



HÉLIO JÚNIOR ALVARENGA GODINHO

**PREDIÇÃO DA ATIVIDADE ANTIOXIDANTE DE FRUTAS ATRÁVES DO USO DE
REDE NEURAL ARTIFICIAL**

LAVRAS – MG

2022

HÉLIO JÚNIOR ALVARENGA GODINHO

**PREDIÇÃO DA ATIVIDADE ANTIOXIDANTE DE FRUTAS ATRÁVES DO USO DE
REDE NEURAL ARTIFICIAL**

Monografia apresentada ao Departamento de
Ciência dos Alimentos da Universidade Federal de
Lavras como parte das exigências do Curso de
Engenharia de Alimentos, para obtenção do título de
Bacharel em Engenharia de Alimentos.

Orientadora

Elisângela Elena Nunes Carvalho

Coorientador

Mário Sérgio Lorenço

LAVRAS – MG

2022

PREDIÇÃO DA ATIVIDADE ANTIOXIDANTE DE FRUTAS ATRÁVES DO USO DE REDE NEURAL ARTIFICIAL

Monografia apresentada ao Departamento de Ciência dos Alimentos da Universidade Federal de Lavras como parte das exigências do Curso de Engenharia de Alimentos, para obtenção do título de Bacharel em Engenharia de Alimentos.

Aprovado em 22/04/2022

Banca: Elisângela Elena Nunes Carvalho

Mário Sérgio Lorenço

Danilo José Machado de Abreu

Orientadora

Elisângela Elena Nunes Carvalho

Coorientador

Mário Sérgio Lorenço

LAVRAS – MG

2022

AGRADECIMENTOS

Primeiramente eu agradeço a Deus, por ter me dado força e proteção por ter chegado até aqui.

Aos meus pais, Maria Ágda Resende Alvarenga e Élio de Almeida Godinho, por terem acreditado na minha capacidade e terem me dado todo o apoio que precisei.

Aos meus amigos que tiveram comigo nessa trajetória, em especial ao Felipe que me acompanhou por todo esse caminho.

A Universidade Federal de Lavras, pela oportunidade de ingresso na instituição.

A minha orientadora, Elisângela Elena Nunes Carvalho, pela oportunidade e confiança em ter aceito meu convite.

Ao meu coorientador, Mário Sérgio Lorenço, pela atenção, ajuda e paciência em todo o processo.

Minha imensa gratidão a todos!

RESUMO

Os compostos antioxidantes são conhecidos por sua alta capacidade antioxidante. Em frutas essas substâncias agem como forma de proteção contra ataques de agentes externos e podem ser usadas a favor da saúde humana. A busca por alimentos funcionais e bem-estar tem crescido dia após dia, e como forma de minimizar processos, a tecnologia pode ser promissora nessa busca. O presente trabalho, teve como objetivo realizar a predição de atividade antioxidante em diferentes frutas utilizando rede neural artificial. As redes neurais artificiais são modelos que possuem unidades de processamento simples, conhecidos como neurônios artificiais, que trabalham com funções matemáticas simulando os neurônios biológicos. O banco de dados do trabalho foi composto por dez artigos, onde foram coletados dados como o perfil fenólico e atividade antioxidante de frutas para serem usados na rede neural de predição. A rede neural artificial se mostrou como uma ferramenta simples, rápida e precisa na previsão da capacidade antioxidante em frutas.

Palavras-chave: compostos fenólicos, alimentos funcionais, aprendizado de máquina, DPPH, HPLC

ABSTRACT

Antioxidant compounds are known for their high antioxidant capacity. In fruits, these substances act as a form of protection against attack by external agents and can be used in favor of human health. The search for functional foods and well-being has grown day by day, and as a way to minimize processes, technology can be promising in this search. The present work aimed to predict the antioxidant activity in different fruits using an artificial neural network. Artificial neural networks are models that have simple processing units, known as artificial neurons, which work with mathematical functions simulating biological neurons. The database of the work consisted of ten articles, where data such as the phenolic profile and antioxidant activity of fruits were collected to be used in the prediction neural network. The artificial neural network proved to be a simple, fast and accurate tool for predicting the antioxidant capacity in fruits.

Keywords: phenolic compounds, functional foods, machine learning, DPPH, HPLC

LISTA DE FIGURAS

Figura 1: Catequinas e Antocianinas

Figura 2: Ácido Hidrocinâmico e Ácido Hidroxibenzóico

Figura 3: Representação de uma RNA

Figura 4: Rede Neural de Multicamadas

Figura 5: Estrutura dos fenólicos estudados

Figura 6: Estrutura da Rede Neural

Figura 7: Comparação entre os valores reais e preditos utilizando a ativação “*tanh*” e o solver “*adam*”

Figura 8: Comparação entre os valores reais e preditos utilizando a ativação “*ReLU*” e o solver “*sgd*”

LISTA DE TABELAS

Tabela 1: Compostos fenólicos encontrados nos artigos utilizados no estudo

Tabela 2: Concentrações dos compostos fenólicos

Tabela 3: Conjunto de dados na forma que foi processado pela rede neural

Tabela 4: Valores obtidos do erro percentual absoluto (MAPE) e erro quadrático médio (MSE)

LISTA DE EQUAÇÕES

Equação 1: Equação da função Sigmóide

Equação 2: Equação da função Tangente Hiperbólica

Equação 3: Equação da função ReLU.

Equação 4: Equação do erro absoluto percentual médio

Equação 5: Equação do erro médio ao quadrado

SUMÁRIO

1. Introdução.....	1
2. Referencial Teórico	2
2.1. Produção e consumo de frutas	2
2.2. Alimentos funcionais	3
2.3. Compostos fenólicos	4
2.4. Redes neurais artificiais	5
2.5. Redes neurais de predição	7
2.6. Redes neurais multicamadas	8
2.7. Função de ativação.....	9
3. Materiais e métodos	10
3.1. Dados de perfil fenólico e atividade antioxidante	10
3.2. Pré-processamento de conjunto de dados	11
3.3. Desing da RNA.....	11
3.4. Erro percentual absoluto médio.....	13
3.5. Erro médio ao quadrado.....	13
4. Resultados e discussão.....	14
5. Conclusão.....	22
6. Referências bibliográficas.....	23

1. INTRODUÇÃO

O Brasil é terceiro maior produtor de frutas do mundo, e possui uma alta diversidade de tipos e espécies. Mesmo estando no pódio de produção, o consumo não satisfaz os padrões necessários para uma saúde equilibrada. As frutas possuem poderosos benefícios ao corpo humano, um deles são os compostos antioxidantes, provenientes de milhares de compostos fenólicos que podem ser encontrados nesses vegetais.

Apesar do consumo de frutas não ser suficiente, a busca por uma alimentação funcional tem ganhado cada vez mais seu papel na sociedade. A importância dos compostos antioxidantes está ligada a essa funcionalidade. Entretanto, as técnicas disponíveis para determinar a capacidade antioxidante são bastante onerosas, pois, demandam muito tempo e dinheiro. Os reagentes utilizados nessas análises possuem um custo elevado e por conta disso, não são todos os laboratórios que podem arcar com essa despesa. Além disso, soma-se a dificuldade dos laboratórios em possuir todos os equipamentos necessários para realizar as análises.

Como forma de estratégia de minimizar processos e obter resultados mais rápidos e precisos, a rede neural artificial (RNA) é uma grande aliada. Uma RNA é composta por um conjunto de neurônios artificiais organizados em camadas interconectadas. Cada neurônio tem uma capacidade específica no processamento da informação. Essas capacidades são calculadas com os pares disponíveis de dados de entrada e saída que constituem o conjunto de treinamento. Enquanto duas dessas camadas estão conectadas ao 'mundo externo' (camada de entrada, onde os dados são apresentados e camada de saída, onde um valor de previsão é obtido), o restante delas (camadas ocultas) são definidas por neurônios conectados entre si, evitando conexões entre neurônios da mesma camada. (CORTES CABRERA,2009)

A utilização de redes neurais é crescente, uma vez que trazem benefícios que impactam nos processos. A redução de custo e a praticidade em resolver problemas complexos são algumas das grandes vantagens, pois uma vez estruturada para determinado objetivo, a rede neural já estará pronta para uso de futuras pesquisas e trabalhos. Dito isso, o objetivo deste trabalho consiste em prever o valor da

atividade antioxidante de frutas usando os compostos fenólicos que as compõem, através do uso de redes neurais artificiais.

