

Previsão do preço do café utilizando redes neurais

Leonardo Marinzek Araujo e Celso Ap.de França
Departamento de Engenharia Elétrica da Universidade Federal de São Carlos

Resumo – O presente artigo tem como objetivo desenvolver um estudo sobre o funcionamento das Redes Neurais de Múltiplas Camadas, a fim de se prever, o preço do café, baseado em determinadas condições climáticas. Dessa forma, podendo assim, ser usado como uma ferramenta de auxílio no momento da venda da cultura. O produto que terá seu preço previsto, é definido como café arábica em saca de 60kg. Para a realização do presente projeto, utilizou-se uma base de dados que contém fatores climáticos relativos aos anos de 2003 a 2008 e os preços relativos à determinadas datas. Para a realização desse projeto, foi implementado uma rede neural de múltiplas camadas do tipo Regressão Linear. A partir disso, foi treinada uma base de dados com algumas características relativas ao tempo, que podem influenciar no preço da saca de café. Ao final do artigo, o leitor terá não somente conhecimento sobre a teoria por trás das redes neurais, mas também um embasamento teórico sobre o uso prático das Redes Neurais de múltiplas camadas a fim de se prever o preço da saca de 60kg de café.

Palavras-Chave – Redes neurais, Previsão do preço do café, Regressão linear, A influência do clima no preço do café

1. INTRODUÇÃO

De acordo com o site Kuhnbrasil [1] “A cadeia produtiva do café gera mais de 8 milhões de empregos diretos e indiretos, segundo dados do Ministério da Agricultura, Pecuária e Abastecimento (MAPA), o que explica a dimensão desse cultivo na economia brasileira. O café é tão importante para o Brasil que fomentou inclusive o crescimento da industrialização no território nacional, ou seja, impacta a diversificação econômica. Vale destacar que as plantações ocupam uma área de 2 milhões de hectares e que existem cerca de 300 mil produtores ligados a esse cultivo no Brasil. Segundo dados divulgados pelo Governo Federal, no período de 2012 a 2017, as exportações do café chegaram a 200 milhões de sacas, o que trouxe US \$35 bilhões para o Brasil.”

O Brasil, é o maior produtor de café do mundo, de acordo com a Organização Internacional do café [2] (2020), tal fato, implica tamanha influência dessa cultura na economia brasileira. Além disso, o café é a segunda bebida mais consumida no Brasil, ficando atrás apenas da água. “De acordo com a Associação Brasileira da Indústria de café (ABIC), considerando o período de novembro 2019 a outubro de 2020 o país consumiu 21.004.430 sacas/ano. O consumo per-capita de café em grão cru (kg/hab.ano) foi de 6,02, e em grão torrado moído 4,82 (kg/hab.ano). Sendo observado no referido período maior consumo em relação ao período anterior.” [2]

O café está presente na vida do brasileiro desde meados de 1927, desde então, foi se tornando cada vez mais influente na economia, chegando a ser responsável por sustentar todo um império e a “República Velha”. Assim como a maioria das *commodities*, o café está sujeito a mudanças de preços, seja por fatores climáticos, políticos, ou mesmo por fatores externos, como guerras e legislações. Tendo em vista isso, por ser um produto de alto impacto econômico, presente na vida da maioria dos brasileiros e sujeito a uma alta volatilidade de seus preços, a previsão do preço futuro do café, torna-se um fator de extrema importância para essa atividade produtiva.

Com base nesse cenário, esse trabalho teve como objetivo, o desenvolvimento de um método, cujo qual, o produtor da cultura de café, ou mesmo estudiosos da área, possam prever o preço da saca do café arábica de 60kg a partir de determinadas características climáticas como: temperatura máxima, temperatura mínima, insolação, evaporação, temperatura média e umidade relativa média. Além disso, mantendo a base desse estudo, pode-se aplicá-lo a outras culturas, apenas mudando as características climáticas. Resumidamente o propósito desse artigo é utilizar tecnologias de inteligência artificial (redes neurais artificiais), a fim de impactar diretamente na economia brasileira, tendo como foco um dos produtos de maior influência na história, o café.

2. TRABALHOS RELACIONADOS

A utilização de Redes Neurais Artificiais a fim de se prever preços, é uma técnica que vem sendo utilizada por diferentes autores, já sendo implementados em diversos projetos ao redor do mundo. O uso dessa tecnologia tem impactado não só no preço do café, mas sim de diferentes culturas de *commodities*, como soja, milho, cana-de-açúcar dentre outros. Além destes, a tecnologia de redes neurais, são bastante utilizadas para previsões de energia elétrica, preços de casas e afins.

HAJIME e SANTANA [3], utilizaram um simulador SNNS (Stuttgart Neural Network Simulator), com o intuito de simular arquiteturas e algoritmos a fim de prever o preço da saca de soja no estado de São Paulo. Já SANTOS [4], fez uma aplicação de rede neural CNN-BLSTM (Convolutional Neural Network - Bidirectional Long Short-Term Memory) com o objetivo de prever o preço do café, porém, este utiliza uma base de dados contendo apenas os preços dos anos anteriores. Uma vertente de utilização de redes neurais para previsão de preços, é utilizado por SILVA [5], no qual a autora, visa a comercialização de energia elétrica, tendo como intuito estimar o preço para a venda no mercado externo. Para isso, ela utilizou-se de uma rede neural com arquitetura *feedforward* em cada uma das regiões selecionadas para o mercado PJM (Pennsylvania, New Jersey, Maryland Interconnection).

Em conjunto com este artigo, todos os trabalhos citados, demonstram a capacidade dos processos de inteligência artificial em fazer a previsão de determinado produto, independente do método, arquitetura ou modelo de tecnologia. Dessa forma, é notável quão útil e impactante -principalmente em setores econômicos- pode vir a ser com o uso das redes neurais.

