

Avaliação de Produtividade e Saúde Vegetal em Mangueira a partir de Dados de Luminosidade de Abertura de Copa

Evaluation of Productivity and Plant Health in Mango Trees Based on Canopy Opening Luminosity Data

Heber do N. Paiva¹

 orcid.org/0009-0001-8651-4584

Jorge L. B. da Silva¹

 orcid.org/0009-0002-8929-6129

Kaandro do N. Teles¹

 orcid.org/0009-0006-1464-444x

Claudemiro Lima Júnior¹

 orcid.org/0000-0002-6640-6340

Alexandre M. A. Maciel¹

 orcid.org/0000-0003-4348-9291

Carmelo Bastos-Filho¹

 orcid.org/0000-0002-0924-5341

¹ Universidade de Pernambuco, Petrolina, Brasil. E-mail: hnpaiva@gmail.com

DOI: 10.25286/rep.v9i1.2787

Esta obra apresenta Licença Creative Commons Atribuição-Não Comercial 4.0 Internacional.

Como citar este artigo pela NBR 6023/2018: Heber do N. Paiva; Jorge L. B. da Silva; Kaandro do N. Teles; Claudemiro Lima Júnior; Alexandre M. A. Maciel; Carmelo Bastos-Filho; Avaliação de Produtividade e Saúde Vegetal em Mangueira a partir de Dados de Luminosidade de Abertura de Copa. Revista de Engenharia e Pesquisa Aplicada, v.9, n. 1, p. 105-114, 2024. DOI: 10.25286/rep.v9i1.2787

RESUMO

Este artigo visa correlacionar dados de luminosidade das mangueiras com a qualidade dos frutos, buscando determinar um nível ideal de luminosidade. Plantas foram divididas em parcelas, e 100 frutos de cada foram avaliados. A análise de dados, realizada em Python, incluiu gráficos de correlação, árvores de decisão e matrizes de confusão. Resultados promissores mostram forte correlação entre maior luminosidade e frutos de melhor qualidade destinados à exportação, apesar de alguns danos observados. A árvore de decisão classificou eficazmente frutos, identificando os de melhor qualidade com menor incidência de pragas e doenças. A matriz de confusão teve desempenho satisfatório na previsão de lotes com maior volume de frutos exportados. Conclui-se que maior luminosidade e abertura da copa estão associadas a melhor qualidade dos frutos. O estudo sugere a possibilidade de estabelecer um valor ideal de luminosidade para otimizar a produção de mangas de alta qualidade com pesquisas adicionais.

PALAVRAS-CHAVE: Ciência de Dados; Mangicultura; Luminosidade.

ABSTRACT

This paper aims to correlate mango tree light data with fruit quality to determine an ideal light level. Plants were divided into plots, and 100 fruits from each were assessed. Data analysis in Python involved correlation graphs, decision trees, and confusion matrices. Despite some observed damage, promising results show a strong correlation between higher light levels and better-quality fruits for export. The decision tree effectively classified fruits, identifying higher-quality samples with lower incidences of pests and diseases. The confusion matrix performed satisfactorily predicting batches with more exported fruits. It is concluded that increased light and canopy openness are associated with better fruit quality. The study suggests establishing an ideal light level to optimize the production of high-quality mangoes with further research.

KEY-WORDS: Data Science; Mango Cultivation; Luminosity.

1 INTRODUÇÃO

A cultura da manga (*Mangifera indica*) sendo uma planta *dicotiledônea*, de porte arbóreo e que pertence à família *Anacardiaceae*. Seu fruto possui diversos usos como seu consumo fresco, para fazer bebidas e sucos, sobremesas, bolos, saladas, cosméticos, indústria de alimentos entre outros. Alguns estudos até sugerem que o uso do farelo de manga pode ser usado para substituir o milho, caracterizando como um alimento alternativo constituindo a base do concentrado energético de ovinos em confinamento [1].