2. REFERENCIAL TEÓRICO

2.1. PRODUÇÃO E CONSUMO DE FRUTAS

A produção mundial de frutas tem apresentado um crescimento contínuo. No triênio 89/91 era de 420 milhões de toneladas, ultrapassou as 500 milhões de toneladas em 1996 e em 2017 colheu-se um volume de 865,2 milhões de toneladas. (DERAL, 2020)

No cenário mundial, o Brasil é o terceiro maior produtor de frutas do mundo (SEBRAE, 2015) ficando atrás da China e Índia. A maior parte desta produção é voltada para o mercado consumidor interno - somente 2,5% é exportada. O mercado brasileiro de hortaliças é altamente diversificado e segmentado, com o volume de produção concentrado em seis espécies – batata, tomate, melancia, alface, cebola e cenoura, sendo a agricultura familiar responsável por mais da metade da produção. (EMBRAPA)

Mesmo com o desenvolvimento da fruticultura, os brasileiros ainda não consomem a quantidade de frutas recomendada pela Organização Mundial da Saúde (OMS), o que revela um mercado em potencial para o setor. Apenas 24,1% dos brasileiros consomem a quantidade recomendado de 400 gramas por dia. (SEBRAE, 2015)

Entretanto, o cenário pandêmico está fazendo com que o consumo de frutas e hortaliças cresçam. Segundo uma pesquisa feita pelo Estudo NutriNet Brasil (2020), “indica crescimento na compra de frutas, hortaliças e feijão (de 40,2% para 44,6%) no Brasil”, uma vez que a preocupação com o fortalecimento imunológico entra em questão para esse crescimento.

2.2. ALIMENTOS FUNCIONAIS

Desde a década de 1960, o processo de transição nutricional vem ocorrendo no Brasil. Essa transição é caracterizada como uma mudança dos hábitos alimentares da população através da diminuição do consumo dos alimentos in natura e aumento dos ultraprocessados que apresentam elevadas concentrações de sódio, açúcares e gorduras. (HENRIQUE... [ET AL.]. ARACAJU, EDIFS, 2018).

Em vista da mudança do perfil epidemiológico e nutricional da população brasileira, tem aumentado a procura pelos alimentos funcionais. Esses alimentos podem ser definidos como alimentos que possuem substâncias bioativas que acarretam benefícios ao organismo, não sendo utilizado na cura de doenças, mas auxiliando na promoção da saúde. (HENRIQUE... [ET AL.]. ARACAJU, EDIFS, 2018).

As frutas, reconhecidas fontes de vitaminas, minerais e fibras, são alimentos nutricionalmente importantes da dieta. O efeito protetor exercido por estes alimentos tem sido atribuído à presença de fitoquímicos com ação antioxidante, dentre os quais se destacam os polifenóis (WANG, CAO, PRIOR, 1996; BRAVO, 1998; MARTINEZ-VALVERDE, PERIAGO, ROS, 2000; KAUR, KAPOOR, 2002).

A capacidade antioxidante dos polifenóis é devida, principalmente, as suas propriedades redutoras, cuja intensidade da ação antioxidante exibida por estes fitoquímicos é diferenciada uma vez que depende, fundamentalmente, do número e posição de hidroxilas presentes na molécula (RICE-EVANS, MILLER, PAGANGA, 1997; OU ET AL., 2002).

A importância relacionada ao desempenho dos antioxidantes in vivo depende dos fatores: tipos de radicais livres formados; onde e como são gerados esses radicais; análise e métodos para a identificação dos danos, e doses ideais para obter proteção. Assim, é perfeitamente possível que um antioxidante atue como protetor em determinado sistema, mas que falhe na proteção, ou mesmo que aumente as lesões induzidas em outros sistemas ou tecidos (HALLIWELL ET AL., 1995).

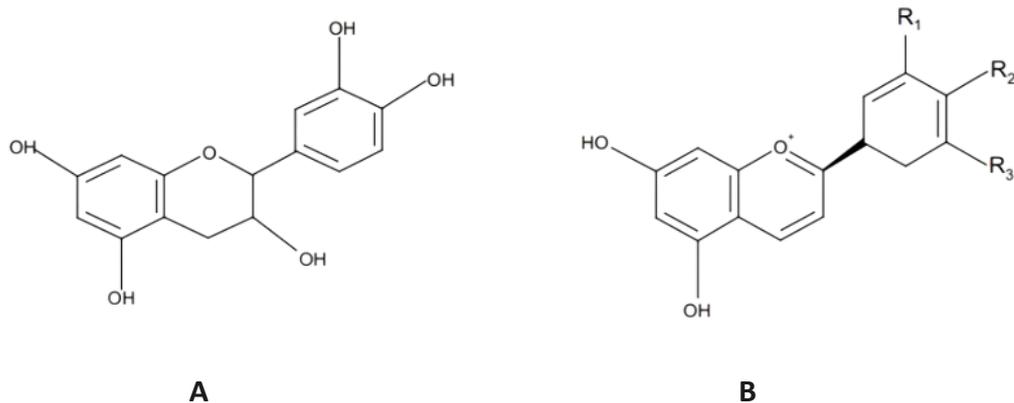
2.3. COMPOSTOS FENÓLICOS

Os compostos fenólicos são substâncias amplamente distribuídas na Natureza, mais de 8000 compostos fenólicos já foram detectados em plantas. Esse grande e complexo grupo faz parte dos constituintes de uma variedade de vegetais, frutas e produtos industrializados. Podem ser pigmentos, que dão a aparência colorida aos alimentos, ou produtos do metabolismo secundário, normalmente derivado de reações de defesa das plantas contra agressões do ambiente. Esses compostos agem como antioxidantes, não somente pela sua habilidade em doar hidrogênio ou elétrons, mas também em virtude de seus radicais intermediários estáveis, que impedem a oxidação de vários ingredientes do alimento, particularmente de lipídios (BRANDWILLIAMS; CUVELIER; BERSET, 1995).

Os compostos fenólicos apresentam, em sua estrutura, vários grupos benzênicos característicos, tendo como substituintes grupamentos hidroxilas (HERNÁNDEZ; PRIETO GONZÁLES, 1999).

Esta classe de compostos apresenta uma grande diversidade e divide-se em flavonóides (polifenóis) e não-flavonóides (fenóis simples ou ácidos). Os flavonóides compreendem um grupo de compostos fenólicos amplamente distribuídos nas frutas e nos vegetais, apresentando-se sob muitas variações como flavonóis, flavonas, flavanonas, catequinas (Figura 1), antocianinas (Figura 1), isoflavonas e chalconas. (GRAHAM, 1992; VAN ACQUIRE, 1996).

Na classe dos não-flavonóides estão os derivados dos ácidos hidroxicinâmico e hidroxibenzóico (Figura 2). Sua atividade antioxidante está relacionada com a posição dos grupos hidroxilas e também com a proximidade do grupo $-CO_2H$ em relação ao grupo fenil. Quanto mais próximo esse grupo estiver do grupo fenil, maior será a capacidade antioxidante do grupo hidroxila na posição meta (HRAZDINA; BORZEL; ROBINSON, 1970).



Fonte: HRAZDINA; BORZEL; ROBINSON, 1970

Figura 1: A: Catequinas e B: Antocianinas.



