

UNIVERSIDADE FEDERAL DE VIÇOSA
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

CAMILO COÊLHO GOMES

**MODELAGEM MATEMÁTICA EM PYTHON PARA A PREVISÃO DO
PREÇO DO CAFÉ: UM ESTUDO DOS MODELOS**

VIÇOSA
2019

CAMILO COÊLHO GOMES

**MODELAGEM MATEMÁTICA EM PYTHON PARA A PREVISÃO DO
PREÇO DO CAFÉ: UM ESTUDO DOS MODELOS**

Monografia apresentada ao Departamento de Engenharia Elétrica do Centro de Ciências Exatas e Tecnológicas da Universidade Federal de Viçosa, para a obtenção dos créditos da disciplina ELT 490 – Monografia e Seminário – e cumprimento do requisito parcial para obtenção do grau de Bacharel em Engenharia Elétrica.

Orientador: Prof. Dr. Mauro de Oliveira Prates

VIÇOSA
2019

CAMILO COELHO GOMES

**MODELAGEM MATEMÁTICA EM PYTHON PARA A PREVISÃO DO
PREÇO DO CAFÉ: UM ESTUDO DOS MODELOS**

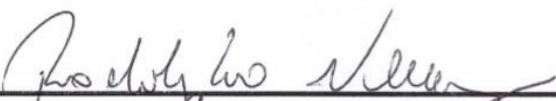
Monografia apresentada ao Departamento de Engenharia Elétrica do Centro de Ciências Exatas e Tecnológicas da Universidade Federal de Viçosa, para a obtenção dos créditos da disciplina ELT 490 – Monografia e Seminário – e cumprimento do requisito parcial para obtenção do grau de Bacharel em Engenharia Elétrica.

Aprovada em 12 de julho de 2019.

COMISSÃO EXAMINADORA



Prof. Dr. Mauro de Oliveira Prates - Orientador
Universidade Federal de Viçosa



Prof. Dr. Rodolpho Vilela Alves Neves - Membro
Universidade Federal de Viçosa



M. Sc. Tiago Zanotelli - Membro
Instituto Federal do Espírito Santo

“O tempo é o bem mais precioso que um homem pode desperdiçar.”
(Autor desconhecido)

Dedico esse trabalho a meus pais, que sempre me apoiaram nas minhas decisões, dedico aos colegas de curso, que me ajudaram e sempre tiveram ao meu lado no decorrer dessa graduação e a Deus, por ter me dado essa oportunidade.

Agradecimentos

Agradeço a todos que me auxiliaram na realização desse trabalho, mas agradeço principalmente aos professores do Departamento por terem me passado seu conhecimento pra que eu pudesse realiza-lo.

Resumo

O café é um importante produto para a economia nacional, possuindo uma parcela expressiva de nossas exportações. Sua cadeia econômica gera inúmeros empregos, sendo Minas Gerais o maior produtor. Entretanto, devido a alta volatilidade dos preços, torna-se difícil para os produtores e exportadores planejarem sua produção e seu período de vendas a fim de maximizarem seus lucros. Dessa forma, esse trabalho tem por objetivo o desenvolvimento de ferramentas matemáticas capazes de analisar o histórico dos preços do café e fazer previsões sobre o mesmo. As ferramentas escolhidas foram: ARIMA, SARIMA, RNN baseada no neurônio de McCulloch and Pitts e RNN baseada em LSTM. Obteve-se um resultado satisfatório na previsão com todas as metodologias, concluindo ser possível realizar previsões nos preços da mesma.

Sumário

Lista de Figuras	10
Lista de Abreviações	11
Lista de Tabelas	12
1 Introdução	13
1.1 Panorama mundial e nacional do Café	13
1.2 O mercado do café	14
1.3 Objetivos	15
1.4 Organização do Trabalho	15
2 Revisão Literária	16
2.1 Fundamentos do Mercado Financeiro	16
2.1.1 Hipótese dos Mercados Eficientes (HME)	16
2.2 Modelos Matemáticos	16
2.2.1 Modelos Estatísticos	16
2.2.1.1 Modelo Auto Regressivo Integrado Média Móvel (ARIMA)	16
2.2.1.2 Modelo Sazonal Auto Regressivo Integrado Média Móvel (SARIMA)	17
2.3 Modelos Baseados em Redes Neuraís Recorrentes (Recurrent Neural Networks – RNN).....	17
2.3.1 Modelo de McCulloch e Pitts	18
2.3.2 Modelo Long Short-Term Memory	19
3 Metodologia	21
3.1 Análise e manipulação dos dados	21
3.2 Etapas para modelagem de modelos estatísticos.....	24
3.2.1 Estimar as ordens do modelo	24
3.2.2 Estimar os coeficientes do modelo	25
3.3 Etapas de modelagem de uma RNN	25
4 Calculo dos Modelos	27
4.1 ARIMA(p,d,q).....	27
4.2 SARIMA(p,d,q)(P,D,Q,S).....	28
4.3 RNN McCulloch and Pitts	29

4.4	<i>RNN LSTM</i>	29
5	Resultados	30
6	Conclusão.....	32
7	<i>Referências</i>	33

Lista de Figuras

Figura 1 - Produção Mundial e Nacional de Café	13
Figura 2 - Variação do preço do café	14
Figura 3 - Diagrama de blocos de uma RNN	18
Figura 4 - Diagrama de blocos do neurônio de McCulloch e Pitts.....	18
Figura 5 - Saída função sigmoide(σ)	19
Figura 6 - Diagrama de blocos do modelo LSTM.....	19
Figura 7 - Saída da função tangente hiperbólica (tanh).....	20
Figura 8 - Preço do café em dólares por saca, intervalo de treinamento e teste.....	21
Figura 9 - Decomposição Clássica da Série	23
Figura 10 - Dados para o treinamento das RNNs	24
Figura 11 - Exemplo de treinamento RNN.....	26
Figura 12 - Autocorreção e correlação parcial da série temporal.....	27
Figura 13 Saída do modelo ARIMA determinado.....	28
Figura 14 Saída do Modelo SARIMA	29
Figura 15 Resultados dos Modelos para o período previsto.....	31

Lista de Abreviações

USDA	United States Department of Agriculture
OEC	Observatory of Economic Complexity
ARIMA	Auto Regressivo Integrado Média Móvel
SARIMA	Sazonal Auto Regressivo Integrado Média Móvel
RNN	Recurrent Neural Networks
LSTM	Long Short-Term Memory
HME	Hipótese dos Mercados Eficientes
ADF	Dickey-Fuller Aumentado
ADAM	Adaptive Moment Estimation

Lista de Tabelas

Tabela 1 - MSE dos Modelos Selecionados	30
---	----

1 Introdução

1.1 Panorama mundial e nacional do Café

Conforme dados do Departamento de Agricultura dos Estados Unidos, *United States Department of Agriculture* (USDA), O Brasil é atualmente o maior produtor de café mundial, com uma safra estimada para o ano de 2019 de 41 milhões de sacas de 60 quilogramas de café, representando assim 36% da produção mundial de café estimada para o ano (United States Department of Agriculture, 2019), tendo também desde 1974 uma representação expressiva na produção, conforme Figura 1.