Segundo Tagore [2], no continente asiático houve uma redução de sua importância devido a sua expansão para outros continentes, com sua participação caindo de 87% nos anos 60 para 77% nos anos 90, e na África e América do Norte, havendo aumento de sua produção a partir dos anos 1960. Em 2018, os maiores produtores de manga eram nessa ordem, Índia, China, Tailândia, México, Indonésia, Paquistão e Brasil [3]. Sendo o Vale do São Francisco uma das regiões mais importantes para produção de manga no país, devidos a diversos fatores como clima, tecnologia agrícola, tecnologia de irrigação, tipo de solo e outras características favoráveis.

A prática da desfolha consiste na eliminação dos ramos vegetativos de uma planta, objetivando uma forma de produção adequada à finalidade requerida, o método deve levar em consideração as diferentes arquiteturas das cultivares, bem como as condições climáticas [4].

A desfolha desempenha um papel crucial na cultura da mangueira (*Mangifera indica*) ao possibilitar o controle do crescimento, aprimorar a qualidade e coloração dos frutos. O manejo adequado da mangueira por meio da desfolha permite prevenir doenças e a melhorar a eficiência da planta na alocação de recursos para a produção de frutos, como também melhorar sua iluminação interna possibilitando o

aumento da taxa fotossintética e arejar a planta para diminuir sua umidade interna.

As mangueiras frequentemente enfrentam diversos problemas quando a atividade de desfolha não é realizada de maneira apropriada. Estes problemas incluem uma inadequada pulverização de defensivos agrícolas em partes específicas da planta, resultando em um aumento na incidência de pragas e doenças. Além disso, observa-se a proliferação de ramos que tanto não produzem frutos, como também geram ausência de exposição à luz solar não favorecendo a fotossíntese. A decisão de abrir a copa da planta por meio da atividade de desfolha é geralmente conduzida através de avaliação visual. Não existe um parâmetro ou padrão estabelecido para tal procedimento

O propósito deste estudo é buscar através do uso de ferramentas de ciência de dados estabelecer uma relação entre os níveis de luminosidade no pomar e sua produção, a fim de determinar se a luminosidade desempenha um papel significativo na definição dos padrões de desfolhamento. Para Van Der Aalst [5], nos últimos tempos, a disciplina emergente conhecida como ciência de dados ganhou destaque, assumindo um papel significativo. Ela pode ser compreendida como uma integração de disciplinas tradicionais, tais como estatística, mineração de dados, gestão de bancos de dados e sistemas distribuídos.

Com algumas ferramentas de ciência de dados essenciais aplicadas como por exemplo por exemplo, árvore de decisão, matriz de confusão e mapa de calor. Na prática, a aprendizagem supervisionada fundamentada em árvore de decisão é caracterizada como uma abordagem de aprendizado supervisionado guiado por regras, utilizando a técnica de construção de árvores binárias. No entanto, sua compreensão torna-se ainda mais acessível ao interpretá-la como uma técnica de subdivisão hierárquica do domínio [6].

Já a matriz de confusão pode conter uma grande quantidade de informações úteis,

podendo levar a melhoria dos sistemas de aprendizado de máquina, apresentando possíveis fontes de erro [7]. E o mapa de calor importante para realização de correlações, onde as cores indicam valores e pode-se visualizar a intensidade e distribuição de valores.

A busca por um padrão de luminosidade para a abertura da copa da mangueira é essencial devido às implicações significativas que a variação na abertura da copa pode ter sobre a saúde e qualidade das frutas. A abertura excessiva da copa da mangueira pode levar à perda de qualidade das frutas, uma vez que isso pode afetar negativamente o teor de °Brix, responsável pelo sabor característico da fruta. Por outro lado, uma copa muito fechada pode comprometer a eficácia dos tratamentos com defensivos agrícolas, uma vez que a distribuição desses produtos pode ser dificultada, resultando em uma proteção inadequada contra pragas e doenças.

Este estudo se concentra na definição de um padrão fixo de abertura da copa da mangueira. Portanto, as variações sazonais na luminosidade não serão investigadas neste contexto. O estudo assume condições de luminosidade estáveis para estabelecer o referencial de abertura.

