

Universidade Estadual Paulista – UNESP
Instituto de Química de Araraquara
Departamento de Engenharia, Física e Matemática

Lucas Janoni dos Santos

**Uso de inteligência artificial para quantificar e classificar impurezas em
cana-de-açúcar para fins industriais**

Araraquara

2021

Autor: Lucas Janoni dos Santos

**Uso de inteligência artificial para quantificar e classificar impurezas em
cana-de-açúcar para fins industriais**

Trabalho de conclusão de curso apresentado ao Instituto de Química de Araraquara, da Universidade Estadual Paulista “Júlio de Mesquita Filho” para obtenção do título de Bacharel em Engenharia Química.

Orientadora: Profa. Dra. Érica Regina Filletti Nascimento

Araraquara

2021

LUCAS JANONI DOS SANTOS

Uso de inteligência artificial para quantificar e classificar impurezas em cana-de-açúcar para fins industriais

Trabalho de conclusão de curso apresentado ao Instituto de Química de Araraquara, da Universidade Estadual Paulista “Júlio de Mesquita Filho” para obtenção do título de Bacharel em Engenharia Química.

Araraquara, 02 de março de 2021

BANCA EXAMINADORA



Profa. Dra. Érica Regina Filletti Nascimento
Instituto de Química – UNESP, Araraquara



Prof. Dr. Fábio Roberto Chavarette
Instituto de Química – UNESP, Araraquara



Profa. Dra. Fabíola Manhas Verbi Pereira
Instituto de Química – UNESP, Araraquara

AGRADECIMENTOS

À Universidade Estadual Paulista “Júlio de Mesquita Filho” Campus Araraquara, corpo docente, direção e administração, por todas as oportunidades que me foram dadas, ensinamentos científicos e projetos.

À FAPESP, pelo apoio financeiro e institucional dado para a realização deste trabalho (2018/03690-1).

À Profa. Dra. Érica Regina Filletti Nascimento, pela sua orientação, amizade, paciência, entusiasmo e incentivo não apenas para esse trabalho, mas muitos outros que desenvolvemos juntos.

À Profa. Dra. Fabíola Manhas Verbi Pereira, pelos ensinamentos, oportunidades, amizade e incentivo durante minha caminhada.

Aos meus pais Aparecido e Marilda, por todo amor, suporte, educação, incentivo e sacrifícios, que me possibilitaram chegar até aqui.

Às minhas irmãs, que sempre me apoiaram durante minha jornada.

Aos meus amigos, que nos meus melhores e piores momentos sempre estiveram ao meu lado para dar suporte, amor, conselhos e força para continuar.

A todos que, direta ou indiretamente, fizeram parte da minha formação.

“O caos é uma ordem por decifrar.”
(José Saramago)

RESUMO

Redes neurais artificiais (RNAs) são modelos computacionais que relembram a estrutura e o funcionamento do cérebro humano. São sistemas em que há distribuição de unidades básicas de processamento (neurônios), interligados por diversas conexões (sinapses), implementadas por funções matemáticas diversas, vetores e matrizes. Apresentam a habilidade de adquirir e registrar informações a partir do processamento de um banco de dados, tendo elevada capacidade de mapear e resolver problemas não-lineares. Em virtude dessas características, foram utilizadas RNAs para quantificar e classificar a porcentagem mássica de cana-de-açúcar e de impurezas tipicamente encontradas em sua colheita, uma vez que elas podem causar aumento no consumo de energia no preparo da cana, perdas de extração, dificuldades para o tratamento do caldo e desgaste de equipamentos. A quantificação e a classificação das 146 amostras foram feitas a partir da utilização de informações numéricas de imagens digitais convertidas em escalas de cores (RGB, HSV, rgb e L), um método de aquisição de dados não invasivo. A classificação envolveu a divisão das amostras em duas classes: 90-100% (classe 1) e 41-89% (classe 2) de cana-de-açúcar, em massa. A implementação das redes neurais artificiais foi realizada no *software* MATLAB R2018a a partir da ferramenta *nntstart*, sendo utilizado o algoritmo de Levenberg-Marquadt (*trainlm*) para os modelos de predição e o algoritmo do gradiente conjugado escalonado (*traincsg*) para os modelos de classificação. O número de neurônios na camada intermediária foi ajustado por tentativa e erro até alcançar o melhor resultado pelas RNAs. Em ambos os processos houve divisão aleatória das amostras, sendo 70% para a realização do treinamento da rede, 15% para a sua validação e os 15% restantes para o teste das RNAs. Os resultados obtidos pelas RNAs foram bastante promissores, uma vez que foi possível prever as quantidades de cana-de-açúcar, material vegetal e solo com baixos erros absolutos, além de altos coeficientes de correlação. Foi possível, também, classificar todas as amostras corretamente, mesmo diante de um conjunto de dados complexo e multivariado como o de parâmetros de imagens digitais.

Palavras-chave: Redes Neurais Artificiais. Imagens Digitais. Cana-de-açúcar. Impurezas. Classificação. Predição.

ABSTRACT

Artificial neural networks (ANNs) are computational models similar to the structure and functioning of the human brain. They are systems in which basic processing units (neurons) are distributed and interconnected by different connections (synapses), implemented by different mathematical functions, vectors and matrices. They present the ability to acquire and record information from the processing of a database, having a high capacity to map and solve non-linear problems. Due to these characteristics, RNAs were used to quantify and classify the mass percentage of sugarcane and impurities typically found in their harvest, since they can cause increased energy consumption in the preparation of the cane, extraction losses, difficulties in the broth's treatment and equipment wear. The quantification and classification of the 146 samples were made using digital information from digital images converted into color scales (RGB, HSV, rgb and L), a non-invasive data acquisition method. The classification involved dividing the samples into two classes: 90-100% (class 1) and 41-89% (class 2) of sugarcane, by weight. The implementation of artificial neural networks were performed in the MATLAB R2018a software with the nnstart tool, using the Levenberg-Marquadt algorithm (trainlm) for the prediction models and the scaled conjugated gradient algorithm (traincsg) for the classification models. The neurons' number in the intermediate layer was adjusted by trial and error until the best result by the ANNs was achieved. In both processes there were a random division of the samples: 70% for the training of the network, 15% for its validation and the remaining 15% for the ANN test. The results obtained by the ANNs were very promising, since it was possible to predict the amounts of sugarcane, plant material and soil with low absolute errors and high correlation coefficients. It was also possible to classify all samples correctly, even in the face of a complex and multivariate data set such as digital image parameters.

Keywords: Artificial Neural Networks. Digital images. Sugarcane. Impurities. Classification. Prediction.

Lista de Figuras

Figura 1. Evolução da produção de cana-açúcar no Brasil.....	16
Figura 2. Histórico da presença de impurezas mineral e vegetal em cana-de-açúcar no Centro Sul do Brasil	17
Figura 3. Estrutura de um neurônio biológico e esquema de sinapse.....	21
Figura 4. Modelo do neurônio proposto por McCulloch e Pitts	23
Figura 5. Modelo de neurônio não linear baseado no <i>perceptron</i>	24
Figura 6. Principais funções de ativação: (a) Degrau, (b) Linear por Partes e (c) Sigmoide.....	25
Figura 7. Exemplo de rede <i>feed-forward</i> com uma única camada	26
Figura 8. Exemplo de rede <i>feed-forward</i> com múltiplas camadas	27
Figura 9. Exemplo de rede neural recorrente	27
Figura 10. Exemplo de rede MLP	29
Figura 11. Cana coletada em Ibaté-SP.....	34
Figura 12. Frações de cana, impureza vegetal e solo das misturas separadas...	35
Figura 13. Imagem digital de amostra contendo 90% de cana, 5% de impureza vegetal e 5% de solo	35
Figura 14. Imagem digital de amostra contendo 90% de cana e 10% de solo ..	36
Figura 15. Imagem digital de amostra contendo 90% de cana e 10% de impureza vegetal.....	36
Figura 16. Interface gráfica da ferramenta NNSStart.....	37
Figura 17. Diagrama de distribuição de erro para RNA de quantificação de cana-de-açúcar	40
Figura 18. Performance do modelo de RNA com base no erro quadrático médio em função do número de épocas para a quantificação de cana-de-açúcar	41
Figura 19. Porcentagem em massa de cana estimada pela RNA para os conjuntos de teste, validação e treinamento	42
Figura 20. Performance do modelo de RNA com base no erro quadrático médio em função do número de épocas para a quantificação de material vegetal	43
Figura 21. Diagrama de distribuição de erro para RNA de quantificação de material vegetal.....	43
Figura 22. Porcentagem em massa de material vegetal estimada pela RNA para os conjuntos de teste, validação e treinamento.....	44

Figura 23. Performance do modelo de RNA com base no erro quadrático médio em função do número de épocas para a quantificação de material mineral	45
Figura 24. Diagrama de distribuição de erro para RNA de quantificação de material mineral.....	46
Figura 25. Porcentagem em massa de material mineral estimada pela RNA para os conjuntos de teste, validação e treinamento.....	47
Figura 26. Matriz de confusão dos conjuntos de treinamento, validação e teste, bem como de todas as amostras para classe cana e RNA com 6 neurônios na camada intermediária.....	49
Figura 27. Desempenho da RNA em função do número de época para classe cana e RNA com 6 neurônios na camada intermediária.....	50
Figura 28. Curva de Característica de Operação do Receptor para classe cana e RNA com 6 neurônios na camada intermediária.....	51
Figura 29. Respostas obtidas pela RNA com 6 neurônios da camada intermediária para classe cana	52
Figura 30. Matriz de confusão dos conjuntos de treinamento, validação e teste, bem como de todas as amostras para classe cana e RNA com 7 neurônios na camada intermediária.....	53
Figura 31. Desempenho da RNA em função do número de época para classe cana e RNA com 7 neurônios na camada intermediária.....	54
Figura 32. Curva de Característica de Operação do Receptor para classe cana e RNA com 7 neurônios na camada intermediária.....	54
Figura 33. Respostas obtidas pela RNA com 7 neurônios da camada intermediária para classe cana	55
Figura 34. Matriz de confusão dos conjuntos de treinamento, validação e teste, bem como de todas as amostras para classe cana e RNA com 10 neurônios na camada intermediária.....	56
Figura 35. Desempenho da RNA em função do número de época para classe cana e RNA com 10 neurônios na camada intermediária.....	57
Figura 36. Curva de Característica de Operação do Receptor para classe cana e RNA com 10 neurônios na camada intermediária.....	57
Figura 37. Respostas obtidas pela RNA com 10 neurônios da camada intermediária para classe cana	58

Lista de Tabelas

Tabela 1. Porcentagem mássica de impurezas mineral e vegetal de acordo com o tipo e colheita.....	17
Tabela 2. Coeficientes de correlação para as 17 arquiteturas de RNAs para predição de porcentagem mássica de cana-de-açúcar	39
Tabela 3. Resumo dos resultados obtidos pelas redes neurais artificiais para cada modelo.....	47
Tabela 4. Respostas obtidas e esperadas pela RNA para o conjunto de treinamento e RNA com 6 neurônios na camada intermediária.....	65
Tabela 5. Respostas obtidas e esperadas pela RNA para o conjunto de validação e RNA com 6 neurônios na camada intermediária.....	67
Tabela 6. Respostas obtidas e esperadas pela RNA para o conjunto de teste e RNA com 6 neurônios na camada intermediária.....	68
Tabela 7. Respostas obtidas e esperadas pela RNA para o conjunto de treinamento e RNA com 7 neurônios na camada intermediária para classe cana	69
Tabela 10. Respostas obtidas e esperadas pela RNA para o conjunto de validação e RNA com 7 neurônios na camada intermediária para classe cana.....	71
Tabela 9. Respostas obtidas e esperadas pela RNA para o conjunto de teste e RNA com 7 neurônios na camada intermediária para classe cana.....	72
Tabela 10. Respostas obtidas e esperadas pela RNA para o conjunto de treinamento e RNA com 10 neurônios na camada intermediária para classe cana	73
Tabela 11. Respostas obtidas e esperadas pela RNA para o conjunto de validação e RNA com 10 neurônios na camada intermediária para classe cana.....	75
Tabela 12. Respostas obtidas e esperadas pela RNA para o conjunto de teste e RNA com 10 neurônios na camada intermediária.....	76
Tabela 13. Porcentagem em massa de cana-de-açúcar, impureza vegetal, solo e soma de impurezas para as 146 amostras	77
Tabela 14. Escala de cores RGB, HSV, rgb e L (luminosidade) das amostras .	80

SUMÁRIO

INTRODUÇÃO	13
1 O PROBLEMA DE PESQUISA.....	15
1.1 Impurezas em carregamentos de cana-de-açúcar.....	16
1.2 Identificação e determinação de impurezas e o uso de imagens digitais	18
2 REDES NEURAIS ARTIFICIAIS (RNAS)	20
2.1 Neurônio biológico	20
2.2 Histórico.....	21
2.3 Modelos de neurônios	22
2.4 Funções de ativação	25
2.5 Arquitetura de Redes Neurais Artificiais	25
2.6 Tipos de aprendizagem	28
2.7 Redes Neurais <i>Perceptrons</i> De Múltiplas Camadas	28
2.8 Algoritmo de retropropagação do erro (<i>backpropagation</i>) e de	
Levenberg-Marquardt.....	30
2.9 Algoritmo do gradiente conjugado escalonado.....	32
3 METODOLOGIA	34
3.1 Amostras de bagaço de cana-de-açúcar	34
3.2 Redes neurais artificiais	37
3.2.1 Modelos de predição	37
3.2.2 Modelos de classificação	38
4 REDES NEURAIS ARTIFICIAIS PARA QUANTIFICAÇÃO DAS	
AMOSTRAS	39
4.1 Porcentagem mássica de cana-de-açúcar	39
4.2 Porcentagem mássica de material vegetal.....	42
4.3 Porcentagem mássica de material mineral.....	45
5 REDES NEURAIS ARTIFICIAIS PARA CLASSIFICAÇÃO DAS	
AMOSTRAS	49
5.1 Camada intermediária com 6 neurônios	49
5.2 Camada intermediária com 7 neurônios	53

5.3	Camada intermediária com 10 neurônios	56
6	CONSIDERAÇÕES FINAIS	60
	REFERÊNCIAS	61
	APÊNDICE.....	65
	ANEXO	77

INTRODUÇÃO

A qualidade de produto final advindo da cana-de-açúcar é fortemente dependente da quantidade de impurezas presente no sistema. À medida que a quantidade dessas impurezas (vegetal ou solo) aumenta, os efeitos na qualidade são mais perceptíveis, já que o aumento de fibra significativo reduz a extração, provoca diminuição da sacarose e pureza do caldo (GUEDES, 2018). Além disso, o transporte de carregamentos desse insumo contendo altas quantidades de impureza pode levar a mais gastos logísticos, pois o espaço ocupado por folhas e solo poderia ser ocupado com a cana-de-açúcar. Dessa forma, identificar a quantidade de impureza presente na matéria-prima tem grande relevância econômica.

Nesse contexto, a utilização de ferramentas computacionais aplicadas ao processo produtivo surge como forma de auxiliar na tomada de decisões cada vez mais assertivas, para que o processamento da cana-de-açúcar atenda aos requisitos de qualidade, depreenda menos gastos energéticos e menos desperdícios.

Devido à sua alta capacidade de reconhecimento de padrões, as Redes Neurais Artificiais surgem como alternativa para determinar as impurezas presentes em carregamentos de cana-de-açúcar, juntamente com um método não invasivo de aquisição de dados a partir de imagens digitais. Linhares, Araújo Jr e Araújo (2007) utilizaram a técnica de redes neurais artificiais de múltiplas camadas para mapear e quantificar o composto penteno na produção de GLP como forma de manter a boa qualidade do produto no processo. Antoneli e Neitzel (2016) desenvolveram redes neurais artificiais para determinar a qualidade de fios de algodão de uma empresa de fios, modelo que ajudou o engenheiro da indústria a melhorar a qualidade do fio e reduzir os índices de quebra e ruptura.

Nesse trabalho, são apresentadas as RNAs para a determinação das quantidades de impurezas presentes na cana-de-açúcar. A metodologia desenvolvida foi construída a partir de um banco de dados contendo informações das escalas de cores RGB, HSV, rgb e L a partir de imagens digitais.

O objetivo geral do estudo consiste na verificação da viabilidade do uso de redes neurais artificiais para quantificar e classificar impurezas em carregamentos de cana-de-açúcar. Já os objetivos específicos consistem em:

- i. Aprender os conceitos básicos de RNAs e alguns algoritmos de aprendizagem;

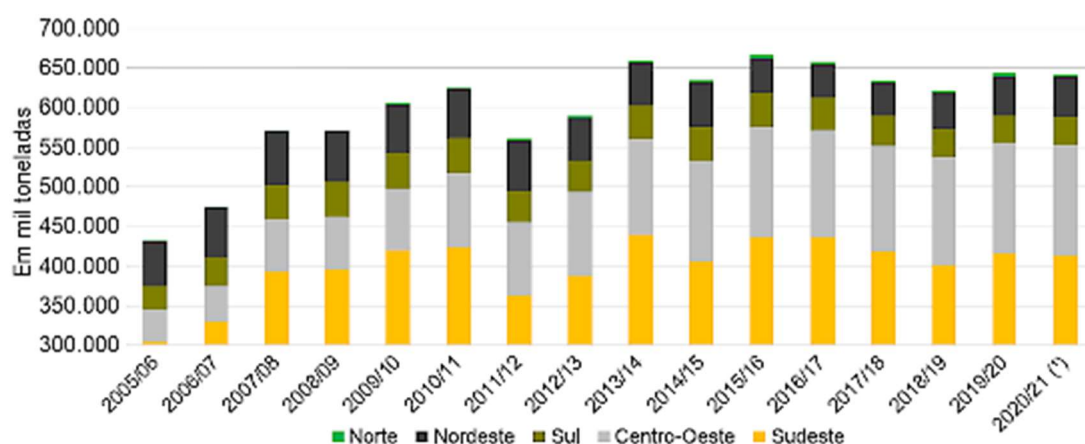
- ii. Desenvolver diversas arquiteturas de RNAs para estimar a porcentagem mássica de cana-de-açúcar, impureza vegetal e impureza mineral a partir de imagens digitais a fim de chegar ao melhor resultado;
- iii. Desenvolver diversas arquiteturas de RNAs para classificar amostras de cana-de-açúcar contendo impurezas vegetal e mineral (solo), com base em sua porcentagem mássica, utilizando imagens digitais, a fim de chegar ao melhor resultado.

1 O PROBLEMA DE PESQUISA

A cana-de-açúcar é um dos principais produtos agrícolas produzidos no mundo, sendo uma das mais importantes culturas no Brasil, que é o líder na produção mundial. Esse setor representa cerca de 2% do PIB brasileiro. São diversos os produtos fabricados a partir do caldo de cana-de-açúcar, bem como dos resíduos (líquidos ou sólidos) gerados durante seu processamento. Além do açúcar e do etanol, destacam-se a cachaça e a rapadura, que são produtos extraídos do caldo; merece destaque, também, a cogeração de energia elétrica a partir da queima do bagaço.

Diferente do padrão internacional, em que a atividade agrícola de produção da cana-de-açúcar se mantém separada da industrial, grande parte da matéria-prima processada no Brasil é produzida diretamente pelas indústrias, sobretudo devido à tradição agrária do país, bem como da grande disponibilidade de terras férteis cultiváveis. Ademais, devido à grande extensão territorial do país e sua diversidade de microclimas, pode-se cultivar a cana em diferentes regiões do Brasil e em diferentes períodos do ano, facilitando o abastecimento de matéria-prima e de seus produtos, sobretudo no que diz respeito à estruturação da logística de distribuição no país. A principal região produtora é a Centro-Sul, com destaque para São Paulo, Goiás, Minas Gerais e Mato Grosso do Sul.

Segundo informações da Conab (Companhia Nacional de Abastecimento), foram coletadas mais de 642,7 milhões de toneladas de cana-de-açúcar na safra de 2019/20, valor superior ao ano anterior em 3,6%. Entretanto, segundo sua estimativa de agosto de 2020, o setor poderá sofrer uma retração de 0,1% na produção de cana-de-açúcar na safra 2020/21, sobretudo em virtude da pandemia do Sars-Cov-2. Ainda assim, é um dos setores que mais empregam (RODRIGUES, 2010), gerando mais de 4,5 milhões de empregos, direta ou indiretamente, abrangendo agricultores, usinas e destilarias por todo território nacional. A Figura 1 apresenta a evolução da produção de cana-de-açúcar no Brasil da safra de 2005/06 até a estimativa para a safra de 2020/21.

Figura 1. Evolução da produção de cana-açúcar no Brasil

Fonte: CONAB, 2020. (1) Estimativa em agosto/2020.

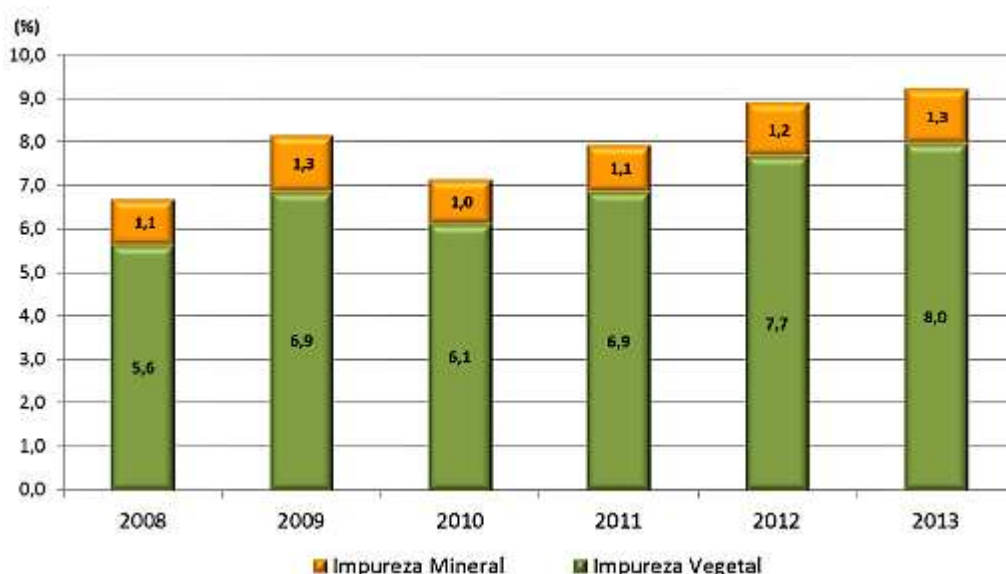
1.1 Impurezas em carregamentos de cana-de-açúcar

No que diz respeito à qualidade de produto final advindo da cana-de-açúcar, sabe-se que ela é fortemente dependente da quantidade de impurezas presentes na matéria-prima. Sua presença é inevitável, sendo tanto a quantidade quanto o tipo de impureza dependentes do sistema de produção, da qualidade do corte da matéria-prima, do carregamento, do transporte e das condições ambientais. Essas impurezas podem ser, majoritariamente, de origem vegetal e mineral.

Dentre as principais impurezas vegetais (FIGUEIREDO; MACIEL; MARQUES, 2008), há as próprias folhas da cana-de-açúcar (verdes, secas ou parcialmente secas), ervas daninhas que infestam o canavial, palmitos e palha. O principal problema desse tipo de impureza é o aumento da quantidade de fibra, que é responsável pela retenção de sacarose no bagaço. Já as impurezas minerais englobam terra, bem como partículas de implementos ou defensivos agrícolas. A presença de solo é uma das principais impurezas minerais presentes na cana a ser moída, levando a problemas como elevação de custos de transporte e de manutenção de equipamentos, bem como redução da eficiência de moagem e de extração de sacarose. O solo pode contaminar o processo com bactérias que causam problemas na fermentação, diminuindo a produção de etanol. Além dos problemas citados, a ocorrência dessas impurezas na matéria-prima pode causar aumento no consumo de energia no preparo da cana, perdas de extração, dificuldades para o tratamento do caldo, desgaste de equipamentos e redução na densidade da carga com consequente aumento de custo de transporte.

A Figura 2 mostra o histórico da porcentagem em massa de impurezas mineral e vegetal em carregamentos de cana-de-açúcar para o Centro-Sul do Brasil. Até 2013, houve um crescimento na presença dessas impurezas na matéria-prima das indústrias de etanol e açúcar, que trazem os problemas já apresentados.

Figura 2. Histórico da presença de impurezas mineral e vegetal em cana-de-açúcar no Centro Sul do Brasil



Fonte: CANAVIEIRA, 2013

Nesse mesmo ano, além disso, a quantidade de cada impureza em carregamentos de matéria-prima variou conforme o tipo de colheita (mecanizada crua, mecanizada queimada e manual queimada). A Tabela 1 indica esses valores.

