

AVALIAÇÃO DA SANIDADE DE SEMENTES DE AMENDOIM (*Arachis hypogaea*) POR MEIO DE PROCESSAMENTO DE IMAGENS.

GABRIEL COSTA PAIÃO¹; OTÁVIO BONACINA BERNADO¹; RENATA BRUNA DOS SANTOS COSCOLIN FAVAN²; JOÃO RICARDO FAVAN².

¹ Discente no curso Big Data no Agronegócio, FATEC Pompéia, Pompéia-SP.

² Docente do curso Big Data no Agronegócio, FATEC Pompéia, Pompéia-SP.

RESUMO:

Este estudo contempla a intersecção de três áreas: germinação de sementes de amendoim contaminadas com fungos, processamento de imagem e treinamento de modelos de aprendizado de máquina. Foi utilizado ferramentas computacionais, como inteligência artificial e processamento de imagens, para analisar a presença de fungos em sementes de amendoim durante o processo germinativo. As sementes foram divididas em grupos, colocadas em sacos diferentes contendo a mesma quantidade de sementes, e fotografadas a cada três dias ao longo de quatro coletas. Essas imagens foram analisadas no *software* ImageJ 1.54g (Java 1.8.0_345 64-bit), seguindo um protocolo específico, com o objetivo de determinar a área do fungo em pixels em cada imagem da semente, estas foram recortadas utilizando o *software* e GIMP, versão 2.10.36. Posteriormente, o *software* Orange Data Mining, versão 3.36.2, foi utilizado para estabelecer uma relação entre as imagens recortadas e as áreas de fungos, com o intuito de prever e classificar a área do fungo em uma determinada semente. Os treinamentos foram divididos com base em cada coleta, ou seja, em cada banco de dados de fotos. No final, foram realizados dois treinamentos, um utilizando toda a base de dados das imagens para um modelo de classificação e outro para cada uma das coletas para o modelo de previsão. Ademais, utilizaram-se ferramentas como Excel, juntamente com os resultados das tabelas do Orange, para apresentar de forma mais intuitiva as análises realizadas nos modelos. Por fim, apresentamos os resultados dos métodos de previsão e classificação.

Palavras-chave: Análise de Imagens; Processamento de Dados; Aprendizado de Máquina; Big Data no Agronegócio; *Aspergillus spp.*

1. INTRODUÇÃO

O amendoim, conhecido como *Arachis hypogaea* L., é uma planta pertencente à família das *Fabaceae*. Suas sementes são desenvolvidas subterraneamente, surgindo a partir de flores que se inclinam em direção ao solo, onde ocorre a frutificação. Essas flores apresentam uma morfologia distinta, com pedúnculos compridos que se estendem em direção ao solo, permitindo a

autofecundação. A planta do amendoim é uma herbácea de porte geralmente pequeno, com caules flexíveis e baixa presença de lenhina, tornando-os propensos a secar após a frutificação (Bashir et al., 2014).

O amendoim, a segunda leguminosa mais importante no mundo, é uma fonte significativa de proteína vegetal e óleo, com grãos que podem conter até 50% de lipídios (Nakagawa & Rosolem, 2011). Cultivado amplamente nas Américas, África e Ásia, é a quarta oleaginosa mais produzida globalmente. No Brasil, até o início da década de 70, destacava-se como um dos principais produtores, com São Paulo e Paraná responsáveis por 90% da produção nacional, utilizada para alimentação animal e fabricação de óleo vegetal e produtos industrializados como margarinas (Silva et al., 2019).

O amendoim auxilia a renovação de canaviais, sendo integrado à prática de rotação de culturas com leguminosas na cultura da cana-de-açúcar. Esta estratégia é recomendada durante a reforma dos canaviais, não interferindo na brotação da cana e resultando em aumentos significativos na produção de cana e açúcar em pelo menos dois cortes. Além disso, a rotação de culturas protege o solo contra a erosão e controla o crescimento de plantas espontâneas (Ambrosano et al., 2011).

O Estado de São Paulo lidera a produção de amendoim no Brasil, sendo responsável por cerca de 77% da área cultivada. A produção concentra-se no Sudeste, seguido pelo Centro-Oeste e Nordeste. Em São Paulo, o amendoim é tradicionalmente cultivado em duas regiões: Alta Mogiana (região de Ribeirão Preto e Jaboticabal), onde é plantado durante a renovação dos canaviais, e Alta Paulista (região de Marília e Tupã), onde se utilizam áreas de pastagens e canaviais para produzir duas safras anuais (safra das águas e safrinha).

O principal problema que afeta a qualidade dos grãos de amendoim é a contaminação por aflatoxina, causada pelo excesso de umidade durante a colheita, especialmente na safra das águas, o que levou a uma redução significativa da área plantada e produção no Brasil entre as décadas de 70 e 90. O amendoim pode ser contaminado por fungos como *Aspergillus spp.*, *Penicillium spp.*, *Fusarium spp.* e *Rhizopus spp.*, que podem infectar as sementes no solo, durante a colheita, secagem, beneficiamento e armazenamento (Rossetto et al., 2003).

A identificação das espécies de fungos contaminantes é um importante indicador da presença de micotoxinas e pode ser usada como método preventivo (Farias et al., 2000). No entanto, Ranjan & Sinha (1991) observa-se que a

identificação dos fungos não garante a detecção de micotoxinas, pois algumas cepas não produzem metabólitos tóxicos. O crescimento de *Aspergillus flavus* durante o armazenamento depende do teor de água dos grãos e da umidade relativa do ar, sendo a temperatura um fator determinante para a velocidade de crescimento (Dhingra & Coelho Neto, 1998). A prevenção da contaminação por aflatoxina em grãos de amendoim pode ser alcançada por meio de práticas como rotação de culturas, irrigação, suplementação de minerais e procedimentos adequados de colheita e secagem (Mixon et al., 1984).

As aflatoxinas, pertencentes à classe das micotoxinas, são substâncias tóxicas produzidas por fungos do gênero *Aspergillus*, especialmente *Aspergillus flavus* e *Aspergillus parasiticus*. Estes fungos proliferam em alimentos e produtos agrícolas em condições de alta umidade e temperatura. Descobertas em 1960, as aflatoxinas são metabólitos secundários que podem contaminar alimentos como amendoim, algodão, arroz, sorgo, milho, feijão, cacau, castanha-do-pará, noz e mandioca, sem alterar suas características sensoriais (Farias et al., 2000; Ranjan & Sinha, 1991). No amendoim, essas toxinas podem estar presentes tanto no farelo quanto nos grãos *in natura* e produtos industrializados destinados ao consumo humano (Dhingra & Coelho NETO, 1998; Mixon et al., 1984).