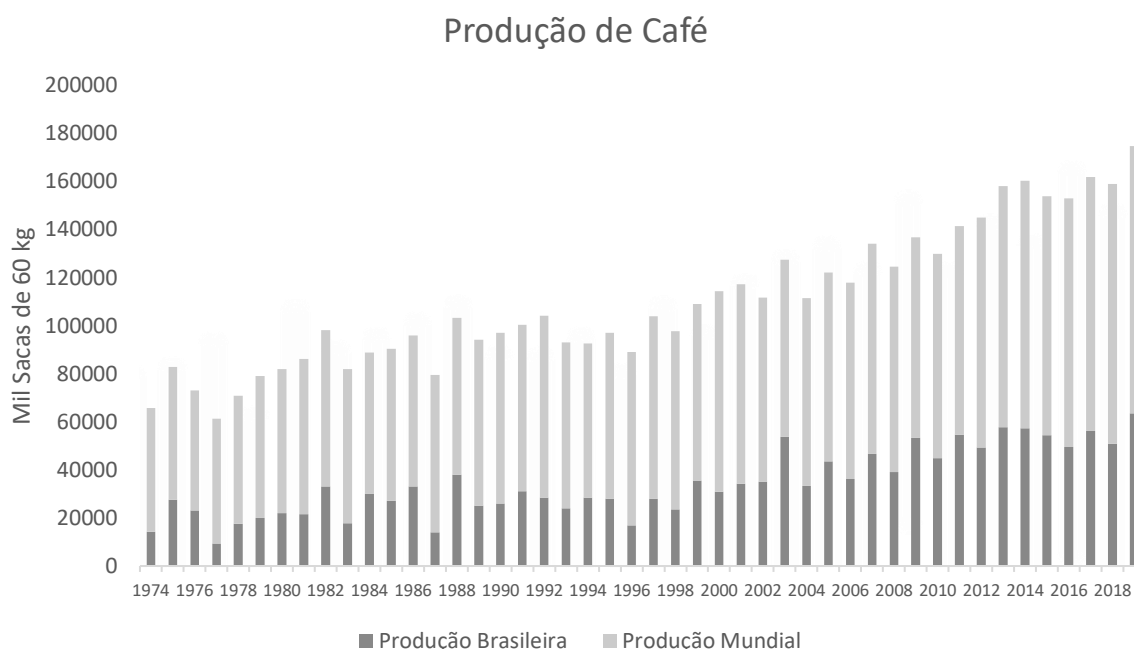


Figura 1 - Produção Mundial e Nacional de Café

Fonte: (United States Department of Agriculture, 2019)

O café também é responsável, segundo o Observatório de Complexidade Econômica – MIT, *Observatory of Economic Complexity* (OEC), pela entrada de 4,86 Bilhões de Dólares, através da exportação da *commodity*, representando assim 2,2% do valor de todas as exportações e sendo a segunda *commodity* agrícola de mais peso em nossa balança comercial, perdendo apenas para a soja, que representa 12% (AJG Simoes, 2019), sendo Minas Gerais o maior produtor (Ministério da Agricultura, Pecuária e Abastecimento, 2018).

Além do peso econômico, o café também é um grande gerador de renda e emprego para diversas famílias, sendo responsável pela geração de 8 milhões de empregos no país, segundo o Ministério da Agricultura, Pecuária e Abastecimento. (Ministério da Agricultura, Pecuária e Abastecimento, 2018).

1.2 O mercado do café

O preço do café é altamente afetado por eventos climáticos, que podem prejudicar muito a produção, ou causar anos de safra recordes, o que influencia direto no preço (DaMatta & Ramalho, 2006). Além dos fatores climáticos, variáveis de mercado também tem um alto impacto na formação dos preços de café, como: níveis de produção, consumo, ritmo de exportações, estoques mundiais e até mesmo o câmbio, tendo em vista que o contrato do café *commodity* é negociado em dólar (Corrêa & Raíces, 2005).

Apesar de todos esses fatores, existe a hipótese do mercado eficiente que diz que os preços de um determinado ativo reflete inteiramente as informações disponíveis sobre os mesmos (Junior & Ikeda, 2004). Dessa forma, neste trabalho somente os preços do café, mostrados na Figura 2, são levados em consideração para a criação dos modelos, adquiridos em (Macrotrends, 2019).

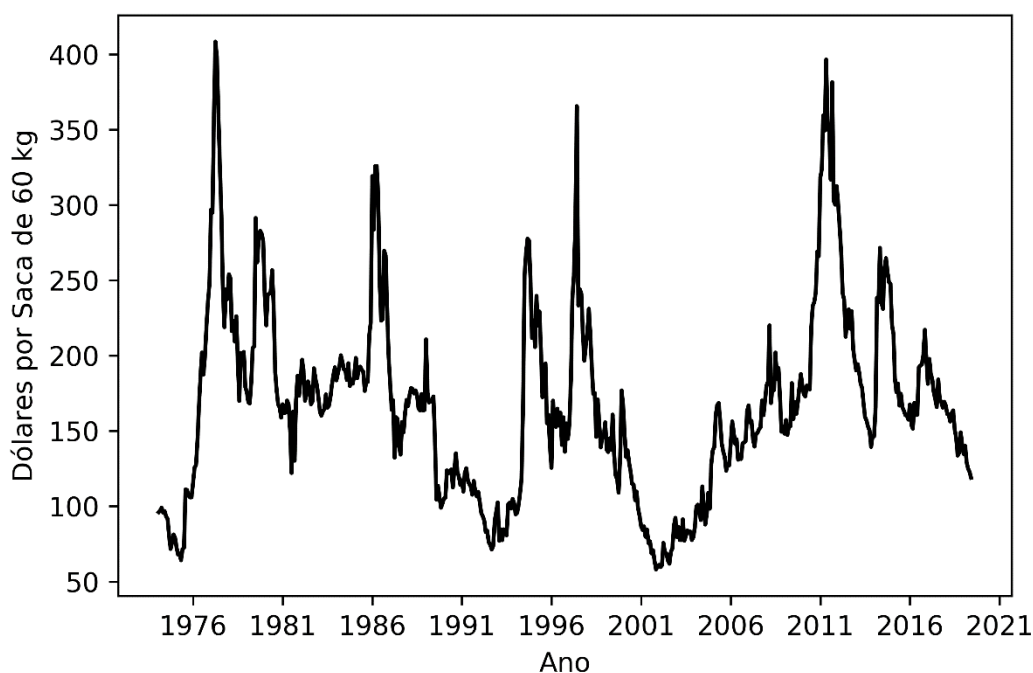


Figura 2 - Variação do preço do café

Portanto, estudos econométricos para a previsão dos preços do café se fazem necessários para que os produtores e exportadores possam desenvolver estratégias de mercados mais adequadas às situações vigentes do mercado, maximizando assim seus lucros e evitando possíveis prejuízos (Gutierrez & Almeida, 2013).