As influências da luminosidade em diferentes regiões geográficas não serão abordadas nesta pesquisa. O foco permanece na determinação de um padrão aplicável às condições específicas do local de estudo, sem explorar as divergências que poderiam ocorrer em diferentes contextos geográficos.

A luminosidade é um fator central neste estudo. Outros elementos climáticos também importantes, como temperatura e umidade, não serão considerados como parte da análise. O estudo limita-se à relação direta entre luminosidade e abertura da copa da mangueira. Embora se reconheça a interação entre a abertura da copa e a aplicação de defensivos agrícolas, este estudo não se aprofundará nos efeitos específicos desses produtos na qualidade das frutas. Essa é uma área que poderia ser explorada em futuras

pesquisas relacionadas. Apesar do uso do sensor de luminosidade seja a abordagem escolhida, outras metodologias alternativas para medir a abertura da copa não serão exploradas neste trabalho. As implicações econômicas da adoção do padrão de abertura da copa da mangueira não serão discutidas neste estudo. A pesquisa se concentra na definição do padrão em si, deixando questões econômicas para investigações posteriores.

2 FUNDAMENTAÇÃO TEÓRICA

Rautenberg e Carmo [8] descreveram a Ciência de Dados como um conceito onde softwares são empregados para transformar dados em informações, dando base para tomada de decisão. Seu objetivo principal é transformar dados brutos em informações valiosas, padrões relevantes e previsões úteis para auxiliar na tomada de decisões estratégicas e na resolução de problemas. A Ciência de Dados é constituída de três domínios do conhecimento: Programação de Computadores; Estatística e Matemática e Domínio do Conhecimento.

A aplicação prática deste estudo encontra um contexto especialmente relevante na cultura da manga (*Mangifera indica*) no Brasil, um dos maiores produtores e exportadores mundiais dessa fruta tropical. Nesta seção, exploramos como a determinação de um padrão de luminosidade para a abertura da copa da mangueira pode contribuir para aprimorar a produção de mangas no país.

A cultura da manga desempenha um papel significativo na economia agrícola brasileira. O país é reconhecido internacionalmente pela variedade e qualidade das suas mangas, que são amplamente consumidas tanto no mercado interno como exportadas para diversos países. A exportação de mangas é uma fonte vital de divisas e renda para muitas regiões do Brasil, especialmente durante as estações de safra.

Apesar da posição proeminente do Brasil como produtor de mangas, a indústria enfrenta desafios diversos. Entre eles, a

otimização da qualidade das frutas e a adoção de práticas agrícolas sustentáveis são de extrema importância. A qualidade das mangas é intrinsecamente ligada a fatores como sabor, aparência e valor nutricional, sendo influenciada por condições de cultivo e manejo.

A determinação de um padrão de luminosidade para a abertura da copa da mangueira pode oferecer soluções concretas para os desafios enfrentados pela indústria de manga no Brasil. Ao otimizar a abertura da copa, os produtores brasileiros podem melhorar a qualidade das frutas, aumentar a produtividade e contribuir para a competitividade no mercado global.

Além disso, a aplicação desses conhecimentos pode resultar em benefícios econômicos diretos, como a possibilidade de alcançar preços mais favoráveis no mercado internacional devido à qualidade aprimorada das mangas. A adoção de práticas agrícolas mais eficientes e sustentáveis também pode gerar impactos positivos no longo prazo, alinhando-se com as demandas crescentes por produtos agrícolas mais sustentáveis e ecologicamente responsáveis.

Em resumo, a pesquisa para determinar um padrão de luminosidade para a abertura da copa da mangueira por meio do estudo da correlação obtido pela análise de dados, traz contribuições significativas para a cultura da manga no Brasil. Ao enfrentar diretamente os desafios de qualidade e sustentabilidade, essa pesquisa não apenas fortalece a posição do Brasil como líder na produção de mangas, mas também contribui para a valorização da produção agrícola nacional e para a promoção de práticas agrícolas mais eficazes e sustentáveis.