A presença de fungos no amendoim é facilitada pela riqueza das sementes em carboidratos e proteínas, criando um ambiente propício para seu desenvolvimento. Durante o cultivo, as vagens subterrâneas são especialmente vulneráveis a infecções fúngicas devido às condições ideais de umidade e temperatura (Nascimento et al., 2018). Além disso, danos mecânicos durante a colheita e o descascamento, bem como a ação de insetos no campo e no armazenamento, aumentam a suscetibilidade das sementes à contaminação fúngica (Fernandes et al., 2020).

A infecção por fungos não só compromete a qualidade das sementes, mas também afeta negativamente o processo de germinação, prejudicando a viabilidade e o crescimento das plantas (Smith & Jones, 2016). Isso resulta em menor produtividade e aumento dos riscos para a saúde de humanos e animais que consomem os produtos contaminados (Silva et al., 2019).

O Aprendizado de Máquina é uma subárea da inteligência artificial focada em desenvolver algoritmos que permitem aos computadores aprenderem a partir de dados e tomar decisões sem programação específica. Analisando grandes volumes

de dados, esses sistemas identificam padrões e fazem previsões baseadas em dados históricos, sendo úteis em diversas aplicações, como previsão de doenças e classificação de imagens (Goodfellow et al., 2016).

Na análise de imagens, algoritmos de aprendizado de máquina, como redes neurais convolucionais (CNNs), têm demonstrado excelente desempenho, especialmente no reconhecimento e classificação de padrões visuais. Essas redes são eficazes na detecção e quantificação de áreas específicas, como infecções fúngicas em sementes de amendoim contaminadas por *Aspergillus*. O processo envolve o treinamento de um modelo com dados rotulados, permitindo que ele aprenda a associar características visuais a diferentes níveis de infecção (Krizhevsky et al., 2012; LeCun et al., 2015).

Um sistema de aprendizado é um programa de computador que toma decisões baseado em experiências acumuladas através da solução bem-sucedida de problemas anteriores (Monard e Baranauskas, 2003, p. 39). Esse conceito é fundamental para a aplicação de aprendizado de máquina na detecção de infecções fúngicas em sementes de amendoim, permitindo quantificar com precisão a extensão da infecção a partir de padrões identificados em imagens. O presente trabalho destaca-se ao utilizar técnicas de aprendizado de máquina, oferecendo uma análise mais detalhada e confiável em comparação com abordagens tradicionais, possibilitando um monitoramento mais eficaz e uma intervenção oportuna no controle dessas infecções.

O objetivo geral deste trabalho é desenvolver um método computacional para processar imagens de amendoins contaminados por fungos, visando estimar a intensidade da infestação. Além disso, o projeto visa avaliar algoritmos de aprendizado de máquina para essa finalidade.

Por fim, modelos de aprendizado de máquina foram treinados com base na relação entre imagens de sementes de amendoim contaminadas e a respectiva área fúngica detectada na semente, visando acompanhar a evolução da proliferação fúngica.

2. MATERIAIS E MÉTODOS

Utilizou-se uma base de dados desenvolvida a partir das imagens de sementes de amendoins e do processamento delas. Com isso, realizou-se o

treinamento de modelos de aprendizado de máquinas cujo objetivo foi a previsão da intensidade fúngica dada uma imagem ou até mesmo a classificação da intensidade.

As sementes utilizadas foram obtidas de um lote colhido na safra de 2023, na região de Ocaçu. A partir destas sementes, cria-se uma metodologia para aquisição de imagens de amostras de sementes de amendoim para identificação de contaminantes durante o teste de germinação e vigor.

2.1. Equipamentos utilizados, preparo das sementes e coleta das imagens

O experimento foi conduzido no Laboratório de Química da FATEC Shunji Nishimura de Pompeia, no período entre os dias 01 e 08 de março de 2024. As sementes foram submetidas ao teste de germinação, no qual foram observados potencial crescimento de contaminantes.

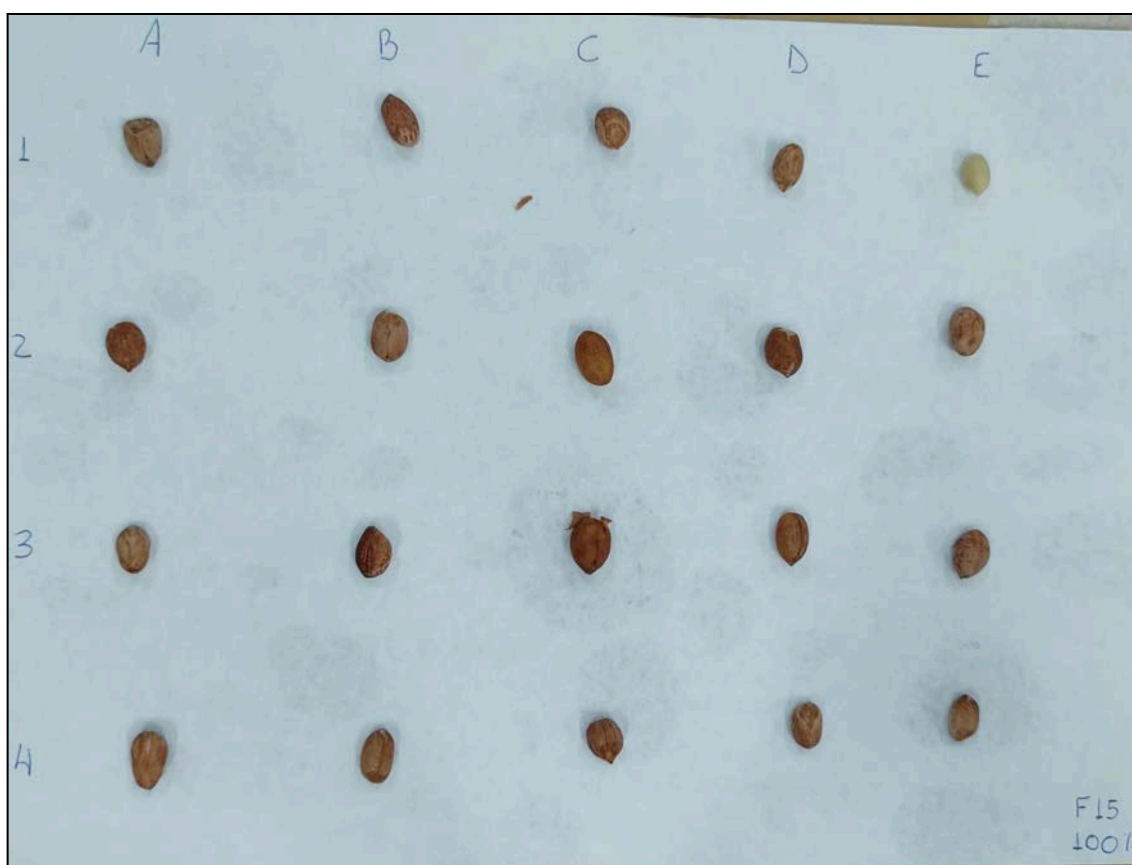
Durante esta fase do projeto, as imagens das sementes foram capturadas utilizando um *Smartphone*, o qual foi posicionado de forma estável em um tripé, complementado por um anel de luz para otimizar a iluminação. O celular utilizado foi um Motorola Moto G 9 Plus, equipado com uma câmera principal de 64 MP (abertura f/1.8), uma lente *ultra wide* de 8 MP (abertura f/2.2), uma lente macro de 2 MP (abertura f/2.2) e um sensor de profundidade de 2 MP (abertura f/2.2). Entre as funcionalidades adicionais do celular, destacam-se as linhas de referência visíveis e um temporizador de 10 segundos. Em relação ao tripé, este foi utilizado com altura máxima de 1,4 m (do solo ao ponto mais alto do tripé), abertura mínima de 30 cm e uma segunda alavanca de 14 cm. A distância focal entre o celular e o papel foi de 39,5 cm, enquanto a distância entre a luz e o papel foi de 50 cm. A mesa onde os equipamentos estavam posicionados estava a 90 cm do solo.