Fonte: HRAZDINA; BORZEL; ROBINSON, 1970

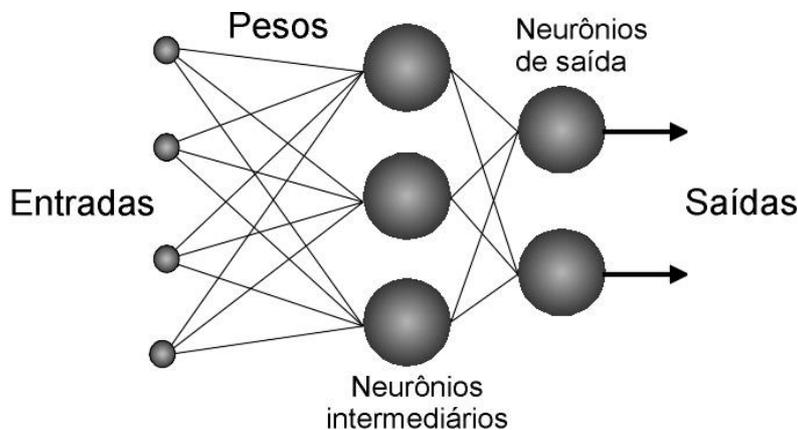
Figura 2: A: Ácido Hidrocínâmico B: Ácido Hidroxibenzóico

2.4. REDES NEURAIS ARTIFICIAIS

As Redes Neurais Artificiais são modelos constituídos por unidades de processamento simples, chamados de neurônios artificiais, que calculam funções matemáticas. Estes modelos são inspirados na estrutura do cérebro e têm como objetivo simular o comportamento humano, tais como: aprendizagem, associação, generalização e abstração quando submetidas a treinamento.

De acordo com HAYKIN (2001) as RNAs são particularmente eficientes para o mapeamento entrada/saída de sistemas não lineares e para realizar processamento paralelo, além de simular sistemas complexos. As RNAs generalizam os resultados obtidos para dados previamente desconhecidos, ou seja, produzem respostas coerentes e apropriadas para padrões ou exemplos que não foram utilizados no seu treinamento.

Uma rede neural artificial (RNA) tem duas facetas elementares: a arquitetura e o algoritmo de aprendizagem. Essa divisão surge naturalmente pelo paradigma como a rede é treinada. Ao contrário de um computador com arquitetura de von Neumann que é programado, a rede é treinada por exemplos de treino. O algoritmo de aprendizagem generaliza os dados e memoriza o conhecimento dentro dos parâmetros adaptáveis da rede, os pesos. A composição da rede é feita pelos neurônios. Normalmente o tipo de processamento de um único neurônio é a combinação linear das entradas com os pesos seguida pela passagem da combinação linear por uma função de ativação.



Fonte: TAFNER,1998

Figura 3: Representação de uma RNA

2.5. REDES NEURAIIS DE PREDIÇÃO

A aplicabilidade das redes neurais artificiais é crescente e a redução no custo de implantação é constante à medida que se popularizam ferramentas computacionais

que permitem a inserção de novas técnicas para solução de problemas complexos, como a predição de dados não-lineares e padrões de comportamento. A entrada e saída de dados dos sistemas produtivos são preditas nas análises múltiplas e paralelas das redes (MURASE, 2000). Elas são aplicadas com sucesso em uma variedade de domínios, como algoritmos de resolução de problemas de mapeamento, regressão, modelagem, análise de dados multivariados, otimização, controle, classificações e diferenciações químicas (NUNES et al., 2011; DEBSKA; GUZOWSKA-SWIDER, 2011; LEMES; JUNIOR, 2008; MARINI; BUCCI; MAGRI, 2008).

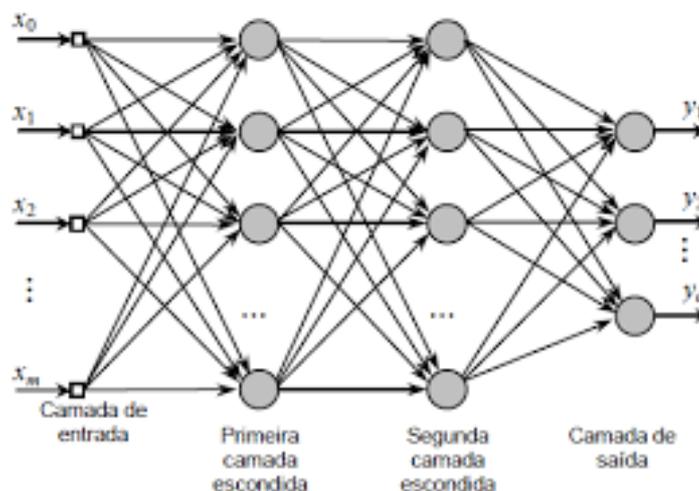
As principais vantagens de utilização dessas redes são a tolerância a falhas, a aplicação em tempo real, a capacidade de auto adaptação e a rápida resolução de problemas práticos, sem necessidade de definição de lista de regras ou de modelos precisos. Logo, o objetivo de definição da rede é resolver o problema com a menor estrutura possível, com potencial de aplicação em situação que requeiram classificação, identificação e associação de padrões, aproximação de funções e aprendizado em áreas em que é difícil criar modelos precisos da realidade e com frequentes mudanças de ambiente (HAYKIN, 2001).

Segundo BONTEMPI G (2012), o processamento de dados, executado através de algoritmos de aprendizagem de máquina, permite a predição da ocorrência através da análise dos dados históricos, possibilitando a compreensão de eventos em diferentes áreas do conhecimento. E para que isso seja possível, os algoritmos são empregados na solução de problemas não lineares e não paramétricos, os quais testes estatísticos clássicos não conseguem solucionar.

2.6. REDES NEURAIIS MULTILAYER PERCEPTRON (MULTICAMADAS)

Uma Rede Neural Artificial Multicamadas (Haykin, 1999) é tipicamente composta de camadas alinhadas de neurônios, conforme é mostrado na Figura 4. Nesse tipo de rede, as entradas são apresentadas na primeira camada, que é chamada camada de entrada. Essa camada distribui as informações para a(s) camada(s) escondida(s) da rede. A última camada é a camada de saída, onde a solução do

problema é obtida. A camada de entrada e a camada de saída podem ser separadas por uma ou mais camadas intermediárias (ou camadas escondidas).



Fonte: OLIVEIRA, 2010

Figura 4: Rede Neural de Multicamadas

A rede do tipo MLP pode ser classificada em dois grupos: parcialmente conectadas e totalmente conectadas. Em uma MLP totalmente conectada, cada neurônio de uma camada é conectado a todos os neurônios da camada anterior e a todos os neurônios da camada posterior. Em uma MLP parcialmente conectada, algumas conexões entre neurônios não acontecem. Entre a camada de entrada e a camada de saída, pode-se ter uma ou mais camadas ocultas. As camadas ocultas proporcionam complexidade e a possibilidade de resolver problemas não linearmente separáveis.

Uma das principais características de uma rede MLP é sua capacidade de resolver problemas não lineares, para isso, é necessário que a função de ativação dos neurônios pertencentes às camadas ocultas seja não-linear.

2.7. FUNÇÕES DE ATIVAÇÃO

A função de ativação é a transformação não linear que se faz ao longo do sinal de entrada. Elas basicamente decidem se um neurônio deve ser ativado ou não. Ou seja, se a informação que o neurônio está recebendo é relevante para a informação

fornecida ou deve ser ignorada (DEEP LEARNING BOOK, 2021). Algumas funções de ativação usadas em RNA's são:

Logistic (Sigmóide)

A equação 1 rege a função Sigmóide (Goodfellow,2016):

$$\sigma(X) = \frac{1}{1 + e^{-x}} \quad (1)$$

Essa função foi por muito tempo uma das principais ativações utilizadas em redes neurais. Em relação a função original (Heaviside), a sigmóide traz vantagens, como a derivada não nula em todo seu domínio, o que possibilita o aprendizado com o algoritmo backpropagation (FOREST, 2014).

Tangente hiperbólica (Tanh)

Uma desvantagem da função sigmóide é o fato de não estar centrada em zero. Essa característica não é desejável pois pode introduzir dinâmicas indesejáveis na busca do mínimo do gradiente (Fei-Fei Li, 2016). A função de ativação tangente hiperbólica visa eliminar essa desvantagem mantendo as demais características da sigmóide. Essa função representada pela equação 2.

$$\sigma(x) = \tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2)$$

Unidade linear retificada – ReLu

A ReLU tornou-se muito popular devido ao pouco esforço computacional necessário para sua implementação. Algumas vantagens da função ReLU são a possibilidade de utilizar o algoritmo *backpropagation* e a aceleração do treinamento com o método do gradiente descendente estocástico (Krizhevsky,2012). A equação 3 representa a função ReLU.

$$\sigma(x) = \max(0, x) \quad (3)$$

3. MATERIAIS E MÉTODOS

3.1. DADOS DE PERFIL FENÓLICOS E ATIVIDADE ANTIOXIDANTE DE FRUTAS

Todos os dados utilizados neste trabalho foram extraídos de estudos já publicados na literatura. Foram selecionados artigos que descreviam o perfil fenólico de frutas através da técnica de HPLC juntamente com a determinação da atividade antioxidante através da técnica de redução do radical DPPH. (BAI,2013; BETTA,2018; FENG,2017; NASCIMENTO,2017; PARIKH,2016; SCHULZ,2015; SCHULZ,2020; SERAGLIO,2017; WANG,2018, WANNES,2012.)