3. FUNDAMENTAÇÃO TEÓRICA

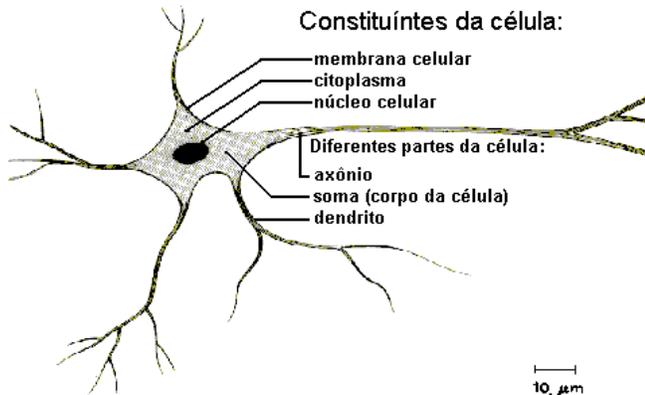
3.1– Rede Neural Artificial

Uma rede neural artificial, é uma técnica computacional que propõe um modelo matemático inspirado na estrutura neural de organismos inteligentes e adquire conhecimento por meio da

experiência. Uma grande rede neural artificial pode ter centenas ou milhares de unidades de processamento; um cérebro de mamífero pode ter bilhões de neurônios.

O sistema nervoso, consiste em um conjunto extremamente complexo de células, os neurônios (Figura 1). Eles desempenham um papel vital na determinação de como o corpo humano funciona e se comporta, bem como no raciocínio. Os neurônios são compostos de dendritos, que são um conjunto de terminais de entrada, um corpo central e axônios, que são terminais de saída longos.

Figura 1 – Constituintes da célula neuronal

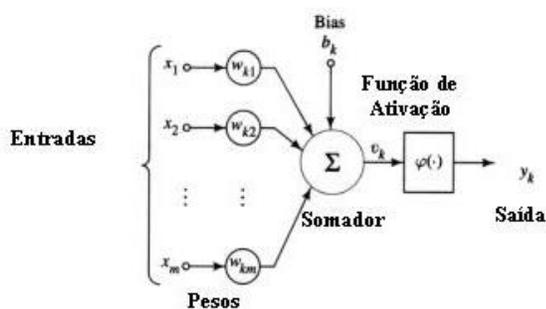


Fonte: [6]

3.2 – Neurônio Artificial

Um neurônio artificial é um meio de simular matematicamente o comportamento de um neurônio real, ou seja, um meio de simular os níveis de transmissão e desempenho de sinais elétricos transmitidos por meio de sinapses. A Figura 2 representa o modelo real de neurônios artificiais usados em redes neurais.

Figura 2 – Neurônio Artificial



Fonte: [7]

De acordo com [8], o neurônio artificial é uma estrutura lógico-matemática que procura simular a forma, o comportamento e as funções de um neurônio biológico. Assim sendo, os dendritos foram substituídos por entradas, cujas ligações com o corpo celular artificial são realizadas através de elementos chamados de peso (simulando as sinapses). Os estímulos captados pelas entradas são processados pela função

de soma, e o limiar de disparo do neurônio biológico foi substituído pela função de transferência.

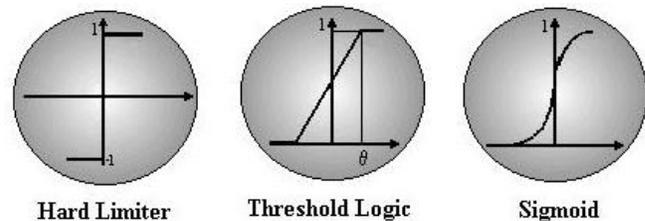
Baseando-se no neurônio real, as entradas (representadas por 'x') simulam os impulsos elétricos passados entre os neurônios. Assim sendo, os sinais têm a capacidade de excitar mais ou menos o neurônio receptor. A fim de se medir essa excitação, utilizam-se pesos, assim, quanto maior o valor desses pesos sinápticos da entrada maior será sua influência no sinal de saída. "Bias" é uma palavra utilizada para descrever um modelo neural matemático, este que aumenta o grau de liberdade do sistema, a fim de se aproximar ainda mais do comportamento real.

O neurônio então, ao possuir os sinais de entrada e os pesos sinápticos, passam por uma combinação linear dada por:

$$\sum_{n=0}^n K_n * W_n$$

A partir disso, o neurônio pode ou não ser ativado, essa ativação vai ter como limitante o fator de ativação. Existem vários tipos de funções de ativação, as mais comuns são: Hard Limiter, Threshold e Sigmoid. (Figura 3)

Figura 3 – Tipos de funções de ativação



Fonte: [6]

Neste presente artigo, a rede neural utiliza a função de ativação do tipo "relu" a mesma pode ser escrita como: $\varphi(U) = \max(0, U)$. Quando não temos uma função de ativação, os pesos e bias simplesmente fazem uma transformação linear. Uma equação linear é simples de resolver, mas é limitada na sua capacidade de resolver problemas complexos. A função de ativação faz a transformação não-linear nos dados de entrada, tornando-o capaz de aprender e executar tarefas mais complexas.[9]

Aprendizagem supervisionada: é um ramo do aprendizado de máquina, um método de análise de dados que usa algoritmos que aprendem iterativamente a partir dos dados para permitir que os computadores encontrem informações ocultas sem serem explicitamente programados para onde procurar. A aprendizagem supervisionada é um dos três métodos usados pelas máquinas para "aprender": supervisionada, não supervisionada e com otimização. [10]

Existem dois tipos principais de aprendizagem supervisionada; classificação e regressão. A classificação é onde um algoritmo é treinado para classificar os dados de entrada em variáveis discretas. Durante o treinamento, os algoritmos recebem dados de entrada de treinamento com um rótulo de 'classe'. Por exemplo, os dados de treinamento podem consistir nas últimas faturas de cartão de crédito de um conjunto de clientes, identificadas se eles fizeram uma compra futura ou não. Quando o saldo de crédito de um novo cliente é apresentado ao algoritmo, ele classifica o cliente no grupo 'comprará' ou 'não

comprará'. [10]

Aprendizagem não supervisionada: Nesse caso, o modelo irá receber somente os dados de entrada, tendo como objetivo a detecção de características em comum entre todos os dados e por fim, agrupá-los baseado nessas características. Por exemplo, os dados de treinamento imobiliário tomariam nota da localização, área e outros parâmetros relevantes. A saída da rede será o preço do imóvel específico.