Tabela 1. Porcentagem mássica de impurezas mineral e vegetal de acordo com o tipo e colheita

Tipo de colheita	Impureza mineral (%)	Impureza vegetal (%)
Mecanizada crua	1,2	7,5
Mecanizada queimada	1,4	6,0
Manual queimada	1,6	6,4

Fonte: Adaptado de CANAVIEIRA, 2013

Apesar de apresentar os menores teores de impurezas, os métodos de colheita que empregam a queima, apesar de eliminar parte de material seco antes do transporte da cana-de-açúcar à usina (SOUZA, 2005; OLIVEIRA, 2014), levam à emissão de gás carbônico para a atmosfera, bem como de outros gases e substâncias particuladas, contribuindo para o agravamento do efeito estufa, poluindo o ar e comprometendo sua qualidade na região em que a atividade é realizada.

1.2 Identificação e determinação de impurezas e o uso de imagens digitais

A determinação e classificação de impurezas em carregamentos de cana-de-açúcar tem sido feita a partir de métodos já estabelecidos. Segundo recomendado pela União Nacional da Bioenergia, UDOP (2014), as impurezas vegetais podem ser determinadas pelo método de limpeza manual e à seco, enquanto as impurezas minerais pelo método da incineração em forno mufla. Entretanto, são operações exaustivas, com custos relativamente elevados e que causam perda de quantidade de amostras. Assim, o desenvolvimento de métodos analíticos alternativos que não requerem solventes ou reagentes, ou que sejam não destrutivos às amostras, além de reduzir custos, pode minimizar ou extinguir a geração de resíduos, diminuindo então os impactos ao meio ambiente (SOUTO et al., 2014). Nesse sentido, o uso de imagens digitais se mostra como uma alternativa vantajosa, já que se trata de um método não invasivo, facilita a aquisição de dados e reduz significativamente os custos.

Souto et al. (2014) propuseram uma metodologia para identificar adulteração por cascas e palitos em cafés moídos utilizando imagens digitais e seleção de variáveis. Foi utilizado como aparelho de captura de imagens um *scanner*, sendo as imagens obtidas registradas e processadas em um *notebook* pessoal. Elas então foram convertidas em histogramas de cores de matiz, luminosidade e saturação (HLS), sendo então usados como dados de entrada e avaliados estatisticamente usando técnicas supervisionadas de reconhecimento de padrões (SPA-LDA, SIMCA e PLS-DA). Dentre as 103 amostras de café moído, os melhores resultados obtidos foram utilizando a técnica SPA-LDA, atingindo uma taxa de precisão média (número de classificações corretas dividido pelo número total de amostras no conjunto, treinamento ou teste, considerado) de 92,5%.

Domínguez et al. (2014) também utilizaram imagens digitais, propondo uma metodologia para classificação de origem geográfica de amostras de mel do sudoeste de Buenos Aires. Para isso, foi empregado um sistema *Flow-Batch* que inclui uma *webcam* para capturar imagens digitais. Essas imagens foram então convertidas a histogramas de cores utilizando o software ImageJ 1.44p, utilizando as escalas de cores RGB (*red*, *green*, *blue*), HSB (*hue-saturation-brightness*) e escala de cinza. O tratamento dos dados envolveu o uso de técnicas de seleção de variável e de classificação, tais como SPA-LDA, SIMCA e PLS-DA. Os resultados obtidos mostraram que a técnica SPA-LDA e o modelo de cores RGB são ferramentas capazes de discriminar as amostras de mel em sua origem geográfica, atingindo uma classificação média correta de 97,3%.

Por fim, Gonçalves (2015) propôs o desenvolvimento de uma metodologia para classificação de plumas de algodão naturalmente colorido de acordo com o tipo de cultivo, utilizando imagens digitais e análise multivariada. As imagens digitais foram adquiridas por meio de uma *webcam*, sendo então convertidas em histogramas nas escalas de cores RGB (*red, green, blue*), HSV (*hue-saturation-value*) e escala de cinza. O tratamento dos dados envolveu o uso de técnicas de seleção de variável (SPA ou SW) e de classificação (PLS-DA ou LDA). Os resultados obtidos mostraram que a técnica LDA/SW conseguiu atingir uma classificação correta de 96% para o conjunto de teste utilizando a combinação HSV.

Assim, essa forma de aquisição de dados aliada a ferramentas de regressão e classificação de informações podem representar uma poderosa alternativa aos métodos de análise tradicionais. Nesse sentido, pode-se explorar o uso de Redes Neurais Artificiais para a resolução de problemas.

2 REDES NEURAIS ARTIFICIAIS (RNAs)

Redes neurais artificiais (RNAs), também conhecidas como sistemas conexionistas ou sistemas de processamento paralelo e distribuído (PDP), são modelos computacionais que relembram a estrutura e o funcionamento do cérebro humano. São sistemas em que há distribuição de unidades básicas de processamento (neurônios), interligados por diversas conexões (sinapses), implementadas por funções matemáticas diversas, vetores e matrizes. Apresentam a habilidade de adquirir e registrar informações a partir do processamento de um banco de dados, tendo elevada capacidade de mapear e resolver problemas não-lineares (BRAGA, CARVALHO, LUDERMIR, 2000; SILVA et al, 2017).

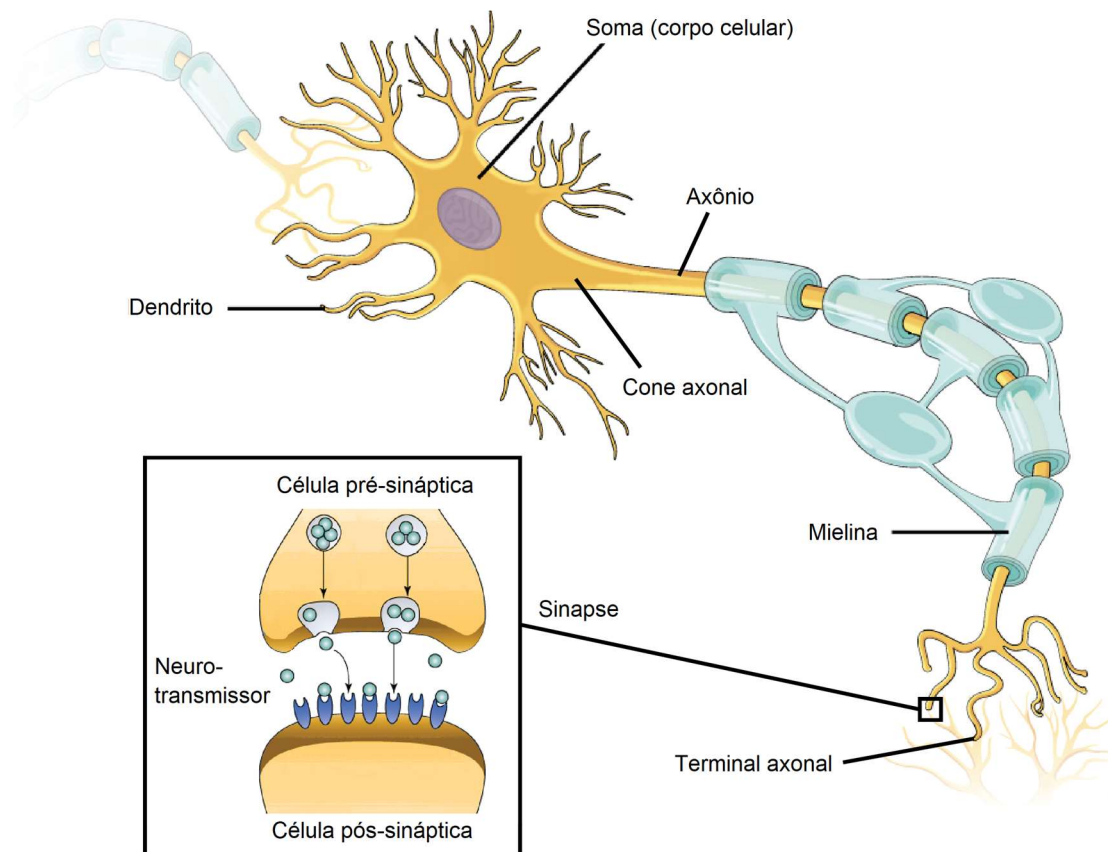
Em RNAs, a resolução de problemas se dá inicialmente por uma fase de aprendizagem e, posteriormente, uma fase de generalização. Na fase de aprendizagem, um conjunto de dados é fornecido à rede, de modo que esta é capaz de extrair as características necessárias para representar as informações contidas nesse conjunto. Na fase de generalização, essas características são utilizadas para gerar respostas para o problema em questão.

2.1 Neurônio biológico

O principal componente do sistema nervoso central é o neurônio, responsável pelo processamento de informações para o cérebro humano (SILVA, 2017). Sua função é conduzir impulsos elétricos em condições específicas para que haja comunicação para determinadas células-alvo no corpo humano. Ele pode ser dividido em três partes principais: dendritos, axônio e corpo celular (Figura 3).

Os dendritos são compostos por várias extensões finas e curtas, que formam as árvores dendríticas. Essa estrutura é responsável pelo recebimento de estímulo (sinais) de outros neurônios ou do ambiente externo e consequente condução dessas informações para o corpo celular. O corpo celular, então, faz o processamento dos sinais recebidos pelos dendritos e verifica se são excitatórios ou inibitórios, isto é, se o neurônio dispara e gera um impulso elétrico ou se não. Além disso, é onde se encontram as organelas citoplasmáticas. Já o axônio, composto por uma única extensão longa a partir do corpo celular, conduz o impulso elétrico para outros neurônios. Ele é coberto por uma substância de proteção chamada mielina, que também tem a função de acelerar a transmissão do impulso nervoso.

Figura 3. Estrutura de um neurônio biológico e esquema de sinapse



Fonte: Khan Academy, 2016

Na terminação do axônio encontram-se vários ramos, conhecidos como terminais nervosos (ou axônicos). A conexão dessa terminação com outro neurônio é conhecida por sinapse, junção que possibilita a transmissão de informação de um neurônio a outro a partir de neurotransmissores. Essas moléculas são então liberadas pelo neurônio pré-sináptico e se ligam a receptores da membrana da célula seguinte (pós-sináptica), carregando estímulos de excitação sináptica ou inibição sináptica. Assim, não há contato físico entre um neurônio e outro.

Estima-se que há mais de 100 bilhões de neurônios na rede neural biológica e que cada um pode estar conectado a cerca de 6000 neurônios. São essas conexões que permitem que o cérebro consiga realizar funções complexas.

2.2 Histórico

Warren McCulloch e Walter Pitts, em 1943, propuseram o primeiro modelo matemático (artificial) de um neurônio biológico. Este modelo ficou conhecido como nó MCP (iniciais dos pesquisadores). Seu trabalho se concentra na descrição de um modelo

artificial, não no processo e técnicas de aprendizado (BRAGA, CARVALHO, LUDERMIR, 2000; FLECK, 2016).

O primeiro trabalho na área de redes neurais cujo processo de aprendizado tentou ser explicado foi apresentado por Donald Hebb, em 1949. O pesquisador propôs uma teoria em que o aprendizado de neurônios biológicos era baseado no reforço das ligações sinápticas entre neurônios excitados (BRAGA, CARVALHO, LUDERMIR, 2000).

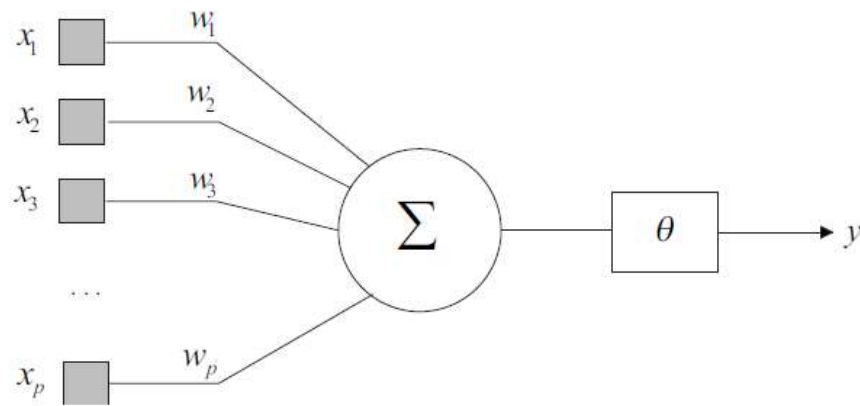
Em 1958, Frank Rosenblatt mostrou que se fossem adicionadas sinapses ajustáveis ao seu modelo de rede (*perceptron*), as RNAs com neurônios MCP poderiam ser treinadas para que alguns tipos de padrões pudessem ser classificados. Ele descreveu uma estrutura para a rede, estruturas de ligação entre os neurônios e propôs um algoritmo para treinar a rede. Rosenblatt demonstrou o teorema de convergência do *perceptron*, que mostra que esse tipo de rede sempre converge para problemas linearmente separáveis. Ele também criou diversas variações do *perceptron*, dentre elas uma rede com uma camada em que pesos e bias poderiam ser treinados para gerar, a partir de vetores de entrada, um vetor de saída correto (BRAGA, CARVALHO, LUDERMIR, 2000; FLECK, 2016).

O trabalho de Rosenblatt possibilitou o aumento de trabalhos relacionados a redes neurais até 1969. Neste ano, Minsky e Papert publicaram um trabalho que mostrou deficiências e limitações ao modelo de *perceptron* de múltiplas camadas, provocando desinteresse pelos estudos na área de RNAs. Apenas a partir da 1982, com a publicação do trabalho de Hopfield, o interesse no estudo das redes neurais voltou a ser despertado (FLECK, 2016).

2.3 Modelos de neurônios

O modelo de rede neural proposto por McCulloch e Pitts (nó MCP) é descrito matematicamente contendo p entradas x_1, x_2, \dots, x_p e apenas uma saída y . Essas entradas são ponderadas por pesos (pesos sinápticos) w_1, w_2, \dots, w_p , podendo estes assumirem valores positivos (sinapses excitatórias) ou negativos (sinapses inibitórias). A Figura 4 ilustra esse neurônio.

Figura 4. Modelo do neurônio proposto por McCulloch e Pitts



Fonte: NASCIMENTO, 2007

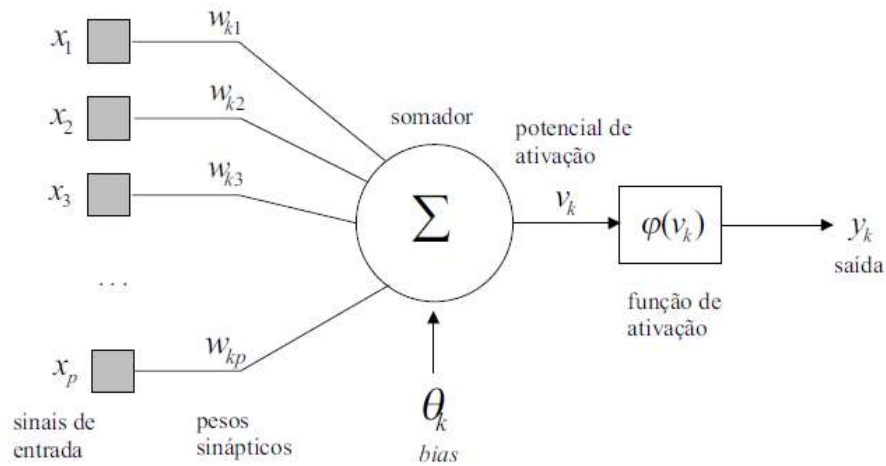
A ativação do neurônio MCP é obtida a partir da aplicação de uma função de ativação, a qual ativa ou não a saída, com base no valor da soma ponderada das entradas. Assim, matematicamente, a saída desse modelo é ativada quando

$$\sum_{i=1}^p x_i w_i \geq \theta \quad 1$$

sendo p o número de entradas do neurônio, w_i o peso associado à entrada x_i e θ o limiar de excitação (*threshold*) do neurônio.

Esse modelo possui algumas limitações, tais como o fato de que redes MCP contendo uma camada são capazes de implementar apenas funções linearmente separáveis, pesos negativos se ajustam melhor a estímulos inibitórios e os pesos são fixos, isto é, são não-ajustáveis (SILVA et al, 2017).

A limitação de pesos fixos (BRAGA, CARVALHO, LUDERMIR, 2000) foi contornada a partir da proposta de novos modelos que permitem a produção de uma saída qualquer e trabalham com diferentes funções de ativação. Esses modelos são inspirados no *perceptron*, proposto por Rosenblatt, que é composto por uma estrutura de rede tendo como unidade básica o MCP, e por uma regra de aprendizado. O modelo desse neurônio não-linear pode ser visto na Figura 5.

Figura 5. Modelo de neurônio não linear baseado no *perceptron***Fonte:** NASCIMENTO, 2007

Podem ser identificados três elementos básicos nesse modelo:

- (i) Um conjunto de sinapses contendo pesos próprios. Assim, um sinal x_i na entrada da sinapse i , conectada ao neurônio k , é multiplicado pelo seu respectivo peso w_{ki} .
- (ii) Um somador, o qual soma os sinais de entrada multiplicados pelos seus respectivos pesos.
- (iii) Uma função de ativação, a qual restringe a amplitude da saída do neurônio.

O modelo da Figura 5 inclui também um limiar de ativação (θ_k) aplicado externamente. Esse limiar, ou *bias*, tem o efeito de aumentar ou diminuir a entrada da função de ativação. Matematicamente, tem-se

$$u_k = \sum_{i=1}^p x_i w_{ki} \quad 2$$

$$v_k = u_k + \theta_k \quad 3$$

$$y_k = \varphi(v_k) \quad 4$$

sendo x_i os sinais de entrada, w_{ki} os pesos do neurônio k , u_k é a saída do somador, v_k é o potencial de ativação, θ_k é o bias, $\varphi(\cdot)$ é a função de ativação e y_k é o sinal de saída do neurônio.

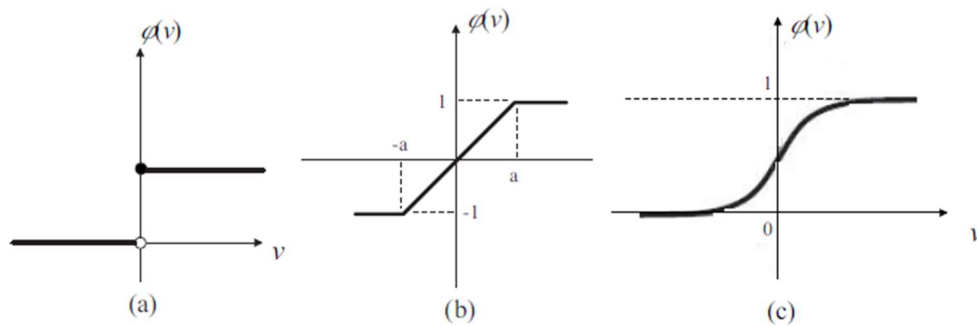
Vale salientar que, nesse modelo, os pesos w_{ki} são parâmetros escalares ajustáveis. Dessa forma, a ideia central por trás da rede é ajustar esses parâmetros até que esta apresente o comportamento de interesse para gerar uma resposta esperada. A

rede, então, pode ser treinada para um determinado problema ajustando-se os pesos ou *bias* ou, ainda, a própria rede ajusta esses parâmetros.

2.4 Funções de ativação

Segundo Fleck (2016), o modelo de cada neurônio da rede pode incluir uma não-linearidade na sua saída, que deve ser reduzida. A função de ativação representa o efeito na saída de um neurônio devido ao potencial de ativação v , isto é, define a saída em termos do potencial de ativação. Existem diversos tipos de funções de ativação, sendo as principais ilustradas na Figura 6.

Figura 6. Principais funções de ativação: (a) Degrau, (b) Linear por Partes e (c) Sigmoide



Fonte: NASCIMENTO, 2007

a) Função Degrau

$$\phi(v) = \begin{cases} 1, & \text{se } v \geq 0 \\ 0, & \text{se } v < 0 \end{cases} \quad 5$$

b) Função Linear por Partes

$$\phi(v) = \begin{cases} 1, & \text{se } v \geq a \\ v, & \text{se } -a < v < a \\ -1, & \text{se } v \leq -a \end{cases} \quad 6$$

c) Função Sigmoide

$$\phi(v) = \frac{1}{1 + \exp(-v)} \quad 7$$

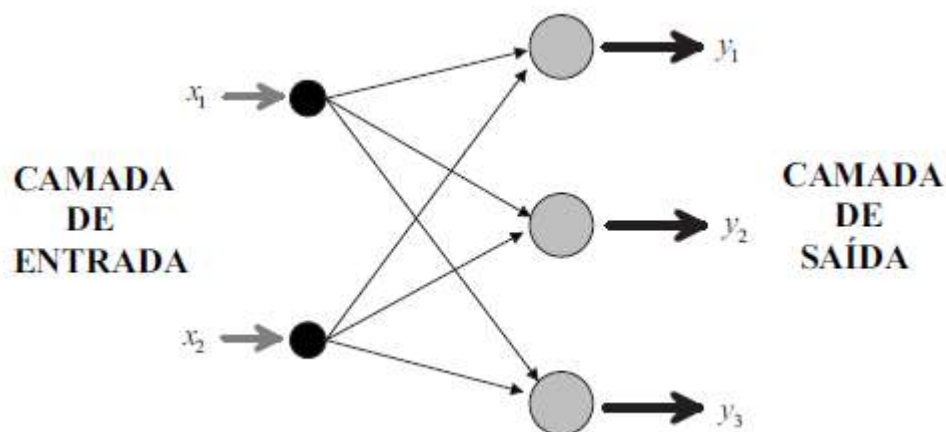
2.5 Arquitetura de Redes Neurais Artificiais

A classificação de RNAs pode ser feita considerando-se diferentes parâmetros, tais como sua arquitetura, o tipo de neurônio e a regra de aprendizagem (SILVA et al, 2017). A arquitetura de uma rede neural é composta por: número de camadas, número de neurônios em cada camada, ligações entre os neurônios e pelo fluxo de informação.

As camadas de uma RNA são divididas em três grupos: camada de entrada (padrões dos dados são apresentados à rede), camadas intermediárias ou ocultas (onde a maior parte do processamento é feito, através das conexões ponderadas e então as informações são extraídas), camada de saída (resultado final é coletado).

Além disso, as redes neurais podem ser divididas em três classes principais: redes *feed-forward* com uma única camada, redes *feed-forward* com múltiplas camadas e redes recorrentes (BRAGA, CARVALHO, LUDERMIR, 2000). Nas redes *feed-forward* com uma única camada, só existe uma única camada de neurônios entre a entrada e a saída da rede, de modo que o fluxo de informação segue estritamente adiante (Figura 7). Nesse tipo de rede, o número de saídas sempre coincide com o número de neurônios. É bastante empregada em problemas lineares e de reconhecimento de padrões.

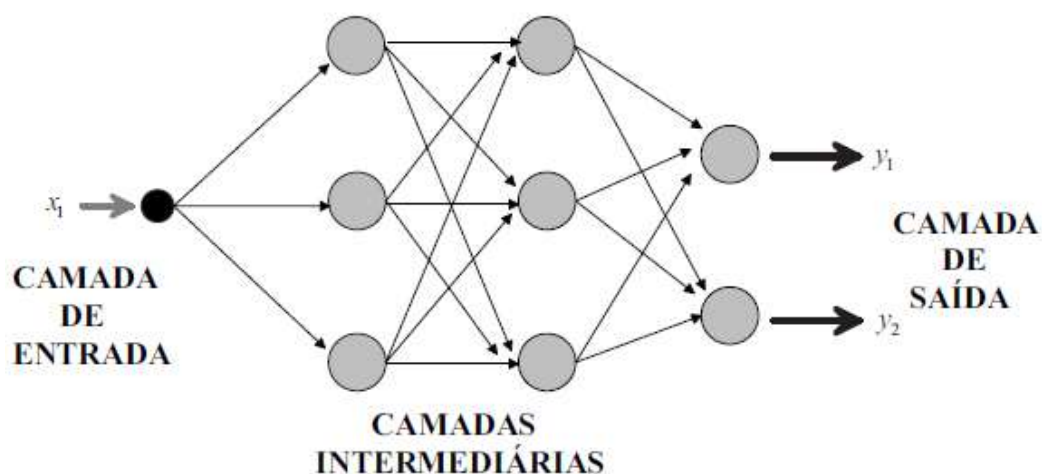
Figura 7. Exemplo de rede *feed-forward* com uma única camada



Fonte: NASCIMENTO, 2007

No caso em que há múltiplas camadas, há mais de uma camada de neurônios entre a entrada e a saída e, da mesma forma, a informação segue estritamente adiante (Figura 8). Essa arquitetura de RNA é empregada na resolução de diversos problemas (SILVA et al, 2017), dentre eles a classificação de padrões, controle de processos, aproximação de funções, otimização, robótica, entre outros. Assim como na rede com uma única camada, o número de saídas (resposta desejada) é o mesmo que o de neurônios na camada de saída.