Devido a alta volatilidade dos preços do café no mercado mundial e o alto impacto que o esse produto possui em nossa balança comercial, um estudo mais detalhado de como seu preço se comporta se faz necessário para o planejamento estratégico de diversos modelos de negócio.

1.3 Objetivos

Esse trabalho tem como objetivo geral desenvolver modelos confiáveis e aplicáveis para à previsão do preço do café, utilizando os métodos ARIMA, SARIMA, RNN com neurônio de McCulloch and Pitts, e RNN com neurônio LSTM.

Tendo como objetivos específicos comparar os modelos utilizados, e identificar a melhor opção dentre eles para a previsão dos preços do café.

1.4 Organização do Trabalho

Este trabalho está dividido em seis capítulos, sendo o primeiro uma contextualização do mercado de café, mostrando os objetivos do estudo.

O segundo capítulo traz a revisão bibliográfica dos modelos e como eles serão definidos, enquanto que no terceiro capítulo mostra a metodologia desenvolvida para se fazer a análise dos dados. Já o quarto capítulo apresenta a criação dos modelos sugeridos, enquanto que o quinto capítulo mostra os resultados obtidos. Por fim, o capítulo 6 apresenta as conclusões finais.

2 Revisão Literária

2.1 Fundamentos do Mercado Financeiro

2.1.1 Hipótese dos Mercados Eficientes (HME)

A HME pressupõe que no processo de tomada de decisões, o homem é capaz de analisar todas as informações disponíveis e considerar todas as hipóteses para a solução do problema (*Junior & Ikeda, 2004*), (*Malkiel & Fama, 1970*).

2.2 Modelos Matemáticos

A análise de séries temporais lida com a avaliação e análise de um determinado dado variante no tempo com o objetivo de realizar previsões futuras para o mesmo (Keiel & Bender, 2018). Os modelos aqui definidos serão subdivididos em duas categorias, os modelos estatísticos e os modelos baseados em redes neurais.

2.2.1 Modelos Estatísticos

Os modelos estatísticos são assim denominados por fazerem auto uso de ferramentas estatísticas, tais como médias móveis e auto regressões. Para a esse trabalho foi considerado os modelo ARIMA e SARIMA, por serem muito difundidos e possuírem uma alta robustez matemática (Andrade, 2009).

2.2.1.1 Modelo Auto Regressivo Integrado Média Móvel (ARIMA)

A metodologia ARIMA (p,d,q) é utilizada para análise de séries temporais no domínio do tempo, onde combina uma componente de auto regressão de ordem p, uma componente integradora de ordem d, e uma componente de médias móveis de ordem q. Sua fórmula geral pode ser descrita conforme (2.1).

$$\sum_{j=0}^p \phi_j Z_{t-j} = \beta_0 X_t + \sum_{j=1}^d \beta_j X_{t-j} + \sum_{j=0}^q \theta_j \alpha_{t-j}, \quad (2.1)$$

onde Z é um vetor de saídas observável, X é um vetor de entradas observáveis, α é um ruído branco, ϕ são as constantes da componente auto regressiva, β as constantes da componente integradora e θ as constantes da componente de médias móveis (Andrade, 2009).

2.2.1.2 *Modelo Sazonal Auto Regressivo Integrado Média Móvel (SARIMA)*

O modelo $SARIMA(p, d, q)(P, D, Q)_S$ é composto por um modelo ARIMA multiplicado por uma componente sazonal, onde P é a ordem da componente auto regressiva da componente sazonal, D é a ordem da componente integrativa da componente sazonal e Q é a ordem da componente de média móvel da componente sazonal, sendo seu modelo geral descrito conforme (2.2).

$$\phi(B)\Phi(B^S)(1 - B^S)^D(1 - B)^d Z_t = \theta(B)\Theta(B^S)\alpha_t, \quad (2.2)$$

onde B^S é um operador de atraso, ou seja $(B^S) N_t = N_{t-S}$, $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ e $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$, $\Phi(B^S) = 1 - \Phi_1 B^S - \dots - \Phi_p B^{pS}$ e $\Theta(B^S) = 1 - \theta_1 B^S - \dots - \theta_p B^{pS}$ (Morettin & Tolo, 2004).

2.3 *Modelos Baseados em Redes Neurais Recorrentes (Recurrent Neural Networks – RNN)*

A rede neural recorrente é uma estrutura capaz de representar uma grande variação de comportamentos dinâmicos e variantes no tempo. A presença da realimentação de informação permite acrescentar dispositivos de memória na mesma, tornando-a excelente no processamento

de sinais temporais sequenciais. Sua representação por diagrama de blocos pode ser verificada na Figura 3 (Zuben, 2005).

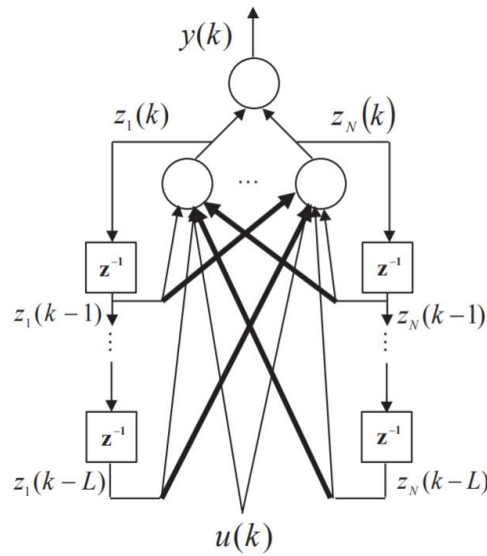


Figura 3 - Diagrama de blocos de uma RNN

Em sua estrutura, as redes neurais são compostas por diversos neurônios distribuídos e separados em camadas. Para esse trabalho, utilizou-se duas topologias de células neurais: o neurônio de McCulloch e Pitts, e o modelo *long short-term memory* (LSTM).

2.3.1 Modelo de McCulloch e Pitts

O modelo da célula neural de McCulloch e Pitts, desenvolvido em 1943, pode ser representado pelo diagrama de blocos da Figura 4, onde X_t é o vetor de entrada, Y_{t-1} é o vetor de saídas anteriores, Y_t é o vetor de saídas, B é o vetor de limiar de ativação. A_1 e A_2 são vetores das constantes dos pesos sinápticos (McCulloch & Pitts, 1943), (Zuben, 2005).