Na regressão, um algoritmo precisa identificar uma relação funcional entre os parâmetros de entrada e a saída. O valor de saída não é discreto como na classificação, em vez disso, é uma função dos parâmetros de entrada. A exatidão de um algoritmo de regressão é calculada com base na variação entre a saída precisa e a saída prevista.

Todo o processo de aprendizagem das redes neurais pode ser realizado por diferentes maneiras:

- **Ajuste de erro (Regra delta):** Normalmente utilizada em algoritmos de treinamento supervisionado, este método tem como objetivo, ajustar os pesos sinápticos por meio de erros obtidos através do valor de saída x desejado

- **Competitiva:** É um método em que os neurônios são expostos aos sinais de entrada e apenas os que obtiveram saída próxima ao valor esperado são ativados enquanto o outro tem seus pesos sinápticos alterados [11]

- **Boltzmann:** É um método estocástico derivado de conceitos da estatística em que os neurônios podem assumir dois estados (0 ou 1), podendo estar agrupados ou não, sendo responsáveis pela interação com o ambiente [11]

- **Hebbiana:** Método no qual os pesos são ajustados baseados nos postulados de Hebb, no qual os mesmos, são baseados via as atividades dos neurônios.

- **Adam:** Estimativa de momento adaptável é um algoritmo para técnica de otimização para descida de gradiente. O método é realmente eficiente ao trabalhar com grandes problemas envolvendo muitos dados ou parâmetros é uma combinação do algoritmo 'descida gradiente com momentum' e o algoritmo 'RMSP – (Root mean square prop)'. [12]

No contexto de treinamento, alguns problemas podem vir a serem encontrados, seja envolvendo os dados de testes (Overfitting), ou mesmo quando o desempenho do modelo já é ruim no próprio treinamento (Underfitting).

Overfitting: O processo de overfitting é quando o modelo aprende demais sobre os dados. Neste caso, o modelo mostra-se adequado apenas para os dados de treino, como se o modelo tivesse apenas decorado os dados de treino e não fosse capaz de generalizar para outros dados nunca vistos antes. Quando isso acontece, os dados de treino apresentam resultados excelentes, enquanto a performance do modelo cai drasticamente com os dados de teste. [13]

Underfitting: Tal processo, ao contrário do overfitting, este se dá pelo “mau” treinamento dos dados, assim, a rede neural não consegue ajustar os pesos da melhor forma a fim de se prever ou tomar determinadas decisões.

3.3– Rede Neural de Múltipla Camada (MLP)

Como é possível fazer uma conexão entre os neurônios – saída com entrada – esse conjunto de ligações forma-se uma arquitetura (rede) de múltiplas camadas (Figura 3), estas são chamadas de *multiayer perceptron (MLP)*, ou perceptrons de múltiplas camadas.

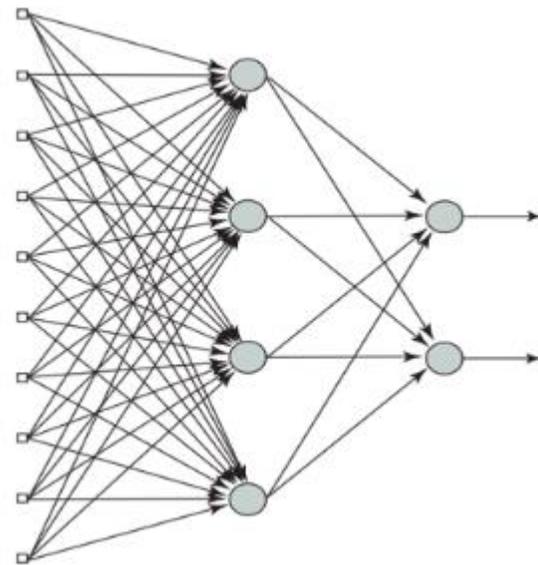
O perceptron de múltiplas camadas, é composto por

mais de três camadas, sendo estas definidas como entrada, saída e ocultas. Sabe-se que as múltiplas conexões junto aos diferentes pesos, estabelecem por uma rede neural de múltiplas camadas.

O aprendizado para esse tipo de rede, normalmente é feito através do algoritmo de retro propagação do erro. Cada camada de toda a rede possui determinada função. As redes neurais de múltiplas camadas, atualmente é uma das mais utilizadas, desde problemas complexos, problemas simples, reconhecimento de objetos ou até mesmo previsões.

Este presente artigo utiliza como base de estudo uma rede neural de múltiplas camadas a partir da retropropagação de erros (backpropagation). Inicialmente, determina-se os pesos sinápticos de forma aleatória, o algoritmo passa por duas fases: processamento direto e processamento reverso, logo após vem os treinamentos e os resultados obtidos pela rede. [14]

Figura 3 – Múltiplas Camadas



Fonte: Autoria própria

3.4– Modelos de Redes Neurais

3.4.1 - Redes Neurais Recorrentes (RNN)

Assim como a MLP, as redes neurais recorrentes possuem uma arquitetura do tipo camada, mas nesse caso, existem conexões recorrentes entre as camadas. Esse tipo de rede neural baseia-se seu resultado em informações previamente ditas, ou seja, basicamente leva em consideração o resultado anterior a fim de se obter o atual. As RNN são comumente utilizadas no processo de reconhecimento de linguagem natural, pois o algoritmo ao reconhecer frases, ele leva em consideração não só as palavras atuais, mas também as anteriores dessa forma ele consegue identificar padrões de formação de frases. [14]

3.4.2 – Redes Neurais Profundas (DNN)

A rede neural profunda é uma MLP que possui múltiplas camadas ocultas podendo conter ou não, milhares de neurônios. Pelo fato desta conter múltiplas camadas ocultas, faz com que a

rede neural consiga se adequar a problemas mais complexos. Em um algoritmo de identificação de palavras, por exemplo, uma ou diferentes camadas poderiam ficar encarregadas de receber os pixels de uma imagem e identificar letras, enquanto outras camadas subsequentes poderiam ser encarregues de a partir dessas letras identificar palavras. Ou seja, é possível delegar diferentes funções por camadas.