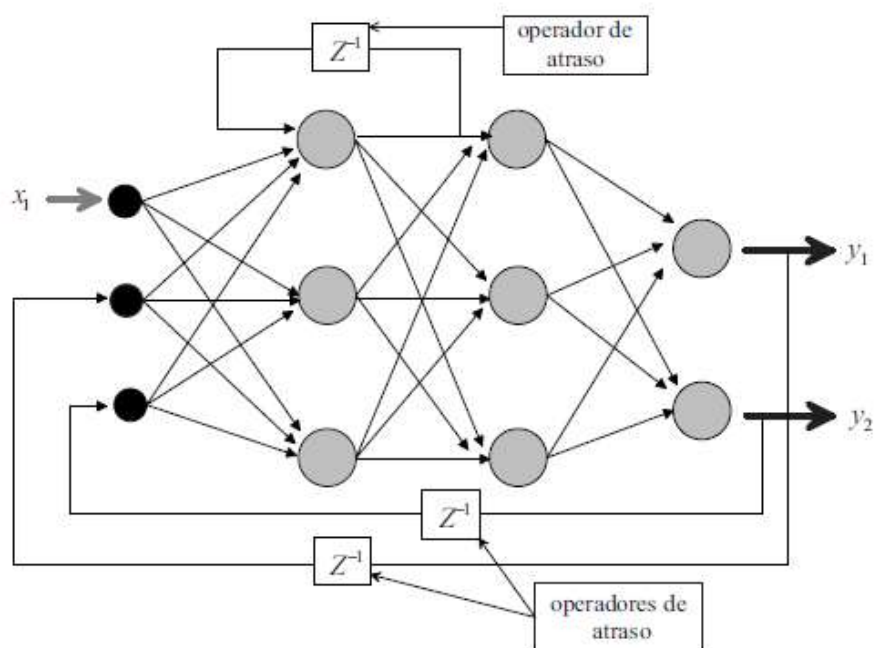
Figura 8. Exemplo de rede *feed-forward* com múltiplas camadas



Fonte: NASCIMENTO, 2007

Nas redes recorrentes, há conexões de alimentação (Figura 9), de modo que um neurônio pode estar conectado a outro de uma mesma camada ou de camadas anteriores (BRAGA, CARVALHO, LUDERMIR, 2000). Por causa dessas conexões, há criação de um comportamento temporal na rede, como uma espécie de memória de curto prazo. Assim, são bastante utilizadas em sistemas de identificação e otimização, controle de processos, entre outros.

Figura 9. Exemplo de rede neural recorrente



Fonte: NASCIMENTO, 2007

2.6 Tipos de aprendizagem

O que torna as redes neurais artificiais tão atrativas é sua capacidade de aprender a partir da apresentação de um grupo de amostras que descrevem alguma característica de interesse. Assim, após a rede aprender as relações existentes entre as entradas e saídas, ela é capaz de generalizar soluções, de modo que as saídas são próximas do que se é esperado ou desejado a partir de valores de entrada dentro do intervalo de dados usados no treinamento.

No processo de aprendizagem da rede neural artificial (SILVA et al, 2017), a cada vez que um conjunto de treinamento tem seus pesos sinápticos e limiar de ativação ajustados, tem-se o que se chama de época.

Os tipos de aprendizado podem ser divididos em três grupos principais: o aprendizado supervisionado, aprendizado não-supervisionado e o aprendizado por reforço. No aprendizado supervisionado, a entrada e a saída da rede são fornecidas pelo operador externo da rede, de modo que se procura ajustar os parâmetros da rede encontrando-se ligações entre os pares de entrada e saída. Dessa forma, a partir do algoritmo de aprendizado, a rede é considerada treinada quando a discrepância entre a saída da rede e a saída real de um determinado problema está dentro de um intervalo aceitável. É a forma de aprendizado mais empregada.

No aprendizado não-supervisionado, não há operador externo à rede acompanhando o processo de aprendizagem, de modo que a própria rede extrai as informações do conjunto de entrada para se chegar à saída. Nesse tipo de aprendizado, a rede começa a identificar as particularidades do conjunto de amostras, num processo de mapeamento das informações.

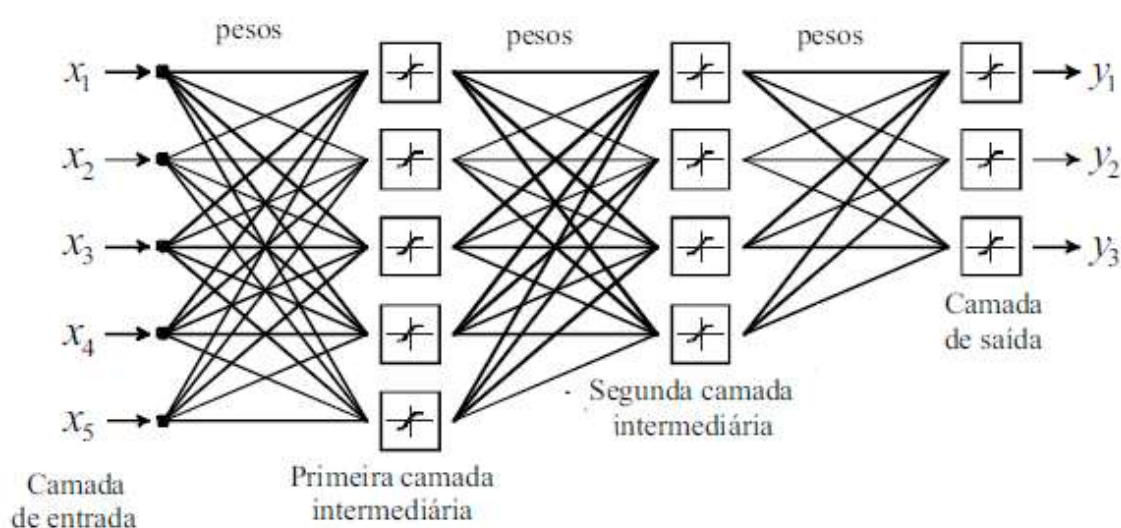
No aprendizado por reforço, algumas vezes designado como uma variação do aprendizado supervisionado, a rede tem como base quaisquer informações qualitativas ou quantitativas adquiridas a partir de interações com o ambiente a ser mapeado. Aqui, o processo é feito por tentativa e erro, na medida em que ações satisfatórias fazem os pesos sinápticos serem incrementados como forma de recompensa, enquanto ações insatisfatórias são penalizadas. Essa forma de aprendizado é bastante utilizada na robótica.

2.7 Redes Neurais *Perceptrons* De Múltiplas Camadas

Problemas mais complexos, especificamente aqueles não-linearmente separáveis, não podem ser resolvidos por uma rede neural de uma camada, como o

modelo tradicional de *Perceptron*. Assim, surgem as redes de *Perceptrons* de Múltiplas Camadas (redes MLP), também conhecidas como redes neurais *feed-forward* de múltiplas camadas, em cuja estrutura há uma ou mais camadas intermediárias entre a camada de entrada e a camada de saída (Figura 10). Assim, nesse tipo de rede, a propagação do sinal de entrada flui em uma direção, camada por camada, isto é, da camada de entrada até a camada de saída. A partir da análise da figura, observa-se que as saídas dos neurônios da primeira camada, por exemplo, serão as entradas dos neurônios da segunda e assim por diante.

Figura 10. Exemplo de rede MLP



Fonte: NASCIMENTO, 2007

As redes MLP apresentam um processo de treinamento supervisionado, e este se dá pela combinação entre o processamento realizado pelos neurônios de cada camada. Assim, essas unidades são como extratoras e detectores de certas características que, em conjunto, conseguem resolver uma diversidade de problemas. Em cada camada há a extração de maior informação possível de determinado comportamento do sistema, sendo codificada pelo conjunto dos pesos e limiares sinápticos dos neurônios, até chegar no sinal de saída da RNA (BRAGA, CARVALHO, LUDERMIR, 2000; SILVA et al, 2017). O algoritmo de aprendizado usado durante o processo de treinamento de um MLP é chamado de retropropagação de erro.

O processo de treinamento de redes MLP utilizando o algoritmo de retropropagação envolve um mecanismo de minimização de erros a partir do ajuste dos pesos sinápticos. Esse treinamento é feito em duas etapas: *forward* e *backward*. Na

etapa *forward*, os sinais de entrada de uma determinada amostra são inseridos na rede e propagados camada por camada até a camada de saída, definindo uma resposta com base nos atuais pesos dos neurônios. Na etapa *backward*, as repostas produzidas são comparadas às respostas desejadas, determinando os erros do modelo. A partir desses erros, propagados da camada de saída até a de entrada, camada por camada, é feito o ajuste dos pesos de todos os neurônios (BRAGA, CARVALHO, LUDERMIR, 2000; SILVA et al, 2017). Assim, esses parâmetros são ajustados a cada iteração, até chegar na melhor solução da rede neural, em que os erros são minimizados.

2.8 Algoritmo de retropropagação do erro (*backpropagation*) e de Levenberg-Marquardt

O algoritmo *backpropagation* propõe uma forma de definir o erro entre os neurônios das camadas intermediárias, o que leva ao ajuste dos pesos.

Para realizar esse ajuste de erro, deve-se, portanto, calculá-lo. Ele é definido como a diferença entre a resposta desejada para o neurônio (valor fornecido à rede) e o valor estimado pela rede.

$$e_i(n) = d_i(n) - y_i(n) \quad 8$$

em que i representa um neurônio qualquer, n a iteração, $d_i(n)$ a resposta esperada e $y_i(n)$ o valor retornado pela RNA.

Além disso, pode-se determinar o erro instantâneo para o neurônio, definido como $(1/2)e_i^2(n)$, bem como o erro total, que é a soma desses erros.

$$E(n) = \frac{1}{2} \sum_{i \in C} e_i^2(n) \quad 9$$

sendo C o conjunto que abrange todos os nodos da camada de saída.

Se existem N exemplos contidos no conjunto de treinamento, pode-se calcular o erro médio E_{med} , que é a soma do erro quadrático total definido acima, e normalizar em relação ao tamanho do conjunto (N).

$$E_{med} = \frac{1}{N} \sum_{i=1}^N E(n) \quad 10$$

O algoritmo *backpropagation* tenta minimizar esse erro médio, fazendo-se ajustes nos pesos sinápticos da rede.

O processamento da rede se dá, após a inclusão dos dados de entrada, com a aplicação do liminar ou potencial de ativação de um dos neurônios da RNA, que é dado por:

$$v_j(n) = \sum_{i=0}^m w_{ji}(n) y_i(n) \quad 11$$

em que j indica o neurônio analisado, i representa o neurônio anterior que envia o sinal recebido para o neurônio seguinte (no caso, o neurônio j), m representa o número de entradas (excluindo o bias), $w_{ij}(n)$ é o peso sináptico que une a saída de um neurônio à entrada do neurônio seguinte, $y_i(n)$ é o sinal funcional ($y_i(n) = \phi_i(v_j(n))$) emitido pelo neurônio i . A entrada associada ao bias é $y_0 = +1$.

O peso sináptico $w_{ij}(n)$ recebe uma correção $\Delta w_{ij}(n)$, que é proporcional à derivada parcial $\frac{\partial E(n)}{\partial w_{ji}(n)}$, sendo calculado pela equação

$$\Delta w_{ji}(n) = -\eta \frac{\partial E(n)}{\partial w_{ji}(n)} \quad 12$$

em que η é o parâmetro de taxa de aprendizado. O sinal negativo indica a direção de mudança para reduzir o valor de $E(n)$.

A partir dessas equações, obtemos a expressão final

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_i(n) \quad 13$$

em que o gradiente local $\delta_j(n)$ dado por

$$\delta_j(n) = e_j(n) \phi'_j(v_j(n)). \quad 14$$

O algoritmo de retropropagação do erro pode ser otimizado de várias formas, tendo destaque o método de Levenberg-Marquadt. Esse algoritmo é um método de segunda ordem, da mesma forma que os métodos do gradiente conjugado e do método quase-Newton, uma vez que se baseia em informações da derivada segunda da função erro. É um método mais eficiente do que o de gradiente conjugado, além de possuir menor custo computacional do que o método quase-Newton (PRONI, HANEDA, FILLETTI, 2020). O algoritmo de Levenberg-Marquardt introduz um parâmetro μ no método iterativo de Gauss-Newton, segundo a equação

$$\Delta x = [J^T(x)J(x) + \mu I]^{-1} J^T(x)e(x) \quad 15$$

em que I é a matriz identidade, $e(x)$ é a função erro e J é a matriz Jacobiana.

A função do parâmetro μ é estabilizar o treinamento, de modo a convergir mais rapidamente do que o método de Newton e evitando passos muito grandes que levassem a um erro de convergência.

2.9 Algoritmo do gradiente conjugado escalonado

Como já visto, de forma simplificada, o treinamento de uma rede neural artificial é equivalente à minimização de uma função de erro global, isto é, uma função multivariada que depende dos pesos da rede. Especificamente a problemas de larga escala, apenas os métodos de otimização mostram-se adequados, aplicáveis e relevantes para sua resolução, especialmente àqueles chamados de Métodos de Gradiente Conjugado. Nesses algoritmos, de modo geral, os pesos são ajustados a partir do método da descida mais íngreme (negativo do gradiente), na direção onde a função de custo (que calcula a média de erros sobre a amostra de treinamento) decresce mais rapidamente (ALMEIDA, 2007).

Segundo MØLLER (1993), algoritmos como CGL (Gradiente Conjugado de busca em linha) e BFGS (Broyden-Fletcher-Goldfarb-Shanno), pertencentes à classe de Métodos de Gradiente Conjugado, melhoram o desempenho de redes neurais em relação ao algoritmo *backpropagation*, mas aumentam significativamente a complexidade da resolução dos problemas, uma vez que tanto a função erro quanto suas derivadas tornam-se mais complexas, aumentando o gasto computacional. Nessa perspectiva, o autor propôs uma variação dos Métodos de Gradiente Conjugado (CG), algoritmo chamado Gradiente Conjugado Escalonado, combinando a abordagem de região de confiança usada no algoritmo Levenberg-Marquardt com a abordagem do CG.

O algoritmo do gradiente conjugado escalonado (GCE ou SCG) é um algoritmo de segunda ordem, já que é utilizada a derivada segunda de sua função objetivo. Além disso, esse algoritmo utiliza uma aproximação do cálculo da matriz hessiana da função objetivo (matriz das derivadas segundas), sendo definida positivamente, de modo que não há taxa de aprendizagem fixa. Por isso, inclui-se um escalar λ para regular essa definição (ALMEIDA, 2007), conforme indicado pela Equação 16.

$$s_k = \frac{\xi'(w_k + \eta_k p_k) - \xi'(w_k)}{\eta_k} + \lambda_k p_k \quad 16$$

Sendo, para uma iteração k , w_k o vetor peso, η_k a taxa de aprendizagem, ξ a função custo e p_k o sentido de busca.

Dada as informações à rede, a cada iteração há o ajuste desse escalar, de modo a chegar na minimização da função erro. Esse escalar é ajustado com base na informação $\delta_k = p_k^T \xi''(w_k) p_k$, fator que revela se a matriz hessiana do problema é ou não positiva (ALMEIDA, 2007). A função *trainscg* do *software* MATLAB compõe esse algoritmo.

Em problemas de classificação usando esse algoritmo, a função de custo (erro) mais utilizada e adequada é a entropia-cruzada (*cross-entropy*). Essa função, ao contrário do erro quadrático médio (MSE), interpreta o sinal de treinamento e as saídas da rede como probabilidades, de modo que o algoritmo, como já descrito, tentará minimizar essa diferença de probabilidades (VIEIRA; RAZENTE; BARIONI, 2017).

3 METODOLOGIA

3.1 Amostras de bagaço de cana-de-açúcar

As amostras de cana-de-açúcar, material vegetal e solo foram adquiridas em uma fazenda localizada em Ibaté-SP, ao longo de 56 dias, como parte do projeto de doutorado do discente Wesley Nascimento Guedes, sob a orientação da profa. Dra. Fabíola Manhas Verbi Pereira. Cada mistura contendo esses materiais foi feita seguindo a seguinte proporção: 41-100% em massa de cana-de-açúcar, 0-40% em massa de material vegetal e 0-20% em massa de solo, seguindo um planejamento fatorial. Essas amostras podem ser vistas nas Figuras 11 e 12.

Foram feitas 122 misturas em proporções definidas. Além disso, 12 amostras de intervalo de 0-10% de impurezas foram triplicadas, totalizando 36 experimentos para este intervalo. Isso foi feito para que houvesse melhor representatividade dos dados. Então, havia 110 misturas variando a porcentagem de impurezas entre 11 e 60% e 36 misturas variando a porcentagem de impureza entre 0-10%, totalizando 146 experimentos. Em seguida, foram obtidas 5 imagens digitais de cada amostra, totalizando 730 imagens digitais.

Por fim, das imagens digitais foram extraídas informações de escalas de cores RGB - Red (R), vermelho; Green (G), verde; Blue (B), azul; HSV - Hue (H), matiz; Saturation (S), saturação; Value (V), valor; as cores relativas da escala RGB - rgb e L - Luminosity, luminosidade utilizando uma rotina no programa MATLAB. Uma vez convertidas nas escalas de cores, os valores finais utilizados para construção do banco de dados foram os da média entre as 5 imagens de uma mesma amostra. Os dados numéricos foram cedidos por Wesley Nascimento Guedes (GUEDES, 2020) e estão reunidos no ANEXO.

Figura 11. Cana coletada em Ibaté-SP



Fonte: GUEDES, 2020

Figura 12. Frações de cana, impureza vegetal e solo das misturas separadas



Fonte: GUEDES, 2020

Os valores de entrada que foram utilizados na RNA são aqueles obtidos a partir da conversão de imagens digitais nas escalas de cores RGB, HSV, rgb e L. As Figuras 13, 14 e 15 mostram alguns exemplos dessas imagens digitais.

Figura 13. Imagem digital de amostra contendo 90% de cana, 5% de impureza vegetal e 5% de solo



Fonte: GUEDES, 2020

Figura 14. Imagem digital de amostra contendo 90% de cana e 10% de solo



Fonte: GUEDES, 2020

Figura 15. Imagem digital de amostra contendo 90% de cana e 10% de impureza vegetal



Fonte: GUEDES, 2020

As Figuras 13, 14 e 15 são muito semelhantes, no entanto, a composição de impureza sólida varia, bem como o conteúdo de folhas verdes, marrons ou secas. Assim, visualmente, é impossível verificar as diferenças entre elas. É relativamente mais fácil ver que a figura 14 é diferente. Entretanto, num cenário em que há um caminhão grande contendo esse material pronto para ser entregue a uma indústria de açúcar e álcool, por

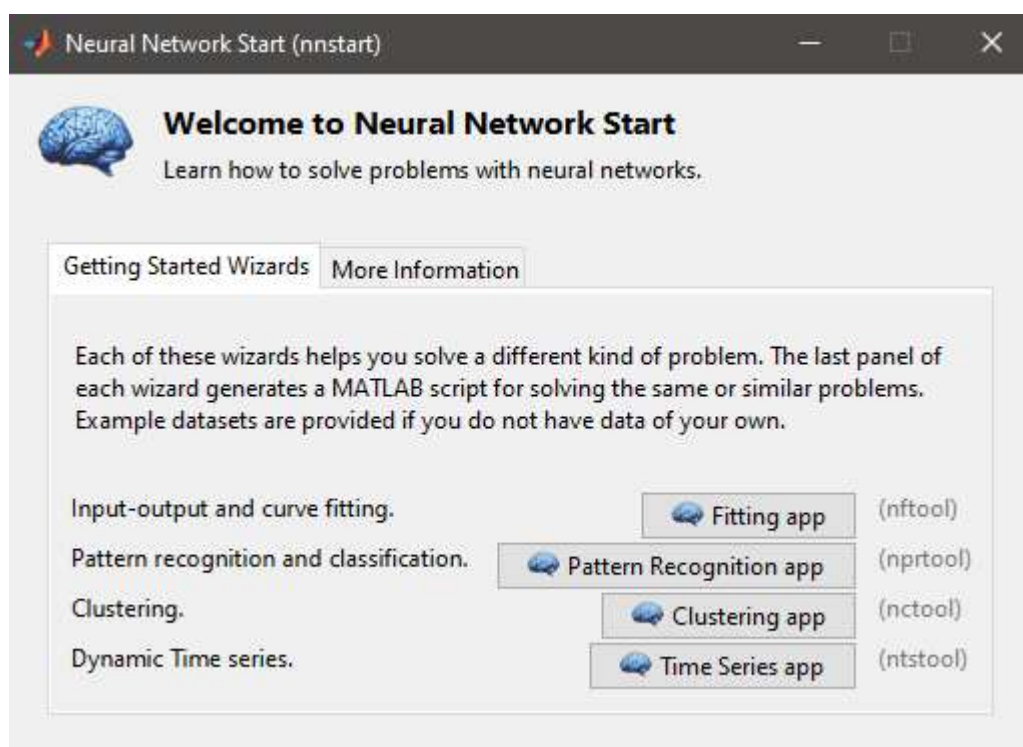
exemplo, é impossível controlar a qualidade da cana crua, considerando o grande volume de matéria-prima nessas fábricas que trabalham o dia todo.

3.2 Redes neurais artificiais

3.2.1 Modelos de predição

O treinamento do modelo de rede neural foi feito utilizando os algoritmos de Levenberg-Marquardt e gradiente descendente, mas apenas o primeiro foi usado em todos os testes. Para isso, foi utilizado o *software* Matlab R2018a. A implementação da RNA foi feita a partir da função NNStart disponível no *software*, escolhendo a ferramenta *fitting app*. cuja camada de entrada (10 cores), número de camadas intermediárias e camada de saída (%m/m de cana-de-açúcar) foram definidos manualmente. A interface da ferramenta NNStart está representada na Figura 16.

Figura 16. Interface gráfica da ferramenta NNStart



Fonte: MATHWORKS, 2019

O processo envolveu divisão aleatória das amostras, usando a função *dividerand* disponível no MATLAB, em três conjuntos: 70% (102 amostras) para o treinamento em si, necessário para atualizar os pesos dos neurônios e *biases* para minimizar o erro; 15% (22 amostras) para a validação, para verificar se a rede está generalizando as informações e para interromper o treinamento antes de ocorrer *overfitting*, que é quando a RNA, em vez de melhorar, piora seus resultados para padrões diferentes daqueles

utilizados no treinamento. Caso o erro quadrático médio para o conjunto de validação aumentasse 5 vezes consecutivas, o treinamento parava, obtendo-se o valor mínimo do erro. Os 15% restantes (22 amostras) foram para testes independentes da generalização da rede. Os resultados foram avaliados a partir de histogramas de erro, minimização do erro quadrático médio e coeficiente de correlação.

3.2.2 Modelos de classificação

O treinamento do modelo de rede neural foi feito utilizando o algoritmo de gradiente conjugado escalonado. Para isso, foi utilizado o *software* MATLAB R2018a. A implementação da RNA foi feita a partir da função NNStart disponível no *software*, escolhendo a ferramenta *pattern recognition app*, em que a camada de entrada (10 cores), número de camadas intermediárias e camada de saída foram definidos manualmente.

Para a classificação com base na porcentagem mássica mínima adequada de cana-de-açúcar (90%), foi feita a seguinte divisão: o intervalo de porcentagem em massa de cana-de-açúcar de 41-89% foi chamado de classe 2 (saída (0, 1)), representando 110 amostras, enquanto o de 90-100% foi chamado de classe 1 (saída (1, 0)), representando 36 amostras.

O processo de desenvolvimento da rede envolveu divisão aleatória das amostras, usando a função *dividerand* disponível no MATLAB, em três conjuntos: 70% (102 amostras) para o treinamento em si; 15% (22 amostras) para a validação, para verificar se a rede está generalizando as informações e para interromper o treinamento antes de ocorrer *overfitting*, que é quando a RNA, em vez de melhorar, piora seus resultados para padrões diferentes daqueles utilizados no treinamento. Caso o erro quadrático médio para o conjunto de validação aumentasse 5 vezes consecutivas, o treinamento parava, obtendo-se o valor mínimo do erro. Os 15% restantes (22 amostras) foram para testes independentes da generalização da rede. Os resultados foram avaliados a partir de matriz de confusão, minimização do erro de entropia cruzada e de curvas de característica de operação do receptor (ROC).