Inicialmente, as sementes foram higienizadas, utilizando hipoclorito de sódio (1,0%) por 5 minutos; em seguida, lavada em água destilada. Os substratos utilizados foram folhas de papel Germitest, sendo estas pesadas e umedecidas com 3 vezes o peso do papel seco. O teste foi conduzido com 20 sementes, distribuídas de forma equidistante em cada rolo, sendo num total de 400 sementes; em seguida, os rolos confeccionados foram acondicionados em câmara germinadora tipo B.O.D., com temperatura constante de $25\pm 2^{\circ}\text{C}$.

As imagens foram adquiridas em 1º de março e repetidas em 4, 6 e 8 de março de 2024. A contagem final ocorreu no décimo dia, seguindo as Regras de Análise de Sementes (Tillmann et al., 2009).

Para identificar cada imagem, adotou-se o seguinte sistema: cada rolo foi designado com um número de F1 a F20 (único para cada rolo), e foram utilizadas identificações para colunas (A, B, C, D e E) e linhas (1, 2, 3 e 4) para cada semente, conforme ilustrado na Figura 1.

**Figura 1 – Fotografia das sementes dispostas no papel para germinação.
Anotação de colunas, linhas e folha**



Fonte: Elaborado pelos autores (2024)

2.2. Recorte das sementes e utilização do software ImageJ

Após a coleta das imagens, cada imagem da semente foi recortada utilizando o software GIMP, versão 2.10.36 sendo redimensionadas para um formato de 360 x 360 pixels. Em seguida, foram identificadas por meio de um conjunto de caracteres,

como, por exemplo, "f1a1", onde "f1" indica a folha, "a" a coluna, e "1" a linha. A Figura 2 ilustra esse modelo.

Figura 2 – Imagem de semente recortada que será utilizada como base de dados para processamento de imagens no *software* ImageJ 1.54g (Java 1.8.0_345 64-bit). Esta imagem será identificada com base no dia, folha, coluna e linha a que pertencia.

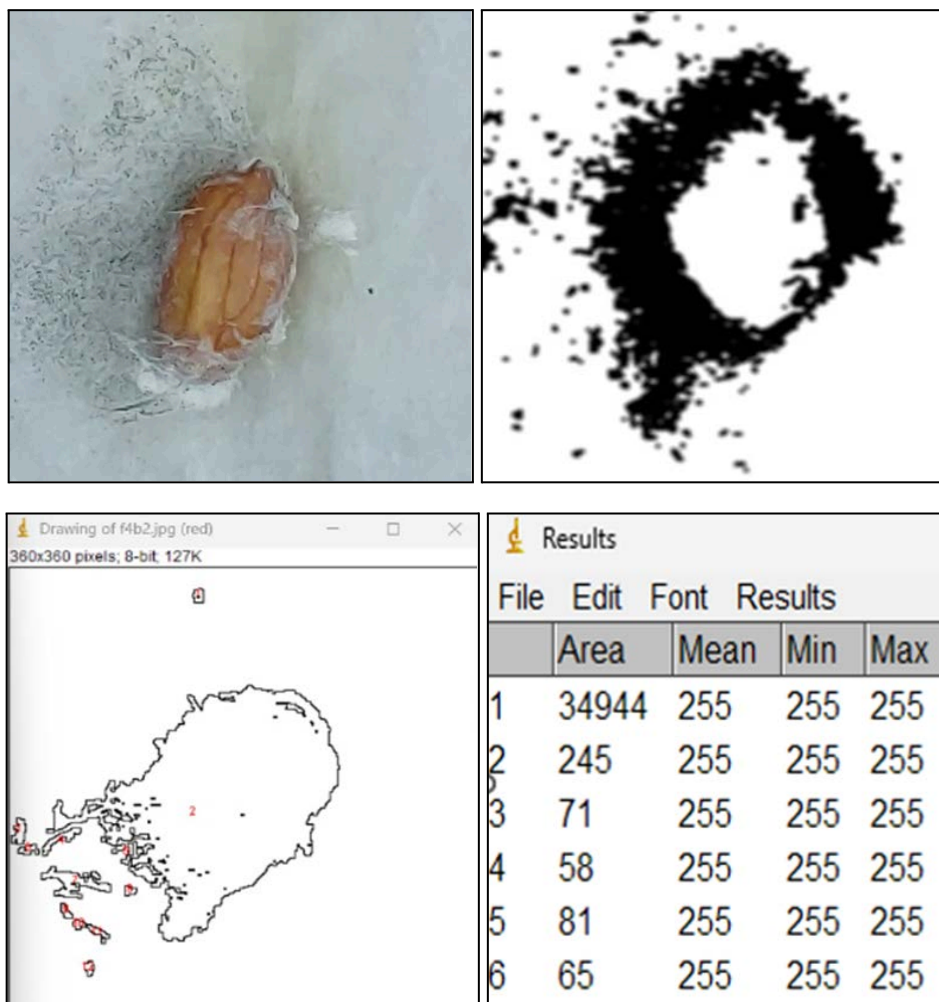


Fonte: Elaborado pelos autores (2024)