A composição fenólica de todas as frutas dos artigos selecionados consistia em 38 diferentes compostos. Tentando evitar o excesso de complexidade da rede neural e minimizar possíveis erros que podem ocorrer na predição dos resultados devido a baixas concentrações de alguns compostos, apenas os 21 compostos com maior concentração e alto poder antioxidante conhecidos foram escolhidos como as variáveis de entrada. (RUBERTO, G., & BARATTA, MT, 2000).

3.2. PRÉ-PROCESSAMENTO DO CONJUNTO DE DADOS

O conjunto de dados passou por dois pré-processamentos antes da sua utilização. A primeira ação realizada nos dados foi a conversão das unidades a fim que todas as concentrações dos compostos fenólicos estivessem aptas a serem comparáveis entre si. Desta forma, todas as concentrações foram convertidas para miligrama de composto fenólico por grama de fruta (mg/g).

O segundo tratamento no banco de dados foi a realização da normalização dos dados de entrada (Concentração de compostos fenólicos). A normalização é extremamente indicada para estudos de aprendizagem de máquina para que se anule os erros relacionados a variação de escala e para garantir uma maior acurácia do modelo. Para isso foi-se utilizado a função *StandardScaler*, que transforma os dados de forma que sua distribuição terá um valor médio de 0 e um desvio padrão de 1.

Assim, cada valor no conjunto de dados terá o valor médio subtraído e dividido entre o desvio padrão de todo o conjunto de dados.

3.3. DESIGN DA REDE NEURAL ARTIFICIAL (RNA)

A implementação da rede neural para predição da atividade antioxidante de frutas foi realizada utilizando a linguagem de programação *Python* (v3.6.9) na plataforma gratuita *Google Colab*. O *Colab* é uma plataforma que permite escrever e executar *Python* no próprio navegador de forma simples e prática. A plataforma é muito utilizada para machine learning pois permite tirar proveito da potência de hardware do Google, independente do poder de processamento da máquina do usuário.

O algoritmo foi escrito empregando a biblioteca *scikit-learn* (v1.0.2). Esta biblioteca de código aberta (*open source*) de aprendizado de máquina possui um acervo robusto e eficiente de algoritmos que permitem o pré-processamento, classificação, regressão, agrupamento e redução de dimensionalidade de dados além de outras recursos. O algoritmo base utilizado foi o *MLPRegressor* (v0.8) e foi escolhido por conta da sua facilidade de uso, sua ampla utilização em outras aplicações, além do fato de se possuir uma série de tutoriais disponíveis na internet.

Para que a rede *MLPRegressor* funcione, é necessário informar os dados de entrada (concentração dos compostos fenólicos) e os dados de saída (atividade antioxidante). Desta forma, será possível treiná-la e testá-la como método de previsão. Entretanto, alguns parâmetros foram alterados e variados para obter a melhor resposta da rede neural artificial.

O primeiro passo foi dividir o conjunto de dados entre o conjunto de testes e conjunto de treinamento. Para a separação dos conjuntos foi utilizado a biblioteca *train_test_split*, também da biblioteca *scikit-learn*, com o parâmetro *test_size = 0.30*, ou seja, o conjunto de dados foi dividido entre 70% para o treinamento e 30% para o teste da rede.

Outro parâmetro fixo foi o desenho da arquitetura da rede neural artificial. Foram 45 neurônios na camada oculta (20,15,10) com um neurônio na camada de

saída (atividade antioxidante). O máximo de interações (*max_inter*) foi fixado em 5000 épocas, ou seja, o conjunto de dados passou 5000 vezes por entre a rede de neurônios. Esses valores foram definidos com base em testes realizados tentando diminuir as métricas de avaliação da rede.

Os parâmetros *activation* e *solver*, o tipo de ativação e resolução, foram variados entre si para se entender como estas funções interferem no resultado aumentando ou diminuindo a capacidade da rede de prever os resultados.

Para o treinamento da RNA foi utilizado o método *fit* do algoritmo *MLPRegressor* tendo como dados de entrada e saída os 70% dos dados de entrada dividido anteriormente. Para realização da predição dos dados, utilizou-se a função *predict* também presente no *MLPRegressor*, utilizando os 30% dos dados separado para os testes. Essa função retorna os valores estimados pelo algoritmo baseado no aprendizado na etapa de treinamento. Os valores preditos são comparados com os valores reais e a partir dessa comparação são calculadas as métricas para avaliação do modelo de predição.

3.4. ERRO PERCENTUAL ABSOLUTO MÉDIO (MAPE)

O erro percentual absoluto médio calcula o percentual médio sobre o desvio absoluto entre o valor real e o valor predito. É uma medida de precisão de um modelo de previsão. Seu valor é definido pela equação 1, onde A_t é o valor real, F_t é o valor previsto e n é o número de amostras.

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (4)$$

3.5. ERRO MÉDIO AO QUADRADO (MSE)

O erro percentual absoluto médio fornece a média dos erros do modelo ao quadrado, ou seja, a diferença quadrática média entre os valores reais e previstos. Assim, o seu cálculo penaliza valores muito distantes da média, onde as menores

diferenças ganham menos importância enquanto diferenças maiores ganham mais importância. Seu valor é definido pela equação 2, onde y_i é o valor esperado, \hat{y}_i é o valor previsto e N é o número de amostras.