4 REDES NEURAIS ARTIFICIAIS PARA QUANTIFICAÇÃO DAS AMOSTRAS

Os modelos de RNAs para quantificar a porcentagem mássica de cana-de-açúcar, material mineral e material vegetal foram arquitetados variando o número de neurônios da camada intermediária entre 5 e 20, para determinar a arquitetura da RNA com a melhor capacidade preditiva. Para os cálculos, o número de camadas, o número de neurônios por camada, a taxa de aprendizado do parâmetro e as funções de ativação foram ajustados por tentativa e erro para obter o melhor resultado. Uma parada antecipada do treinamento foi realizada para evitar *overfitting* pela RNA, o que poderia levar a um mau desempenho de generalização. O algoritmo de retropropagação de erro com gradiente descendente também foi testado, variando a taxa de aprendizado de 0,1 a 1. Entretanto, os resultados obtidos não foram satisfatórios, além de que sua convergência é mais lenta quando comparada ao algoritmo de Levenberg-Marquardt.

4.1 Porcentagem mássica de cana-de-açúcar

A Tabela 2 resume os resultados para cada arquitetura de RNA testada, com base no coeficiente de correlação. Esses resultados foram publicados em Guedes et al. (2019).

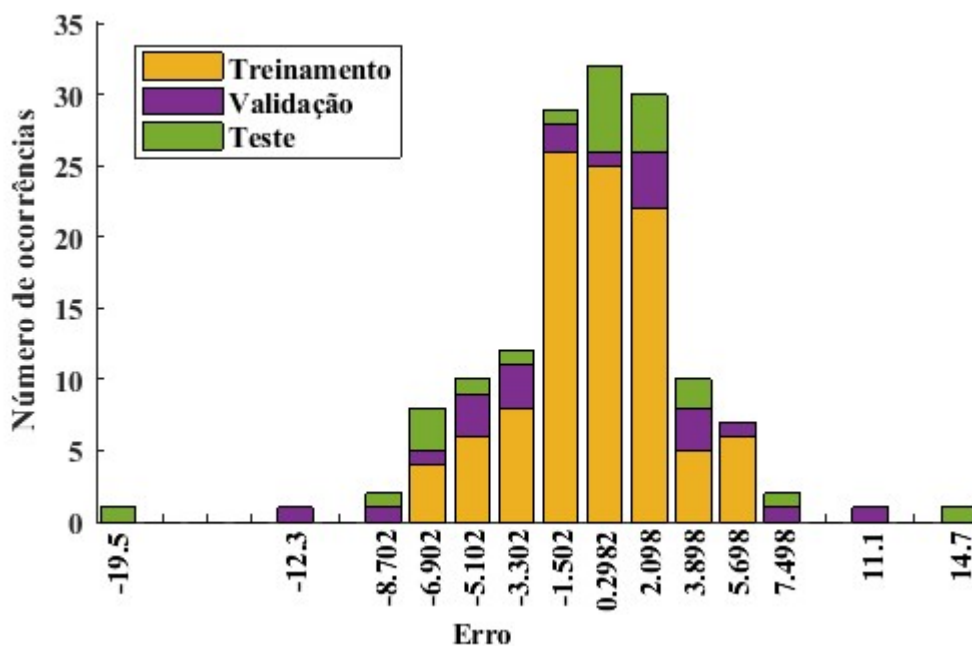
Tabela 2. Coeficientes de correlação para as 17 arquiteturas de RNAs para predição de porcentagem mássica de cana-de-açúcar

Arquitetura	Treinamento	Validação	Teste
1	0,96	0,97	0,93
2	0,96	0,96	0,96
3	0,95	0,97	0,92
4	0,96	0,91	0,92
5	0,98	0,94	0,92
6	0,98	0,93	0,91
7	0,93	0,92	0,93
8	0,93	0,95	0,95
9	0,92	0,93	0,92
10	0,94	0,95	0,91
11	0,94	0,91	0,90
12	0,96	0,91	0,93
13	0,98	0,89	0,95
14	0,95	0,97	0,89
15	0,97	0,93	0,94
16	0,96	0,94	0,95
17	0,94	0,93	0,96

Fonte: O autor, 2020

Uma das estratégias para decidir qual delas era a melhor era conciliar o maior coeficiente de correlação juntamente com a distribuição de erros para os testes. Entre as 17 arquiteturas testadas, aquela que contribuiu para os menores erros (Figura 17), bem como obteve distribuição normal de frequência do erro, foi a arquitetura 6.

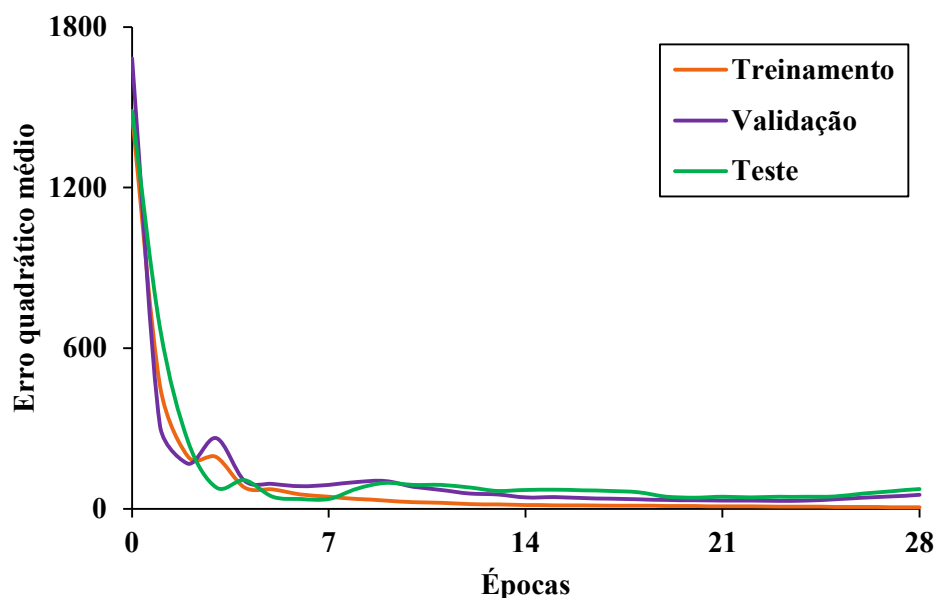
Figura 17. Diagrama de distribuição de erro para RNA de quantificação de cana-de-açúcar



Fonte: O autor, 2020

Esse modelo de RNA, com 8 neurônios na camada intermediária utilizando a função sigmoide e uma camada de saída utilizando uma função linear com o algoritmo de Levenberg-Marquardt, obteve um treinamento muito rápido (menos de 1 minuto), obtendo o melhor resultado em 23 épocas com um erro quadrático médio (MSE) de aproximadamente 8, 30 e 45 para os conjuntos de treinamento, validação e teste, respectivamente, conforme indicado na Figura 18. Os erros absolutos para o treinamento variaram de 0 a 8, os valores de validação variaram entre 1 e 13 e os valores de teste variaram entre 0 e 20. O valor de erro igual a 20 foi observado apenas para uma amostra.

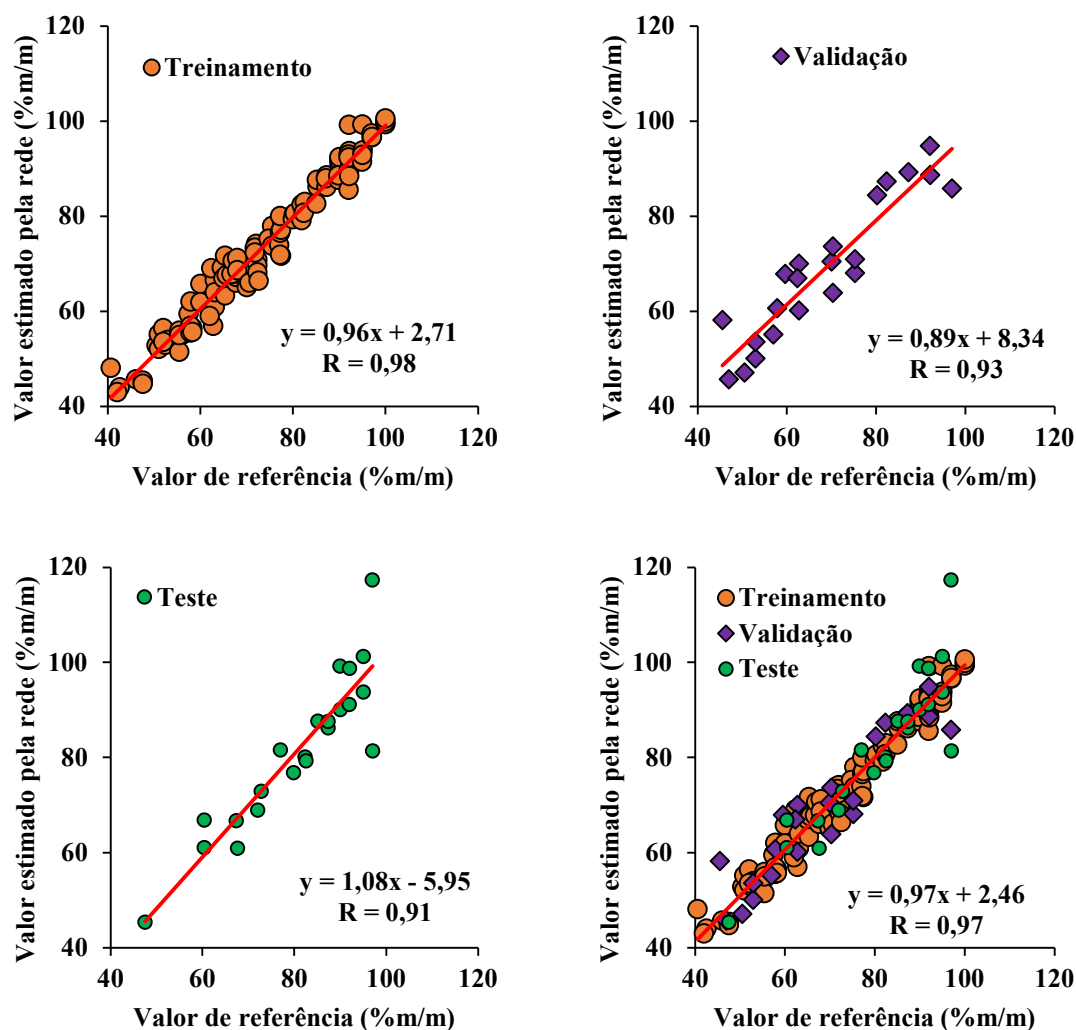
Figura 18. Performance do modelo de RNA com base no erro quadrático médio em função do número de épocas para a quantificação de cana-de-açúcar



Fonte: O autor, 2020

A Figura 19 mostra a alta correlação entre os valores do conteúdo de cana de referência e os valores previstos pelo modelo de RNA, sendo o intervalo real dos valores observados de porcentagem em massa de cana-de-açúcar de [41;100]. As retas que melhor representam os resultados obtidos pelos conjuntos de treinamento, validação e teste são $y = 0,96x + 2,82$, $y = 0,89x + 8,34$ e $y = 1,08x - 5,95$, respectivamente. Já os coeficientes de correlação obtidos foram 0,98, 0,93 e 0,91 para os conjuntos de treinamento, validação e teste. Os resultados obtidos pela RNA em virtude dos altos coeficientes de correlação mostram que o modelo teve alta capacidade em relacionar a natureza multivariada das entradas da rede com a respectiva saída, gerando respostas coerentes com os dados de saída reais. Das 146 amostras, nenhuma das estimativas feitas tiveram valores menores do que 40% em massa de cana-de-açúcar. Entretanto, 3 estimativas da rede foram maiores do que 100%, sendo duas delas em 101% e uma em 117%.

Figura 19. Porcentagem em massa de cana estimada pela RNA para os conjuntos de teste, validação e treinamento

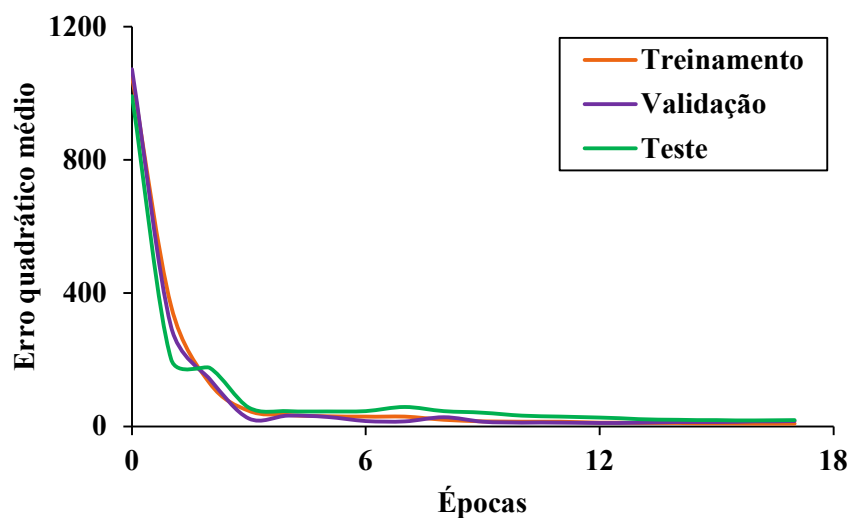


Fonte: O autor, 2020

4.2 Porcentagem mássica de material vegetal

A melhor RNA para a estimativa de material vegetal apresenta 10 neurônios na camada intermediária utilizando a função sigmoide, além de uma camada de saída utilizando uma função linear com o algoritmo de Levenberg-Marquardt. Ela obteve um treinamento muito rápido (menos de 1 minuto), obtendo o melhor resultado em 12 épocas com um erro quadrático médio (MSE) de aproximadamente 12, 10 e 27 para os conjuntos de treinamento, validação e teste, respectivamente, conforme indicado na Figura 20.

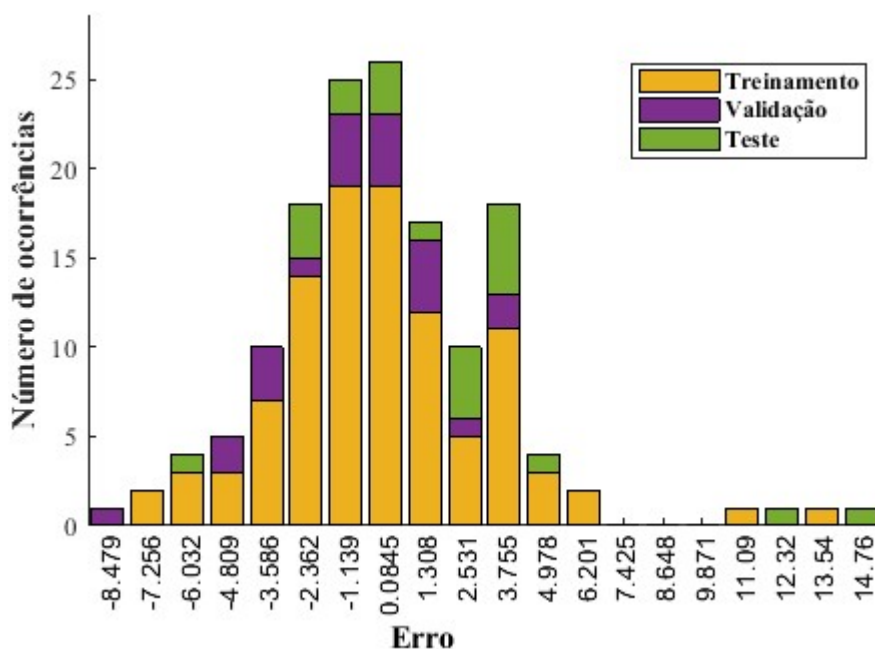
Figura 20. Performance do modelo de RNA com base no erro quadrático médio em função do número de épocas para a quantificação de material vegetal



Fonte: O autor, 2019

Os erros absolutos para o treinamento variaram de 0 a 14, os valores de validação variaram entre 0 e 9 e os valores de teste variaram entre 0 e 15. O histograma de erros pode ser observado na Figura 21, tendo apresentado comportamento de distribuição normal.

Figura 21. Diagrama de distribuição de erro para RNA de quantificação de material vegetal

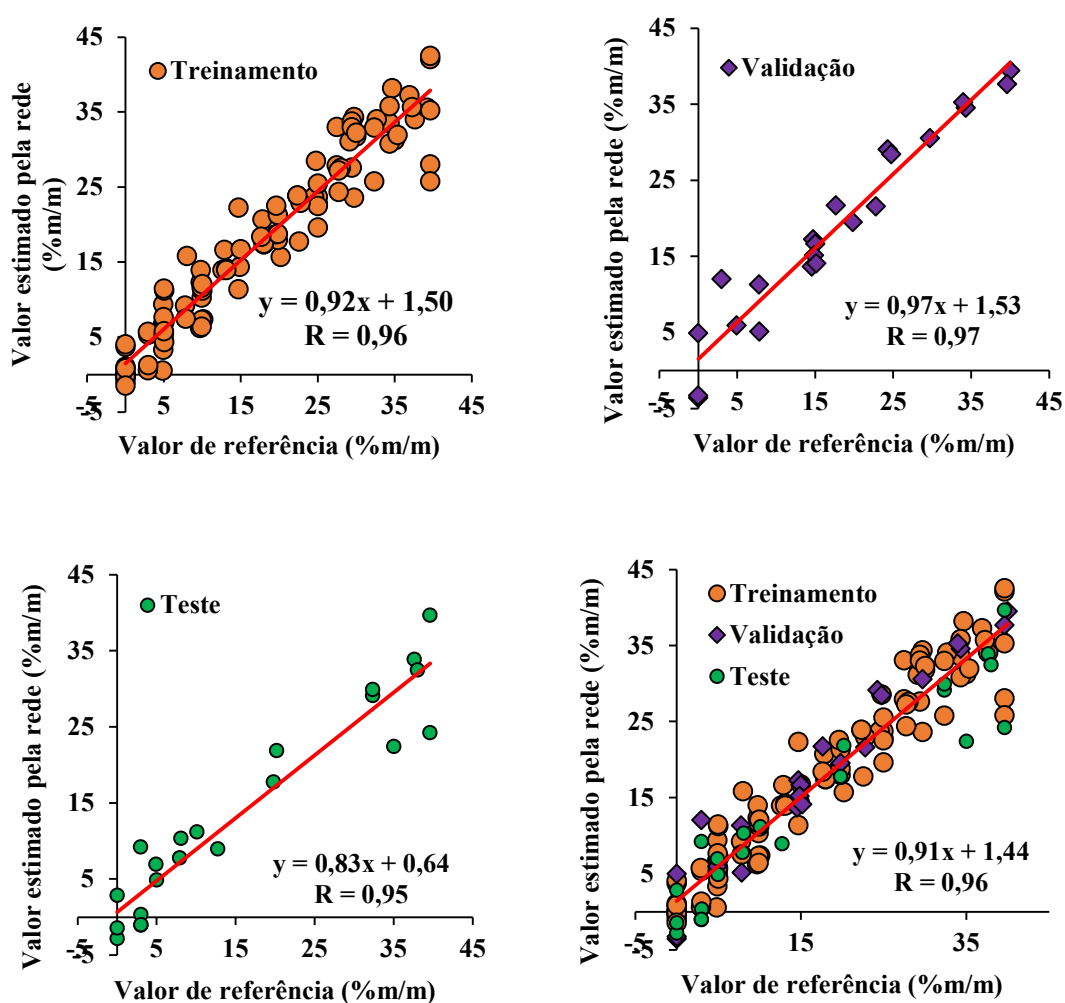


Fonte: O autor, 2020

A Figura 22 mostra a alta correlação entre os valores do conteúdo de material vegetal de referência e os valores previstos pelo modelo de RNA, sendo o intervalo real dos valores observados de porcentagem em massa de material vegetal de [0;40]. As

retas que melhor representam os resultados obtidos pelos conjuntos de treinamento, validação e teste são $y = 0,92x + 1,50$, $y = 0,97x + 1,53$ e $y = 0,83x + 0,64$, respectivamente. Já os coeficientes de correlação obtidos foram 0,96, 0,97 e 0,95 para os conjuntos de treinamento, validação e teste. Os resultados obtidos pela RNA em virtude dos altos coeficientes de correlação mostram que o modelo teve alta capacidade em relacionar a natureza multivariada das entradas da rede com a respectiva saída, gerando respostas coerentes com os dados de saída reais. Das 146 amostras, 7 estimativas feitas tiveram valores menores do que 0% em massa de material vegetal, tendência observada para as amostras em que não há presença desse tipo de impureza. Além disso, 2 estimativas da rede foram maiores do que 40%, sendo justamente amostras contendo 40% de material vegetal.

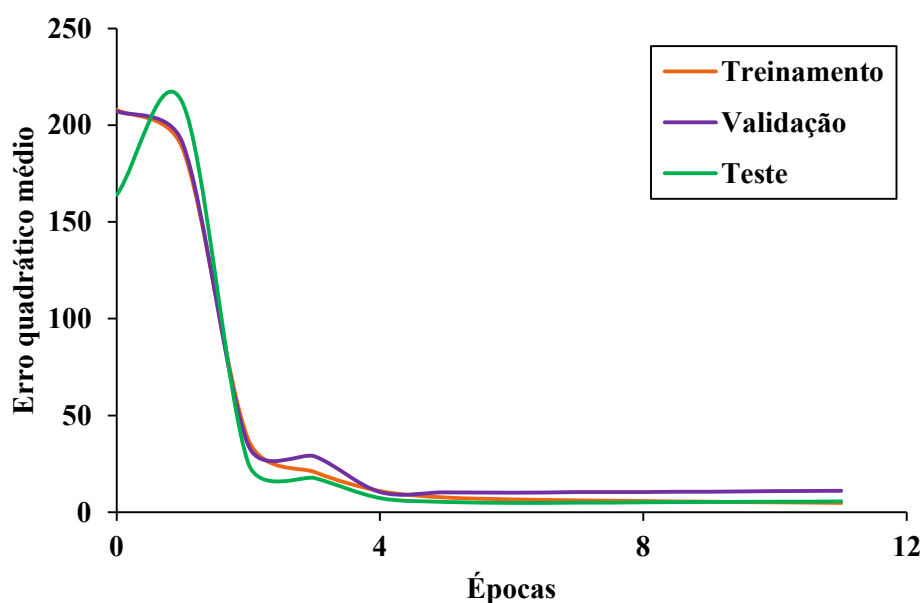
Figura 22. Porcentagem em massa de material vegetal estimada pela RNA para os conjuntos de teste, validação e treinamento



4.3 Porcentagem mássica de material mineral

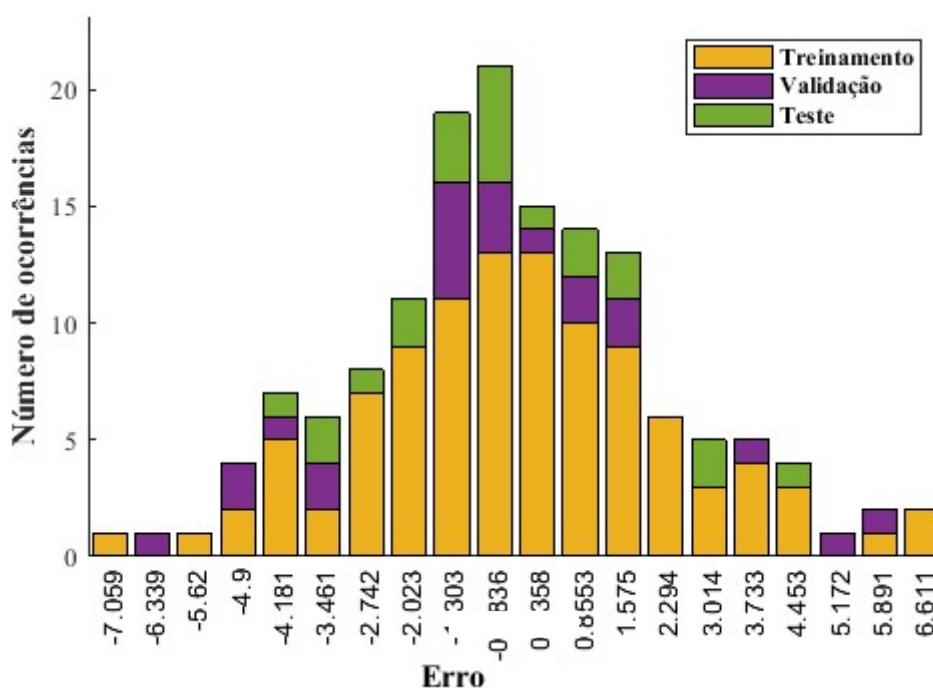
A RNA para a estimativa de material vegetal apresenta 17 neurônios na camada intermediária utilizando a função sigmoide, além de uma camada de saída utilizando uma função linear com o algoritmo de Levenberg-Marquardt. Ela obteve um treinamento muito rápido (menos de 1 minuto), obtendo o melhor resultado em 6 épocas com um erro quadrático médio (MSE) de aproximadamente 5, 7 e 10 para os conjuntos de treinamento, validação e teste, respectivamente, conforme indicado na Figura 23.