Após recortar e identificar todas as imagens das sementes de amendoim, aplicou-se um protocolo de processamento de imagens utilizando o *software* ImageJ 1.54g (Java 1.8.0_345 64-bit). Este protocolo utiliza a imagem recortada, submetendo-a a um tratamento que faz uma descoloração, deixando com que o fungo fique destacado na cor preta e o resto da imagem em branco. Após o destaque, a área fúngica afetada é contornada e, em seguida, numerada e descrita em uma tabela, indicando suas respectivas dimensões em pixels. Finalmente, somam-se os valores da coluna "Área" da tabela para obter a área fúngica total, em pixels, do fungo presente na semente específica. Vale ressaltar que o protocolo usado no processamento das imagens por meio do *software* ImageJ ainda não foi publicado pelo autor.

Um exemplo desse procedimento está ilustrado na sequência de imagens da Figura 3.

Figura 3 – Protocolo Software ImageJ. Semente de amendoim disposta com fungo. Fungo identificado e destacado em cor preta. Identificação e numeração das áreas fúngicas. Área em pixel de cada região portadora do fungo.



Fonte: Elaborado pelos autores (2024)

Assim, cada semente identificada é vinculada à sua área de infestação, calculada pela soma das regiões detectadas como fungos, como indicado na última imagem da Figura 3.

Uma planilha eletrônica foi criada utilizando o software Microsoft Excel, versão 2404 Build 16.0.17531.20152, para registrar a identificação de cada semente de amendoim e a respectiva intensidade de contaminação fúngica em relação à área total da imagem.

2.3. Processamento, treinamento de modelos de *machine learning* e análise de dados

Um fluxo de trabalho de análise de dados foi estabelecido utilizando a ferramenta Orange Data Mining, versão 3.36.2 (Demsar et al., 2013), no qual inclui o carregamento e processamento dos dados e ajuste de modelos de aprendizado de máquinas para predição da intensidade fúngica com base nas imagens das sementes de amendoim, como ilustrado pela Figura 4.

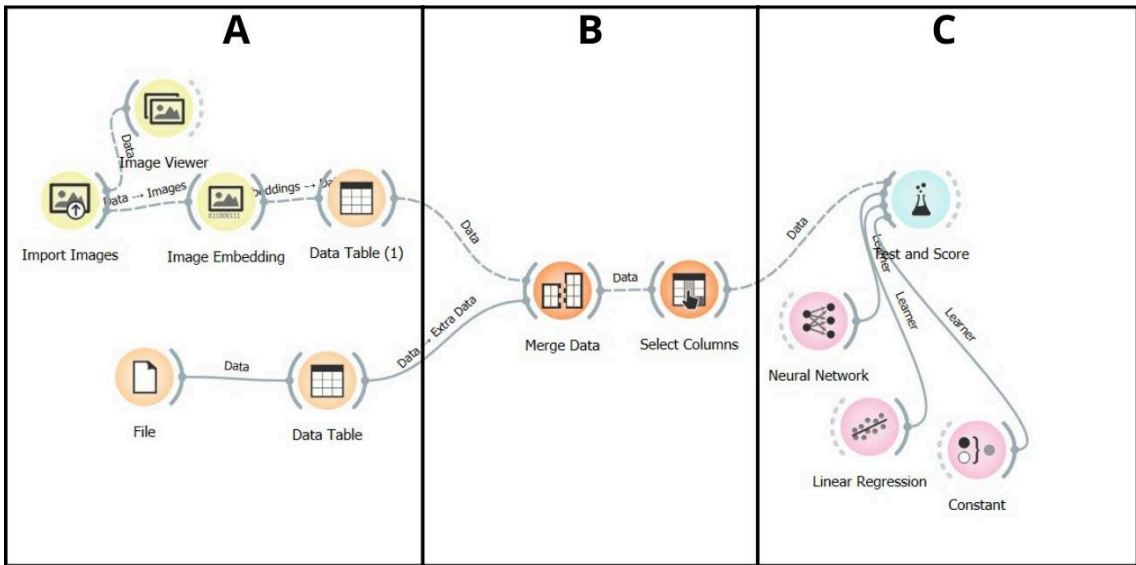
O carregamento dos dados foi feito por meio da importação das imagens individuais da semente de amendoim atacada pelos fungos e da planilha eletrônica com os dados da área fúngica para cada imagem, estes dois atributos foram combinados formando o conjunto de dados utilizados para análise.

As imagens individuais de sementes de amendoim foram processadas utilizando o algoritmo de Redes Residuais (ResNet) (Koonce, 2021), implementado pelo pacote *Image Analysis* do Orange Data Mining, no qual, as imagens são transformadas em um conjunto de 999 variáveis independentes de acordo o processo de convolução utilizado no processamento da imagem (He et al., 2016)

Para comparação, foram ajustados 3 modelos de aprendizado de máquinas utilizando os algoritmos de Perceptron de multicamadas (Taud; Mas, 2018), regressão linear (Malud; Abdulazeez, 2020) e valor constante (Demsar et al., 2013), todos eles em suas implementações no programa Orange Data Mining. Os modelos ajustados foram avaliados por meio dos indicadores MSE, RMSE, MAE, MAPE e R^2 (Kwon et al., 2021). Este processo é destacado na Figura 4.

Pode-se observar que a imagem foi separada em 3 quadrantes nomeados em “A”, “B” e “C”. Sendo assim, a parte nomeada em “A” fica responsável pelo carregamento dos dados, tanto das imagens quanto das planilhas eletrônicas do Excel. No meio, descrito por “B”, há a junção destes dados através da nomenclatura da semente, a qual é a mesma responsável por identificar a intensidade fúngica de uma determinada semente. Por último, o quadrante “C” realiza o treinamento e retorna os valores obtidos dos aprendizados de máquinas.