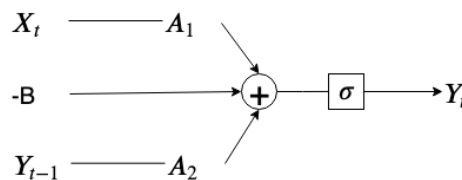


Figura 4 - Diagrama de blocos do neurônio de McCulloch e Pitts

A resposta final do neurônio é passada por uma função de ativação. Existem diversas funções de ativação na literatura, conforme demonstra (Carvalho, 2019), sendo que devido a

natureza sempre positiva dos preços do café, escolheu-se a função sigmoide, equação 2.3, pois sua saída é contínua, derivável e sempre positiva, conforme demonstra a Figura 5. Fazendo com que esse *perceptron* seja chamado de sigmoide.

$$\sigma(u) = \frac{1}{1+e^{-u}}. \quad (2.3)$$

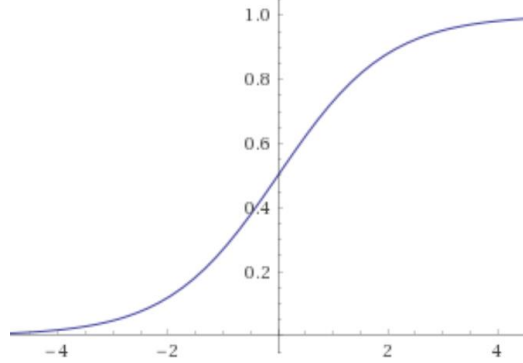


Figura 5 - Saída função sigmoide(σ)

2.3.2 Modelo Long Short-Term Memory

Foi proposto por (S. Hochreiter, 1987) um modelo onde a saída atual depende da saída anterior, mais um fator de memória de longo prazo, definido por C_t , como demonstra o diagrama de blocos da Figura 6. Definindo assim o modelo de neurônio LSTM.

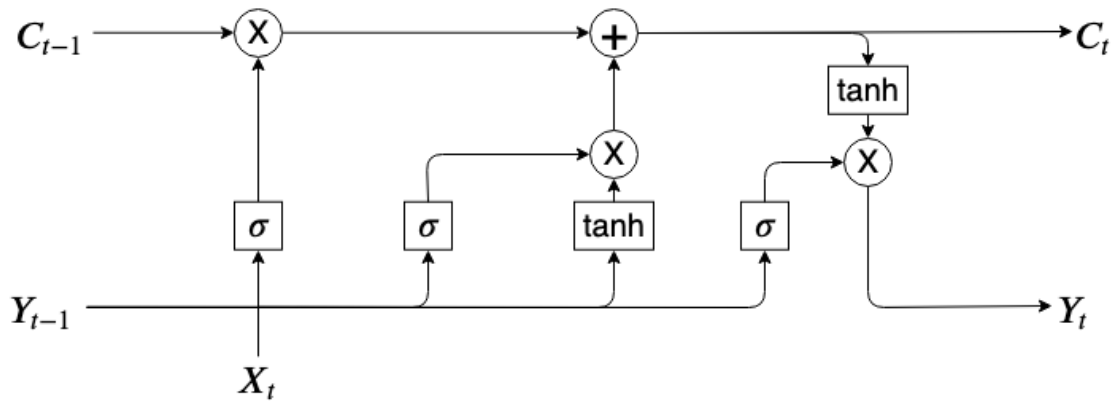


Figura 6 - Diagrama de blocos do modelo LSTM

Para o modelo LSTM, além da função de ativação sigmoide(σ), descrita em (2.3), o modelo apresenta uma segunda função de ativação, a tangente hiperbólica ($\tanh$), definida em (2.4) e sua saída pode ser verificada na Figura 7

$$\tanh(u) = \frac{e^u - e^{-u}}{e^u + e^{-u}}. \quad (2.4)$$

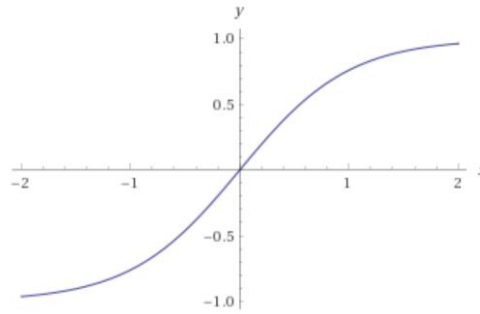


Figura 7 - Saída da função tangente hiperbólica ($\tanh$)

3 Metodologia

Esse capítulo mostra a metodologia do trabalho. Ou seja, são apresentados como foi feita a análise e manipulação dos dados e as etapas para se chegar aos parâmetros corretos dos modelos.

3.1 Análise e manipulação dos dados

Primeiro precisa-se garantir um número de dados iguais para todos os modelos matemáticos, separando uma porção deles para realizar o desenvolvimento dos modelos matemáticos e outra para testar os modelos desenvolvidos.

Conforme mostra a Figura 8, os dados foram separados entre um intervalo de treinamento, para efetuar a modelagem, que são os dados de janeiro de 1974 até dezembro de 2018, e os dados de teste, referentes a janeiro de 2019 até maio de 2019, ambos com periodicidade mensal.

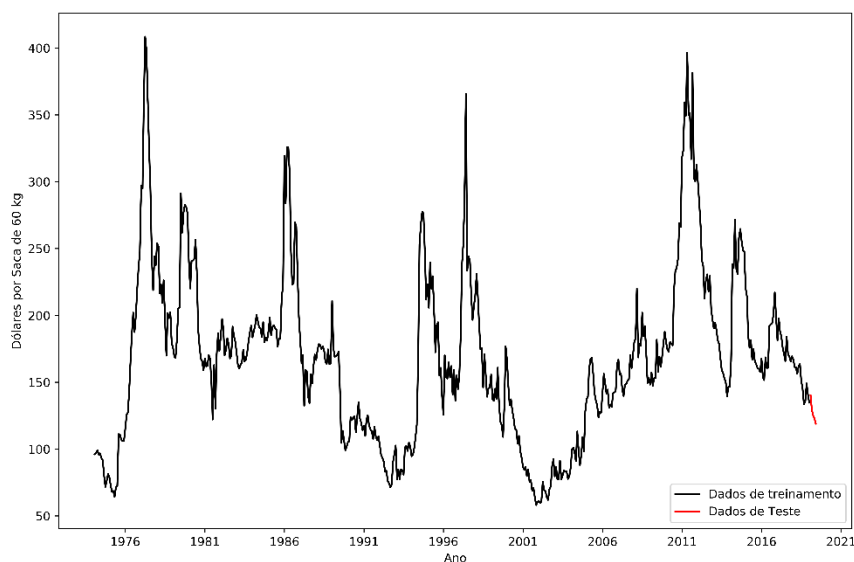


Figura 8 - Preço do café em dólares por saca, intervalo de treinamento e teste

Após separado, se fez necessário verificar a sazonalidade e a tendência geral da componente de treinamento, tendo em vista que para realizar a modelagem por meio dos

modelos estatísticos ARIMA e SARIMA, se faz necessária que a série seja estacionária. Para isso, aplicou-se o teste de Augment Dickey-Fuller (ADF) (*Margarido & Junior, 2009*).