Figura 23. Performance do modelo de RNA com base no erro quadrático médio em função do número de épocas para a quantificação de material mineral



Fonte: O autor, 2020

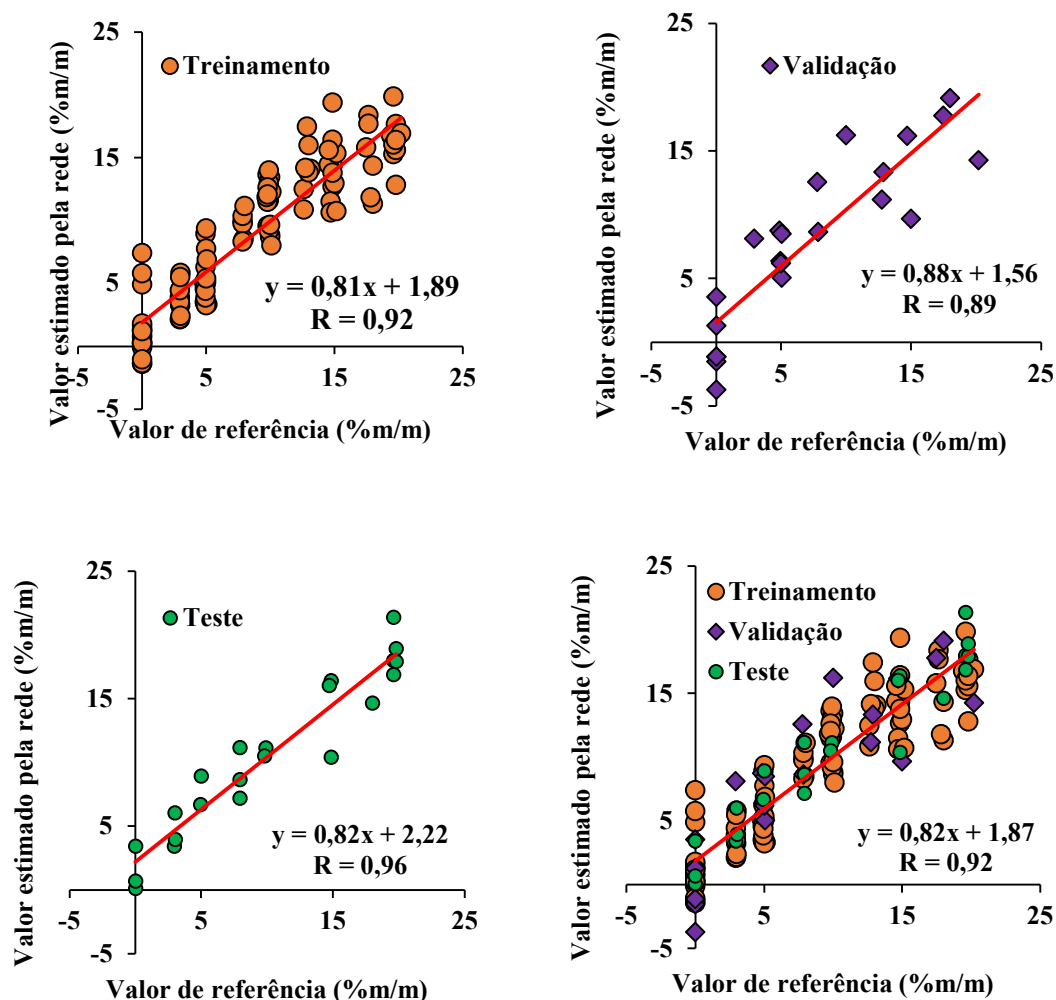
Os erros absolutos para o treinamento variaram de 0 a 7, os valores de validação variaram entre 0 e 6 e os valores de teste variaram entre 0 e 5. O histograma de erros pode ser observado na Figura 24, apresentando um comportamento de distribuição normal.

Figura 24. Diagrama de distribuição de erro para RNA de quantificação de material mineral

Fonte: O autor, 2020

A Figura 25 mostra a alta correlação entre os valores do conteúdo de material mineral de referência e os valores previstos pelo modelo de RNA, sendo o intervalo real dos valores observados de porcentagem em massa de material vegetal de [0;20]. As retas que melhor representam os resultados obtidos pelos conjuntos de treinamento, validação e teste são $y = 0,81x + 1,89$, $y = 0,88x + 1,56$ e $y = 0,82x + 2,22$, respectivamente. Já os coeficientes de correlação obtidos foram 0,92, 0,89 e 0,96 para os conjuntos de treinamento, validação e teste. Os resultados obtidos pela RNA em virtude dos altos coeficientes de correlação mostram que o modelo teve alta capacidade em relacionar a natureza multivariada das entradas da rede com a respectiva saída, gerando respostas coerentes com os dados de saída reais. Das 146 amostras, 7 estimativas feitas tiveram valores menores do que 0% em massa de material mineral, tendência observada para as amostras em que o valor real é de 0% para essa classe. Além disso, 1 estimativa da rede foi maior do que 20%, sendo justamente uma amostra contendo 20% de material mineral.

Figura 25. Porcentagem em massa de material mineral estimada pela RNA para os conjuntos de teste, validação e treinamento



Fonte: O autor, 2020

A Tabela 3 apresenta um resumo dos resultados obtidos para cada etapa dos modelos de redes neurais artificiais, considerando o coeficiente de correlação (R), raiz do erro quadrático médio (RMSE) e a inclinação da reta de ajuste (m).

Tabela 3. Resumo dos resultados obtidos pelas redes neurais artificiais para cada modelo

Classe	Treinamento			Validação			Teste			Geral		
	R	RMSE	m	R	RMSE	m	R	RMSE	m	R	RMSE	m
Caná-de-açúcar	0,98	2,86	0,96	0,93	5,47	0,89	0,91	6,70	1,08	0,97	4,12	0,97
Impureza vegetal	0,96	3,41	0,92	0,97	3,20	0,97	0,95	5,15	0,83	0,96	3,70	0,91
Impureza mineral (solo)	0,92	2,59	0,81	0,89	3,19	0,88	0,96	2,21	0,82	0,92	2,64	0,82

Fonte: O autor, 2020

De modo geral, é observado que o parâmetro RMSE decresce gradativamente da classe cana-de-açúcar até impureza mineral, ao mesmo tempo em que o intervalo de valores de porcentagem mássica de cada classe também decresce (41-100% para cana, 0-40% para impureza vegetal e 0-20% para impureza mineral). Essa diferença de intervalos e amplitude de valores pode estar diretamente associada a esses erros, já que a rede possui uma maior diversificação de saídas possíveis para a predição. Além disso, ao analisar os resultados conjuntamente, observa-se que tanto a correlação quando a inclinação da resta de ajuste apresentam valores maiores na classe cana-de-açúcar. Uma inclinação próxima de 1 indica que as respostas do modelo estão cada vez mais próximas do que seria a reta ideal $y = x$, enquanto um coeficiente de correlação próximo de 1 indica alta relação entre os valores previstos pela rede e os valores reais.

Nesse sentido, a construção de um modelo que preveja a porcentagem mássica de cana-de-açúcar pode apresentar vantagem, uma vez que conhecido esse valor, a correspondente porcentagem de impureza sólida é definida pela diferença entre 100% e o valor estimado para cana. Assim, uma vez que industrial e analiticamente a porcentagem de interesse é a de cana, o correspondente valor de impureza é automaticamente determinado e, então, a decisão de como prosseguir o processo industrial consegue ser tomada (limpeza ou continuidade do processo).

5 REDES NEURAIS ARTIFICIAIS PARA CLASSIFICAÇÃO DAS AMOSTRAS

Os neurônios foram arquitetados de acordo com a seguinte estrutura: camada de entrada contendo 10 neurônios e de saída, 2 neurônios. O número de neurônios na camada intermediária foi modificado até se determinar os melhores resultados. As melhores arquiteturas foram as com 6, 7 e 10 neurônios na camada intermediária.

5.1 Camada intermediária com 6 neurônios

A Figura 26 ilustra as matrizes de confusão geradas para os conjuntos de treinamento, validação e teste.

Figura 26. Matriz de confusão dos conjuntos de treinamento, validação e teste, bem como de todas as amostras para classe cana e RNA com 6 neurônios na camada intermediária

		Treinamento		
Classe prevista	1	28 27.5%	0 0.0%	100% 0.0%
	2	0 0.0%	74 72.5%	100% 0.0%
		100% 0.0%	100% 0.0%	100% 0.0%
		1	2	
		Classe de referência		
		Validação		
Classe prevista	1	3 13.6%	0 0.0%	100% 0.0%
	2	0 0.0%	19 86.4%	100% 0.0%
		100% 0.0%	100% 0.0%	100% 0.0%
		1	2	
		Classe de referência		
		Teste		
Classe prevista	1	5 22.7%	0 0.0%	100% 0.0%
	2	0 0.0%	17 77.3%	100% 0.0%
		100% 0.0%	100% 0.0%	100% 0.0%
		1	2	
		Classe de referência		
		Todos		
Classe prevista	1	36 24.7%	0 0.0%	100% 0.0%
	2	0 0.0%	110 75.3%	100% 0.0%
		100% 0.0%	100% 0.0%	100% 0.0%
		1	2	
		Classe de referência		

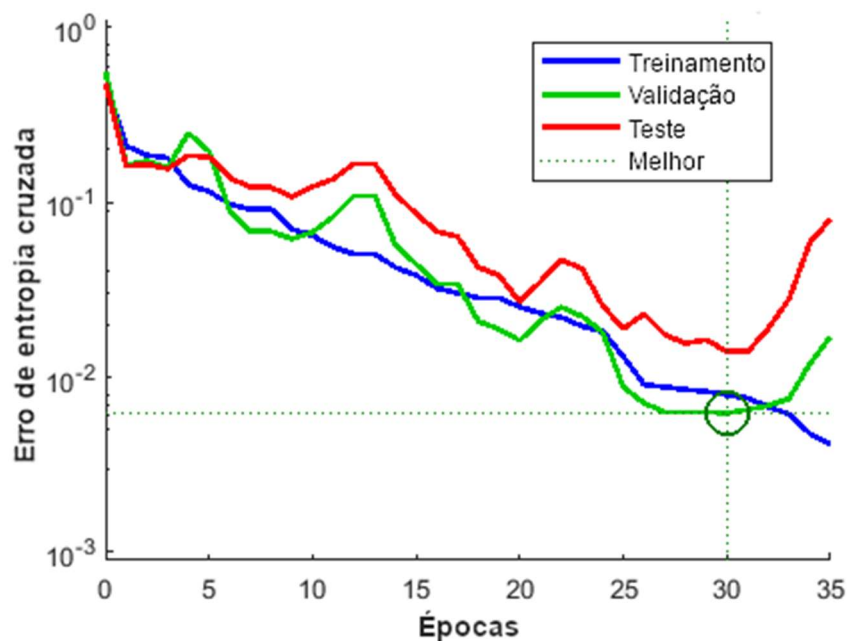
Fonte: O autor, 2020

Para o conjunto de treinamento, pode-se observar que das 28 amostras da classe 1 e das 74 amostras da classe 2 todas foram classificadas corretamente. Para o conjunto de validação, pode-se observar que das 3 amostras da classe 1, nenhuma foi classificada como classe 2; já dentre as 19 amostras da classe 2, nenhuma foi classificada erroneamente. Por fim, para o conjunto de teste, pode-se observar que das 5 amostras da classe 1 e das 17 amostras da classe 2, nenhuma foi classificada erroneamente. Logo,

das 146 amostras, nenhuma foi classificada erroneamente, representando uma taxa de acerto média de 100%.

O algoritmo de gradiente conjugado escalonado propõe a substituição do erro quadrático médio comum (MSE) pela função de erro de entropia cruzada, na qual o sinal de erro associado à camada de saída é diretamente proporcional à diferença entre os valores de saída reais e desejados. É comprovado que a função de entropia fornece bom desempenho geral da rede com períodos de estagnação relativamente curtos. Assim, a rede neural obteve melhor desempenho em 30 épocas, com erro de entropia cruzada na validação de 0,0062. A Figura 27 ilustra esse comportamento.

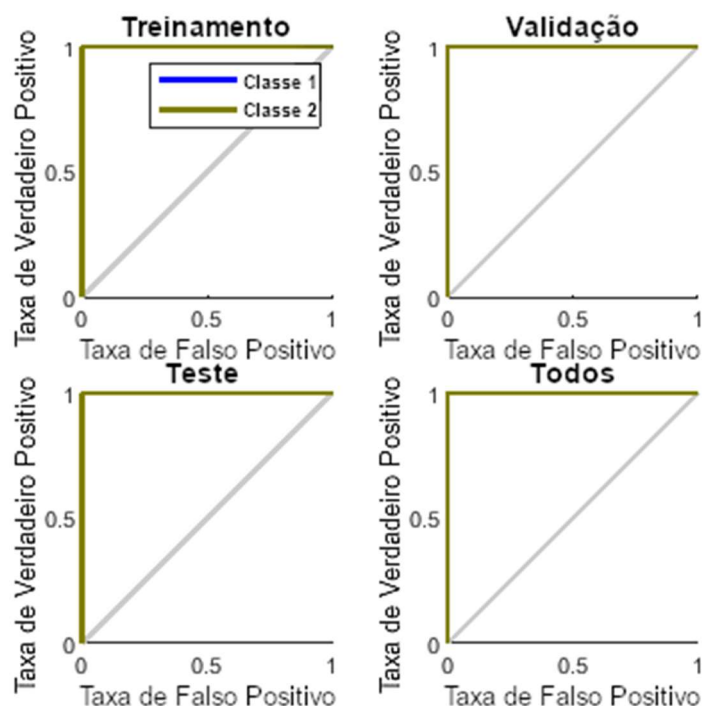
Figura 27. Desempenho da RNA em função do número de época para classe cana e RNA com 6 neurônios na camada intermediária



Fonte: O autor, 2020

Por fim, pode-se construir a curva característica de operação do receptor para a classe cana, mostrando o quão específico e sensível é o modelo. Conforme indicado pela Figura 28, a RNA obteve um ótimo desempenho (curvas no canto superior esquerdo do gráfico) na classificação de ambas as classes. Isso também pode ser reforçado pelos valores da área abaixo da curva, todos numericamente iguais a 1 para os conjuntos de treinamento, validação e teste, de modo que o modelo não apresentou nenhum falso positivo.

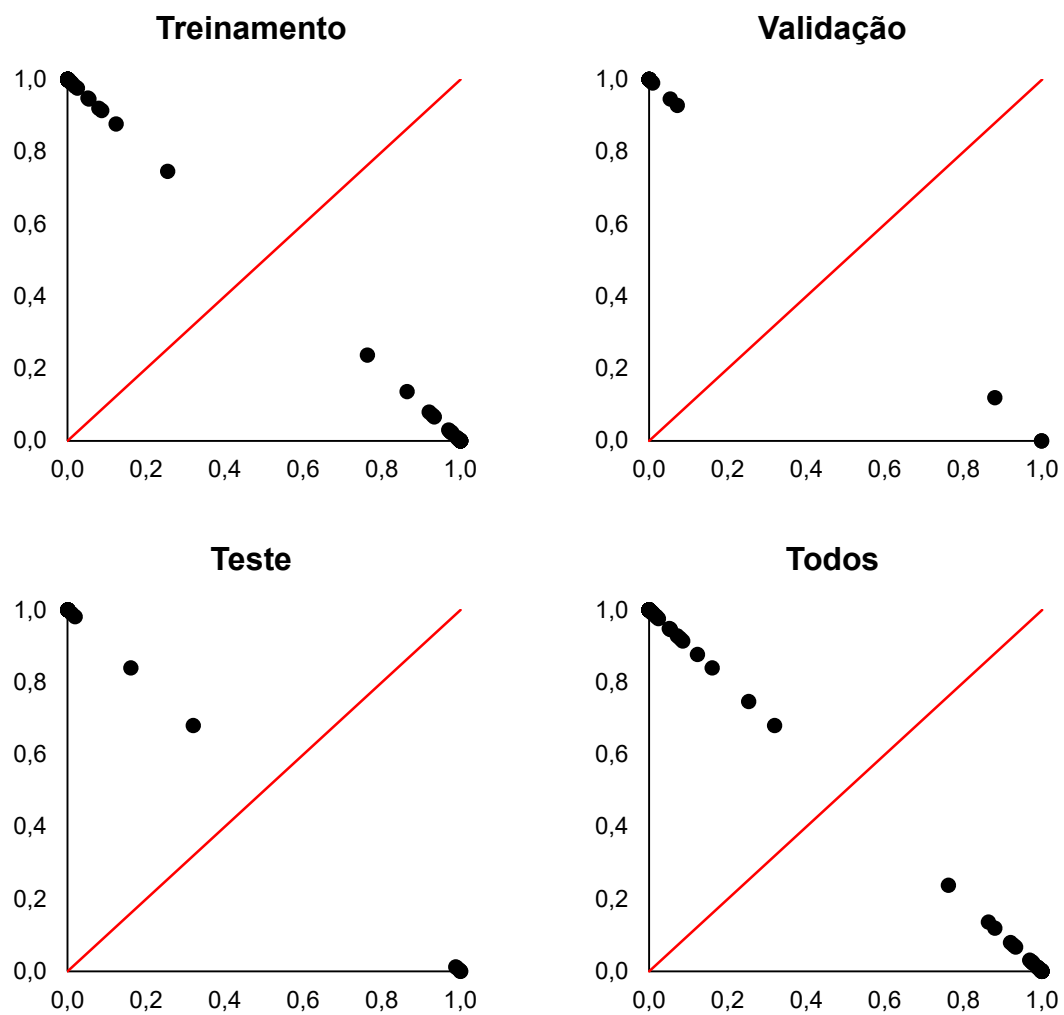
Figura 28. Curva de Característica de Operação do Receptor para classe cana e RNA com 6 neurônios na camada intermediária



Fonte: O autor, 2020

As coordenadas calculadas pela RNA estão reunidas na Figura 29, mostrando o comportamento para os conjuntos de treinamento, validação e teste, bem como o comportamento de todas as amostras. As amostras na parte superior à curva vermelha, próximas ao ponto (0,1), são as da classe 2 (40-89% em massa de cana-de-açúcar); já às próximas ao ponto (1,0), abaixo da curva, são da classe 1 (90-100% em massa de cana-de-açúcar). Note que não há pontos em destaque (vermelhos), uma vez que não houve erros de classificação. As amostras estão identificadas no Apêndice.

Figura 29. Respostas obtidas pela RNA com 6 neurônios da camada intermediária para classe cana

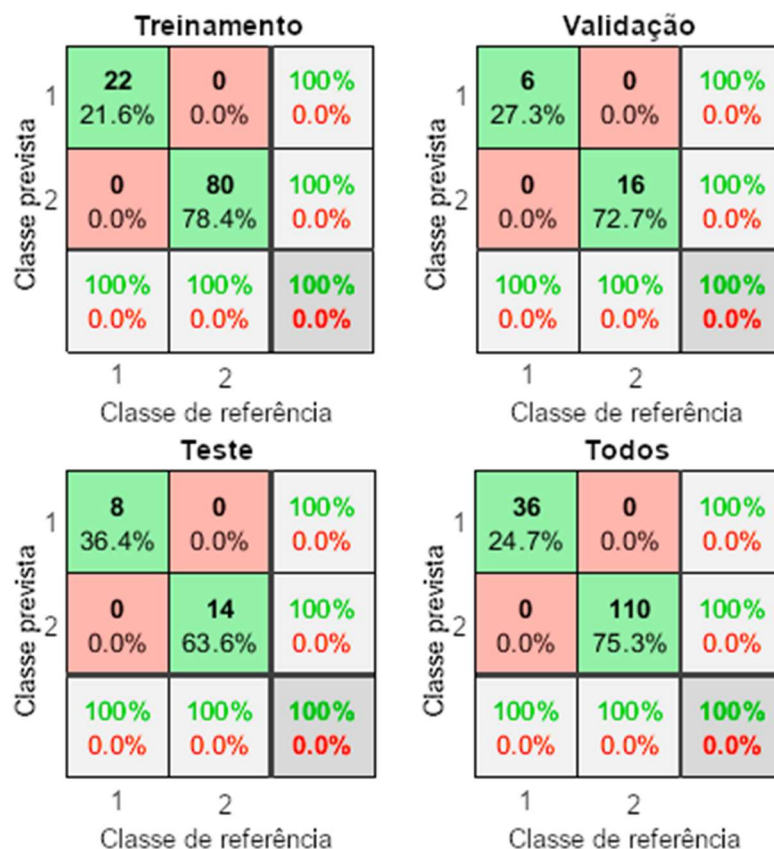


Fonte: O autor, 2020

5.2 Camada intermediária com 7 neurônios

A Figura 30 ilustra as matrizes de confusão geradas para cada os conjuntos de treinamento, validação e teste.

Figura 30. Matriz de confusão dos conjuntos de treinamento, validação e teste, bem como de todas as amostras para classe cana e RNA com 7 neurônios na camada intermediária

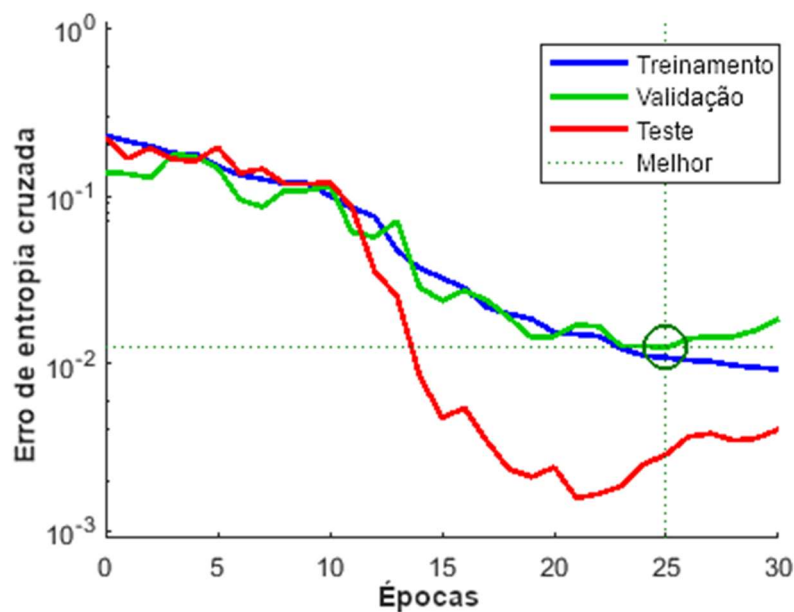


Fonte: O autor, 2020

Para o conjunto de treinamento, pode-se observar que das 22 amostras da classe 1, nenhuma foi classificada como classe 2; já dentre as 80 amostras da classe 2, nenhuma foi classificada erroneamente. Para o conjunto de validação, pode-se observar que das 6 amostras da classe 1 e das 16 amostras da classe 2, nenhuma foi classificada erroneamente. Por fim, para o conjunto de teste, pode-se observar que das 8 amostras da classe 1, nenhuma foi classificada como classe 2; já dentre as 14 amostras da classe 2, nenhuma foi classificada erroneamente. Logo, das 146 amostras, nenhuma foi classificada erroneamente, representando uma taxa de acerto média de 100%.

A rede neural obteve melhor desempenho em 25 épocas, com erro de entropia cruzada na validação de 0,0124. A Figura 31 ilustra esse comportamento.

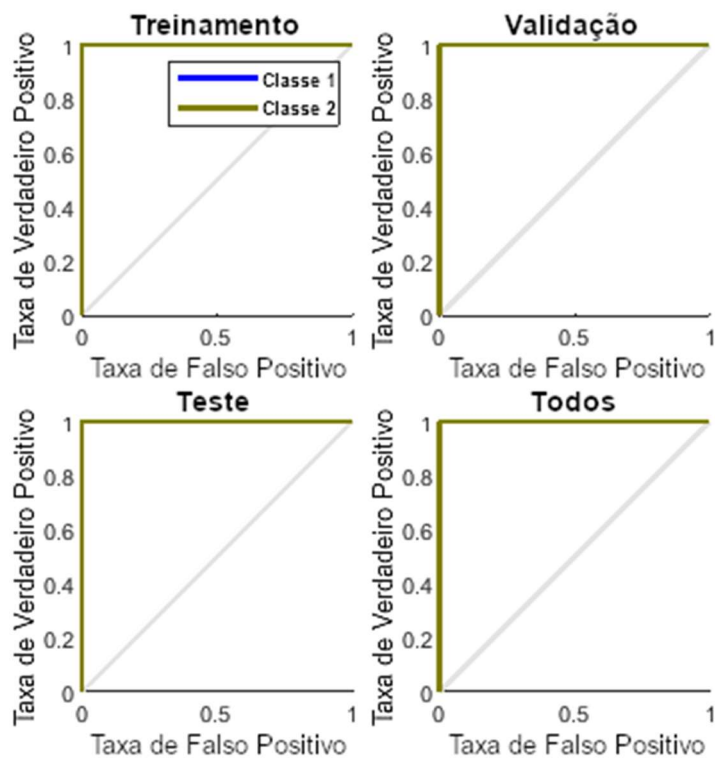
Figura 31. Desempenho da RNA em função do número de época para classe cana e RNA com 7 neurônios na camada intermediária



Fonte: O autor, 2020

Por fim, pode-se construir a curva característica de operação do receptor para a classe cana, indicada pela Figura 32.