Uma pesquisa feita por Sobrinho (2002) [9], realizada em Mossoró-RN em 2001/2002 teve como objetivo melhorar a produtividade e qualidade da manga Tommy Atkins, considerando fatores como arquitetura do dossel vegetativo, radiação solar e temperatura do ar. Foram estudadas 30 árvores com diferentes aberturas de copa, medindo a

penetração de radiação solar nos dosséis. Foram registrados dados de radiação global, infravermelha e fotossinteticamente ativa (PAR) no topo e base dos dosséis, além de temperatura do ar próxima aos frutos. As colheitas foram subdivididas em quadrantes para análise pós-colheita da coloração, SST, ATT e pH dos frutos. Diferentes arquiteturas de poda foram testadas, impactando a radiação solar e a produtividade da manga. Outros trabalhos relacionados não foram encontrados.

3 MATERIAIS E MÉTODOS

O uso de um sensor de luminosidade e a aplicação de ciência de dados são elementos cruciais para a coleta de dados e a análise das correlações entre a abertura ideal da copa da mangueira e a qualidade da fruta. O MVP é uma ferramenta preliminar que nos permite iniciar a exploração dessas relações de maneira direcionada.

A base da proposta é a utilização de um sensor de luminosidade, um dispositivo eletrônico capaz de medir a intensidade de luz em um ambiente específico. Este sensor foi escolhido devido à sua capacidade de fornecer informações valiosas sobre o ambiente de cultivo das mangueiras. O sensor de luminosidade nos permite coletar dados objetivos sobre os níveis de luz ao redor das mangueiras, o que é essencial para avaliar a luminosidade relacionada à abertura da copa.

A ciência de dados desempenha um papel crucial no sistema proposto, permitindo-nos analisar as informações coletadas pelo sensor de luminosidade e identificar padrões, tendências e correlações. Ao reunir esses dados com informações sobre a abertura da copa da mangueira e a qualidade das frutas produzidas, podemos explorar de forma sistemática a relação entre os níveis de luminosidade e a qualidade das mangas. Isso envolve o uso de algoritmos e técnicas de análise de dados para extrair *insights*

significativos a partir das informações coletadas.

O sistema representa uma abordagem pragmática para começar a explorar as correlações entre a abertura ideal da copa da mangueira e a qualidade das frutas, utilizando tecnologia acessível e metodologias analíticas. Enquanto este estudo não pretende ser uma análise exaustiva, a proposta fornece um ponto de partida valioso para investigações mais aprofundadas.

A fonte de dados para o presente estudo foi coletada em campo através da atividade laboral dos monitores de pragas. Para efetuar a coleta destes dados de luminosidade foi criado uma plataforma de apontamento personalizada e com capacidade de registrar a geolocalização para garantir a qualidade da amostra, como observado nas Figuras 2 e 3.

Figura 2: Plataforma virtual de coleta dedados de luminosidade.

Fonte: Os autores.

Essa plataforma de apontamento é preenchida pelos próprios monitores que coletam estes dados e sua resposta é gravada na nuvem assim que é encontrado sinal para transmissão. A medição da luminosidade foi feita usando um Luxímetro de LED digital HLX-912 (Figura 4). Foram coletados dados de

luminosidade do ambiente externo e do ambiente interno da copa da planta, próximo ao tronco, em todas as parcelas da fazenda C2 localizada em Petrolina-PE. As parcelas totalizam 240 hectares. Cada uma das parcelas terá 5 amostras de plantas. As coletas foram feitas entre 06h00 e 11h00 da manhã.

Figura 3: Dados de georreferenciamento das coletas dedados de luminosidade.

Fonte: Os autores.

Figura 4: Luxímetro digital de LED HLX-912.



Fonte: Os autores.

Os valores de luminosidade coletados foram tabulados em planilha de excel (como mostrado na Tabela 1) e usados para fazer as correlações com as variáveis de saída. As planilhas contêm a data, o horário, a parcela e os valores da coleta externa e interna.

Quanto aos dados de variáveis resposta, estes foram obtidos após a colheita. Todos os caminhões colhidos no mesmo dia e parcela recebem o mesmo lote ao ingressar no Packing House. São amostrados 100 frutos de cada lote na recepção, para identificação de características de qualidade deste lote.

Tabela 1: Dados de luminosidade tabulados.