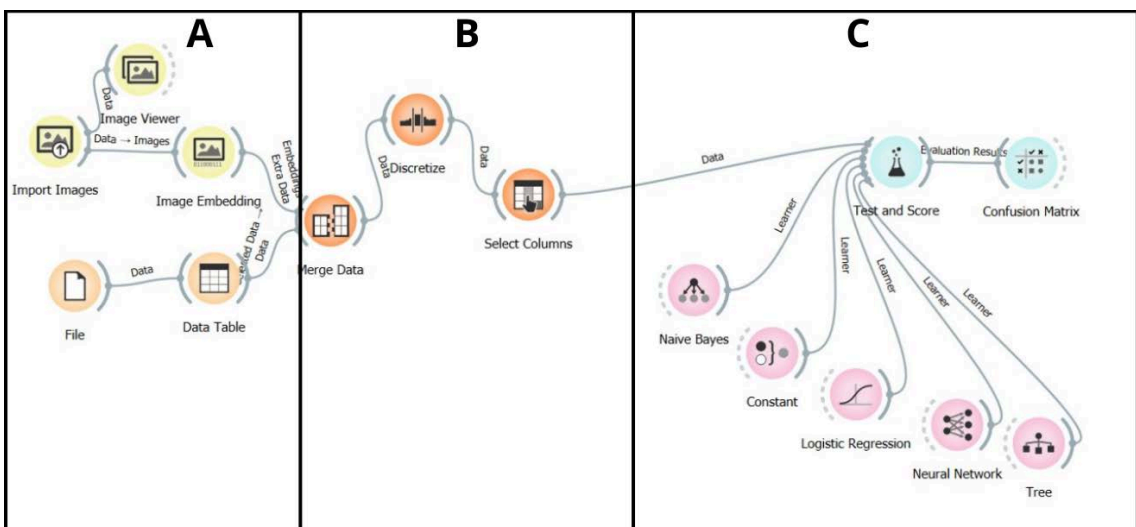
Figura 4 - Protocolo Software Orange Data Mining. Utilizado para análise separada das Coletas 1, 2 e 3



Fonte: Elaborado pelos autores (2024)

Em seguida, o mesmo procedimento foi utilizado, porém agora o treinamento foi com base em todos os dados disponíveis. Além disso, foram treinados diferentes tipos de *machine learning*. A Figura 5 evidencia isto.

Figura 5 – Protocolo Software Orange Data Mining. Utilizado para análise conjunta de todas as coletas



Fonte: Elaborado pelos autores (2024)

3. RESULTADOS E DISCUSSÃO

Para as Coletas 1, 2 e 3, a análise preditiva foi utilizada para identificar a área fúngica em pixels para cada imagem de semente analisada. Em seguida, tem-se a implementação do método de classificação, utilizando todo o dataset das imagens capturadas, revelou-se eficaz ao determinar a severidade do fungo nas sementes, classificando-as em baixa, moderada ou alta intensidade. Os resultados obtidos demonstraram que a classificação forneceu números mais eficazes durante o treinamento, comprovando a eficiência do modelo em categorizar corretamente a intensidade fúngica e proporcionando uma melhor compreensão do comportamento dos contaminantes nas amostras.

Este modelo foi feito com base nos dados das Coletas 1, 2 e 3 de forma separada. Utilizando os algoritmos *Neural Network*, *Linear Regression* e *Constant*, observou-se que as métricas para avaliar o desempenho dos modelos foram o *Mean Squared Error* (MSE) e o Coeficiente de Determinação (R^2). O MSE representa a média dos quadrados dos erros entre os valores previstos e os valores reais, sendo que valores menores indicam um melhor ajuste do modelo aos dados. O R^2 , por sua vez, mede o quão bem os resultados previstos se ajustam aos dados reais. Para evidenciar a discussão em relação a previsão tem-se a Tabela 1.

Tabela 1- Resultado do processamento para previsão utilizando o *software* Orange Data Mining para as Coletas 1, 2 e 3

Model	Coleta 1		Coleta 2		Coleta 3	
	MSE	R^2	MSE	R^2	MSE	R^2
Rede Neural	0,027	-0,888	0.034	-0.168	0,159	-6,793
Constante	0,015	-0,003	0.029	-0.005	0,021	-0,005
Regressão Linear	0,042	-1,927	0.103	-2.539	0,028	-0,353

Fonte: Elaborado pelos autores (2024)

Valores negativos de R^2 indicam que os modelos não são adequados para o método de previsão utilizado, demonstrando que o sistema não foi eficaz na

previsão da área do fungo dado uma semente. Portanto, esses modelos não podem ser utilizados para essa finalidade.

Em uma segunda análise, foi submetido todo o conjunto de dados a um procedimento para desenvolver um modelo de classificação. Aplicou-se os algoritmos *Neural Network*, *Constant*, *Logistic Regression*, *Naive Bayes* e *Tree*. As métricas para avaliar o desempenho dos modelos foram *Classification Accuracy* (CA) e F1 Score.

A CA é uma medida da proporção de previsões corretas feitas pelo modelo em relação ao total de previsões realizadas. Quanto maior o valor da CA, melhor é o desempenho do modelo. O F1 Score, por sua vez, é uma métrica que combina precisão e *recall* em uma única medida. Ele varia de 0 a 1, sendo 1 a melhor pontuação possível.

Tabela 2 - Processamento de Dados Análise Conjunta

MODEL	AUC	CA	F1	PREC	RECALL
NEURAL NETWORK	0,805	0,658	0,654	0,653	0,658
CONSTANT	0,438	0,453	0,282	0,205	0,453
LOGISTIC REGRESSION	0,803	0,657	0,657	0,658	0,657
NAIVE BAYES	0,732	0,510	0,505	0,578	0,510
TREE	0,636	0,572	0,573	0,574	0,572

Fonte: Elaborado pelos autores (2024)

De acordo com a Tabela 2, a rede neural e a regressão logística apresentam os melhores resultados, com taxas de acurácia superiores a 65%, indicando um bom desempenho na classificação do fungo. O modelo Naive Bayes tem um desempenho moderado, com métricas gerais em torno de 0,51. O modelo de árvore de decisão possui um desempenho intermediário, com um AUC de 0,636 e outras métricas em torno de 0,57. Por outro lado, o modelo constante tem o pior desempenho, com um

AUC de 0,438 e uma baixa precisão de 0,205, mostrando-se ineficaz para este conjunto de dados.