O teste ADF possui hipótese nula. Quando a série possui raiz unitária é dita estacionária, caso não tenha, é dita não estacionária. O teste é definido pela equação (3.1)

$$y_t = py_{t-1} + u_t, -1 \leq p \leq 1, \quad (3.1)$$

onde u_t é um ruído branco. Quando $p = 1$ é dito que a série é não estacionária, quando $p < 1$ a série é considerada estacionária (*Fracaro, 2018*).

Calculando o p para a série em estudo, encontra-se um valor de 0,00072, portanto, segundo o teste de ADF, a série é estacionária.

Além disso para que possa ser aplicado o modelo SARIMA, é necessário garantir ainda que a série tenha uma componente sazonal. Para tal, realiza-se uma decomposição sazonal na série apresentada por (*Bouzada, 2012*), que pode ser verificada na Figura 9. Uma componente sazonal de periodicidade anual pode ser verificada no teste.

Pode ser observado no gráfico de sazonalidade que o preço possui uma componente sazonal de periodicidade anual bem expressiva, como o intervalo de tempo trabalhado é mensal considera-se uma sazonalidade de 12 amostras na série

Garantido que os dados estão preparados para a aplicação dos modelos ARIMA e SARIMA, precisou prepará-los para a aplicação dos modelos por redes neurais. Para isso, é necessário que os mesmos estejam normalizados entre 0 e 1, de acordo com transformação descrita na equação 3.2 (Luciano Ferreira, 2005).

$$x_n = \frac{X_n - \min(X)}{\max(X) - \min(X)}, \quad (3.2)$$

onde x_n é o novo valor do elemento da série, X_n é o valor correspondente daquele elemento, $\min(X)$ é o menor valor da série, e $\max(X)$ é o maior valor da série. Dessa forma os dados são apresentados à rede neural para treinamento conforme a Figura 10.