Data	Horário	Parcela	Coleta externa	Coleta interna
21/08/2023	08:00:00.000-03:00	PA-07	27700	1224
21/08/2023	08:23:00.000-03:00	PA-07	14400	1386
21/08/2023	08:27:00.000-03:00	PA-07	12300	2134
21/08/2023	08:29:00.000-03:00	PA-07	12300	1815
21/08/2023	08:32:00.000-03:00	PA-07	15100	3463
21/08/2023	08:37:00.000-03:00	PA-07	14800	1188
21/08/2023	08:38:00.000-03:00	PA-07	19700	1482
21/08/2023	08:40:00.000-03:00	PA-07	25400	2448
21/08/2023	08:42:00.000-03:00	PA-07	27800	4827
21/08/2023	08:44:00.000-03:00	PA-07	25900	2725
21/08/2023	08:50:00.000-03:00	PA-07	27700	2675
21/08/2023	08:51:00.000-03:00	PA-07	29400	2105
21/08/2023	08:53:00.000-03:00	PA-07	21200	1198
21/08/2023	08:55:00.000-03:00	PA-07	15200	2492

Fonte: Os autores.

Além destas informações também foram coletadas o aproveitamento geral de cada lote. Estes dados foram filtrados e tabulados em planilhas de excel, como mostrado na Tabela 2, para posterior processamento e cruzamento de dados com os resultados obtidos da luminosidade de cada parcela. As

informações tratadas nesta planilha incluem a quantidade de frutas sem danos, frutos com danos mecânicos, frutos afetados por pragas e doenças, coloração acima de 40%, média da luminosidade, quantidade de quilogramas colhidos, percentual de quilogramas exportados, percentual de quilogramas vendidos no mercado interno, percentual de quilogramas perdidos e classificação do lote.

Tabela 2: Dados tabulados de variáveis-resposta.

MEDIA_LUM	SEM_DANOS	DANOS_MECHANICOS	PRAGAS_DOENCAS	COLORACAO_ACIMA_40%	KG	EXP(%)	MI(%)	PERDA(%)	CLASSIFICACAO
5804	28	29	42	64	44.28	74,90	0,00	25,10	1
5804	33	28	39	14	44.13	70,50	17,60	11,90	1
5804	45	34	21	46	22.13	87,00	4,80	8,20	1
5804	36	29	35	51	21.68	62,00	25,20	12,80	2
5804	21	32	47	99	43.4	72,50	16,40	11,20	1
5804	24	32	43	86	44.43	72,40	18,30	9,30	1
5804	36	38	26	60	22.23	62,40	26,70	10,90	2
5804	26	37	37	67	33.9	64,00	23,60	12,40	2
5804	40	28	33	57	22.49	50,10	42,80	7,10	2
5804	37	23	40	34	11.08	37,80	50,70	11,50	3

Fonte: Os autores.

Foi realizado a limpeza da base de dados, redução dos dados por aplicação de média de luminosidade, verificação de outliers, dados em branco e escolha das features para realização da árvore de decisão.

Os dados que estão presentes na planilha final da base para este estudo são:

- FRUTOS_SEM_DANOS: Quantidade de frutos sem danos na amostra.
- FRUTOS_COM_DANOS: Quantidade de frutos com danos mecânicos.
- FRUTOS_COM_PRAGAS_DOENCAS: Quantidade de frutos encontrados com doenças ou danos causados por pragas.
- COLORACAO>40%: Quantidade de frutos com coloração acima de 40% encontrados na amostra.
- MEDIA_LUM: Média das amostras de luminosidade encontrada no pomar da parcela.
- QTD_PROCESSADA: Quantidade de quilos do lote.
- QTD_EXPORTADA: Percentual de frutos para exportação embalados do lote.
- QTD_INTERNA : Percentual de frutos para mercado interno embalados do lote.
- QTD_PERDIDA: Percentual de frutos de refugo ou perda do lote.
- CLASSIFICACAO: Indicador de qualidade que será nossa referência para o treinamento do processamento de dados.