3.4.3 – Redes Neurais Convolucionais (CNN)

A rede neural Convolucional é um algoritmo de aprendizado profundo, que capta uma imagem de entrada, atribui pesos a determinados aspectos/objetos da imagem fazendo uma diferenciação um do outro.[15]

Esse tipo de rede neural é capaz de aplicar filtros em dados visuais mantendo a relação de vizinhança entre os pixels de uma imagem ao longo de todo o processamento da rede. Uma rede neural convolucional é constituída por múltiplas camadas com funções distintas. Uma das formas de avaliar o aprendizado desse tipo de rede é verificar as ativações dos filtros obtidos ao fim do treinamento. [15].

4 – METODOLOGIA

4.1 – Base de dados

Para a realização desse presente estudo, foi utilizado duas diferentes bases de dados. A primeiro momento, foi extraído da Estação Meteorológica da Universidade de São Paulo [16] uma base que continha as seguintes informações: temperatura máxima, temperatura mínima, insolação, evaporação, temperatura média e umidade relativa, durante o período de 2003 a 2008. Posteriormente, foi extraído do Centro de estudos avançados em Economia Aplicada (CEPEA) [17] o valor diário em reais, da saca de 60kg do café arábica também no mesmo intervalo de tempo, (2003 a 2008).

4.2 – Informações computacionais do projeto

O projeto em questão, foi realizado com o uso de linguagem de programação Python 3.7. Colaborando com isso, utilizou-se um computador de processador Intel(R) Core (TM) i5-6400U CPU @ 1.60GHz 2.71 GHz, 8GB de memória RAM, sistema operacional Windows 64 e por fim uma placa NVIDIA 960 TI.

4.3 – Uso dos dados

O preço do café possui variações diárias, podendo atingir diferentes valores ao longo de um único dia. Tal variação pode ser acarretada por diferentes motivos, como: política interna/externa, condições climáticas, pragas, doenças, queimadas, dentre outros. Assim sendo, o projeto em questão utilizou-se de bases de dados (já citadas anteriormente) relativas a determinadas questões climáticas atreladas ao preço da saca do café arábica.

Após um longo estudo feito, além de consultas a profissionais da área de agronomia, sabe-se que os 6 fatores climáticos citados acima, são os que mais determinam o preço da saca de café. Dessa forma, o presente projeto faz a utilização de uma rede neural que utiliza desses dados confrontando-os diretamente com o preço do café, a ponto de

utilizar uma rede do tipo *backpropagation* e prever, baseando-se em pesos distintos a cada fator externo o preço do café.

4.4 – Definição de todos os parâmetros

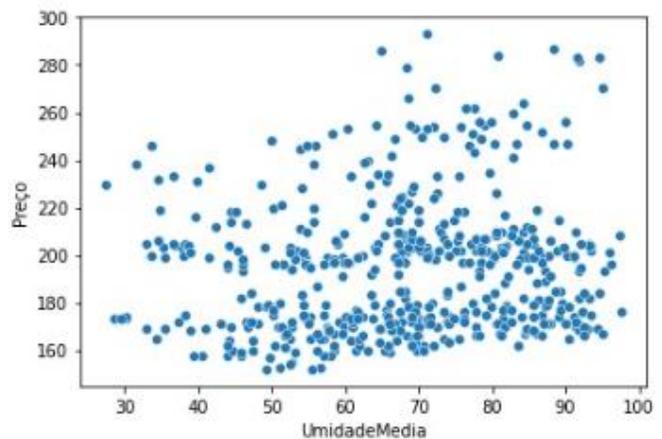
- Preço da saca do café arábica 60kg: reais
- Temperatura máxima: graus celsius
- Temperatura mínima: graus celsius
- Temperatura média: graus celsius
- Insolação: watt por metro quadrado
- Evaporação: milímetros por hora
- Umidade relativa: grama por quilograma

Para a determinação do preço do café, as características climáticas listadas acima foram tratadas através de uma regressão linear múltiplas, cujo algoritmo, ao final dos treinamentos é capaz de prever o preço futuro do café.

4.5 – Processos analisados durante a Codificação

Após a manipulação dos dados citados, foi plotado um gráfico (Figura 4) que traz os valores da Umidade Média x Preço do café. Tal imagem é presente neste artigo pois um dos fatores determinantes/influentes no preço da saca de 60kg do café arábica é a umidade relativa do ar. De forma simples e direta, pode-se observar pelo gráfico da Figura 4 que o preço da cultura em questão tende a ser mais alto em períodos cuja umidade permeia entre 60 a 70, isso se deve pois o café ao mesmo tempo que necessita de água para sua produção, ele precisa de “estresses naturais” para amadurecer e se fortalecer como planta. Ou seja, nem em períodos extremamente chuvosos quanto seca extrema favorecem a produção e consequentemente o preço do produto.