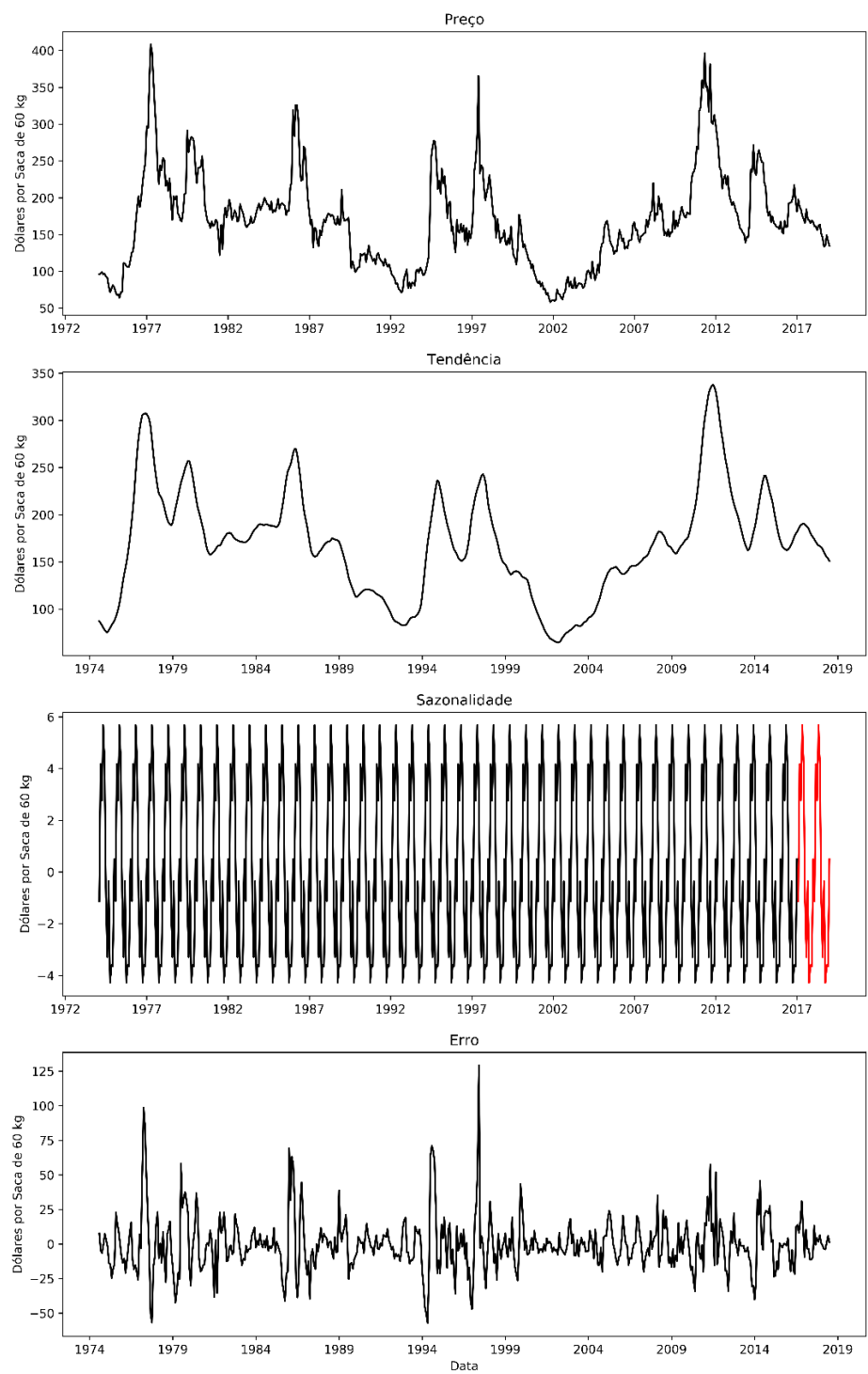


Figura 9 - Decomposição Clássica da Série

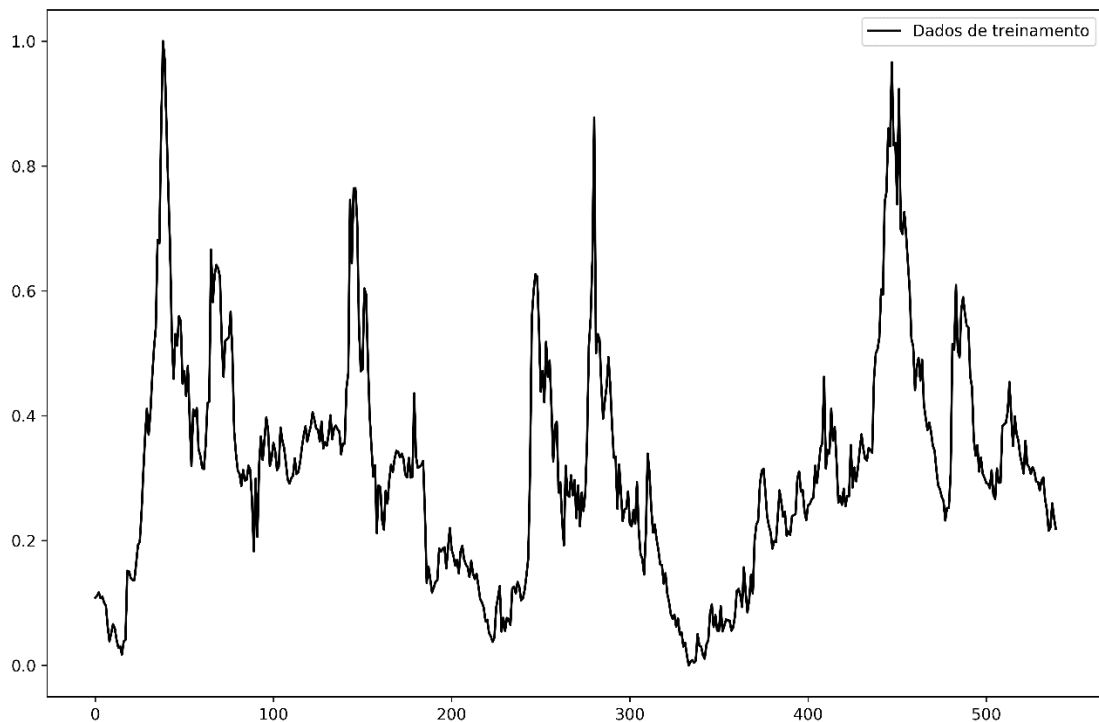


Figura 10 - Dados para o treinamento das RNNs

3.2 Etapas para modelagem de modelos estatísticos

3.2.1 Estimar as ordens do modelo

Segundo (Morettin & Toloi, 2004), é necessário garantir que as séries utilizadas na análise sejam estacionárias, caso não sejam, faz-se a derivada das mesmas, a fim de se obter séries estacionárias que não possuam tendência.

Garantido que as séries temporais analisadas sejam estacionárias, precisa-se estimar os valores de p, d e q para o modelo ARIMA; e P, D, Q e S para o modelo SARIMA. Para isso, foi realizado uma análise das funções de auto correlação e auto correlação parcial das séries temporais em estudo (Universidade de Duke, 2019).

Caso a série tenha uma auto correlação positiva para um grande número de atrasos, considere fazer uma diferenciação na mesma, na qual o número de diferenciações será o valor de d . Caso a auto correlação parcial tenha uma queda brusca, a partir do atraso K , o número de termos da componente auto regressiva será igual a K , portanto $p = K$. Caso a função de

correlação apresente uma queda brusca a partir de um atraso D , considere D termos na componente de médias móveis, portanto $q = D$ (*Universidade de Duke, 2019*).

3.2.2 Estimar os coeficientes do modelo

Existem diversos algoritmos para estimar os coeficientes de um modelo estatístico. Nesse trabalho, escolheu-se o Método maximum likelihood estimation (MLE), para determinar os coeficientes (*White, 1982*).

3.3 Etapas de modelagem de uma RNN

Para realizar o treinamento da RNN é utilizado um algoritmo de *deep learning*, ou seja, primeiramente deve-se definir o número de camadas escondidas, e em cada camada o seu número de neurônios.

Os dados do treinamento devem ser divididos entre dados de modelagem e dados de verificação, sendo que, neste trabalho, os dados de verificação para ambas as redes foram os preços do café no período de janeiro de 2019 a maio de 2019, enquanto que os dados de treinamento foram de janeiro de 1974 até dezembro de 2018.

Diferentemente de uma RNA (Rede Neural Artificial), a RNN tem um processo de treinamento diferente, pois nem todos os dados são apresentados de uma única vez. Neste caso, apresenta-se um número de dados pré-definidos, chamado de dados de análise, e a partir deles é pedido os valores dos dados subsequentes.

Nesse trabalho, analisou-se os dados de 12 amostras de sequências, um ano, e foram solicitados os valores das 5 amostras consequentes àquele período. Esses valores foram

definidos a partir dos dados de teste e da característica sazonal anual da série. Um exemplo pode ser visto na Figura 11.



Figura 11 - Exemplo de treinamento RNN

Com os dados separados e com a rede definida a rede, utiliza-se a seguir um algoritmo de treinamento, e caso ela falhe em convergir, deve-se aumentar os neurônios nas camadas escondidas. Caso ela convirja, deve-se aplicar à rede os dados do intervalo de teste e verificar o erro da mesma, repetindo assim o processo até a mesma convergir (**Nelson, 2017**).

O algoritmo de treinamento escolhido foi o ADAM (Adaptive Moment Estimation), pois como demonstra os autores em (**Kingma & Ba, 2015**), ele converge mais rápido para sistemas de *deep learning*, que nesse trabalho terá por objetivo minimizar o erro quadrático médio do modelo.

4 Cálculo dos Modelos

4.1 ARIMA(p,d,q)

Como observado no capítulo anterior, a série estudada já é estacionária, não sendo necessário a efetuação de nenhuma manipulação nos dados. Assim sendo, ao plotar a autocorrelação e a correlação parcial, observa-se que o modelo deve ter uma componente auto regressiva presente, enquanto a integral e a média móvel são mais fracas, conforme mostra a Figura 12.

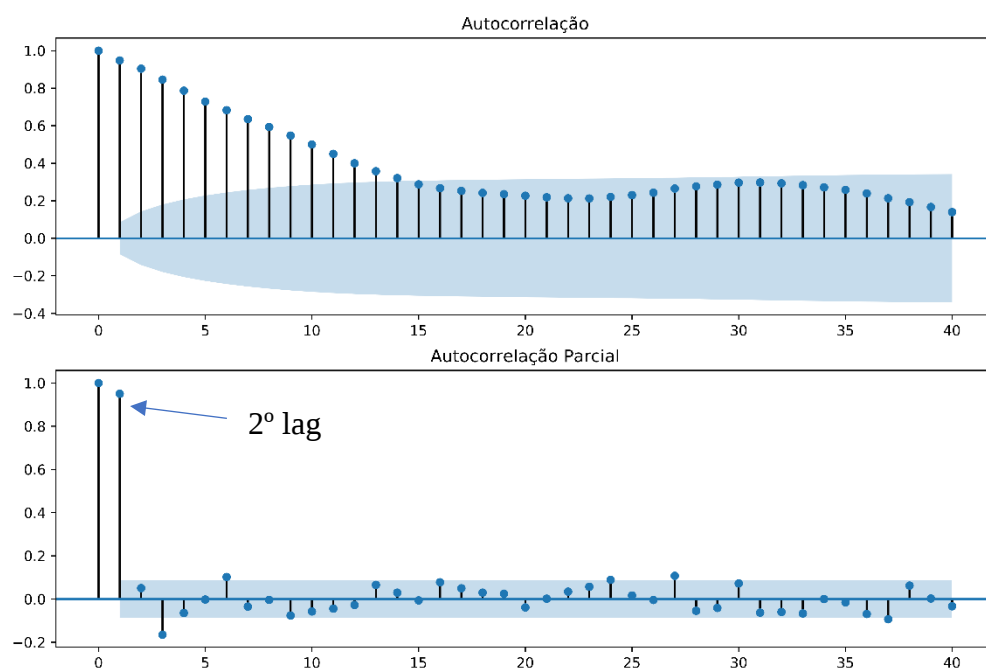


Figura 12 - Autocorreção e correlação parcial da série temporal

Como a auto correlação não apresentou um declínio súbito em seus valores, pode-se concluir pouca influência de uma componente de médias móveis. Já a auto correlação parcial mostrou um declínio súbito já no 2º lag, portanto a série pode ser descrita por um modelo ARIMA(2,1,1).