Para a análise de correlação, utilizamos a média de luminosidade como variável independente e diversas variáveis dependentes, tais como frutos sem danos, com pragas e doenças, danos mecânicos, coloração acima de 40%, quantidade de quilos por lote, quantidade de frutos exportados por lote, percentagem de frutos destinados ao mercado interno, percentagem de frutos considerados refugo e a categoria de classificação. No que concerne à classificação, esta é estabelecida com base na percentagem de frutos exportados, em que acima de 65% de frutos exportados corresponde à classe 1, entre 65% e 50% à classe 2 e, por fim, abaixo de 50% à classe 3.

A limpeza de dados é um processo fundamental na análise de dados e no processamento em Python, assim como em qualquer outra linguagem de programação. Neste trabalho, utilizamos os processos tradicionais de limpeza, como imputação, tratamento de valores faltantes e normalização.

De acordo com Silva Junior et al. [10], o Python representa uma alternativa viável para a automação de processos. Nessa linguagem de programação, dispomos de uma biblioteca, que constitui um conjunto de métodos e funções utilizados na otimização do processamento de dados. Para a realização deste estudo, procedemos com a importação da base de dados para o Python, e posteriormente foram executados uma série de comandos.

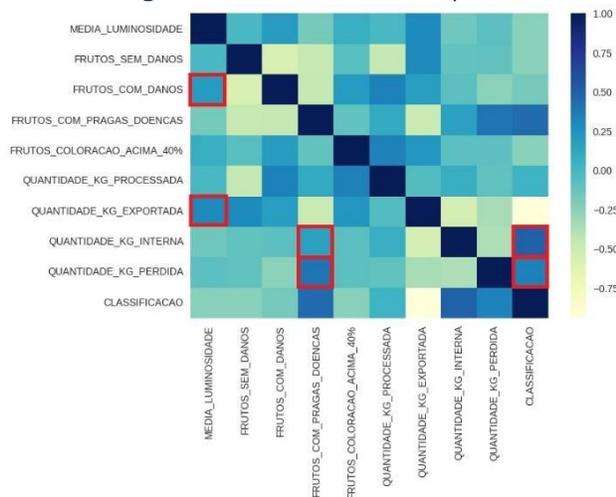
Após a importação da base de dados, o próximo passo consistiu na instalação das bibliotecas essenciais para a análise. Após a conclusão dessas etapas de preparação, foi realizado a execução dos códigos destinados à geração dos gráficos, que posteriormente foram apresentados na Seção 4 (Resultados) para uma análise detalhada dos dados.

4 RESULTADOS

Nas subseções a seguir serão apresentados os resultados das três análises feitas, sendo elas: correlação, árvore de decisão e matriz de confusão. Também será apresentado a discussão dos resultados obtidos.

Analisar correlações através de mapa de calor é muito usual por diminuir o problema de sobreposição de visualização dos dados, principalmente se tratando de muitos dados [11]. Se trata de uma ferramenta muito prática e que mostra a correlação entre dados de diversas variáveis diferentes.

Figura 5: Resultado – Heat Map.



Fonte: Os autores.

A Figura 5 mostra os resultados da análise entre os diferentes atributos, onde foi possível observar que existe correlação entre a média de luminosidade e a quantidade de frutos exportados, ou seja, quando houve maior média de luminosidade nas parcelas avaliadas houve maior quantidade de frutos exportados. Isso pode significar que a atividade de desfolha bem-feita potencializa a formação do fruto e ajuda a acumular nutrientes nas plantas. Outra correlação importante encontrada, é a média de luminosidade e frutos com danos mecânicos, característica esta que também é fator determinante para um melhor aproveitamento dos frutos para o mercado de exportação. A quantidade de frutas vendidas no mercado interno e as frutas perdidas têm uma forte correlação com a classificação. Em outras palavras, quanto maior a classificação, menor é a quantidade de frutas vendidas para o mercado externo. O mesmo ocorre quando são correlacionadas com frutos que possuem doenças e pragas.