Tabela 3 - Matriz de Confusão sobre análise conjunta

	BAIXA	MEDIA	ALTA	
BAIXA	75.8 %	25.5 %	6.2 %	405
MEDIA	22.5 %	61.0 %	47.8 %	378
ALTA	1.8 %	13.5 %	46.0 %	111
	396	385	113	894

Fonte: Elaborado pelos autores (2024)

A matriz de Confusão, é uma tabela utilizada para descrever o desempenho de um modelo de classificação em um conjunto de dados de teste para os quais os valores verdadeiros são conhecidos.

A matriz compara os valores alvo reais com aqueles previstos pelo modelo. As linhas representam as instâncias em uma classe real e as colunas representam instâncias em uma classe prevista.

Esta matriz de confusão específica tem três classes: BAIXA, MÉDIA e ALTA. Cada célula dentro da matriz mostra a porcentagem de previsões feitas pelo modelo, com totais para cada classe fornecidos nas margens.

Por exemplo, 75.8% de BAIXA foi corretamente previsto como BAIXA, enquanto 25.5% foram incorretamente previstos como MÉDIA e 6.2% como ALTA.

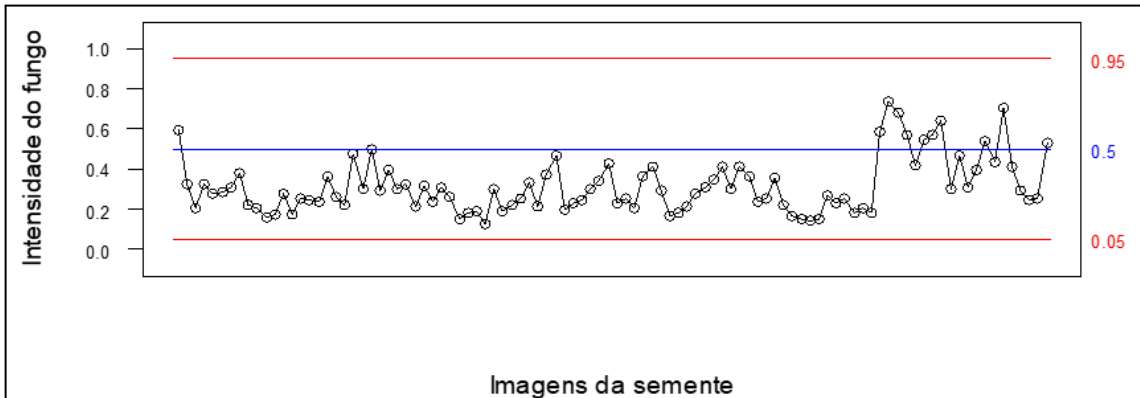
A matriz de confusão é interessante porque fornece insights não apenas sobre quantas previsões estavam corretas (a diagonal do canto superior esquerdo para o canto inferior direito onde as classes verdadeiras se alinham com as classes previstas), mas também sobre quais tipos de erros estavam sendo cometidos pelo modelo de classificação.

3.1. Análise dos resultados por meio de Gráficos de Dispersão

O gráfico da Figura 6 exibe a intensidade do fungo em função das imagens. Definiu-se uma escala de 0.0. a 1.0 e uma média de 0.5. Analisando os resultados, observa-se que a tendência da intensidade se concentra em sua maioria na parte inferior à média, ou seja, os resultados são classificados com baixa intensidade para

a primeira coleta. Pode-se destacar a existência de *outliers*, contudo há estabilidade do processo na parte inferior do gráfico.

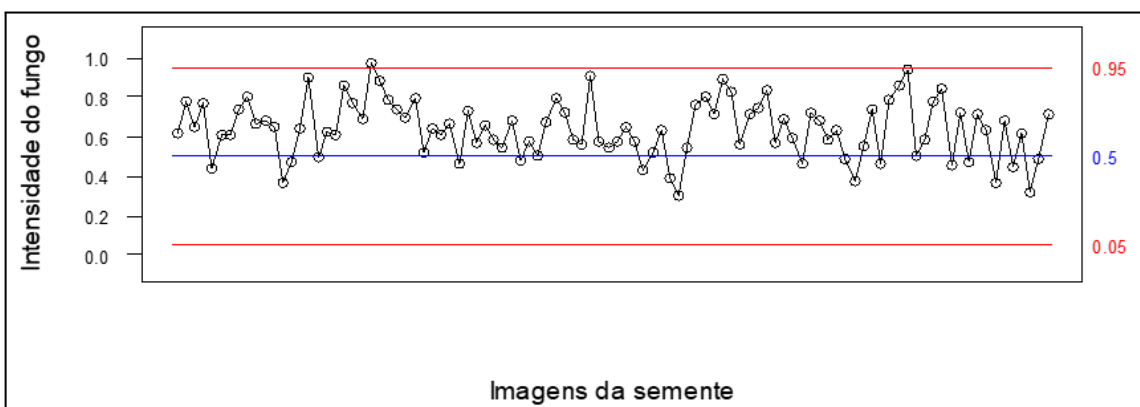
Figura 6 - Descrição gráfica da intensidade fúngica sobre a base de imagens em relação à Coleta 1



Fonte: Elaborado pelos autores (2024)

O gráfico de dispersão referente à Coleta 2 está apresentado na Figura 7. Pode-se observar que a intensidade do fungo por semente é maior em comparação com a primeira coleta. Verifica-se que a intensidade está, na maioria dos casos, em torno ou acima da média de 0,5. Conclui-se, portanto, que o fungo se desenvolveu no período entre a Coleta 1 e a Coleta 2.