Figura 32. Curva de Característica de Operação do Receptor para classe cana e RNA com 7 neurônios na camada intermediária

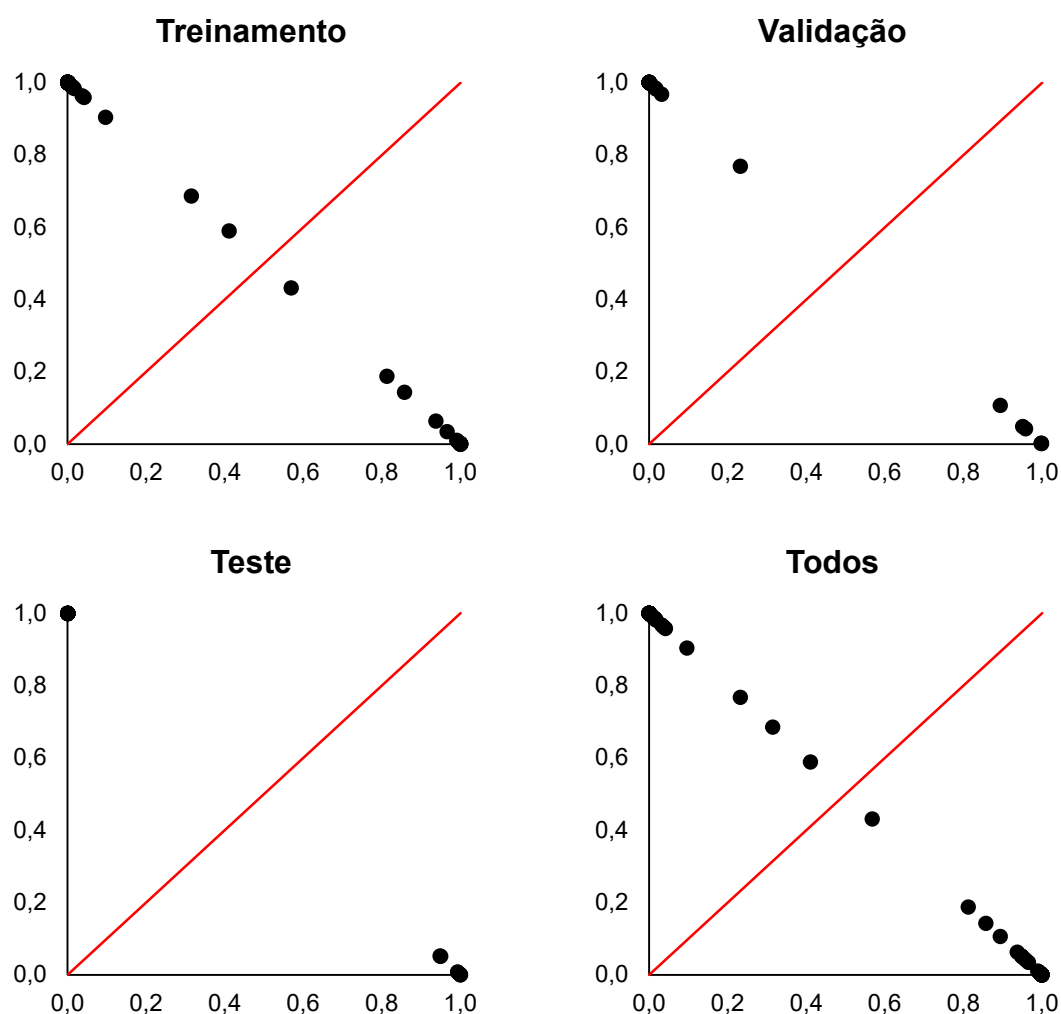


Fonte: O autor, 2020

Com base na Figura 32, a RNA obteve um ótimo desempenho (curvas no canto superior esquerdo do gráfico) na classificação de ambas as classes. Isso também pode ser reforçado pelos valores da área abaixo da curva, todos numericamente iguais a 1 para os conjuntos de treinamento, validação e teste, de modo que o modelo não apresentou nenhum falso positivo.

As coordenadas calculadas pela RNA estão reunidas na Figura 33, mostrando o comportamento para os conjuntos de treinamento, validação e teste, bem como o comportamento de todas as amostras. As amostras na parte superior à curva vermelha, próximas ao ponto (0,1), são as da classe 2 (40-89% em massa de cana-de-açúcar); já às próximas ao ponto (1,0), abaixo da curva, são da classe 1 (90-100% em massa de cana-de-açúcar). Note que não há pontos em destaque (vermelhos), uma vez que não houve erros de classificação. As amostras estão identificadas no Apêndice.

Figura 33. Respostas obtidas pela RNA com 7 neurônios da camada intermediária para classe cana



Fonte: O autor, 2020

5.3 Camada intermediária com 10 neurônios

A Figura 34 ilustra as matrizes de confusão geradas para cada os conjuntos de treinamento, validação e teste.

Figura 34. Matriz de confusão dos conjuntos de treinamento, validação e teste, bem como de todas as amostras para classe cana e RNA com 10 neurônios na camada intermediária

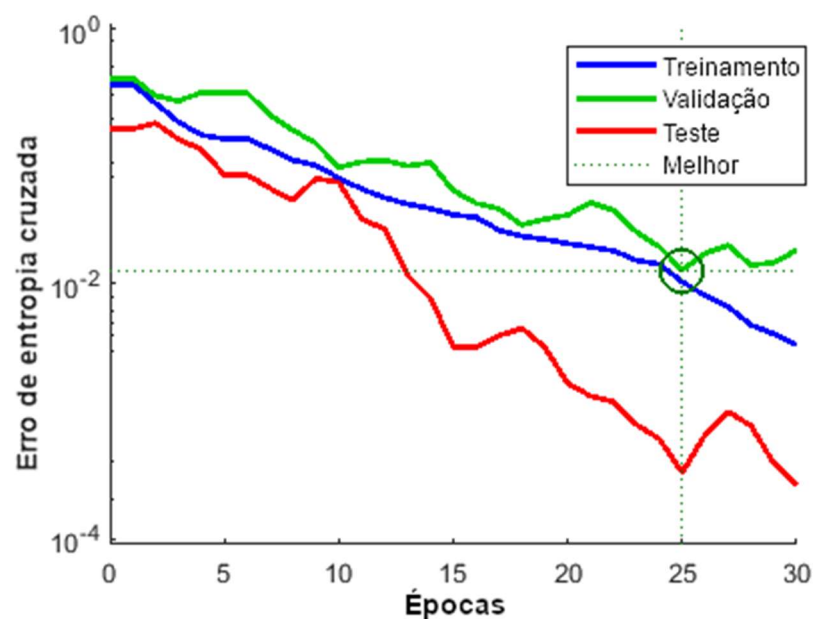
		Treinamento					Validação		
Classe prevista	1	27 26.5%	0 0.0%	100% 0.0%	Classe prevista	1	7 31.8%	0 0.0%	100% 0.0%
	2	0 0.0%	75 73.5%	100% 0.0%		2	0 0.0%	15 68.2%	100% 0.0%
		100% 0.0%	100% 0.0%	100% 0.0%			100% 0.0%	100% 0.0%	100% 0.0%
		1 2					1 2		
		Classe de referência					Classe de referência		

		Teste					Todos		
Classe prevista	1	2 9.1%	0 0.0%	100% 0.0%	Classe prevista	1	36 24.7%	0 0.0%	100% 0.0%
	2	0 0.0%	20 90.9%	100% 0.0%		2	0 0.0%	110 75.3%	100% 0.0%
		100% 0.0%	100% 0.0%	100% 0.0%			100% 0.0%	100% 0.0%	100% 0.0%
		1 2					1 2		
		Classe de referência					Classe de referência		

Fonte: O autor, 2020

Para o conjunto de treinamento, pode-se observar que das 27 amostras da classe 1 das 75 amostras da classe 2, nenhuma foi classificada erroneamente. Para o conjunto de validação, pode-se observar que das 7 amostras da classe 1, nenhuma foi classificada como classe 2; já dentre as 15 amostras da classe 2, nenhuma foi classificada erroneamente. Por fim, para o conjunto de teste, pode-se observar que das 2 amostras da classe 1 e das 20 amostras da classe 2, nenhuma foi classificada incorretamente. Logo, das 146 amostras, nenhuma foi classificada erroneamente, representando uma taxa de acerto média de 100%. Além disso, a rede neural obteve melhor desempenho em 25 épocas, com erro de entropia cruzada na validação de 0,0128. A Figura 35 ilustra esse comportamento.

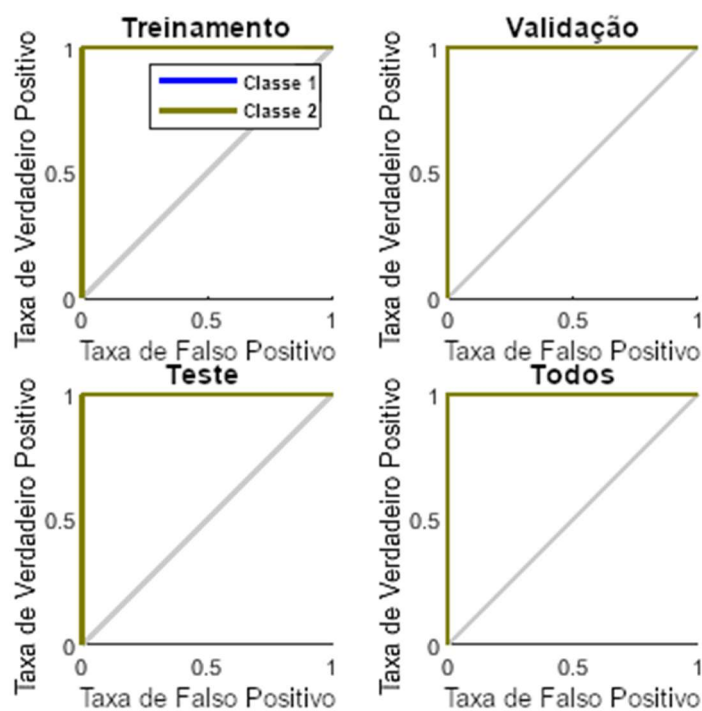
Figura 35. Desempenho da RNA em função do número de época para classe cana e RNA com 10 neurônios na camada intermediária



Fonte: O autor, 2020

Pode-se, finalmente, construir a curva característica de operação do receptor para a classe cana, indicada pela Figura 36.

Figura 36. Curva de Característica de Operação do Receptor para classe cana e RNA com 10 neurônios na camada intermediária

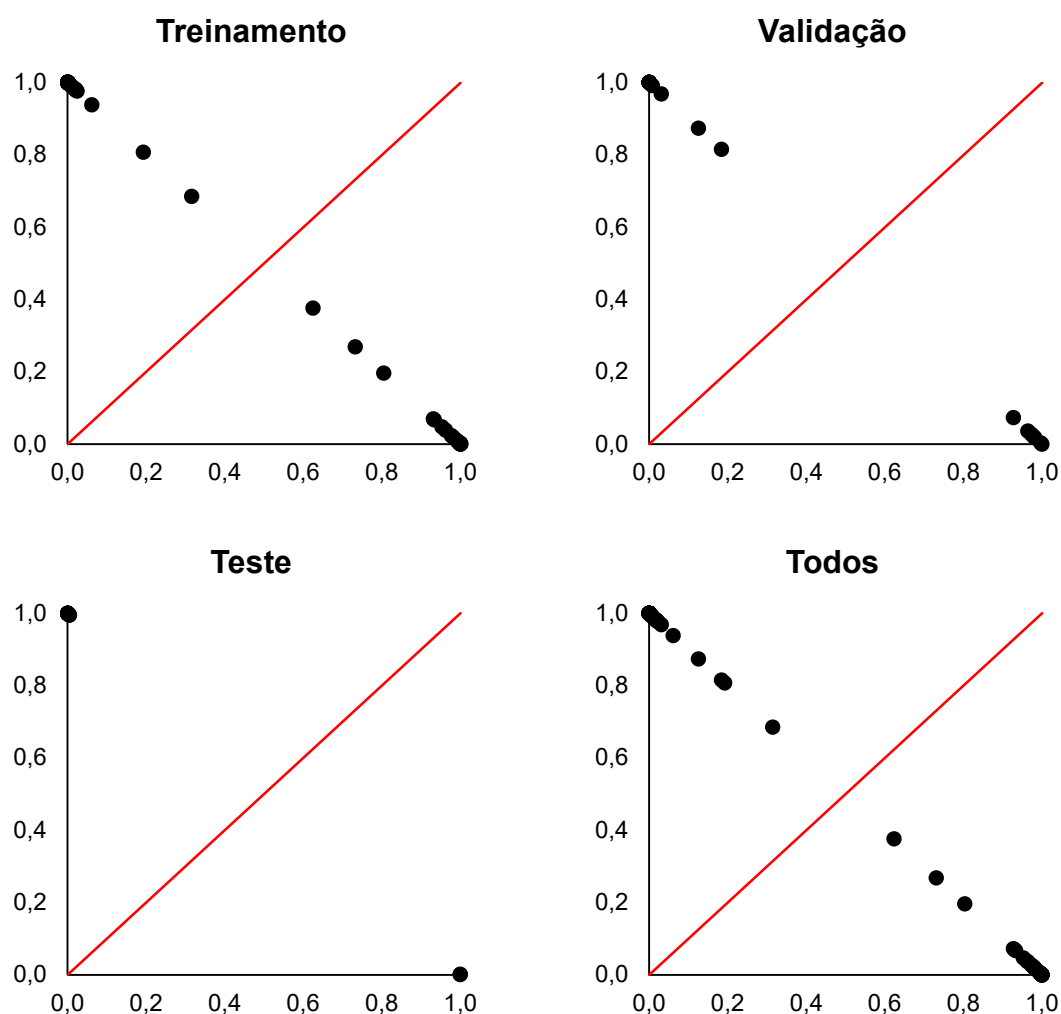


Fonte: O autor, 2020

Com base na Figura 36, a RNA obteve um ótimo desempenho (curvas no canto superior esquerdo do gráfico) na classificação de ambas as classes. Isso também pode ser reforçado pelos valores da área abaixo da curva, todos numericamente iguais a 1 para os conjuntos de treinamento, validação e teste, de modo que o modelo não apresentou nenhum falso positivo.

As coordenadas calculadas pela RNA estão reunidas na Figura 37, mostrando o comportamento para os conjuntos de treinamento, validação e teste, bem como o comportamento de todas as amostras. As amostras na parte superior à curva vermelha, próximas ao ponto (0,1), são as da classe 2 (40-89% em massa de cana-de-açúcar); já às próximas ao ponto (1,0), abaixo da curva, são da classe 1 (90-100% em massa de cana-de-açúcar). Note que não há pontos em destaque (vermelhos), uma vez que não houve erros de classificação. As amostras estão identificadas no Apêndice.

Figura 37. Respostas obtidas pela RNA com 10 neurônios da camada intermediária para classe cana



Assim, mesmo que todas as arquiteturas de RNA descritas tenham atingindo 100% de acerto para os conjuntos de treinamento, validação e teste, a rede com 6 neurônios na camada intermediária obteve o menor erro de entropia cruzada na validação (0,0062). Isso pode ser explicado pelas amostras que foram selecionadas pelo algoritmo, sendo bastante representativas para cada conjunto e gerando então respostas mais certas por parte dessa RNA. Devido a esses fatores, esse foi considerado o melhor modelo. Além disso, durante os testes foi observado que o comportamento da RNA ao aumentar a quantidade de neurônios na camada intermediária faz com que a rede não consiga generalizar o problema e classificar as amostras adequadamente, aumentando-se os erros consideravelmente. Assim, os modelos do problema proposto se limitam a menos de 15 neurônios na camada intermediária.

6 CONSIDERAÇÕES FINAIS

A partir da utilização de características de histogramas de cores como banco de dados, foi possível construir arquiteturas de redes neurais artificiais capazes de prever a quantidade de cana-de-açúcar, material vegetal e material mineral em amostras de composições diversas. Para os modelos de predição das três classes, os erros apresentaram comportamento de distribuição normal. O coeficiente de correlação considerando todas as amostras em cada classe foi 0,97, 0,96 e 0,92 para cana-de-açúcar, impureza vegetal e impureza mineral (solo), respectivamente, com a raiz do erro quadrático médio sendo igual a 4,12, 3,70 e 2,64.

Dentre todas as arquiteturas de RNAs testadas para classificar o conjunto de amostras com base na porcentagem mássica de cana-de-açúcar desejada, aquelas contendo 6, 7 e 10 neurônios na camada intermediária foram capazes de classificar corretamente todas as amostras, sendo a de 6 neurônios a que forneceu melhor resultado, com erro de entropia cruzada para o conjunto de validação de 0,0062.

O trabalho desenvolvido é pioneiro na área de imagens digitais e redes neurais artificiais. Além do seu caráter inovador na abordagem do problema de impurezas em carregamentos de cana-de-açúcar, ele também tem potencial para reduzir custos relacionados a sua análise química, principalmente por ser aplicado diretamente na amostra sem tratamento prévio. Dessa forma, o método tem potencial para ser implementado em indústrias sucroalcooleiras, servindo de norteador para decisões no processamento da cana.

Assim, a utilização de redes neurais artificiais para quantificar e classificar parâmetros de qualidade em carregamentos de cana-de-açúcar, com base em imagens digitais, mostra-se ser um processo promissor, por apresentar baixo custo computacional, rápida aquisição de dados e por ser um método não invasivo.

REFERÊNCIAS

- ACADEMY, Khan. **Visão geral da estrutura do neurônio e a sua função**. 2016. Disponível em: <https://pt.khanacademy.org/science/biology/human-biology/neuron-nervous-system/a/overview-of-neuron-structure-and-function>. Acesso em: 27 out. 2020.
- ALMEIDA, Leandro Maciel. **Uma metodologia de busca por redes neurais artificiais quase-ótimas**. 2007. 102 f. Dissertação (Mestrado) - Curso de Ciência da Computação, Universidade Federal de Pernambuco, Recife, 2007.
- ANDRADE, Daniel Fernandes; GUEDES, Wesley Nascimento; PEREIRA, Fabiola Manhas Verbi. **Detection of chemical elements related to impurities leached from raw sugarcane: Use of laser-induced breakdown spectroscopy (LIBS) and chemometrics**. Microchemical Journal, v. 137, p.443-448, mar. 2018.
- ANTONELI, G. C.; NEITZEL, I. **Aplicação de redes neurais artificiais na indústria de fios de algodão**. GEPROS. Gestão da Produção, Operações e Sistemas, Bauru, Ano 11, nº 2, abr-jun/2016, p. 1-20.
- BOVI, Roberto; SERRA, Gil Eduardo. **Impurezas fibrosas da cana-de-açúcar e parâmetros tecnológicos do caldo extraído**. Scientia Agrícola, v. 56, n. 4, p.885-896, out. 1999.
- BRAGA, Antônio de Pádua; CARVALHO, André Ponce de Leon F. de; LUDERMIR, Teresa Bernarda. **Redes Neurais Artificiais: teoria e aplicações**. Rio de Janeiro: Ltc - Livros Técnicos e Científicos Editora S.a., 2000. 263 p.
- CANAVIEIRA, CTC - Centro de Tecnologia. **Boletins do Produtor: De olho na safra**. 2013. Disponível em: <<https://ctc.com.br>>. Acesso em: 01 dez. 2019.
- CONAB. **Boletim da Safra de Cana-de-açúcar**. Disponível em: <https://www.conab.gov.br/info-agro/safras/cana/boletim-da-safra-de-cana-de-acucar>. Acesso em: 31 maio 2020.
- DOMÍNGUEZ, Marina Alejandra et al. **Geographical origin classification of Argentinean honeys using a digital image-based flow-batch system**. Microchemical Journal, Buenos Aires, v. 112, p.104-108, jan. 2014.
- FIGUEIREDO, Izabel Cristina; MACIEL, Bruno Ferreira; MARQUES, Marcos Omir. **A qualidade da cana-de-açúcar como matéria-prima para produção de álcool**. Nucleus, Jaboticabal, v. 5, Edição especial, p.82-92, jan. 2008.

FLECK, Leandro et al. **Redes neurais artificiais: princípios básicos**. Revista Eletrônica Científica Inovação e Tecnologia, Medianeira, v. 1, n. 13, p.47-57, jun. 2016.

GONÇALVES, Maria Ivanda Silva. **Aplicações de imagens digitais e análise multivariada para classificação e determinação de parâmetros de qualidade em plumas de algodão**. 2015. 110 f. Dissertação (Mestrado) - Curso de Química, Universidade Federal da Paraíba, João Pessoa, 2015.

GUEDES, Wesley Nascimento et al. **Sugarcane Stalk Content Prediction in the Presence of a Solid Impurity Using an Artificial Intelligence Method Focused on Sugar Manufacturing**. Food Analytical Methods, Araraquara, v. 13, n. 1, p. 140-144, 8 jun. 2019.

GUEDES, Wesley Nascimento. **Applicable analytical methods for the determination of solid impurities content in raw sugarcane**. 2020. 98 f. Tese (Doutorado) - Curso de Química, Universidade Estadual Paulista, Instituto de Química, Araraquara, 2020.

GUEDES, Wesley Nascimento; PEREIRA, Fabíola Manhas Verbi. **Classifying impurity ranges in raw sugarcane using laser-induced breakdown spectroscopy (LIBS) and sum fusion across a tuning parameter window**. Microchemical Journal, Araraquara, v. 143, p. 331-336, dez. 2018.

LINHARES, L. L. S.; ARAÚJO Jr, J. M.; ARAÚJO, F. M. U. (2007). **Redes neurais artificiais para identificação da fração molar de pentano na composição do GLP**. Anais do VIII Simpósio Brasileiro De Automação Inteligente (SBAI). Florianópolis – SC, Brasil.

MATHWORKS. **Makers of MATLAB & Simulink**. Disponível em: <<https://www.mathworks.com>>. Acesso em: 01 dez. 2019.

MECHATRONICSUSP. **Importância da cana-de-açúcar**. Disponível em: <http://www.mecatronica.eesc.usp.br/wiki/index.php/Importância_da_cana-de-açúcar>. Acesso em: 19 nov. 2018.

MØLLER, Martin Fodsette. **A scaled conjugate gradient algorithm for fast supervised learning**. Neural Networks, Aarhus, v. 6, n. 4, p. 525-533, jan. 1993.

NASCIMENTO, Erica Regina Filletti. **Desenvolvimento de Modelos Neurais para o Processamento de Sinais Acústicos Visando a Medição de Propriedades**

Topológicas em Escoamentos Multifásicos. 2007. 144 f. Tese (Doutorado) - Curso de Engenharia Meânica, Universidade de São Paulo, São Carlos, 2007.

OLIVEIRA, Ana P.P. de et al. **Sistemas de colheita da cana-de-açúcar: Conhecimento atual sobre modificações em atributos de solos de tabuleiro.** Revista Brasileira de Engenharia Agrícola e Ambiental, Campina Grande, v. 18, n. 9, p. 939-947, abr. 2014.

PRONI,C.; HANEDA,R.N.; FILLETTI,É.R. **Desenvolvimento de redes neurais artificiais para análise do fluxo de permeado de uma bebida à base de açaí no processo de micro-filtração tangencial.**C.Q.D.–Revista Eletrônica Paulista de Matemática, Bauru,v. 17, p. 189-205,fev. 2020.

RODRIGUES, Luciana Deotti. **A cana-de-açúcar como matéria-prima para a produção de biocombustíveis: impactos ambientais e o zoneamento agroecológico como ferramenta para mitigação.** 2010. 64 f. TCC (Graduação) - Curso de Engenharia, Universidade Federal de Juiz de Fora, Juiz de Fora, 2010.