$$MAPE = \frac{100\%}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (5)$$

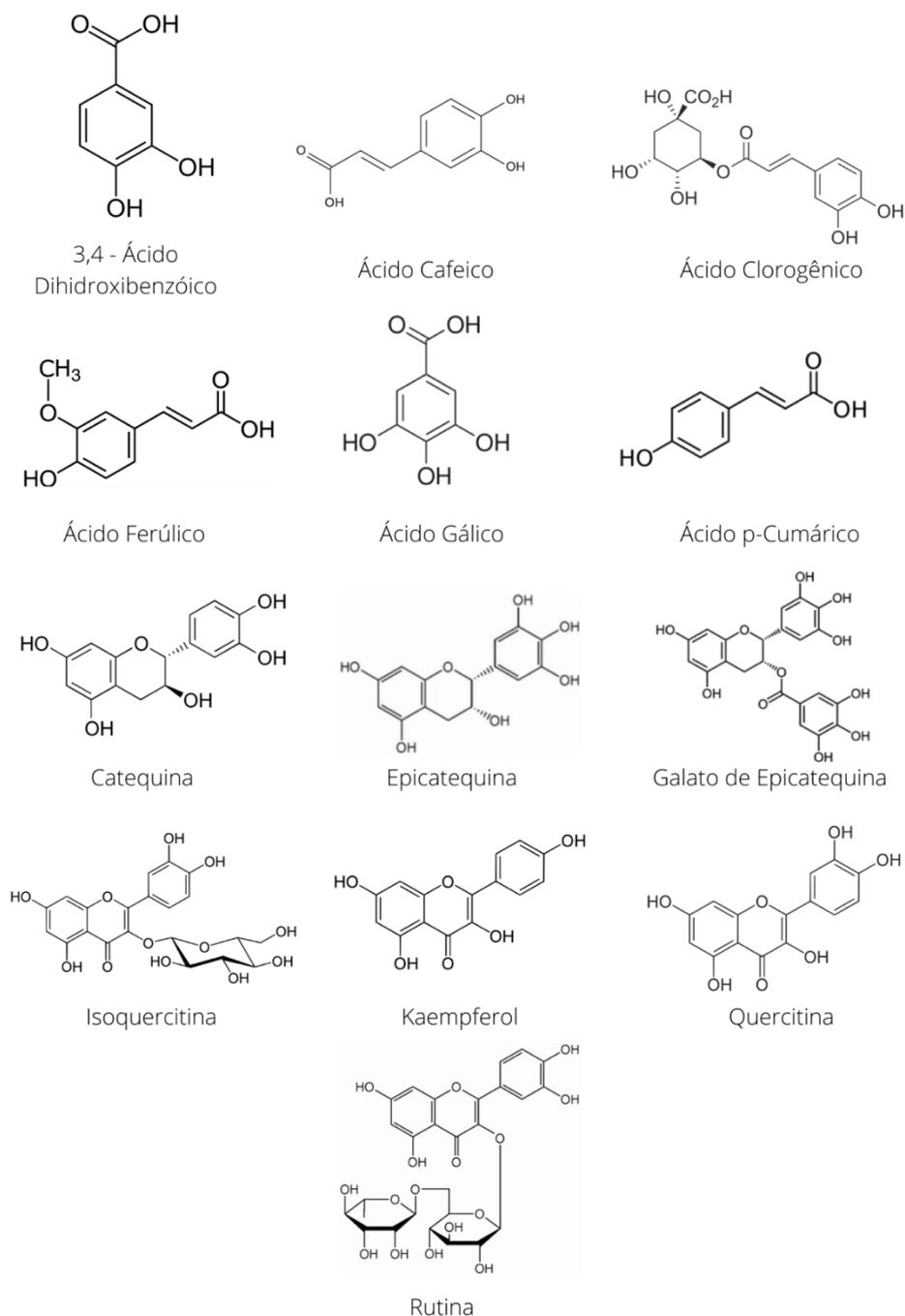
4. RESULTADOS E DISCUSSÃO

Os trinta e oito compostos encontrados nos trabalhos podem ser observados na tabela 1 abaixo.

Tabela 1: Compostos fenólicos encontrados nos artigos utilizados no estudo

Composto fenólico	Composto fenólico
3,4 – Ácido Dihidroxibenzoico	Epicatequina
Ácido Benzoico	Florizina
Ácido Cafeico	Galato de Epicatequina
Ácido clorogênico	Hispidulina
Ácido Cumárico	Isoquercitina
Ácido Elágico	Isorhamnetina
Ácido Ferúlico	Kaempferol
Ácido Gálico	L-Epicatequina
Ácido p-Cumárico	Luteolina
Ácido Protocatecuico	Miricetina
Ácido Salicílico	Naringenina
Ácido Sinapínico	Pinobanksina
Ácido Sirínico	Quercetina
Ácido Vanilínico	Quercitrina
Aldeído Coniferílico	Rutina
Aromadendria	Sinapaldeído
Catecol	Siringaldeído
Catequina	Taxifolina
Crisina	Vanilina

Como já foi dito, o conjunto de dados foi reduzido para vinte e um compostos para facilitar a implementação da rede. Entretanto, durante o processamento dos testes com o banco de dados, percebeu-se que as baixas concentrações de compostos fenólicos e atividade antioxidante da fruta Jussara (SCHULZ,2015) estava diminuindo as métricas de avaliação da predição da RNA a níveis inaceitáveis. Por isso, os dados relativos à fruta Juçara Fruit (*Euterpe edulis* Martius) foram retirados. Em consequência, o número de compostos fenólicos utilizados diminuiu para 13 compostos fenólicos que podem ser visualizados na figura 5.



Fonte: Do autor adaptado

Figura 5: Estrutura dos fenólicos estudados

O conjunto de dados das concentrações dos compostos fenólicos obtido pode ser visualizado na tabela 2, juntamente com os dados de atividade antioxidante que foram utilizados como dados de saída na rede neural. Além disso, o data set padronizado utilizado na rede pode ser observado na tabela 3.

Tabela 2: Concentrações dos compostos fenólicos

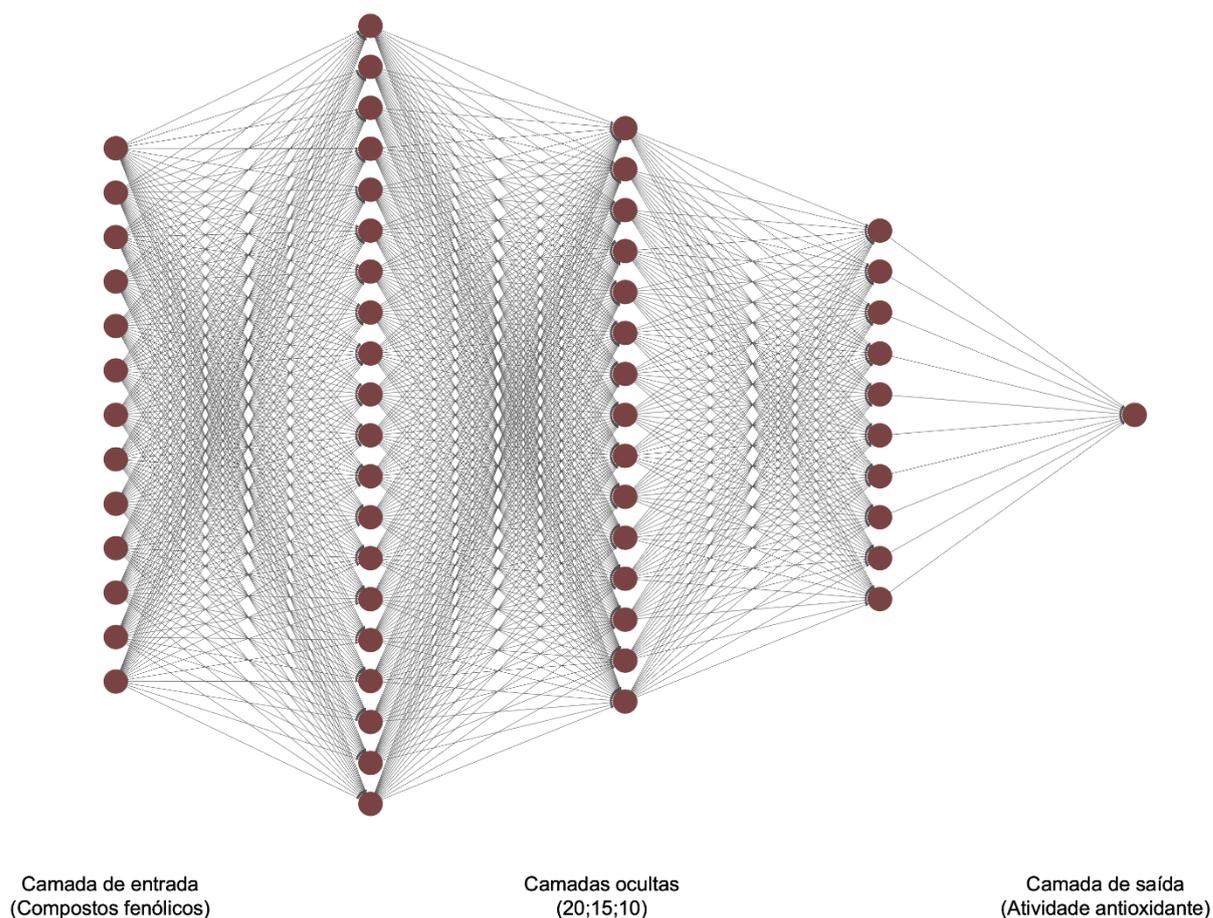
Amostras	3,4 ácido dihidroxibenzoico	Ácido caféico	Ácido clorogênico	Ácido Ferúlico	Ácido Gálico	Ácido p-cumárico	Catequina	Epicatequina	Galato de epicatequina	Isoquercetina	Kaempferol	Quercetina	Rutina	DPPH
Jaboticaba (intermediária)	1709.15	70.4	348.71	502.26	7680.83	531.3	12.84	0	0	677.46	78.56	6795.33	0	5066.35
Jaboticaba (madura)	1299.74	25.41	193.27	199.74	4164.51	358.47	0	0	0	0	33.48	5211.41	0	4662.09
Guabiju (intermediária)	127.13	41.97	11.08	128.87	7780.34	170.15	194.81	36.22	1577.83	1417.87	1079.66	78	84.81	3105.99
Guabiju (madura)	82.88	25.13	5.86	0	4379.98	111.51	48.37	11.22	971.72	2160.35	901.11	8240.85	0	2569.28
Jamelão (intermediária)	30.64	0	0	0	21494.1	70.36	0	0	25.75	0	0	98.45	0	2761.32
Jamelão (madura)	36.03	0	0	0	11203.1	50.07	0	0	0	13.16	0	44.72	0	2760.5
Goiaba amarela (verde)	64.06	0	18.13	0	532.61	53.69	11424.2	0	32.99	3158.31	32.85	1833.67	0	19.69
Goiaba amarela (intermediária)	88.81	0	11.87	124.56	781.93	90.31	10880.74	0	0	2622.13	228.97	1118.72	0	20.01
Goiaba amarela (madura)	75.6	54.28	8.74	79.28	1542.01	63	12795.5	22.34	0	1194.6	32.7	1038.77	0	18.46
Blackberry (estágio I)	4.3	0	0	0	6.8	0	101	381.4	0	189.1	0.9	53.5	128.1	248.8
Blackberry (estágio II)	12	0	0	0	6.8	0	3.9	8.6	0	191.8	1.8	10.2	96.3	331.1
Acerola (estágio I)	3.1	9.5	0	12.7	3.8	9.8	8	7	0	431.9	16.1	120.8	0	112.4
Acerola (estágio II)	3.9	7.3	0	6.2	4	7.9	6.4	6.3	0	331.6	23.8	119.8	0	96.33