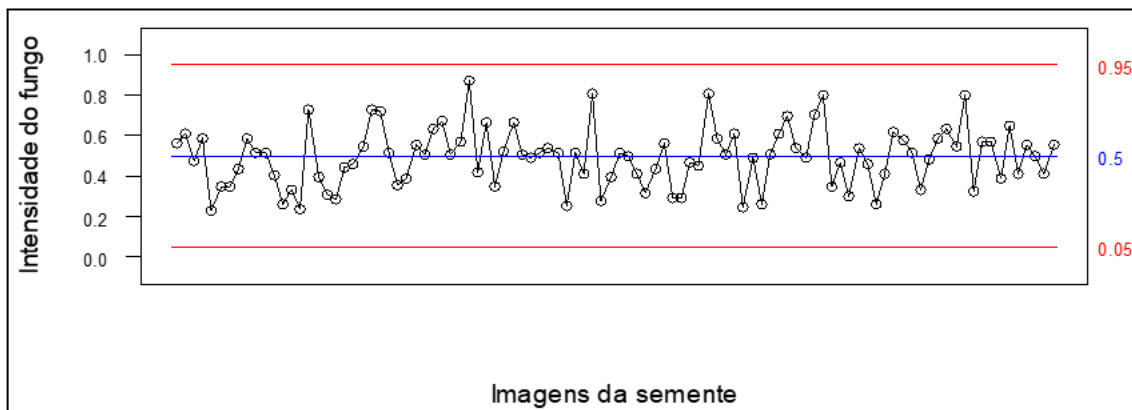
Figura 7 - Descrição gráfica da intensidade fúngica sobre a imagem em relação à Coleta 2



Fonte: Elaborado pelos autores (2024)

Para a Coleta 3, o gráfico de dispersão mostrou intensidades concentradas entre os pontos 0,2 e 0,8. Esta distribuição se concentra ao redor da média, sem a presença de pontos fora dos limites.

Figura 8 - Descrição gráfica da intensidade fúngica sobre a imagem em relação à Coleta 3



Fonte: Elaborado pelos autores (2024)

3.2. Análise dos resultados por meio de Gráficos de Box Plot

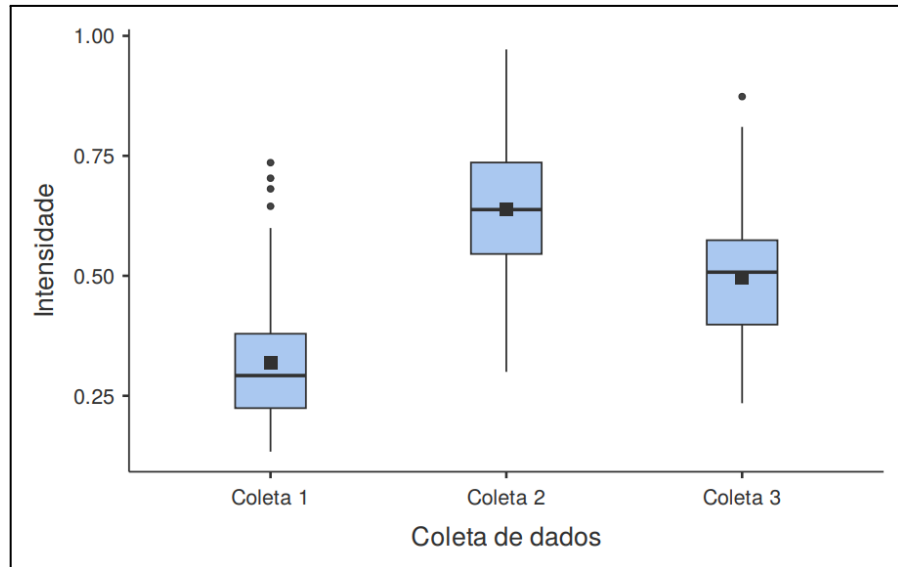
O gráfico de box plot apresentado pela Figura 9 ilustra a distribuição de intensidades em três diferentes coletas de dados. Na Coleta 1, tem-se uma mediana ligeiramente abaixo de 0,5, com a maior dispersão entre os dados, indicada por vários outliers acima de 0,75.

A Coleta 2 demonstra uma mediana em torno de 0,75, o que indica que a maioria dos dados está concentrada em níveis de intensidade mais elevados em comparação com as outras coletas. A caixa vai de cerca de 0,5 a 0,85, e os bigodes se estendem de 0,25 até próximo de 1,0, sem a presença de outliers visíveis. Isso sugere que os dados da Coleta 2 são mais consistentes e menos variáveis do que os da Coleta 1, apresentando uma dispersão menor.

Por fim, a Coleta 3 apresenta uma mediana ligeiramente abaixo de 0,75, com a caixa se estendendo de 0,5 até um pouco abaixo de 0,75. Há um outlier acima de 1,0, indicando um valor excepcionalmente alto. Os bigodes se estendem de cerca de 0,25 até próximo de 1,0, mostrando uma dispersão intermediária entre as Coletas 1 e 2.

No geral, a Coleta 3 possui níveis de intensidade intermediários e uma variabilidade moderada, situando-se entre a alta variabilidade da Coleta 1 e a consistência da Coleta 2.

Figura 9 - Gráfico Box Plot da intensidade fúngica com base nas coletas 1, 2 e 3



Fonte: Elaborado pelos autores (2024)

4. CONCLUSÃO

Diante dos resultados obtidos nos processos empregados, pode-se concluir que embora o treinamento dos modelos de machine learning não tenha alcançado o nível desejado de precisão para o processo de previsão, eles se mostraram eficazes para tarefas de classificação. Com uma precisão de 65%, o modelo é viável para classificação. Os modelos de *Machine Learning* conseguiram diferenciar entre essas classes com métricas de assertividade relevantes, demonstrando a capacidade de classificar a intensidade fúngica em sementes de amendoim por meio de imagens digitais.

A metodologia empregada contribuiu para a comunidade científica, ao explorar o uso de processamento de imagens na detecção de fungos em amendoins. Apesar das limitações encontradas no processo de previsão, a abordagem de classificação demonstrou ser promissora, abrindo novas possibilidades para aplicações futuras neste contexto.

REFERÊNCIAS

AMBROSANO, Edmilson José et al. **Produtividade da cana-de-açúcar após o cultivo de leguminosas**. *Bragantia*, v. 70, p. 810-818, 2011. Disponível em: <https://www.scielo.br/j/brag/a/gH8gGWp39kMDHCG8kQMn3Wn/?lang=pt> Acesso em: 27.mar.2024

Barbosa, R. M., Homem, B. F. M., & Tarsitano, M. A. A.. (2014). **Custo de produção e lucratividade da cultura do amendoim no município de Jaboticabal, São**

Paulo. Revista Ceres, 61(4), 475–481. 2014. Disponível em:
<https://doi.org/10.1590/0034-737X201461040005>. Acesso em: 02.abr.2024

BASHIR, M., ISSA KHAN, M., & IQBAL, A. **Advances in Groundnut (Arachis hypogaea L.) Research and Development**. Nova York, NY: Springer. (2014).