Segundo Monard e Baranauskas (2003) [12] uma árvore de decisão nada mais é que uma estrutura de dados que pode ser definida de duas formas: Um nó folha que corresponde a uma classe; um nó de decisão que apresenta um teste de algum atributo. Para cada

resultado sendo positivo ou negativo (sim ou não) existe uma aresta para uma sub-árvore.

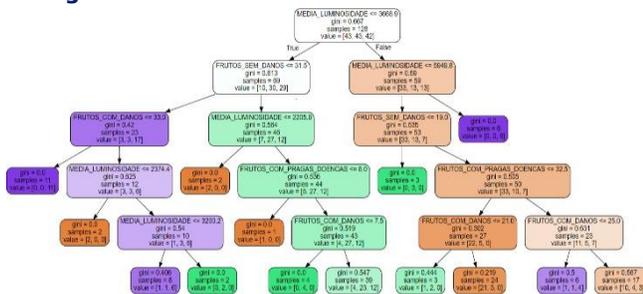
Na Figura 6 é apresentado uma árvore de decisão para análise de dados agrícolas, visando classificar amostras com base em características específicas. A árvore é construída em vários níveis, cada um representando uma divisão dos dados.

Inicialmente, o critério de divisão é estabelecido pela condição "MEDIA_LUM" \leq 3668,9, resultando em um índice Gini de 0,667. Este nível inicial contém 128 amostras, com distribuição equilibrada entre as classes 1, 2 e 3, com 43, 43 e 42 amostras, respectivamente. A primeira divisão ocorre com base na condição "SEM_DANOS" \leq 31,5, levando a um índice Gini de 0,613. No caminho "True", com 69 amostras, a distribuição de classes é de 10 amostras da classe 1, 30 amostras da classe 2 e 29 amostras da classe 3. No caminho "False", a árvore continua a se ramificar com base na característica "MEDIA_LUM" e um novo limite. A divisão subsequente considera a condição "MEDIA_LUM" \leq 5949,8, resultando em um índice Gini de 0,59. Este ramo "False" leva a 59 amostras, com distribuição de classes de 33 amostras da classe 1, 13 amostras da classe 2 e 13 amostras da classe 3. A árvore continua com novas divisões, incluindo considerações sobre "COM_DANOS_MECANICOS" e limites específicos. Uma das ramificações mais detalhadas ocorre no caminho "False" da condição "MEDIA_LUM" \leq 3200,2. Nesse ponto, a árvore se ramifica múltiplas vezes com base em diferentes características, como "PRAGAS_DOENCAS" e "COM_DANOS_MECANICOS". Cada subdivisão é avaliada com seu índice de Gini respectivo, considerando a pureza das amostras resultantes.

Essa abordagem hierárquica possibilita a construção de uma árvore de decisão que revela padrões complexos nos dados agrícolas. Cada nível da árvore representa uma decisão fundamentada em características

específicas, visando maximizar a pureza das classes nas amostras resultantes. A análise detalhada das decisões tomadas pela árvore proporciona *insights* valiosos para a classificação de dados agrícolas. Destaca-se que, no contexto luminosidade coletada, variando de um mínimo de 2046 lumens a um máximo de 6069 lumens, identificamos que frutas provenientes de plantas mais arejadas, com índices mais elevados de luminosidade, apresentam um percentual de exportação superiora 65% do total produzido.

Figura 6: Resultado – Tree Decision.



Fonte: Os autores.

A Figura 7 apresenta a análise da matriz de confusão. Na tabela de valores verdadeiros, observa-se que para a classe 1 (primeira linha), houve 14 instâncias corretamente classificadas, 4 falsos negativos e 3 falsos positivos. Isso sugere que o modelo tem uma boa precisão ao identificar a classe 1, mas ainda comete alguns erros.

Figura 7: Resultado – Matriz de confusão.



Fonte: Os autores.

Analisando a classe 2, nota-se que houve 7 previsões corretas, 2 falsos negativos e 2 falsos positivos. Semelhante à classe 1, o modelo tem uma precisão razoável, mas há espaço para melhorias, especialmente na redução dos falsos positivos.