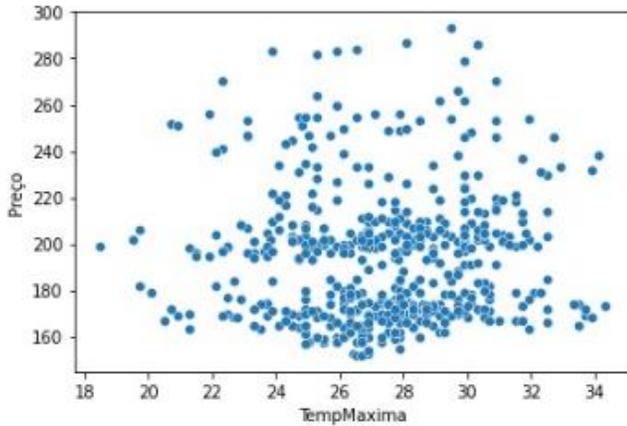
Figura 4 – Gráfico Preço x Umidade Média



Fonte: Autoria própria

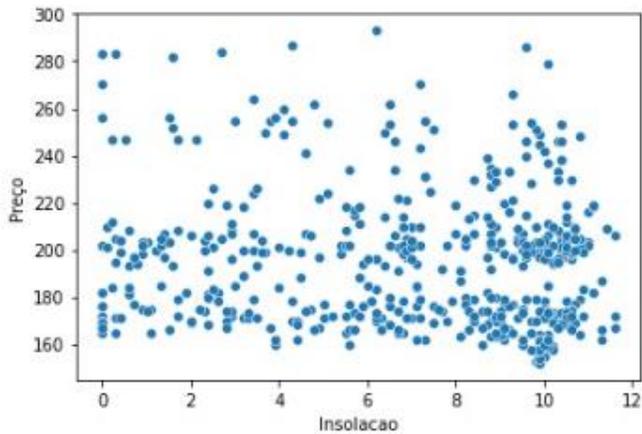
A fim de se fazer um estudo completo, relacionando todas os atributos com o preço, utilizando a função em python “pairplot” [18] foi gerado os gráficos abaixo (Figura 5 a 9). Tais gráficos trazem a relação das entradas com o preço final (assim como na Figura 4).

Figura 5 – Preço x Temperatura Máxima



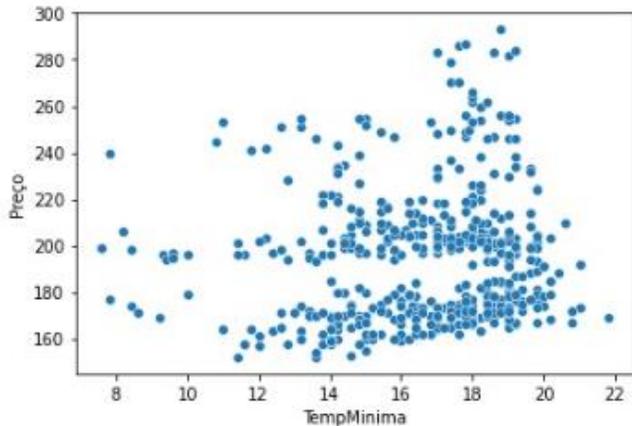
Fonte: Autoria própria

Figura 6 – Preço x Insolação



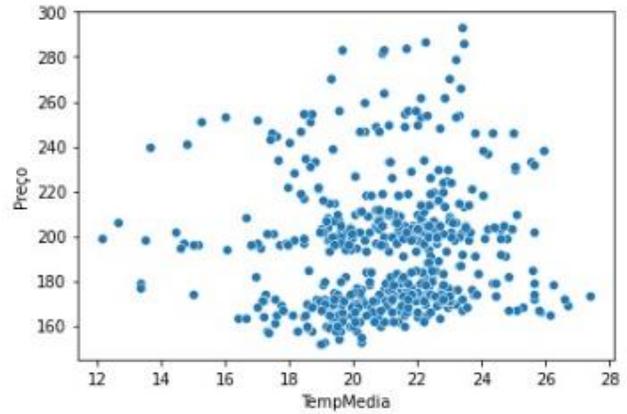
Fonte: Autoria própria

Figura 7 – Preço x Temperatura Mínima



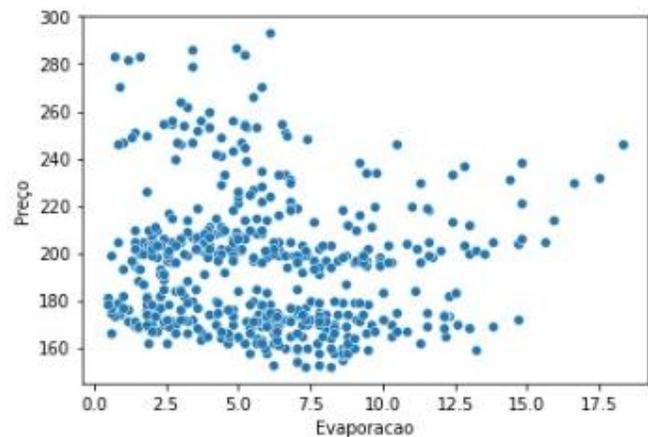
Fonte: Autoria própria

Figura 8 – Preço x Temperatura Média



Fonte: Autoria própria

Figura 9 – Preço x Evaporação



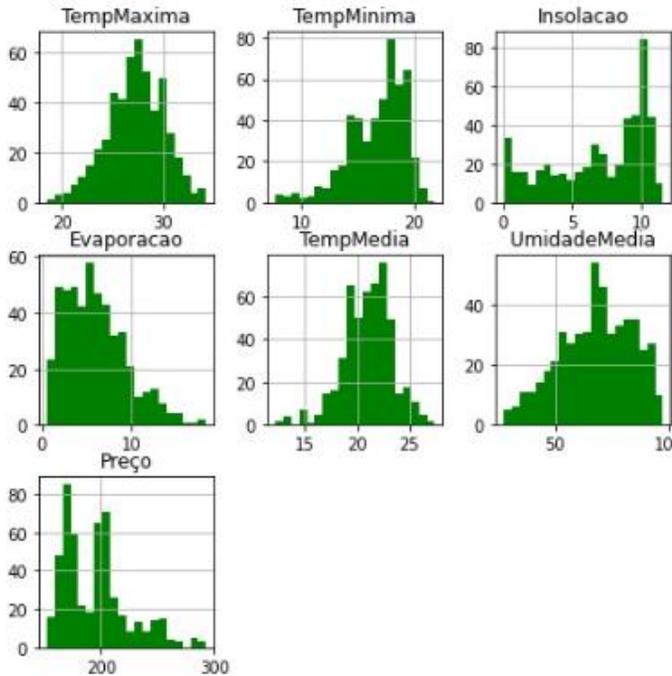
Fonte: Autoria própria

As Figuras 4 a 9 mostram como as características de entrada selecionadas influenciam no preço do produto, mostrando a tendência de alta no preço do café. Na Figura 9, por exemplo, o preço tende a ser mais alto se a evaporação estiver entre 0 e 7,5 mm/dia. Nota-se que os preços variam bastante demonstrando que não é um problema de simples solução.