Aplicando então o método MLE para estimar as componentes auto regressivas, integradoras e de média móvel, obtém-se o resultado do modelo e a curva de densidade do erro, que podem ser observadas na Figura 13.

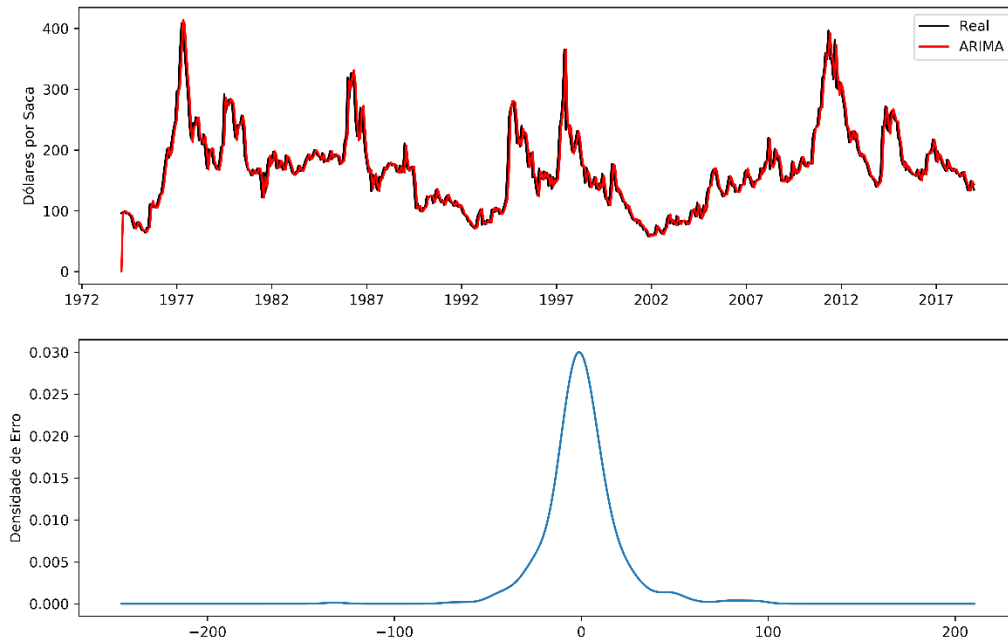


Figura 13 Saída do modelo ARIMA determinado

4.2 $SARIMA(p,d,q)(P,D,Q,S)$

A série ARIMA já foi definida como sendo um $ARIMA(2,1,1)$, assim sendo, e também pela decomposição sazonal, observou-se uma componente sazonal de periodicidade anual, com período igual a 12. Dessa forma, calcula-se os coeficientes para um modelo $SARIMA(2,1,1)(2,1,1,12)$ para a série estudada, obtendo uma saída e uma densidade de erro em relação aos dados de modelagem conforme Figura 14.

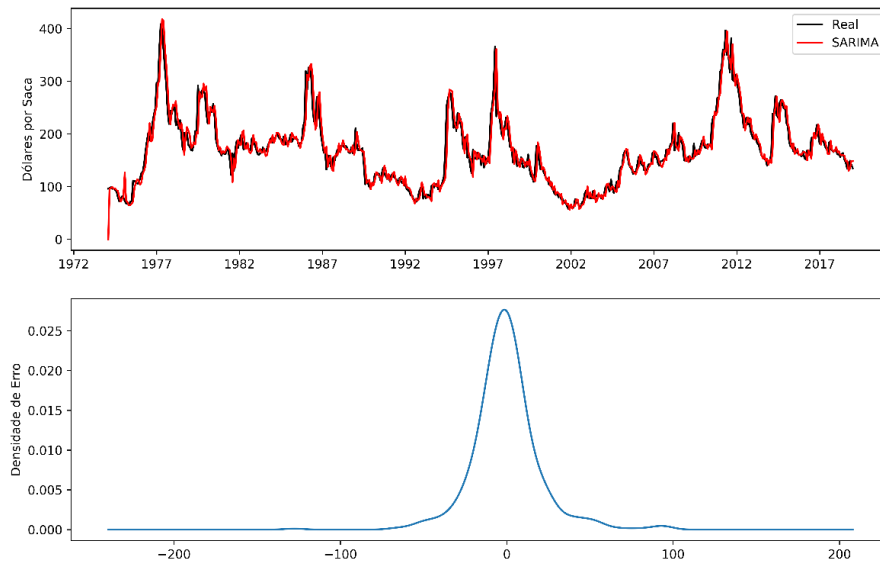


Figura 14 Saída do Modelo SARIMA

4.3 RNN McCulloch and Pitts

Antes de iniciar o treinamento da rede neural, precisa-se definir o critério de parada do mesmo. Para esse trabalho, utilizou-se dois critérios: erro quadrático médio menor que 10^{-3} e taxa de variação do erro quadrático médio menor que 10^{-5} .

O número de neurônios escolhido é igual ao número de dados que pretende-se observar no futuro, portanto, para esse trabalho definiu-se como sendo 5 neurônios em série, cada um dando a saída de $t+1$ até $t+5$.

Assim sendo, após 5767 iterações obteve-se um erro quadrático médio de $6,7 \times 10^{-4}$ e uma taxa de variação no erro de $7,6 \times 10^{-7}$.

4.4 RNN LSTM

Considerando o mesmo número de neurônios que o *McCulloch and Pitts*, e também o mesmo critério de treinamento, com o mesmo algoritmo, observou-se um erro quadrático médio de $8,7 \times 10^{-4}$ e uma taxa de variação do erro quadrático médio de $1,2 \times 10^{-6}$, sendo que esses valores foram encontrados apenas após 9872 iterações.

5 Resultados

Aplicando os modelos para previsão dos dados no período selecionado, obtém-se a saída conforme a Figura 15, observa-se o modelo neural LSTM foi o mais assertivo na tendência geral do ativo, como pode ser observado na Tabela 1, entretanto, todos os outros modelos foram mais assertivos a curto prazo. Isso ocorre devido a característica de memória de longa prazo e fator de esquecimento presente na célula.

Tabela 1 - MSE dos Modelos Selecionados

Modelo	MSE
ARIMA	1,84%
SARIMA	2,28%
McCulloch and Pitts	9,12%
LSTM	0,09%

Seguindo a análise de sazonalidade, seria esperado um aumento dos preços do café no período analisado, entretanto, devido a tendência de queda mais recente nos preços, a componente de esquecimento do LSTM estava desconsiderando a informação de sazonalidade do período.