Quanto à classe 3, a matriz de confusão acertou 10 previsões, com 3 falsos negativos e 4 falsos positivos. Isso sugere que a classe 3 é relativamente bem identificada pelo modelo, mas novamente há margem para aprimoramentos, especialmente na redução dos falsos positivos.

Em resumo, a matriz de confusão parece ter um desempenho razoável, mas há áreas específicas que precisam de atenção, como a redução de falsos positivos para as classes 1 e 2, e a melhoria da sensibilidade para a classe 2. A interpretação desses resultados pode orientar ajustes no modelo para otimizar sua capacidade de previsão.

5 CONCLUSÕES

Diante dos resultados obtidos ao longo do trabalho, foi possível observar como a abertura da copa, medida através de sua luminosidade interna, impacta na qualidade de frutos da mangueira. Como uso de softwares e ferramentas como o mapa de calor pode-se ver a alta correlação entre média de luminosidade e quantidade de frutos exportados devido à atividade de desfolha bem-feita. A árvore de decisão confirmou os resultados das correlações, pois ela sempre aponta um número maior de frutos exportados quando tem a presença maior de luminosidade.

A matriz de confusão demonstrou que o modelo apresentou bom desempenho em relação a todas as classificações dos frutos (1, 2 e 3), onde em todas as classes houve mais acertos na predição da classe a qual pertenciam as amostras (onde o modelo teve mais verdadeiros positivos), indicando que o modelo de classificação está desempenhando bem em identificar corretamente os casos positivos em um conjunto de dados.

Apesar da base de dados ter sido relativamente pequena, o uso deste tipo de ferramentamostra potencial em aplicações na agricultura, no papel de base de treinamento de máquina para determinar uma luminosidade ideal em planta de mangueira. Recomenda-se para trabalhos futuros a ampliação das coletas de luminosidade para maisfazendas, em número maior de plantas e ao longode diferentes ciclos de produção, para se obter uma base ideal de treinamento para desenvolvimento de uma futura inteligência artificial.

REFERÊNCIAS

- [1] ARAGÃO, A. S. L. et al. Farelo de manga na dieta de cordeiros em confinamento. Arquivo Brasileiro de Medicina Veterinária e Zootecnia, v. 64, p. 967-973, 2012.
- [2] SIQUEIRA, Tagore Villarim de. A cultura da manga: desempenho no período 1961-2001. 2003.
- [3] PARIONA, A. The Top Mango Producing Countries In The World. April 9, 2018. Disponível em: . Acesso em: 11 nov. 2020.
- [4] DE ALBUQUERQUE, João Antônio Silva et al. Sistemas de poda. 2002.
- [5] VAN DER AALST, Wil; VAN DER AALST, Wil. Data science in action. Springer Berlin Heidelberg, 2016.
- [6] SUTHAHARAN, Shan; SUTHAHARAN, Shan. Decision tree learning. Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning, p. 237-269, 2016.
- [7] KADERABEK, Eva; SUWANNAJANG, Panittha. Confusion Matrix Viz.
- [8] RAUTENBERG, Sandro; DO CARMO, Paulo Ricardo Viviurka. Big data e ciência de dados: complementariedade conceitual no processo de tomada de decisão. Brazilian Journal of Information Science, v. 13, n. 1, p. 56-67, 2019.
- [9] ESPÍNOLA SOBRINHO, José et al. Influência do tipo de arquitetura do dossel na absorção de radiação solar, na produtividade e na quantidade do fruto da mangueira. 2003.
- [10] SILVA JUNIOR, Claudiomar da Fonseca et al. USO DO PYTHON NA AUTOMAÇÃO EMPRESARIAL. Encontro Internacional de Gestão, Desenvolvimento e Inovação (EIGEDIN), v. 6, n. 1,2022.
- [11] PERROT, Alexandre et al. HeatPipe: High throughput, low latency big data heatmap with spark streaming. In: 2017 21st International Conference Information Visualisation (IV). IEEE, 2017. p. 66-71.
- [12] MONARD, Maria Carolina; BARANAUSKAS, José Augusto. Indução de regras e árvores de decisão. Sistemas Inteligentes-fundamentos e aplicações, v. 1, p. 115- 139, 2003.