Esta função (python “pairplot”) mostra várias distribuições bivariadas dos atributos no conjunto de dados, ou seja, uma combinação de variável em um DataFrame como uma matriz de gráficos, tudo a fim de visualizar melhor, além de estabelecer assim uma relação entradas x saída. [18]

Após análises das correlações entre os atributos e o preço, foi gerado um histograma (Figura 10), ele traz a distribuição de frequência de cada um dos atributos baseado na quantidade de registros e em seus respectivos valores. Na figura é possível notar uma vasta distribuição dos valores perante todas as entradas, isso mostra uma similaridade entre a influência das entradas na saída final, o preço.

Figura 10 – Histograma



Fonte: Autoria própria

4.6 – Tratamento da base de dados

Com o objetivo de melhor compreender, treinar, testar e otimizar toda a base de dados via rede neural, foi preciso através da função em python “MinMaxScaler” transformar todos os atributos em uma escala entre 0 e 1 (Figura 8), isso devido à discrepância entre seus valores. Além disso essa função serve para processar melhor toda a rede neural e também para que ela não considere um atributo muito mais importante do que o outro, por isso é bastante importante realizar essa normalização em redes que utilizam a regressão.

A divisão da base de dados foi composta por treinamento e teste e para isso, foi utilizada a função em python “train_test_split” sendo utilizado 162 dados (30%) da base para testar e 378 (70%) para treinar.

Figura 8 – Valores das entradas entre 0 e 1

```
array([[0.63291139, 0.83098592, 0.25, 0.16292135, 0.70526316,
0.76785714],
[0.58227848, 0.8028169, 0.03448276, 0.12359551, 0.64210526,
0.86428571],
[0.27848101, 0.85915493, 0., 0.03370787, 0.59210526,
1.],
...,
[0.74683544, 0.70422535, 0.82758621, 0.16292135, 0.74210526,
0.53571429],
[0.69620253, 0.78873239, 0.53448276, 0.31460674, 0.73947368,
0.62142857],
[0.60759494, 0.71830986, 0.37068966, 0.24719101, 0.66315789,
0.86785714]])
```

Fonte: Autoria própria

4.7 – Criação e construção do modelo

A seguir, na Tabela 1 estão listadas todas as especificações da rede neural criada. A primeira coluna mostra a especificação do tipo da camada da rede sendo a mesma especificada como “Dense”, na segunda coluna indica o formato de saída de cada camada, no caso, a rede neural é composta por 3 camadas ocultas contendo 200 neurônios em cada camada e uma última camada, a de saída, com um único neurônio. Com isso, há 82001 parâmetros que devem ser analisados, nesse número estão inclusos pesos e os bias.

Como já citado anteriormente, o peso está ligado às entradas e pode significar a relevância de uma determinada entrada para o neurônio. O bias é como o intercepto adicionado em uma equação linear, ou seja, ele pode funcionar como uma forma de transladar a função de ativação. [19]

Tabela 1 – Especificação da Rede

Model: "sequential_24"

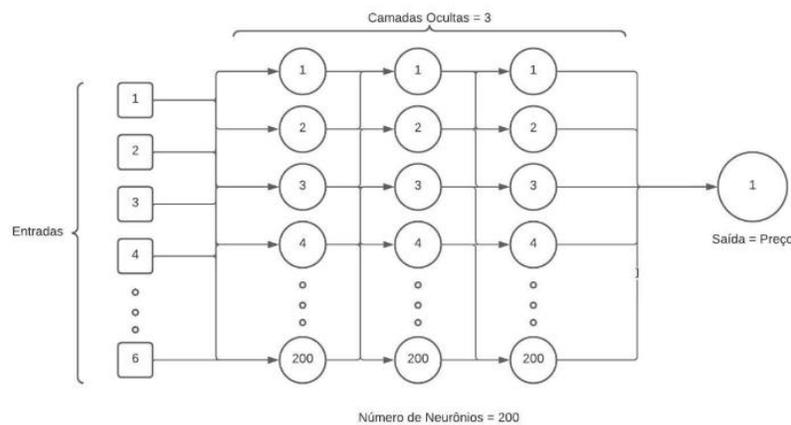
Layer (type)	Output Shape	Param #
dense_100 (Dense)	(None, 200)	1400
dense_101 (Dense)	(None, 200)	40200
dense_102 (Dense)	(None, 200)	40200
dense_103 (Dense)	(None, 1)	201

Total params: 82,001
Trainable params: 82,001
Non-trainable params: 0

Fonte: Autoria própria

A fim de representar fielmente a rede neural estudada neste presente artigo, foi feita uma breve ilustração (Figura 11) da sua forma.

Figura 11 – Representação da Rede Neural



Fonte: Autoria própria

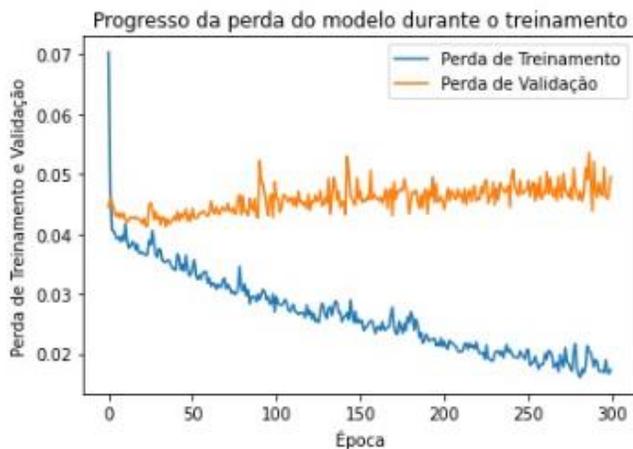
Nota-se que a rede neural de regressão especificada durante todo este artigo, possui 6 atributos de entrada: temperatura máxima, temperatura mínima, insolação, evaporação, temperatura média e umidade relativa média. 3 camadas escondidas contendo 200 neurônios em cada camada e 1 camada de saída contendo o preço.

