.