DEMŠAR, Janez et al. **Orange: data mining toolbox in Python**. *the Journal of machine Learning research*, v. 14, n. 1, p. 2349-2353, 2013. Disponível em:
<https://www.jmlr.org/papers/volume14/demsar13a/demsar13a.pdf> Acesso em: 27.mai.2024

DHINGRA, O.D.; COELHO NETO, R.A. Micotoxinas em grãos. **Revisão Anual de Patologia de Plantas**, Passo Fundo, v.6, n.1, p.49-101, 1998. Disponível em:
<https://repositorio.inpa.gov.br/handle/1/35535>. Acesso em: 17.abr.2024

FARIAS, A. X. D., ROBBS, C. F., BITTENCOURT, A. M., ANDERSEN, P. M., & CORRÊA, T. B. S.. (2000). **Contaminação endógena por Aspergillus spp. em milho pós-colheita no Estado do Paraná**. Pesquisa Agropecuária Brasileira, 35(3), 617–621. Disponível em: <https://doi.org/10.1590/S0100-204X2000000300018>. Acesso em: 10.abr.2024

FERNANDES, M. C., et al. **Effect of mechanical damage on peanut (Arachis hypogaea L.) seed colonization by Aspergillus flavus**. *Journal of Stored Products Research*, 88, 101656. (2020).

FORNARI, Gabriel. **Caracterização e desenvolvimento de um sistema autoadaptativo para estimação da posição de veículos aéreos não tripulados baseado em imagens**. 2020. 183 p. Tese (Doutorado em Computação Aplicada) - Instituto Nacional de Pesquisa Espaciais, São José Dos Campos, 2020. Disponível em:
<http://mtc-m21c.sid.inpe.br/col/sid.inpe.br/mtc-m21c/2021/02.05.01.02/doc/publicacao.pdf> Acesso em: 1.mai.2024

Goodfellow, I., Bengio, Y. and Courville, A., 2016. *Deep learning*. MIT press. Disponível em:
<https://books.google.com.br/books?id=omivDQAAQBAJ&lpg=PR5&ots=MON5ansFVT&dq=%20Deep%20Learning.%20MIT%20Press&lr&pg=PR5#v=onepage&q=Deep%20Learning.%20MIT%20Press&f=false>. Acesso em: 18.mar.2024

KOONCE, Brett; KOONCE, Brett. ResNet 50. **Convolutional neural networks with swift for tensorflow: image recognition and dataset categorization**, p. 63-72, 2021. Disponível em: https://link.springer.com/chapter/10.1007/978-1-4842-6168-2_6 Acesso em: 27.mai.2024

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. *nature*, v. 521, n. 7553, p. 436-444, 2015. [Deep learning | Nature](#)

MAULUD, Dastan; ABDULAZEEZ, Adnan M. A review on linear regression comprehensive in machine learning. **Journal of Applied Science and Technology**

Trends, v. 1, n. 2, p. 140-147, 2020..Disponível em:
<https://doi.org/10.38094/jastt1457> Acesso em: 27.mai.2024

(MIXON et al., 1984).MIXON, A.C. et al. **Effect of chemical and biological agents on the incidence of *Aspergillus flavus* and aflatoxin contamination of peanut seed**. *Phytopathology*, v.74, n.2,p.1440-1444, 1984 Disponível em:
https://www.apsnet.org/publications/phytopathology/backissues/Documents/1984Articles/Phyto74n12_1440.PDF Acesso em 30.abr.2024

MONARD, M. C., & BARANAUSKAS, J. A. **Sistemas inteligentes: fundamentos e aplicações**. São Paulo, SP: Editora Unicamp. (2003).

NAKAGAWA, J., & ROSOLEM, C. A. **O amendoim: tecnologia de produção**. Botucatu, SP: FEPAF. p. 325, 2011.

NASCIMENTO, J. S., et al. **Impact of environmental conditions on the growth and aflatoxin production by *Aspergillus flavus* strains isolated from stored peanut**. *Food Control*, 94, 248–255. (2018).

RANJAN, K.S.; SINHA, A.K. **Occurrence of mycotoxigenic fungi and mycotoxins in animal feed from Bihar, India**. *Journal of the Science of Food and Agriculture*, Essex, v.56, p.39-47, 1991. Disponível em: <https://doi.org/10.1002/jsfa.2740560105>. Acesso em: 7.mai.2024

SILVA, E. F. da, et al. **Cultura do Amendoim**. Brasília, DF: Embrapa Informação Tecnológica. (2019). Disponível em:
<https://www.embrapa.br/agencia-de-informacao-tecnologica/tematicas/agroenergia/biodiesel/materias-primas/amendoim>. Acesso em: 10.abr.2024

SILVA, R. B., et al. **Mycotoxin contamination in peanuts and peanut products in Brazil: an overview**. *Food Control*, 98, 1–9. (2019).

Rossetto, C. A. V., Silva, O. F., & Araújo, A. E. da S.. (2005). **Influência da calagem, da época de colheita e da secagem na incidência de fungos e aflatoxinas em grãos de amendoim armazenados**. *Ciência Rural*, 35(2), 309–315. Disponível em: <https://doi.org/10.1590/S0103-84782005000200010> Acesso em: 15.mar.2024

SMITH, A. B., & JONES, C. D. **Seed-borne fungi: diversity, pathogenicity, and management**. *Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, 11(045). (2016).

TAUD, Hind; MAS, Jean-Francois. Multilayer perceptron (MLP). **Geomatic approaches for modeling land change scenarios**, p. 451-455, 2018. Disponível em: https://link.springer.com/chapter/10.1007/978-3-319-60801-3_27 Acesso em: 27.mai.2024

TILLMANN, M. A. A., & MELLO, V. D. C. **Análise de sementes**. In Curso de ciência e tecnologia de sementes (Módulo 4, p. 37). Brasília, DF: ABEAS; Pelotas, RS: Universidade Federal de Pelotas/Departamento de Fitotecnia. (2009).

UNIVERSITY OF LJUBLJANA (Eslovénia) et al, (org.). **Orange: Software de machine learning e visualização de dados**. Versão 3.27.0. [S. l.], 2020. Disponível em: <https://orange.biolab.si/download/#windows>. Acesso em: 26.mar. 2024.