Para se fazer o treinamento da base de dados, através da função em python “model.fit” foi ajustado para 300 épocas (épocas nada mais são do que os ciclos do evento), “batch_size” de 15 (aqui, o ajuste dos pesos são feitos de 15 em 15 registros) e uma “validation_split” de 25% ou seja, de todos os registros feitos irá selecionar apenas 25% para fazer a validação época por época enquanto o treinamento é feito.

5 - RESULTADOS E DISCUSSÕES

Após a realização do treinamento e do teste da base de dados, foi gerado um gráfico (Figura 10), este gráfico representa a evolução do modelo tanto no treinamento quanto na validação baseado nas épocas. Esse método tem como objetivo acompanhar a evolução do aprendizado nas curvas correspondentes aos subconjuntos de dados de treinamento e de validação (Figura 12). Assim sendo, o treinamento é interrompido, quando a curva de validação decresce a um erro mínimo, e antes de começar a crescer, conforme o treinamento, continua. [20]

Figura 12 – Gráfico de perda (Treinamento x Validação)



Fonte: Autoria própria

Pelo gráfico (Figura 12) pode-se notar que ambas as curvas tendem a uma estabilidade a partir da época 250 (valor aproximado), logo após isso ambas as curvas “tendem” a uma estabilidade.

5.1 – Manipulação dos resultados

A partir de todo o estudo feito até então, a fim de se fazer uma previsão, foi criada uma variável em forma de vetor contendo valores aleatórios relacionados a todos os atributos de entrada. Para melhor visualização do experimento, a Tabela 2, consegue ilustrar perfeitamente valores reais de entrada x preço (saída).

Na tabela a seguir (Tabela 2), as entradas estão representadas como:

- Temperatura Máxima (I)
- Temperatura Mínima (II)
- Insolação (III)
- Evaporação (IV)
- Temperatura Média (V)
- Umidade (VI)
- Preço real (VII)
- Preço previsto (VIII)

Tabela 2 – Valores aleatórios de entrada x saída

I	II	III	IV	V	VI	VII	VIII
24	16.4	0.7	5	18.64	88.75	244	249
27.7	17.4	8.7	1	21.18	76.25	257	232
28.5	16.4	10.2	5.4	21.26	61.75	243	264
26.7	13.2	10.2	11.2	18.94	45.25	202	204
32.2	17.6	9.8	9.5	24.24	58	277	245
30.3	19.2	9.4	4.4	24.38	67.75	283	273

Fonte: Autoria própria

A partir da Tabela 2, mostrada anteriormente, pode-se concluir que para determinados valores de entrada, a rede consegue prever o preço de saída e este está totalmente condizente com o período analisado em questão.

Para finalização do estudo, foram calculados o RMSE (Root mean square error) e o MAE (Mean absolute error), a fim de verificar a proximidade das previsões relacionadas aos valores reais.

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_i - \hat{y}_i)^2} = 8.5648$$

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_i - \hat{y}_i| = 5.4766$$

O RMSE dá-se pelo cálculo da raiz quadrada média dos erros, entre valores observados (reais) e previsões (hipóteses), já o MSE calcula o “erro absoluto médio” dos erros entre valores observados e previsões. Para ambos, o valor está entre 0 (ideal) até infinito.[21]

O RMSE mede a diferença entre valores previstos por um modelo e os valores realmente observados no ambiente que está sendo modelado. Ou seja, quanto mais próximo de 0 melhor a sua precisão.

Já o MAE traz como informação o nível de proximidade de uma linha de regressão em relação a um

conjunto de valores de dados de teste, considerando as distâncias dos pontos até a linha de regressão.

Quanto maior for a diferença entre os erros, menor a acurácia da rede [22].

6 – CONCLUSÕES

Neste presente documento foi dada uma visão sobre os diferentes tipos de Redes Neurais Artificiais, com foco nas Redes Neurais de Múltiplas Camadas, utilizando o método de Regressão Linear. O modelo traz boas similaridades (vide Tabela 2) e coerências na previsão do preço do café utilizando determinadas entradas, mas vale lembrar que a base de dados se restringe de 2003 a 2008.

Apesar de uma acurácia não ideal (erros com valores diferentes de zero), a rede conseguiu prever preços com entradas aleatórias e esses valores estão completamente dentro dos valores esperados (valor máximo e mínimo exposto na base de dados)

Após todo o processo de codificação da rede, junto aos gráficos plotados, foi possível observar uma semelhança com relação à influência das entradas na saída, ou seja, todas as condições climáticas utilizadas influenciam de forma igual no preço final do café.

Tomando-se este estudo como comparativo aos artigos aqui citados, o mesmo possui uma assertividade similar. Fazendo uma comparação com o estudo feito por José Airton Azevedo Santos, no qual utiliza uma rede CNN-BLSTM [4], o autor chega a um RMSE de 4.86 e um MAE de 4.31, porém vale lembrar que nesse e nos demais artigos feitos pelos outros autores, eles utilizam a base de dados de forma diferente, pois nelas contém apenas as datas e não fatores como condições climáticas, o que também torna este estudo único.

Além disso, o projeto por se tratar de uma grande rede neural com diversos dados, faz com que o processo de treinamento do modelo proposto seja lento, sendo de extrema importância um sistema computacional avançado e de qualidade para fazer uma análise maior e conseqüentemente mais assertiva de todo o sistema.

A partir disso depreende-se que, o modelo em si funciona bem, porém vale ressaltar que o preço do café é ditado pela quantidade do produto no mercado, além do mesmo ser referente ao ano posterior, ou seja, o preço de março de 2006 tem como principal indicador a quantidade de café que terá no mercado em 2007. Essa quantidade é baseada principalmente em fatores climáticos. O estudo aqui feito, consegue a partir de determinadas condições climáticas prever o preço sim, porém existem fatores externos (principalmente políticos) que influenciam mais ainda no preço da *commoditie*.