Tabela 3: Conjunto de dados na forma que foi processado pela rede neural.

ADB ¹	AC ²	ACL ³	AF ⁴	AG ⁵	APC ⁶	CA ⁷	EC ⁸	GEC ⁹	IQ ¹⁰	KL ¹¹	QU ¹²	RU ¹³	DPPH ¹⁴
0.0171	0.0007	0.0035	0.0050	0.0768	0.0053	0.0001	0.0000	0.0000	0.0068	0.0008	0.0680	0.0000	50.6635
0.0130	0.0003	0.0019	0.0020	0.0416	0.0036	0.0000	0.0000	0.0000	0.0000	0.0003	0.0521	0.0000	46.6209
0.0013	0.0004	0.0001	0.0013	0.0778	0.0017	0.0019	0.0004	0.0158	0.0142	0.0108	0.0008	0.0008	31.0599
0.0008	0.0003	0.0001	0.0000	0.0438	0.0011	0.0005	0.0001	0.0097	0.0216	0.0090	0.0824	0.0000	25.6928
0.0003	0.0000	0.0000	0.0000	0.2149	0.0007	0.0000	0.0000	0.0003	0.0000	0.0000	0.0010	0.0000	27.6132
0.0004	0.0000	0.0000	0.0000	0.1120	0.0005	0.0000	0.0000	0.0000	0.0001	0.0000	0.0004	0.0000	27.6050
0.0006	0.0000	0.0002	0.0000	0.0053	0.0005	0.1142	0.0000	0.0003	0.0316	0.0003	0.0183	0.0000	19.6900
0.0009	0.0000	0.0001	0.0012	0.0078	0.0009	0.1088	0.0000	0.0000	0.0262	0.0023	0.0112	0.0000	20.0100
0.0008	0.0005	0.0001	0.0008	0.0154	0.0006	0.1280	0.0002	0.0000	0.0119	0.0003	0.0104	0.0000	18.4600
0.0043	0.0000	0.0000	0.0000	0.0068	0.0000	0.1010	0.3814	0.0000	0.1891	0.0009	0.0535	0.1281	24.8000
0.0120	0.0000	0.0000	0.0000	0.0068	0.0000	0.0039	0.0086	0.0000	0.1918	0.0018	0.0102	0.0963	33.1000
0.0031	0.0095	0.0000	0.0127	0.0038	0.0098	0.0000	0.0070	0.0000	0.4319	0.0161	0.1208	0.0000	112.4000
0.0039	0.0073	0.0000	0.0062	0.0040	0.0079	0.0064	0.0063	0.0000	0.3316	0.0238	0.1198	0.0000	96.3300

¹ 3,4 – ácido dihidroxibenzóico² Ácido caféico³ Ácido clorogênico⁴ Ácido Ferúlico⁵ Ácido Gálico⁶ Ácido p-cumárico⁷ Catequina⁸ Epicatequina⁹ Galato de epicatequina¹⁰ Isoquercetina¹¹ Kaempferol¹² Quercetina¹³ Rutina¹⁴ Atividade antioxidante

A topologia da rede neural artificial então possuiu 13 neurônios na camada de entrada (compostos fenólicos), 45 neurônios nas camadas ocultas e 1 neurônio na saída (atividade antioxidante). A estrutura da rede neural pode ser observada na Figura 6.



Fonte: Do autor

Figura 6: Estrutura da Rede Neural

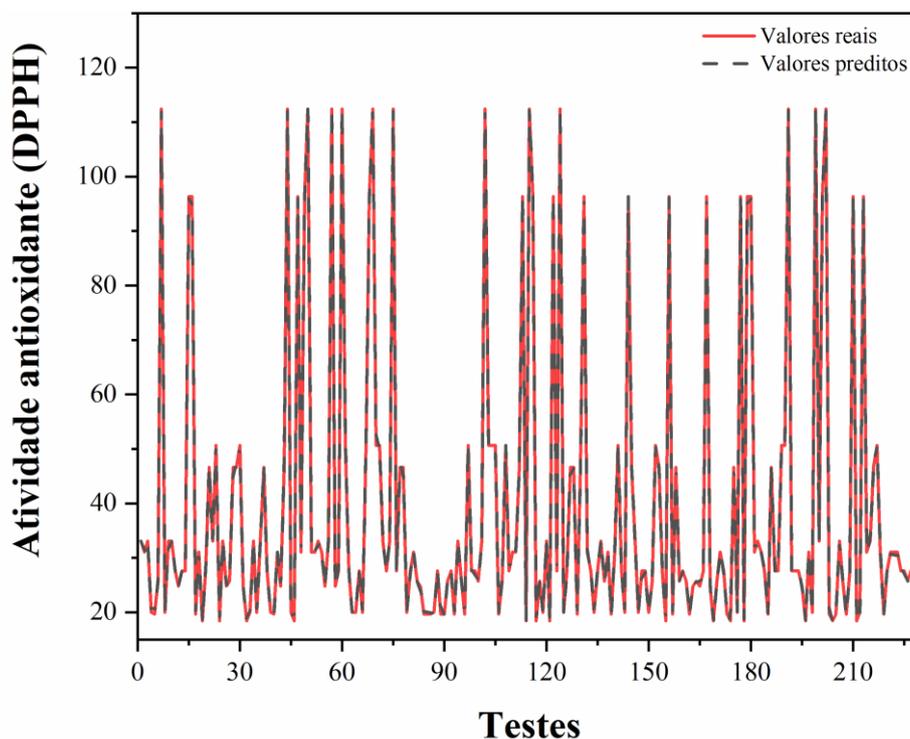
Na tabela 4 encontramos os resultados das métricas da RNA para a predição da atividade antioxidante de frutas. Analisando os resultados das métricas da RNA podemos perceber que os tipos de função de ativação e resolução são capazes de gerar grande influência da predição dos dados pela rede neural.