SILVA, Ivan Nunes da et al. **Artificial Neural Networks: A Practical Course.** São Carlos: Springer, 2017. 327 p.

SOUTO, Urijatan Teixeira de Carvalho Polar et al. **Screening for Coffee Adulteration Using Digital Images and SPA-LDA.** Food Analytical Methods, João Pessoa, v. 8, n. 6, p.1515-1521, 8 nov. 2014.

SOUZA, Zigomar Menezes de et al. **Sistemas de colheita e manejo da palhada de cana-de-açúcar.** Pesquisa Agropecuária Brasileira, Jaboticabal, v. 40, n. 3, p. 271-278, mar. 2005.

UDOP. **Determinação das impurezas minerais em carregamentos de cana-de-açúcar pelo método da incineração em forno mufla.** Disponível em: <https://www.udop.com.br/download/legislacao/bioenergia/institucional_site_juridico/impurezas_minerais_cana_objetivo_equipamentos_procedimentos.pdf>. Acesso em: 01 dez. 2019.

UDOP. **Determinação das impurezas vegetais e totais em carregamentos de cana-de-açúcar pelo método de limpeza manual e à seco.** Disponível em: <https://www.udop.com.br/download/legislacao/bioenergia/institucional_site_juridico/impurezas_vegetais_totais_objetivo_equipamentos_procedimentos.pdf>. Acesso em: 01 dez. 2019.

VIEIRA, Vaninha; RAZENTE, Humberto Luiz; BARIONI, Maria Camila Nardini. **Tópicos em gerenciamento de dados e informações 2017**. Uberlândia: Sociedade Brasileira de Computação - Sbc, 2017. 94 p.

APÊNDICE

Camada intermediária com 6 neurônios

As Tabelas 4, 5 e 6 reúnem os resultados obtidos pela RNA para os conjuntos de treinamento, validação e teste, respectivamente.

Tabela 4. Respostas obtidas e esperadas pela RNA para o conjunto de treinamento e RNA com 6 neurônios na camada intermediária

Amostra	Resposta obtida pela RNA		Classificação obtida pela RNA	Resposta esperada		Classificação esperada
1	0,2538	0,7462	2	0	1	2
4	0,1232	0,8768	2	0	1	2
5	0,0001	0,9999	2	0	1	2
7	0,0001	0,9999	2	0	1	2
9	0,0003	0,9997	2	0	1	2
10	0,0000	1,0000	2	0	1	2
11	0,0000	1,0000	2	0	1	2
12	0,0000	1,0000	2	0	1	2
13	0,0000	1,0000	2	0	1	2
14	0,0513	0,9487	2	0	1	2
15	0,0000	1,0000	2	0	1	2
17	0,0000	1,0000	2	0	1	2
18	0,0000	1,0000	2	0	1	2
19	0,0000	1,0000	2	0	1	2
20	0,0542	0,9458	2	0	1	2
23	0,0005	0,9995	2	0	1	2
24	0,0000	1,0000	2	0	1	2
26	0,0000	1,0000	2	0	1	2
30	0,0241	0,9759	2	0	1	2
31	0,0000	1,0000	2	0	1	2
32	0,0000	1,0000	2	0	1	2
33	0,0000	1,0000	2	0	1	2
34	0,0000	1,0000	2	0	1	2
35	0,0000	1,0000	2	0	1	2
36	0,0000	1,0000	2	0	1	2
39	0,0000	1,0000	2	0	1	2
40	0,0000	1,0000	2	0	1	2
41	0,0000	1,0000	2	0	1	2
43	0,0000	1,0000	2	0	1	2
47	0,0000	1,0000	2	0	1	2
48	0,0000	1,0000	2	0	1	2
49	0,0000	1,0000	2	0	1	2
50	0,0000	1,0000	2	0	1	2
53	0,0000	1,0000	2	0	1	2

55	0,0000	1,0000	2	0	1	2
56	0,0000	1,0000	2	0	1	2
57	0,0000	1,0000	2	0	1	2
58	0,0000	1,0000	2	0	1	2
59	0,0000	1,0000	2	0	1	2
61	0,0000	1,0000	2	0	1	2
62	0,0000	1,0000	2	0	1	2
63	0,0000	1,0000	2	0	1	2
64	0,0000	1,0000	2	0	1	2
66	0,0000	1,0000	2	0	1	2
67	0,0000	1,0000	2	0	1	2
70	0,0000	1,0000	2	0	1	2
71	0,0001	0,9999	2	0	1	2
72	0,0018	0,9982	2	0	1	2
74	0,0000	1,0000	2	0	1	2
75	0,0001	0,9999	2	0	1	2
76	0,0000	1,0000	2	0	1	2
77	0,0013	0,9987	2	0	1	2
78	0,0182	0,9818	2	0	1	2
79	0,0000	1,0000	2	0	1	2
80	0,0000	1,0000	2	0	1	2
81	0,0000	1,0000	2	0	1	2
83	0,0000	1,0000	2	0	1	2
84	0,0073	0,9927	2	0	1	2
87	0,0026	0,9974	2	0	1	2
90	0,0001	0,9999	2	0	1	2
93	0,0027	0,9973	2	0	1	2
94	0,0000	1,0000	2	0	1	2
96	0,0000	1,0000	2	0	1	2
97	0,0000	1,0000	2	0	1	2
98	0,0004	0,9996	2	0	1	2
99	0,0093	0,9907	2	0	1	2
100	0,0058	0,9942	2	0	1	2
101	0,0106	0,9894	2	0	1	2
103	0,0006	0,9994	2	0	1	2
104	0,0033	0,9967	2	0	1	2
105	0,0792	0,9208	2	0	1	2
106	0,0017	0,9983	2	0	1	2
107	0,0012	0,9988	2	0	1	2
108	0,0860	0,9140	2	0	1	2
112	0,9775	0,0225	1	1	0	1
114	0,9201	0,0799	1	1	0	1
116	0,9998	0,0002	1	1	0	1
117	0,9739	0,0261	1	1	0	1
118	0,9769	0,0231	1	1	0	1
119	0,9991	0,0009	1	1	0	1

121	0,7622	0,2378	1	1	0	1
122	0,9979	0,0021	1	1	0	1
123	0,9693	0,0307	1	1	0	1
124	0,9930	0,0070	1	1	0	1
125	0,8637	0,1363	1	1	0	1
127	0,9996	0,0004	1	1	0	1
128	0,9997	0,0003	1	1	0	1
129	0,9969	0,0031	1	1	0	1
130	0,9995	0,0005	1	1	0	1
131	0,9996	0,0004	1	1	0	1
132	0,9977	0,0023	1	1	0	1
136	0,9978	0,0022	1	1	0	1
137	0,9972	0,0028	1	1	0	1
138	0,9898	0,0102	1	1	0	1
139	0,9964	0,0036	1	1	0	1
140	0,9995	0,0005	1	1	0	1
141	0,9738	0,0262	1	1	0	1
142	0,9948	0,0052	1	1	0	1
143	0,9277	0,0723	1	1	0	1
144	0,9332	0,0668	1	1	0	1
145	0,9995	0,0005	1	1	0	1
146	0,9997	0,0003	1	1	0	1

Fonte: O autor, 2020

Tabela 5. Respostas obtidas e esperadas pela RNA para o conjunto de validação e RNA com 6 neurônios na camada intermediária

Amostra	Resposta obtida pela RNA		Classificação obtida pela RNA	Resposta esperada		Classificação esperada
6	0,0001	0,9999	2	0	1	2
8	0,0001	0,9999	2	0	1	2
22	0,0000	1,0000	2	0	1	2
27	0,0000	1,0000	2	0	1	2
38	0,0000	1,0000	2	0	1	2
42	0,0000	1,0000	2	0	1	2
44	0,0000	1,0000	2	0	1	2
45	0,0000	1,0000	2	0	1	2
46	0,0000	1,0000	2	0	1	2
52	0,0000	1,0000	2	0	1	2
60	0,0000	1,0000	2	0	1	2
69	0,0000	1,0000	2	0	1	2
85	0,0001	0,9999	2	0	1	2
88	0,0000	1,0000	2	0	1	2
89	0,0062	0,9938	2	0	1	2
91	0,0093	0,9907	2	0	1	2
95	0,0001	0,9999	2	0	1	2
102	0,0716	0,9284	2	0	1	2

109	0,0537	0,9463	2	0	1	2
120	0,9993	0,0007	1	1	0	1
126	0,8804	0,1196	1	1	0	1
134	0,9995	0,0005	1	1	0	1

Fonte: O autor, 2020

Tabela 6. Respostas obtidas e esperadas pela RNA para o conjunto de teste e RNA com 6 neurônios na camada intermediária

Amostra	Resposta obtida pela RNA		Classificação obtida pela RNA	Resposta esperada		Classificação esperada
2	0,0001	0,9999	2	0	1	2
3	0,0006	0,9994	2	0	1	2
16	0,1604	0,8396	2	0	1	2
21	0,0000	1,0000	2	0	1	2
25	0,0000	1,0000	2	0	1	2
28	0,0001	0,9999	2	0	1	2
29	0,0000	1,0000	2	0	1	2
37	0,0000	1,0000	2	0	1	2
51	0,0000	1,0000	2	0	1	2
54	0,0000	1,0000	2	0	1	2
65	0,0001	0,9999	2	0	1	2
68	0,0000	1,0000	2	0	1	2
73	0,0191	0,9809	2	0	1	2
82	0,0001	0,9999	2	0	1	2
86	0,0007	0,9993	2	0	1	2
92	0,0097	0,9903	2	0	1	2
110	0,3197	0,6803	2	0	1	2
111	0,9876	0,0124	1	1	0	1
113	0,9904	0,0096	1	1	0	1
115	0,9993	0,0007	1	1	0	1
133	0,9968	0,0032	1	1	0	1
135	0,9998	0,0002	1	1	0	1

Fonte: O autor, 2020

Camada intermediária com 7 neurônios

As Tabelas 7, 8 e 9 reúnem os resultados obtidos pela RNA para os conjuntos de treinamento, validação e teste, respectivamente.

Tabela 7. Respostas obtidas e esperadas pela RNA para o conjunto de treinamento e RNA com 7 neurônios na camada intermediária para classe cana

Amostra	Resposta obtida pela RNA		Classificação obtida pela RNA	Resposta esperada		Classificação esperada
1	0,4107	0,5893	2	0	1	2
2	0,0000	1,0000	2	0	1	2
3	0,0066	0,9934	2	0	1	2
4	0,3144	0,6856	2	0	1	2
5	0,0000	1,0000	2	0	1	2
6	0,0000	1,0000	2	0	1	2
7	0,0000	1,0000	2	0	1	2
9	0,0000	1,0000	2	0	1	2
12	0,0000	1,0000	2	0	1	2
14	0,0962	0,9038	2	0	1	2
15	0,0000	1,0000	2	0	1	2
16	0,0007	0,9993	2	0	1	2
17	0,0000	1,0000	2	0	1	2
18	0,0000	1,0000	2	0	1	2
19	0,0000	1,0000	2	0	1	2
21	0,0000	1,0000	2	0	1	2
24	0,0000	1,0000	2	0	1	2
25	0,0000	1,0000	2	0	1	2
26	0,0000	1,0000	2	0	1	2
27	0,0000	1,0000	2	0	1	2
28	0,0000	1,0000	2	0	1	2
30	0,0122	0,9878	2	0	1	2
31	0,0000	1,0000	2	0	1	2
32	0,0000	1,0000	2	0	1	2
34	0,0000	1,0000	2	0	1	2
35	0,0000	1,0000	2	0	1	2
36	0,0000	1,0000	2	0	1	2
39	0,0000	1,0000	2	0	1	2
40	0,0000	1,0000	2	0	1	2
41	0,0000	1,0000	2	0	1	2
42	0,0000	1,0000	2	0	1	2
43	0,0000	1,0000	2	0	1	2
44	0,0000	1,0000	2	0	1	2
45	0,0000	1,0000	2	0	1	2
46	0,0000	1,0000	2	0	1	2
47	0,0000	1,0000	2	0	1	2

48	0,0000	1,0000	2	0	1	2
49	0,0000	1,0000	2	0	1	2
50	0,0000	1,0000	2	0	1	2
51	0,0000	1,0000	2	0	1	2
52	0,0000	1,0000	2	0	1	2
54	0,0000	1,0000	2	0	1	2
55	0,0000	1,0000	2	0	1	2
58	0,0000	1,0000	2	0	1	2
59	0,0000	1,0000	2	0	1	2
60	0,0000	1,0000	2	0	1	2
61	0,0000	1,0000	2	0	1	2
63	0,0000	1,0000	2	0	1	2
64	0,0000	1,0000	2	0	1	2
67	0,0000	1,0000	2	0	1	2
69	0,0000	1,0000	2	0	1	2
72	0,0003	0,9997	2	0	1	2
73	0,0011	0,9989	2	0	1	2
75	0,0003	0,9997	2	0	1	2
76	0,0000	1,0000	2	0	1	2
77	0,0000	1,0000	2	0	1	2
79	0,0000	1,0000	2	0	1	2
80	0,0000	1,0000	2	0	1	2
81	0,0000	1,0000	2	0	1	2
84	0,0004	0,9996	2	0	1	2
86	0,0000	1,0000	2	0	1	2
87	0,0005	0,9995	2	0	1	2
88	0,0000	1,0000	2	0	1	2
89	0,0011	0,9989	2	0	1	2
90	0,0000	1,0000	2	0	1	2
91	0,0024	0,9976	2	0	1	2
92	0,0002	0,9998	2	0	1	2
93	0,0140	0,9860	2	0	1	2
94	0,0000	1,0000	2	0	1	2
95	0,0000	1,0000	2	0	1	2
97	0,0000	1,0000	2	0	1	2
98	0,0000	1,0000	2	0	1	2
100	0,0005	0,9995	2	0	1	2
102	0,0166	0,9834	2	0	1	2
103	0,0000	1,0000	2	0	1	2
104	0,0009	0,9991	2	0	1	2
106	0,0030	0,9970	2	0	1	2
107	0,0003	0,9997	2	0	1	2
108	0,0417	0,9583	2	0	1	2
110	0,0370	0,9630	2	0	1	2
111	0,9988	0,0012	1	1	0	1
113	0,9925	0,0075	1	1	0	1

114	0,8574	0,1426	1	1	0	1
115	0,9987	0,0013	1	1	0	1
116	0,9654	0,0346	1	1	0	1
117	0,9999	0,0001	1	1	0	1
121	0,5684	0,4316	1	1	0	1
122	0,9986	0,0014	1	1	0	1
123	1,0000	0,0000	1	1	0	1
125	0,8124	0,1876	1	1	0	1
126	0,9896	0,0104	1	1	0	1
129	0,9998	0,0002	1	1	0	1
133	0,9989	0,0011	1	1	0	1
134	0,9991	0,0009	1	1	0	1
135	0,9999	0,0001	1	1	0	1
137	0,9988	0,0012	1	1	0	1
139	0,9968	0,0032	1	1	0	1
140	0,9975	0,0025	1	1	0	1
142	0,9370	0,0630	1	1	0	1
144	0,9999	0,0001	1	1	0	1
145	1,0000	0,0000	1	1	0	1
146	1,0000	0,0000	1	1	0	1

Fonte: O autor, 2020

Tabela 8. Respostas obtidas e esperadas pela RNA para o conjunto de validação e RNA com 7 neurônios na camada intermediária para classe cana

Amostra	Resposta obtida pela RNA		Classificação obtida pela RNA	Resposta esperada		Classificação esperada
10	0,0000	1,0000	2	0	1	2
13	0,0000	1,0000	2	0	1	2
20	0,2321	0,7679	2	0	1	2
22	0,0000	1,0000	2	0	1	2
23	0,0000	1,0000	2	0	1	2
53	0,0000	1,0000	2	0	1	2
56	0,0000	1,0000	2	0	1	2
57	0,0000	1,0000	2	0	1	2
65	0,0001	0,9999	2	0	1	2
78	0,0042	0,9958	2	0	1	2
82	0,0000	1,0000	2	0	1	2
96	0,0000	1,0000	2	0	1	2
99	0,0008	0,9992	2	0	1	2
101	0,0146	0,9854	2	0	1	2
105	0,0323	0,9677	2	0	1	2
109	0,0185	0,9815	2	0	1	2
118	0,9513	0,0487	1	1	0	1
119	0,9982	0,0018	1	1	0	1
130	0,9987	0,0013	1	1	0	1
132	0,8936	0,1064	1	1	0	1

136	0,9979	0,0021	1	1	0	1
143	0,9581	0,0419	1	1	0	1

Fonte: O autor, 2020

Tabela 9. Respostas obtidas e esperadas pela RNA para o conjunto de teste e RNA com 7 neurônios na camada intermediária para classe cana

Amostra	Resposta obtida pela RNA		Classificação obtida pela RNA	Resposta esperada		Classificação esperada
8	0,0000	1,0000	2	0	1	2
11	0,0000	1,0000	2	0	1	2
29	0,0000	1,0000	2	0	1	2
33	0,0000	1,0000	2	0	1	2
37	0,0000	1,0000	2	0	1	2
38	0,0000	1,0000	2	0	1	2
62	0,0000	1,0000	2	0	1	2
66	0,0000	1,0000	2	0	1	2
68	0,0000	1,0000	2	0	1	2
70	0,0000	1,0000	2	0	1	2
71	0,0000	1,0000	2	0	1	2
74	0,0000	1,0000	2	0	1	2
83	0,0000	1,0000	2	0	1	2
85	0,0000	1,0000	2	0	1	2
112	0,9923	0,0077	1	1	0	1
120	0,9477	0,0523	1	1	0	1
124	0,9926	0,0074	1	1	0	1
127	0,9984	0,0016	1	1	0	1
128	0,9988	0,0012	1	1	0	1
131	0,9988	0,0012	1	1	0	1
138	0,9488	0,0512	1	1	0	1
141	1,0000	0,0000	1	1	0	1

Fonte: O autor, 2020

Camada intermediária com 10 neurônios

As Tabelas 10, 11 e 12 reúnem os resultados obtidos pela RNA para os conjuntos de treinamento, validação e teste, respectivamente.

Tabela 10. Respostas obtidas e esperadas pela RNA para o conjunto de treinamento e RNA com 10 neurônios na camada intermediária para classe cana

Amostra	Resposta obtida pela RNA		Classificação obtida pela RNA	Resposta esperada		Classificação esperada
1	0,3149	0,6851	2	0	1	2
2	0,0003	0,9997	2	0	1	2
4	0,1925	0,8075	2	0	1	2
6	0,0003	0,9997	2	0	1	2
7	0,0002	0,9998	2	0	1	2
8	0,0001	0,9999	2	0	1	2
11	0,0000	1,0000	2	0	1	2
12	0,0000	1,0000	2	0	1	2
16	0,0193	0,9807	2	0	1	2
17	0,0000	1,0000	2	0	1	2
18	0,0000	1,0000	2	0	1	2
19	0,0000	1,0000	2	0	1	2
20	0,0184	0,9816	2	0	1	2
22	0,0000	1,0000	2	0	1	2
23	0,0004	0,9996	2	0	1	2
26	0,0000	1,0000	2	0	1	2
28	0,0000	1,0000	2	0	1	2
29	0,0000	1,0000	2	0	1	2
31	0,0000	1,0000	2	0	1	2
32	0,0000	1,0000	2	0	1	2
33	0,0000	1,0000	2	0	1	2
34	0,0000	1,0000	2	0	1	2
36	0,0000	1,0000	2	0	1	2
37	0,0000	1,0000	2	0	1	2
38	0,0000	1,0000	2	0	1	2
39	0,0000	1,0000	2	0	1	2
40	0,0000	1,0000	2	0	1	2
41	0,0000	1,0000	2	0	1	2
44	0,0000	1,0000	2	0	1	2
46	0,0000	1,0000	2	0	1	2
47	0,0000	1,0000	2	0	1	2
48	0,0000	1,0000	2	0	1	2
49	0,0000	1,0000	2	0	1	2
51	0,0000	1,0000	2	0	1	2
52	0,0000	1,0000	2	0	1	2
53	0,0000	1,0000	2	0	1	2

54	0,0000	1,0000	2	0	1	2
55	0,0000	1,0000	2	0	1	2
60	0,0000	1,0000	2	0	1	2
61	0,0000	1,0000	2	0	1	2
63	0,0000	1,0000	2	0	1	2
65	0,0000	1,0000	2	0	1	2
66	0,0000	1,0000	2	0	1	2
67	0,0000	1,0000	2	0	1	2
68	0,0000	1,0000	2	0	1	2
69	0,0000	1,0000	2	0	1	2
71	0,0001	0,9999	2	0	1	2
74	0,0000	1,0000	2	0	1	2
75	0,0001	0,9999	2	0	1	2
76	0,0000	1,0000	2	0	1	2
77	0,0001	0,9999	2	0	1	2
79	0,0000	1,0000	2	0	1	2
80	0,0000	1,0000	2	0	1	2
83	0,0000	1,0000	2	0	1	2
85	0,0000	1,0000	2	0	1	2
86	0,0000	1,0000	2	0	1	2
87	0,0007	0,9993	2	0	1	2
88	0,0000	1,0000	2	0	1	2
89	0,0057	0,9943	2	0	1	2
90	0,0002	0,9998	2	0	1	2
91	0,0020	0,9980	2	0	1	2
92	0,0023	0,9977	2	0	1	2
94	0,0000	1,0000	2	0	1	2
96	0,0000	1,0000	2	0	1	2
97	0,0000	1,0000	2	0	1	2
99	0,0144	0,9856	2	0	1	2
100	0,0012	0,9988	2	0	1	2
101	0,0240	0,9760	2	0	1	2
102	0,0064	0,9936	2	0	1	2
104	0,0007	0,9993	2	0	1	2
105	0,0202	0,9798	2	0	1	2
106	0,0002	0,9998	2	0	1	2
107	0,0001	0,9999	2	0	1	2
108	0,0191	0,9809	2	0	1	2
109	0,0615	0,9385	2	0	1	2
112	0,9762	0,0238	1	1	0	1
114	0,9302	0,0698	1	1	0	1
116	0,9965	0,0035	1	1	0	1
117	0,9940	0,0060	1	1	0	1
118	0,9526	0,0474	1	1	0	1
119	0,9968	0,0032	1	1	0	1
120	0,9616	0,0384	1	1	0	1

121	0,7312	0,2688	1	1	0	1
122	0,9959	0,0041	1	1	0	1
124	0,9808	0,0192	1	1	0	1
125	0,8038	0,1962	1	1	0	1
127	0,9982	0,0018	1	1	0	1
128	0,9988	0,0012	1	1	0	1
129	0,9985	0,0015	1	1	0	1
130	0,9993	0,0007	1	1	0	1
131	0,9991	0,0009	1	1	0	1
132	0,6237	0,3763	1	1	0	1
133	0,9978	0,0022	1	1	0	1
135	0,9999	0,0001	1	1	0	1
136	0,9967	0,0033	1	1	0	1
139	0,9897	0,0103	1	1	0	1
140	0,9977	0,0023	1	1	0	1
141	0,9998	0,0002	1	1	0	1
142	0,9988	0,0012	1	1	0	1
143	0,9328	0,0672	1	1	0	1
144	0,9999	0,0001	1	1	0	1
146	1,0000	0,0000	1	1	0	1

Fonte: O autor, 2020

Tabela 11. Respostas obtidas e esperadas pela RNA para o conjunto de validação e RNA com 10 neurônios na camada intermediária para classe cana

Amostra	Resposta obtida pela RNA		Classificação obtida pela RNA	Resposta esperada		Classificação esperada
10	0,0001	0,9999	2	0	1	2
14	0,1260	0,8740	2	0	1	2
15	0,0000	1,0000	2	0	1	2
21	0,0000	1,0000	2	0	1	2
24	0,0000	1,0000	2	0	1	2
27	0,0002	0,9998	2	0	1	2
30	0,0019	0,9981	2	0	1	2
35	0,0000	1,0000	2	0	1	2
45	0,0000	1,0000	2	0	1	2
58	0,0000	1,0000	2	0	1	2
59	0,0000	1,0000	2	0	1	2
78	0,0087	0,9913	2	0	1	2
84	0,0316	0,9684	2	0	1	2
98	0,0000	1,0000	2	0	1	2
110	0,1847	0,8153	2	0	1	2
111	0,9790	0,0210	1	1	0	1
113	0,9823	0,0177	1	1	0	1
123	0,9729	0,0271	1	1	0	1
126	0,9273	0,0727	1	1	0	1
137	0,9968	0,0032	1	1	0	1

138	0,9641	0,0359	1	1	0	1
145	1,0000	0,0000	1	1	0	1

Fonte: O autor, 2020

Tabela 12. Respostas obtidas e esperadas pela RNA para o conjunto de teste e RNA com 10 neurônios na camada intermediária