7. REFERÊNCIAS

- [1] QUAL o valor do café para a economia brasileira? **KUHN Brasil**, 2022. Disponível em: <<https://www.kuhnbrasil.com.br/noticias/qual-o-valor-do-caffe-para-economia-brasileira>>. Acesso em: 26 jun. 2022.
- [2] CENÁRIO e importância do café no Brasil. **Rehagro ensino**, 2022. Disponível em:

<<https://rehagro.com.br/blog/cenario-e-importancia-do-caffe-no-brasil/>>. Acesso em: 26 jun. 2022.

- [3] YONENAGA, W; FIGUEIREDO, R. Previsão do preço da soja utilizando redes neurais – ENGEP. UNIVERSIDADE FEDERAL DE SÃO CARLOS, 1999.
- [4] SANTOS, J. Previsão do café arábica: uma aplicação de redes neurais CNN-BLSTM – Universidade Tecnológica Federal do Paraná, 2022
- [5] SILVA, C. Aplicação de Redes Neurais Artificiais à Previsão do preço de energia elétrica para distintas zonas de mercados desregulamentados – DEE. Universidade Federal de São Carlos, 2018.
- [6] REDES neurais artificiais **ICMS USP**, 2022. Disponível em: < <https://sites.icmc.usp.br/andre/research/neural/>>. Acesso em: 07 ago. 2022.
- [7] O NEURÔNIO artificial **GSIGMA UFSC**, 2015. Disponível em: < https://www.gsigma.ufsc.br/~popov/aulas/rna/neuronio_artificial/index.html/>. Acesso em: 07 ago. 2022
- [8] NEURÔNIOS artificiais – uma breve introdução **BLOG GLT**, 2022. Disponível em: < <https://blog.gft.com/br/2022/02/08/neuronios-artificiais-uma-breve-introducao/>>. Acesso em: 10 ago. 2022
- [9] FUNÇÃO de ativação **DEEPLARNING BOOK**, 2022. Disponível em: < <https://www.deeplearningbook.com.br/funcao-de-ativacao/#:~:text=Uma%20rede%20neural%20sem%20fun%C3%A7%C3%A3o,e%20executar%20tarefas%20mais%20complexas>>. Acesso em: 08 ago. 2022
- [10] WHATS is learning **TIBCO**, 2022. Disponível em: < <https://www.tibco.com/pt-br/reference-center/what-is-supervised-learning>>. Acesso em: 10 ago. 2022
- [11] ROSA, J.L.G. Redes Neurais – SCC-5809. ICMC USP São Carlos, 2011.C
- [12] ADAM optimizer **ACERVO LIMA**, 2022. Disponível em: < <https://acervolima.com/intuicao-de-adam-otimizador/#:~:text=Adam%20Optimizer,menos%20mem%C3%B3ria%20e%20C3>>
- [13] OVERFITTING e underfitting **ABRACD ORG**, 2022. Disponível em: < <https://abracd.org/overfitting-e-underfitting-em-machine-learning/>>. Acesso em: 10 ago. 2022
- [14] PERCEPTRON de múltiplas camadas **EDISCIPLINAS USP**, 2022. Disponível em: < https://edisciplinas.usp.br/pluginfile.php/4445475/mod_resource/content/1/rn_5_mlp_1.pdf>. Acesso em: 20 ago. 2022
- [15] MOURA, G. Redes neurais recorrentes para a classificação de estruturas retóricas – CTDI. UNIVERSIDADE ESTADUAL DE MARINGÁ 2018.
- [16] ESTAÇÃO Meteorológica Seção Técnica de Serviços Meteorológico **IAG-USP**, 2022. Disponível em: < <http://www.estacao.iag.usp.br/>>. Acesso em 28 ago. 2022.
- [17] CEPEA, **ESALQ | USP**, 2022. Disponível em: < <https://www.cepea.esalq.usp.br/br/>>. Acesso em 28 ago. 2022.
- [18] INTRODUÇÃO as redes neurais **DEEPLARNING BOOK**, 2022. Disponível em: < [https://www.deeplearningbook.com.br/introducao-as-redes-neurais-convolucionais/#:~:text=Uma%20Rede%20Neural%20Convolutacional%20\(ConvNet,de%20diferenciar%20um%20do%20o](https://www.deeplearningbook.com.br/introducao-as-redes-neurais-convolucionais/#:~:text=Uma%20Rede%20Neural%20Convolutacional%20(ConvNet,de%20diferenciar%20um%20do%20o)>. Acesso em: 17 ago. 2022
- [19] O QUE É BIAS em redes neurais **STACK OVERFLOW**, 2022. Disponível em: < <https://pt.stackoverflow.com/questions/338825/o-que-%C3%A9-bias-em-redes-neurais#:~:text=O%20peso%20mostra%20a%20efic%C3%A1cia,adicionado%20em%20uma%20equa%C3%A7%C3%A3o%20linear>>. Acesso em 17 ago. 2022
- [20] GUIMARÃES, A; MATHIAS, I; DIAS, A; JONES, F;

JUNIOR, C. Módulo de variação cruzada para treinamento de redes neurais artificiais com algoritmos backpropagation e resiliente propagation – Universidade Estadual de Ponta Grossa, 2008.

- [21] AVALIAR o modelo ML.NET com métricas, **MICROSOFT**, 2022. Disponível em: <<https://docs.microsoft.com/pt-br/dotnet/machine-learning/resources/metrics>> Acesso em 30 ago. 2022.
- [22] RMSE ou MAE? Como avaliar meu modelo de machine learning? **LINKEDIN**, 2022. Disponível em: <<https://www.linkedin.com/pulse/rmse-ou-mae-como-avaliar-meu-modelo-de-machine-learning-rezende/?originalSubdomain=pt>>. Acesso em 30 ago. 2022.