Tabela 4. Valores obtidos do erro percentual absoluto (MAPE) e erro quadrático médio (MSE) pela rede neural na predição da atividade antioxidante

Activation	Solver	MAPE (%)	MSE (%)
ReLU	sgd	3,04	19,09
	adam	0,0021	$6,37 \times 10^{-6}$
tanh	sgd	$3,93 \times 10^{-6}$	$2,28 \times 10^{-10}$
	adam	1,10	8,93
logistic	sdg	0,08	0,03
	adam	2,33	47,41

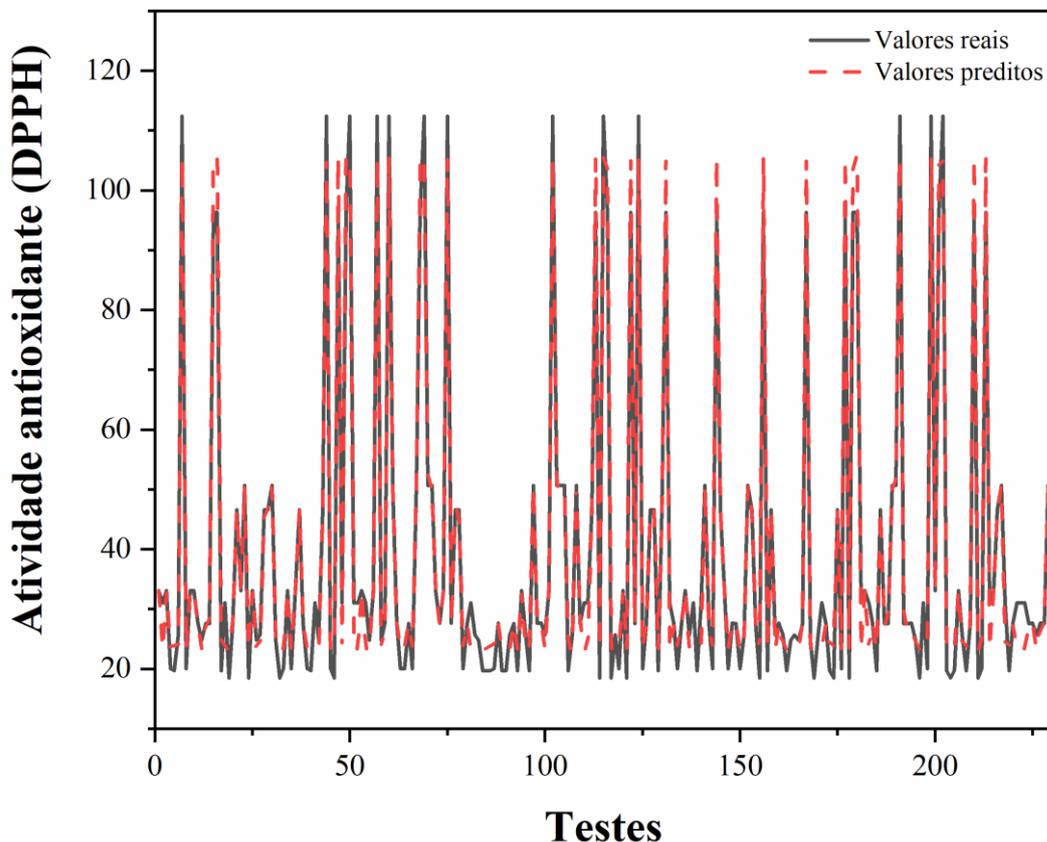
Conforme pudemos verificar, a função de ativação tangente hiperbólica juntamente com a resolução utilizando o gradiente descendente estocástico foram os parâmetros que melhor se adaptaram ao conjunto de dados e, portanto, apresentou os melhores resultados dos erros percentual absoluto e erro quadrático médio. Na figura 7 podemos visualizar uma comparação entre os valores reais e os valores preditos pela a RNA utilizando os parâmetros das melhores métricas.

Figura 7: Comparação entre os valores reais e preditos utilizando a ativação “*tanh*” e o solver “*sgd*”



Fonte: Do autor

Figura 8: Comparação entre os valores reais e preditos utilizando a ativação “ReLU” e o solver “sgd”



Fonte: Do autor

A importância em fazer testes com diferentes funções e solver é que para cada objetivo final essas funções dará uma precisão adequada. E de acordo com os resultados encontrados, observamos como a RNA fluiu e entregou um bom resultado utilizando a função de ativação tangente hiperbólica (tanh) juntamente com o solver gradiente descendente estocástico (sgd), trazendo uma margem de erro pouco significativa de $3,93 \times 10^{-6}$ para o percentual absoluto (MAPE) e $2,28 \times 10^{-10}$ para o erro quadrático médio (MSE).

Podemos observar nos gráficos, que a combinação que mais se adequou traz os resultados preditos muito próximos aos resultados reais (Figura 7), e percebe-se que a outra combinação entre a função unidade linear retificada (ReLU) e o solver sgd possui algumas diferenças e erros maiores em seus resultados (Figura 8).

Os resultados confirmam que a predição de antioxidante em frutas fazendo o uso de RNA é confiável, com grande potencial e baixo custo, o que é capaz de ganhar mais espaço na área de alimentos, podendo prever compostos, ingredientes com rapidez e precisão, ou seja, avaliando a qualidade dos alimentos.

Além disso, as RNA's podem ser muito utilizada e explorada em outras áreas da Engenharia de Alimentos, com objetivos de pesquisa (BARBON,2014), processamento térmico (GONÇALVES, 2003, embalagens, saúde (JANUÁRIO; GUEDES; SILVA, 2019) e até mesmo em análise sensorial (CARVALHO, 2011).

5. CONCLUSÃO

Concluimos, portanto, que apesar da redução do número de compostos feita ao longo do processo, com o objetivo de alinhar os valores de atividade antioxidante, a rede neural utilizada mostrou eficiência, provando nos resultados finais quando comparados com os resultados já conhecidos. Provando também ser algo promissor na área de alimentos, uma vez que possui baixo custo, eliminação de análises laboratoriais complexas, consequentemente otimizando tempo.

6. REFERÊNCIAS BIBLIOGRÁFICAS

AGÊNCIA BRASIL. Brasileiro aumentou consumo de frutas, hortaliças e feijão na pandemia, aponta estudo. **AGÊNCIA BRASIL**, Revista Globo Rural, 20 ago. 2020. Disponível em: <https://revistagloborural.globo.com/Noticias/Agricultura/Hortifruiti/noticia/2020/08/brasileiro-aumentou-consumo-de-frutas-hortalicas-e-feijao-na-pandemia-aponta-estudo.html>. Acesso em: 13 fev. 2022.

ANDRADE, Paulo Fernando de Souza. FRUTICULTURA Análise da Conjuntura. **Departamento de Economia Rural ? DERAL**, [s. l.], 2020. Disponível em: https://www.agricultura.pr.gov.br/sites/default/arquivos_restritos/files/documento/2020-01/fruticultura_2020.pdf. Acesso em: 10 fev. 2022.

BAI, Xuelian. Antioxidant activity and HPLC analysis of polyphenol-enriched extracts from industrial apple pomace. **Research Article**, SCL, 4 mar. 2013. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/23460126/>. Acesso em: 12 jul. 2021.

BETTA, Fabiana Della. Phenolic Compounds Determined by LC-MS/MS and In Vitro Antioxidant Capacity of Brazilian Fruits in Two Edible Ripening Stages. **Plant Foods for Human Nutrition**, ORIGINAL PAPER, 14 jul. 2018. Disponível em <https://link.springer.com/article/10.1007/s11130-018-0690-1>. Acesso em: 30 ago. 2021.

CARVALHO, Naiara Barbosa. Artificial neural networks applied to the sensory analysis of light cream cheese. 2011. 116 f. Dissertação (Mestrado em Ciência de Alimentos; **Tecnologia de Alimentos; Engenharia de Alimentos**) - Universidade Federal de Viçosa, Viçosa, 2011. Disponível em: <http://locus.ufv.br/handle/123456789/2898>. Acesso em: 12 abr. 2022

CORTES CABRERA, Álvaro. Application of artificial neural networks to the prediction of the antioxidant activity of essential oils in two experimental in vitro models. **Food Chemistry**, Elsevier, 18 abr. 2009. Disponível em: <https://www.journals.elsevier.com/food-chemistry>. Acesso em: 13 abr. 2022.

DEEP LEARNING BOOK, Capítulo 8 – Função de ativação. **Data Science Academy**. 2021. Disponível em: <https://www.deeplearningbook.com.br/funcao-de-ativacao/> Acesso em: 12 abr. 2022

EMBRAPA. Frutas e hortaliças. **Ciência que transforma**, EMBRAPA, 2015. Disponível em: <https://www.embrapa.br/grandes-contribuicoes-para-a-agricultura-brasileira/frutas-e-hortalicas>. Acesso em: 10 fev. 2022.

FENG, Cheng-Yong. Polyphenol profile and antioxidant activity of the fruit and leaf of *Vaccinium glaucoalbum* from the Tibetan Himalayas. **Food Chemistry**, Elsevier, 15 mar. 2017. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0308814616314662?via%3Dihub>. Acesso em: 12 jul. 2021.