Amostra	Resposta obtida pela RNA		Classificação obtida pela RNA	Resposta esperada		Classificação esperada
3	0,0014	0,9986	2	0	1	2
5	0,0006	0,9994	2	0	1	2
9	0,0007	0,9993	2	0	1	2
13	0,0000	1,0000	2	0	1	2
25	0,0000	1,0000	2	0	1	2
42	0,0000	1,0000	2	0	1	2
43	0,0000	1,0000	2	0	1	2
50	0,0000	1,0000	2	0	1	2
56	0,0000	1,0000	2	0	1	2
57	0,0000	1,0000	2	0	1	2
62	0,0000	1,0000	2	0	1	2
64	0,0000	1,0000	2	0	1	2
70	0,0000	1,0000	2	0	1	2
72	0,0003	0,9997	2	0	1	2
73	0,0044	0,9956	2	0	1	2
81	0,0000	1,0000	2	0	1	2
82	0,0002	0,9998	2	0	1	2
93	0,0044	0,9956	2	0	1	2
95	0,0000	1,0000	2	0	1	2
103	0,0000	1,0000	2	0	1	2
115	0,9991	0,0009	1	1	0	1
134	0,9984	0,0016	1	1	0	1

Fonte: O autor, 2020

ANEXO

Tabela 13. Porcentagem em massa de cana-de-açúcar, impureza vegetal, solo e soma de impurezas para as 146 amostras

Amostra	Cana (%)	Impurezas vegetais (%)	Solo (%)	Soma de impurezas (%)
im1	41	40	20	59
im2	43	38	20	57
im3	42	40	18	58
im4	46	40	15	54
im5	46	34	20	54
im6	48	40	13	52
im7	48	38	15	52
im8	47	35	18	53
im9	48	33	20	52
im10	50	40	10	50
im11	50	35	15	50
im12	51	29	20	49
im13	51	39	10	49
im14	52	40	8	48
im15	52	38	10	48
im16	53	34	13	47
im17	53	32	15	47
im18	52	30	18	48
im19	53	27	20	47
im20	55	40	5	45
im21	56	34	10	44
im22	55	30	15	45
im23	55	25	20	45
im24	57	40	3	43
im25	58	37	5	42
im26	58	34	8	42
im27	58	32	10	42
im28	58	29	13	42
im29	58	27	15	42
im30	57	25	18	43
im31	58	23	20	42
im32	60	40	0	40
im33	60	35	5	40
im34	60	30	10	40
im35	60	25	15	40
im36	60	20	20	40
im37	63	37	0	37
im38	63	34	3	37
im39	63	32	5	37
im40	63	29	8	37
im41	62	28	10	38

im42	63	24	13	37
im43	63	22'	15	37
im44	62	20	18	38
im45	62	18	20	38
im46	65	35	0	35
im47	65	30	5	35
im48	65	25	10	35
im49	65	20	15	35
im50	66	15	20	34
im51	68	32	0	32
im52	68	29	3	32
im53	67	28	5	33
im54	68	25	8	32
im55	67	23	10	33
im56	67	20	13	33
im57	67	18	15	33
im58	68	15	17	32
im59	68	13	19	32
im60	70	30	0	30
im61	70	25	5	30
im62	70	20	10	30
im63	71	15	15	29
im64	70	10	20	30
im65	72	28	0	28
im66	72	25	3	28
im67	72	23	5	28
im68	72	20	8	28
im69	73	18	10	27
im70	72	15	13	28
im71	72	13	15	28
im72	73	10	17	27
im73	72	8	20	28
im74	75	25	0	25
im75	75	20	5	25
im76	75	15	10	25
im77	75	10	15	25
im78	75	5	20	25
im79	78	22	0	22
im80	77	20	3	23
im81	77	18	5	23
im82	77	15	8	23
im83	77	13	10	23
im84	77	10	13	23
im85	77	8	15	23
im86	77	5	18	23

im87	77	3	20	23
im88	80	20	0	20
im89	80	0	20	20
im90	80	15	5	20
im91	80	5	15	20
im92	80	10	10	20
im93	82	18	0	18
im94	82	15	3	18
im95	82	13	5	18
im96	82	10	8	18
im97	83	8	10	17
im98	82	5	13	18
im99	83	3	15	17
im100	82	0	18	18
im101	85	15	0	15
im102	85	0	15	15
im103	85	10	5	15
im104	85	5	10	15
im105	87	13	0	13
im106	87	10	3	13
im107	87	8	5	13
im108	87	5	8	13
im109	87	3	10	13
im110	87	0	13	13
im111r1	90	10	0	10
im111r2	90	10	0	10
im111r3	90	10	0	10
im112r1	90	0	10	10
im112r2	90	0	10	10
im112r3	90	0	10	10
im113r1	90	5	5	10
im113r2	90	5	5	10
im113r3	90	5	5	10
im114r1	92	8	0	8
im114r2	92	8	0	8
im114r3	92	8	0	8
im115r1	92	5	3	8
im115r2	92	5	3	8
im115r3	92	5	3	8
im116r1	92	3	5	8
im116r2	92	3	5	8
im116r3	92	3	5	8
im117r1	92	0	8	8
im117r2	92	0	8	8
im117r3	92	0	8	8

im118r1	95	5	0	5
im118r2	95	5	0	5
im118r3	95	5	0	5
im119r1	95	0	5	5
im119r2	95	0	5	5
im119r3	95	0	5	5
im120r1	97	3	0	3
im120r2	97	3	0	3
im120r3	97	3	0	3
im121r1	97	0	3	3
im121r2	97	0	3	3
im121r3	97	0	3	3
im122r1	100	0	0	0
im122r2	100	0	0	0
im122r3	100	0	0	0

Fonte: GUEDES, 2020

Tabela 14. Escala de cores RGB, HSV, rgb e L (luminosidade) das amostras

Amostra	R	G	B	H	S	V	r	g	b	L
im1	154,9162	147,3158	137,4333	0,2996	0,2076	0,6342	0,3610	0,3360	0,3030	439,6653
im2	143,9142	135,6698	128,1944	0,3275	0,2082	0,5889	0,3618	0,3324	0,3058	407,7784
im3	142,1436	137,1422	129,8396	0,3476	0,2241	0,5950	0,3559	0,3374	0,3067	409,1254
im4	147,2201	137,5343	129,9310	0,3486	0,2259	0,6048	0,3655	0,3316	0,3029	414,6854
im5	142,7213	133,4719	127,7583	0,3354	0,2100	0,5848	0,3631	0,3299	0,3070	403,9516
im6	141,7537	133,9533	128,1625	0,3367	0,2068	0,5842	0,3591	0,3325	0,3084	403,8695
im7	142,4059	133,5725	127,0953	0,3282	0,2040	0,5817	0,3621	0,3309	0,3070	403,0738
im8	144,9757	135,4878	129,4399	0,3210	0,2010	0,5900	0,3632	0,3301	0,3067	409,9034
im9	144,9426	136,5359	130,9436	0,3435	0,2035	0,5948	0,3608	0,3308	0,3083	412,4221
im10	139,8415	133,1789	126,2045	0,3244	0,2022	0,5754	0,3569	0,3352	0,3079	399,2249
im11	134,7919	127,1114	118,2927	0,2340	0,1997	0,5426	0,3625	0,3351	0,3023	380,1961
im12	135,7731	125,0752	115,4200	0,2046	0,2083	0,5402	0,3709	0,3327	0,2965	376,2683
im13	132,1368	123,4014	113,3971	0,2087	0,1950	0,5259	0,3668	0,3354	0,2978	368,9354
im14	153,5977	149,5033	134,6143	0,2418	0,2128	0,6316	0,3579	0,3447	0,2974	437,7154
im15	133,7081	124,8476	113,4545	0,1866	0,2044	0,5309	0,3691	0,3364	0,2945	372,0102
im16	153,0303	144,2511	132,9440	0,2343	0,2050	0,6165	0,3663	0,3356	0,2981	430,2254
im17	137,3673	128,7162	117,8291	0,2006	0,2047	0,5493	0,3670	0,3364	0,2966	383,9126
im18	135,5859	126,5002	115,4315	0,1979	0,2060	0,5404	0,3685	0,3360	0,2955	377,5176
im19	133,7813	124,2907	114,8359	0,2097	0,1971	0,5310	0,3684	0,3336	0,2980	372,9079
im20	131,2650	122,3170	118,7331	0,3973	0,1727	0,5316	0,3567	0,3284	0,3150	372,3151
im21	132,3481	123,8612	114,5928	0,2143	0,2026	0,5296	0,3662	0,3350	0,2988	370,8021
im22	133,9837	124,9560	115,6017	0,2213	0,2044	0,5365	0,3667	0,3343	0,2990	374,5413
im23	137,2276	127,1302	122,7038	0,2962	0,1581	0,5471	0,3627	0,3268	0,3105	387,0616
im24	133,4815	126,0081	115,9158	0,2052	0,1766	0,5276	0,3633	0,3371	0,2996	375,4053
im25	133,7274	125,7138	116,1206	0,2085	0,1978	0,5344	0,3648	0,3365	0,2988	375,5618
im26	136,6571	128,0091	117,5462	0,1998	0,1994	0,5435	0,3674	0,3360	0,2966	382,2125

im27	147,8637	137,3773	123,9073	0,1612	0,2291	0,5883	0,3748	0,3373	0,2879	409,1484
im28	130,0921	120,8878	110,0609	0,2057	0,2167	0,5189	0,3710	0,3355	0,2936	361,0407
im29	132,2895	122,9904	110,7746	0,1683	0,2155	0,5238	0,3725	0,3363	0,2912	366,0546
im30	133,7281	122,3881	118,1713	0,3402	0,1759	0,5338	0,3643	0,3262	0,3095	374,2875
im31	136,4861	124,2925	112,7971	0,1483	0,2400	0,5390	0,3815	0,3332	0,2853	373,5756
im32	135,0374	126,8202	115,8552	0,1795	0,2028	0,5362	0,3682	0,3377	0,2941	377,7128
im33	137,2472	127,0173	112,4503	0,1817	0,2459	0,5454	0,3766	0,3389	0,2845	376,7148
im34	139,7504	129,0128	116,6198	0,1662	0,2251	0,5546	0,3742	0,3360	0,2898	385,3829
im35	134,8932	124,4210	112,7983	0,1920	0,2324	0,5374	0,3749	0,3354	0,2898	372,1126
im36	137,5033	124,6736	110,5419	0,1415	0,2567	0,5444	0,3830	0,3347	0,2823	372,7188
im37	135,7987	128,0999	114,2750	0,1782	0,2156	0,5404	0,3682	0,3410	0,2908	378,1736
im38	130,4821	121,8833	111,8520	0,1836	0,1977	0,5176	0,3674	0,3360	0,2966	364,2174
im39	131,1324	122,0722	108,4161	0,1564	0,2336	0,5203	0,3744	0,3393	0,2863	361,6207
im40	138,6091	129,7900	118,4099	0,1876	0,2018	0,5507	0,3684	0,3364	0,2951	386,8090
im41	136,8966	126,7551	115,5199	0,1776	0,2087	0,5415	0,3716	0,3349	0,2936	379,1716
im42	142,0646	132,0870	119,5684	0,1620	0,2137	0,5639	0,3712	0,3363	0,2925	393,7201
im43	135,7736	124,8374	113,4045	0,1659	0,2167	0,5369	0,3738	0,3344	0,2918	374,0156
im44	134,3983	123,0990	109,9123	0,1823	0,2482	0,5348	0,3795	0,3354	0,2851	367,4096
im45	141,6037	129,9951	118,7115	0,1658	0,2206	0,5622	0,3741	0,3337	0,2922	390,3102
im46	134,9896	126,1143	113,1001	0,1522	0,2173	0,5338	0,3712	0,3396	0,2892	374,2041
im47	133,0380	123,8715	112,5969	0,1617	0,2069	0,5268	0,3702	0,3365	0,2933	369,5064
im48	141,2476	130,8125	117,1586	0,1678	0,2317	0,5623	0,3741	0,3372	0,2887	389,2187
im49	136,6388	125,3214	111,8345	0,1691	0,2523	0,5456	0,3787	0,3364	0,2849	373,7946
im50	139,8086	127,7790	116,4395	0,1685	0,2333	0,5555	0,3776	0,3329	0,2896	384,0271
im51	133,9439	125,1297	111,6875	0,1660	0,2240	0,5315	0,3719	0,3400	0,2881	370,7611
im52	132,6112	124,4448	112,1332	0,1720	0,2106	0,5277	0,3683	0,3389	0,2929	369,1892
im53	137,6138	127,6544	113,6257	0,1652	0,2368	0,5479	0,3748	0,3386	0,2866	378,8939
im54	139,2595	128,9723	115,8680	0,1705	0,2302	0,5545	0,3735	0,3372	0,2893	384,0998
im55	135,7450	125,6392	112,6320	0,1755	0,2415	0,5429	0,3755	0,3373	0,2872	374,0163
im56	144,2518	132,9057	120,9410	0,1614	0,2228	0,5732	0,3740	0,3345	0,2915	398,0985
im57	140,9486	129,6416	114,2620	0,1711	0,2593	0,5620	0,3808	0,3373	0,2819	384,8522
im58	145,2738	132,5291	119,2314	0,1733	0,2496	0,5778	0,3811	0,3340	0,2850	397,0344
im59	142,3061	129,3576	117,5169	0,1650	0,2412	0,5647	0,3798	0,3326	0,2876	389,1807
im60	139,1530	129,9082	116,7661	0,1614	0,2130	0,5518	0,3701	0,3382	0,2917	385,8272
im61	130,5405	120,4794	107,2397	0,1597	0,2379	0,5170	0,3765	0,3373	0,2861	358,2596
im62	140,5327	129,8436	115,5325	0,1855	0,2505	0,5595	0,3790	0,3374	0,2836	385,9088
im63	144,2189	132,8497	121,3658	0,1711	0,2148	0,5716	0,3734	0,3333	0,2933	398,4344
im64	144,2110	130,0982	116,7276	0,1661	0,2547	0,5712	0,3843	0,3318	0,2839	391,0368
im65	127,4040	118,8092	105,5741	0,1557	0,2338	0,5039	0,3749	0,3401	0,2850	351,7874
im66	132,2127	122,6599	109,3696	0,1580	0,2305	0,5225	0,3751	0,3383	0,2866	364,2422
im67	136,5281	126,7424	111,0452	0,1631	0,2548	0,5447	0,3769	0,3408	0,2823	374,3156
im68	140,6056	129,7097	114,7923	0,1643	0,2476	0,5589	0,3775	0,3381	0,2843	385,1076
im69	135,5704	125,0430	112,2854	0,1631	0,2385	0,5379	0,3771	0,3366	0,2863	372,8988
im70	133,5887	120,9867	107,9955	0,1418	0,2428	0,5292	0,3801	0,3331	0,2868	362,5709
im71	132,2544	119,7586	103,5424	0,1564	0,2843	0,5271	0,3872	0,3365	0,2764	355,5555

im72	145,3699	128,4540	110,3416	0,1608	0,3100	0,5798	0,3949	0,3329	0,2722	384,1654
im73	137,8238	121,3273	104,4782	0,1625	0,3169	0,5494	0,3975	0,3324	0,2701	363,6293
im74	129,8374	121,8600	105,0031	0,1519	0,2417	0,5148	0,3733	0,3437	0,2829	356,7006
im75	127,5698	118,4710	100,9961	0,1708	0,2797	0,5137	0,3801	0,3417	0,2782	347,0370
im76	136,5226	124,2452	109,5821	0,1552	0,2585	0,5429	0,3823	0,3348	0,2829	370,3499
im77	140,4436	124,1389	108,9490	0,1678	0,2947	0,5605	0,3920	0,3312	0,2768	373,5315
im78	139,3060	121,8514	105,5136	0,1713	0,3220	0,5571	0,3989	0,3307	0,2704	366,6710
im79	133,5800	125,5701	106,3287	0,1725	0,2675	0,5370	0,3750	0,3452	0,2798	365,4789
im80	136,1831	126,4587	110,6531	0,1726	0,2476	0,5440	0,3751	0,3400	0,2849	373,2948
im81	135,5112	124,0874	109,0091	0,1590	0,2535	0,5397	0,3792	0,3366	0,2842	368,6077
im82	132,0582	119,5187	104,6085	0,1630	0,2797	0,5273	0,3857	0,3358	0,2785	356,1855
im83	134,4214	122,3842	109,3264	0,1602	0,2505	0,5352	0,3803	0,3343	0,2855	366,1320
im84	131,9121	116,7864	103,6720	0,1692	0,2872	0,5260	0,3909	0,3303	0,2787	352,3705
im85	141,7775	126,1024	112,8214	0,1660	0,2737	0,5652	0,3880	0,3299	0,2821	380,7013
im86	146,3562	129,0456	115,4478	0,1668	0,2811	0,5826	0,3912	0,3287	0,2802	390,8496
im87	148,3057	128,2641	112,7860	0,1739	0,3155	0,5912	0,4010	0,3268	0,2722	389,3558
im88	131,8215	122,5687	103,5146	0,1683	0,2848	0,5304	0,3786	0,3445	0,2769	357,9048
im89	148,6331	123,7767	108,7123	0,1810	0,3426	0,5938	0,4096	0,3217	0,2687	381,1221
im90	129,6617	118,7963	103,3459	0,1711	0,2739	0,5199	0,3817	0,3378	0,2804	351,8038
im91	141,2495	124,4372	107,5377	0,1825	0,3136	0,5635	0,3975	0,3315	0,2710	373,2244
im92	135,2224	121,4928	105,1003	0,1812	0,3022	0,5419	0,3914	0,3347	0,2739	361,8156
im93	139,1427	128,0304	106,0148	0,1994	0,3097	0,5598	0,3848	0,3455	0,2696	373,1879
im94	137,0988	125,3815	108,3918	0,1837	0,2809	0,5513	0,3811	0,3393	0,2797	370,8720
im95	143,3158	131,6494	114,4290	0,1873	0,2684	0,5735	0,3800	0,3386	0,2814	389,3942
im96	143,7874	130,7689	115,1779	0,1606	0,2578	0,5735	0,3803	0,3352	0,2846	389,7342
im97	141,6943	128,6233	114,4972	0,1759	0,2591	0,5662	0,3809	0,3340	0,2851	384,8147
im98	147,8281	130,8376	117,7495	0,1784	0,2737	0,5917	0,3866	0,3289	0,2845	396,4152
im99	144,2431	123,1932	109,1474	0,1763	0,3213	0,5760	0,4031	0,3243	0,2726	376,5837
im100	153,8407	129,1677	115,2642	0,1969	0,3246	0,6126	0,4069	0,3217	0,2714	398,2726
im101	143,7567	131,1785	110,4950	0,2114	0,3019	0,5762	0,3853	0,3422	0,2725	385,4302
im102	148,0845	123,7536	104,9883	0,1901	0,3652	0,5870	0,4180	0,3250	0,2569	376,8264
im103	140,9190	127,3986	110,8139	0,2054	0,2867	0,5653	0,3850	0,3356	0,2794	379,1315
im104	144,3411	128,1037	111,9275	0,1937	0,3045	0,5770	0,3946	0,3317	0,2737	384,3723
im105	142,2087	127,9195	104,6605	0,2157	0,3330	0,5670	0,3944	0,3433	0,2623	374,7887
im106	137,1404	125,5266	106,9760	0,2092	0,2968	0,5512	0,3846	0,3407	0,2747	369,6430
im107	139,8275	126,9561	108,5774	0,2006	0,2944	0,5607	0,3861	0,3379	0,2760	375,3609
im108	144,2467	125,3429	108,2961	0,2236	0,3245	0,5742	0,4026	0,3297	0,2677	377,8856
im109	144,1479	125,0182	110,5391	0,2241	0,3162	0,5762	0,4011	0,3268	0,2721	379,7052
im110	149,5330	123,5503	108,3026	0,2237	0,3550	0,5948	0,4179	0,3202	0,2620	381,3859
im111r1	148,8958	135,9641	111,0686	0,2676	0,3159	0,5912	0,3903	0,3467	0,2630	395,9285
im111r2	152,9088	144,9436	132,5582	0,2688	0,2471	0,6326	0,3645	0,3394	0,2961	430,4106
im111r3	151,7046	141,4206	125,7487	0,2422	0,2630	0,6214	0,3707	0,3389	0,2904	418,8739
im112r1	146,3378	123,2537	108,1758	0,2404	0,3338	0,5793	0,4138	0,3219	0,2643	377,7674
im112r2	161,8149	147,2444	138,4752	0,2997	0,2479	0,6551	0,3792	0,3277	0,2931	447,5345
im112r3	157,6537	140,8170	128,4025	0,2744	0,2839	0,6363	0,3884	0,3289	0,2827	426,8732

im113r1	135,3249	120,0027	102,9395	0,2393	0,3228	0,5406	0,3983	0,3339	0,2678	358,2671
im113r2	156,0455	145,8101	136,0668	0,2881	0,2270	0,6376	0,3666	0,3333	0,3001	437,9224
im113r3	154,7906	143,6840	130,2258	0,2666	0,2528	0,6296	0,3728	0,3364	0,2909	428,7003
im114r1	144,3850	131,9722	111,9599	0,2909	0,2877	0,5705	0,3901	0,3420	0,2679	388,3171
im114r2	154,1307	146,2904	135,4872	0,2667	0,2315	0,6366	0,3621	0,3375	0,3004	435,9083
im114r3	149,8524	139,9670	124,5954	0,2723	0,2700	0,6135	0,3726	0,3401	0,2873	414,4149
im115r1	134,8769	119,9375	100,2725	0,2319	0,3305	0,5383	0,3987	0,3374	0,2639	355,0869
im115r2	152,8408	142,8675	130,8374	0,2766	0,2483	0,6284	0,3680	0,3363	0,2956	426,5457
im115r3	150,3544	139,8078	129,7947	0,2910	0,2393	0,6151	0,3696	0,3336	0,2968	419,9568
im116r1	139,9199	122,2909	106,1603	0,2383	0,3143	0,5561	0,4004	0,3299	0,2696	368,3712
im116r2	154,8621	142,5458	131,2583	0,2954	0,2576	0,6304	0,3764	0,3326	0,2910	428,6663
im116r3	154,8885	142,2292	126,9792	0,2661	0,2781	0,6298	0,3792	0,3367	0,2841	424,0970
im117r1	143,8996	122,7863	107,5321	0,2743	0,3269	0,5682	0,4121	0,3238	0,2641	374,2180
im117r2	164,6968	147,9004	136,3058	0,2789	0,2617	0,6635	0,3838	0,3283	0,2880	448,9029
im117r3	161,0871	146,2028	135,0627	0,2798	0,2581	0,6519	0,3802	0,3302	0,2896	442,3526
im118r1	141,0059	127,6304	103,4204	0,2452	0,3319	0,5567	0,3981	0,3461	0,2559	372,0568
im118r2	155,8899	145,7869	134,3856	0,2972	0,2518	0,6416	0,3682	0,3362	0,2956	436,0624
im118r3	152,1441	141,3084	126,6567	0,2914	0,2633	0,6180	0,3743	0,3388	0,2870	420,1092
im119r1	144,3134	126,5520	116,3136	0,3514	0,2627	0,5678	0,3972	0,3234	0,2794	387,1790
im119r2	157,5026	144,8891	138,6602	0,3451	0,2248	0,6406	0,3719	0,3274	0,3008	441,0520
im119r3	153,0945	141,3743	134,8022	0,3833	0,2166	0,6191	0,3705	0,3298	0,2997	429,2709
im120r1	142,2774	131,2786	115,3430	0,3064	0,2506	0,5626	0,3814	0,3403	0,2783	388,8990
im120r2	151,8265	141,4726	132,1864	0,3165	0,2464	0,6234	0,3697	0,3334	0,2970	425,4856
im120r3	151,1627	138,9123	128,0794	0,3307	0,2557	0,6144	0,3755	0,3331	0,2913	418,1543
im121r1	139,4235	126,2120	117,7074	0,4135	0,2109	0,5492	0,3797	0,3280	0,2923	383,3429
im121r2	144,9966	135,4925	133,0509	0,4605	0,1837	0,5883	0,3609	0,3277	0,3115	413,5400
im121r3	149,0220	138,9622	134,7094	0,4206	0,1862	0,6013	0,3639	0,3289	0,3072	422,6936
im122r1	133,6426	130,1842	124,7690	0,4522	0,0934	0,5258	0,3490	0,3371	0,3139	388,5959
im122r2	151,3383	145,6791	143,7907	0,6342	0,1039	0,5996	0,3484	0,3330	0,3186	440,8081
im122r3	146,5738	140,1647	138,6752	0,6699	0,1092	0,5805	0,3492	0,3318	0,3190	425,4138

Fonte: GUEDES, 2020