FERNEDA, Edberto. Redes neurais e sua aplicação em sistemas de recuperação de informação. **Documentação da Faculdade de Filosofia Ciências e Letras de Ribeirão Preto ? USP**, SciELO, 2006. Disponível em: <https://www.scielo.br/j/ci/a/SQ9myjZWLxnyXfstXMgCdch/abstract/?lang=pt#>. Acesso em: 10 fev. 2022.

FEI-FEI LI, Andrej Karpathy, and Justin Johnson. Cs231n: Convolutional neural networks for visual recognition 2016. Acesso em 03 abr. 2022

FERREIRA, Arthur. UM ESTUDO SOBRE PREVISÃO DA DEMANDA DE ENCOMENDAS UTILIZANDO UMA REDE NEURAL ARTIFICIAL. **XVIII Simpósio de Pesquisa Operacional & Logística da Marinha**, ResearchGate, Agosto 2016. Disponível em: https://www.researchgate.net/publication/306062626_UM_ESTUDO SOBRE_PREVISAO_DA_DEMANDA_DE_ENCOMENDAS_UTILIZANDO_UM_A_REDE_NEURAL_ARTIFICIAL. Acesso em: 12 fev. 2022.

FOREST AGOSTINELLI, Matthew Hoffman, Peter Sadowski, and Pierre Baldi. **Learning activation functions to improve deep neural networks**. 12 2014. Acesso em: 03 abr. 2022.

GONÇALVES, Eliane Calomino. Redes neurais artificiais aplicadas ao processamento térmico de alimentos. 2003. 71f. **Tese (Doutorado em Ciência e Tecnologia de Alimentos)** - Universidade Federal de Viçosa, Viçosa. 2003 Disponível em: <https://www.locus.ufv.br/handle/123456789/8992> Acesso em: 12 abr 2022.

HAYKIN, S. **Redes neurais: princípios e práticas**. Tradução ENGEL, P.M. 2.ed. Porto Alegre: Bookman, 2001. Disponível em: <https://www.docsity.com/pt/redes-neurais-principios-e-pratica-2a-ed/4848223/> Acesso em: 30 ago. 2021

JANUÁRIO, Jailson; GUEDES, Elloá; SILVA, Fabio de. Classificação do Índice Glicêmico a partir de Imagens de Alimentos com Redes Neurais

Convolucionais. *In*: ENCONTRO NACIONAL DE INTELIGÊNCIA ARTIFICIAL E COMPUTACIONAL (ENIAC), 16. , 2019, Salvador. **Anais** [...]. Porto Alegre: Sociedade Brasileira de Computação, 2019

KRIZHEVSKY, Alex ,Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, **Advances in Neural Information Processing Systems** 25, pages 1097–1105. Curran Associates, Inc., 2012. Acesso em: 03 abr. 2022

M.M. NASCIMENTO, Eidla. HPLC and in vitro evaluation of antioxidant properties of fruit from *Malpighia glabra* (Malpighiaceae) at different stages of maturation. **Food and Chemical Toxicology**, Elsevier, 26 nov. 2017. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0278691517306956?via%3Dihub>. Acesso em: 5 jul. 2021.

MELO, Enayde de Almeida. Capacidade antioxidante de frutas. **Universidade Federal Rural de Pernambuco**, Revista Brasileira de Ciências Farmacêuticas, 2008. Disponível em: <https://www.scielo.br/j/rbcf/a/54sB3W4GSKPLwTJZ8xxN6pL/?format=pdf&lang=pt>. Acesso em: 13 fev. 2022.

SERAGLIO, Siluana Katia Tischer. Nutritional and bioactive potential of Myrtaceae fruits during ripening. **Food Chemistry**, Elsevier, 21 jun. 2017. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0308814617311007?via%3Dihub>. Acesso em: 30 ago. 2021.

OLIVEIRA , Hélder da Silva. Atitudes e hábitos de consumo de alimentos funcionais - Um estudo exploratório. -, B Digital, 18 maio 2012. Disponível em: <https://bdigital.ufp.pt/bitstream/10284/3153/3/83-94.pdf>. Acesso em: 12 fev. 2022.

PANDORF, Héilton. Uso de redes neurais artificiais para predição de índices zootécnicos nas fases de gestação e maternidade na suinocultura. **Sociedade Brasileira de Zootecnia**, Revista Brasileira de Zootecnia, 2011. Disponível em: <https://www.rbz.org.br/pt-br/article/uso-de-redes-neurais-artificiais-para-predicao-de-indices-zootecnicos-nas-fases-de-gestacao-e-maternidade-na-suinocultura/>. Acesso em: 10 fev. 2022.

PARIKH, Bhumi. Quantification of phenolic compounds and antioxidant capacity of an underutilized Indian fruit: Rayan [*Manilkara hexandra* (Roxb.) Dubard]. **Food Science and Human Wellness**, ScienceDirect, 6 dez. 2016. Disponível em:

<https://www.sciencedirect.com/science/article/pii/S2213453016300982?via%3Dihub>. Acesso em: 20 set. 2021.

PASSOS, Iago dos. Uso de Redes Neurais Artificiais para a Detecção das Doenças Olho de Boi e Mancha de Sarna em Maçãs. **UNIVERSIDADE DE CAXIAS DO SUL**, [s. l.], 2017. Disponível em: <https://repositorio.uces.br/xmlui/bitstream/handle/11338/3724/TCC%20Iago%20dos%20Passos.pdf?sequence=1&isAllowed=y#page=55&zoom=100,113,296>. Acesso em: 12 fev. 2022.

RUBERTO, G., & BARATTA, MT (2000). Atividade antioxidante de óleo essencial selecionado componentes em dois sistemas modelo lipídico. **Química de Alimentos**. Disponível em: <https://www.journals.elsevier.com/food-chemistry>. Acesso em: 13 abr 2022.

SCHULZ, Mayara. Determination of Phenolic Compounds in Three Edible Ripening Stages of Yellow Guava (*Psidium cattleianum* Sabine) after Acidic Hydrolysis by LC-MS/MS. **Plant Foods for Human Nutrition**, ORIGINAL PAPER, 7 jan. 2020. Disponível em: <https://link.springer.com/article/10.1007/s11130-019-00792-0>. Acesso em: 30 ago. 2021.

SCHULZ, Mayara. Chemical composition, bioactive compounds and antioxidant capacity of juçara fruit (*Euterpe edulis* Martius) during ripening. **Food Research International**, Elsevier, 6 ago. 2015. Disponível em: <https://www.sciencedirect.com/journal/food-research-international>. Acesso em: 30 ago. 2021.

SILVA, Lordêlo Cardoso. Compostos fenólicos, carotenóides e atividade antioxidante em produtos vegetais. **Universidade Estadual de Londrina**, Redalyc, 26 jan. 2009. Disponível em: <https://www.redalyc.org/pdf/4457/445744097017.pdf>. Acesso em: 12 fev. 2022.

VIANNA, Leonardo Silva. Predição de casos de COVID-19 nos municípios de Santa Catarina utilizando redes neurais recorrentes. **Artigos Científicos**, Saúde Coletiva, 2 ago. 2021. Disponível em: <http://www.revistas.mpmcomunicacao.com.br/index.php/saudecoletiva/article/view/1734>. Acesso em: 10 fev. 2022.

WANG, Yutang. Bioactive compounds and in vitro antioxidant activities of peel, flesh and seed powder of kiwi fruit. **Food Science and Technology**, IFST, 9 abr. 2018. Disponível em: <https://ifst.onlinelibrary.wiley.com/doi/abs/10.1111/ijfs.13812>. Acesso em: 12 jul. 2021.

WANNES, WISSEM AIDI. DIFFERENCES BETWEEN MYRTLE FRUIT PARTS (MYRTUS COMMUNIS VAR. ITALICA) IN PHENOLICS AND ANTIOXIDANT CONTENTS. **Journal of Food Biochemistry** , ResearchGate, 6 nov. 2012. Disponível em: https://www.researchgate.net/publication/258510763_Differences_between_Myrtle_Fruit_Parts_Myrtus_communis_var_italica_in_Phenolics_and_Antioxidant_Contents. Acesso em: 20 set. 2021.