Os modelos ARIMA e SARIMA obtiveram respostas similares, desconsiderando a componente sazonal de início, mas logo em seguida levaram a mesma em consideração, o que causou que seus valores se distanciassem do resultado esperado.

Seguindo a componente sazonal da série, que esperava-se um aumento natural dos preços para os meses de junho, a RNN baseada no neurônio de McCulloch and Pitts atribui um peso ainda mais que a ARIMA e a SARIMA a componente sazonal, que previa um aumento nos preços para o período, como pode ser observado em seu comportamento na Figura 9.

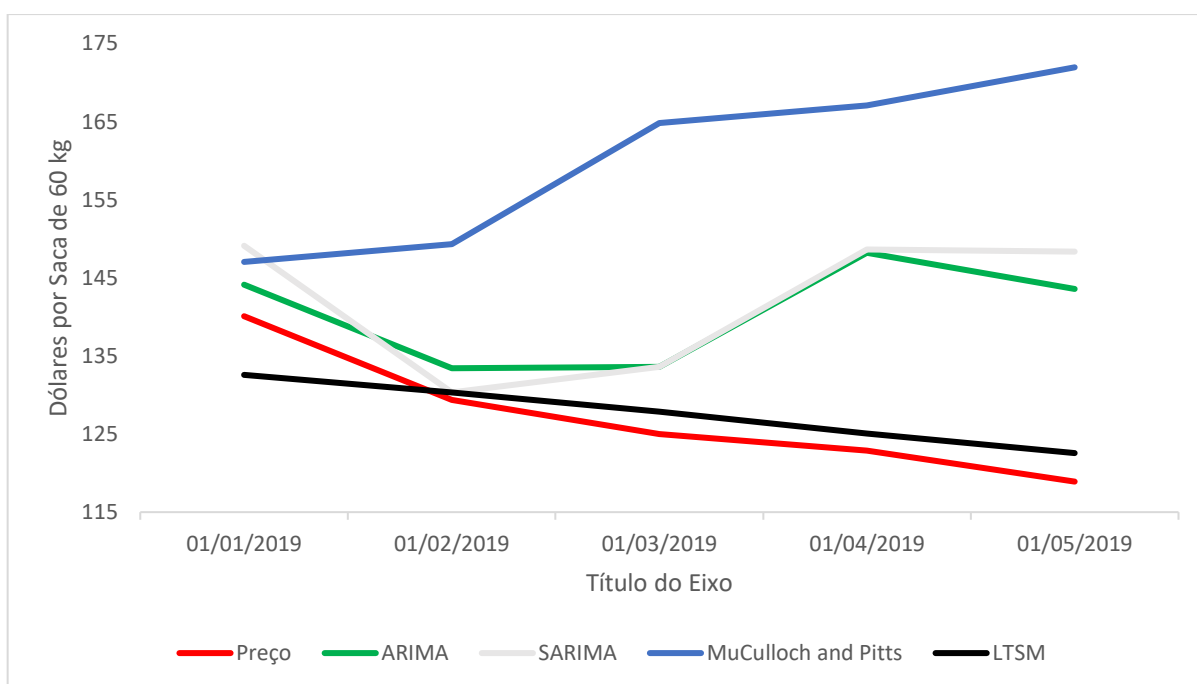


Figura 15 Resultados dos Modelos para o período previsto

O MSE para cada modelo pode ser verificado na Tabela 1, demonstrando que o modelo com um menor erro foi o baseado no algoritmo LTSM. Os modelos ARIMA e SARIMA tiveram resultados similares, e o modelo de McCulloch and Pitts, apesar de ter sido o mais assertivo na primeira amostra, foi o que teve um desempenho pior no decorrer do tempo.

6 Conclusão

Este trabalho apresentou um estudo sobre as variações do preço do café, utilizando quatro técnicas de modelagem matemática conhecidas na literatura: ARIMA, SARIMA, RNN baseada no neurônio de McCulloch and Pitts, e RNN baseada no neurônio LSTM.

Os quatro modelos estudados tiveram performances diferentes em relação à previsão do preço, sendo que a RNN baseada no neurônio de LSTM obteve um melhor resultado de modo geral. Isso ocorreu devido ao seu fator de esquecimento de dados, e também pelo fato do café estar em uma tendência muito forte de queda, já que nos últimos meses o preço estava caindo muito rápido. A rede neural decidiu por desconsiderar a componente sazonal do preço, prevendo um declínio acentuado no preço.

O modelo ARIMA também obteve bons resultados, apesar da mesma não levar em consideração a componente sazonal do preço, prevendo uma queda, porém menos acentuada do que a que de fato ocorreu.

O modelo SARIMA, devido a sua componente sazonal, e pelo período analisado, ser um período historicamente falando de alta natural do preço, portanto a componente sazonal teve um peso superior fazendo com que as previsões fossem maiores do que as que de fato ocorreram.

A RNN baseada no neurônio de McCulloch and Pitts, assim como a SARIMA, esperava uma alta para o período analisado, devido a componente sazonal do preço.

De uma maneira geral, pode-se concluir que todos os métodos de modelagem obtiveram resultados satisfatórios. No entanto, devido a um comportamento anormal dos preços no período, os modelos que não levavam em consideração a componente sazonal, ou que podiam ignorá-la em determinados momentos, obtiveram resultados mais próximos do que de fato ocorreu. Portanto cabe ao produtor ou exportador usar de seu conhecimento e experiência no mercado, para entender como estão as outras variáveis do mercado e a partir desse conceito decidir qual ferramenta será melhor aplicada para a situação atual.

Devido à forte queda momentânea no preço, acredito que o modelo LSTM seja mais adequado.

Para estudos futuros pode-se considerar aplicar essas técnicas nas variações de outras commodities, como a soja, o milho ou o boi gordo, criando modelos de previsão de preço para que os agentes do mercado desses setores do agronegócio possam se preparar para as possíveis oscilações no preço.

7 Referências

- AJG Simoes, C. H. (02 de Maio de 2019). Acesso em 12 de Junho de 2019, disponível em The Economic Complexity Observatory: An Analytical Tool for Understanding the Dynamics of Economic Development.: <https://atlas.media.mit.edu/en/profile/country/bra/>
- Andrade, W. Ç. (2009). *Estimação de Modelos ARIMA/ARIMAX e Aplicação em Inferência de Perdas de Propano*. Natal: Universidade Federal do Rio Grande do Norte.
- Bouzada, M. A. (2012). Aprendendo Decomposição Clássica: Tutorial para um Método de Análise de Séries Temporais. *ANPAD*, 1-18.
- Carvalho, A. P. (4 de Março de 2019). *Redes Neurais Artificiais*. Acesso em 05 de Maio de 2019, disponível em ICMC USP: <http://conteudo.icmc.usp.br/pessoas/andre/research/neural/>
- Corrêa, A. L., & Raíces, C. (2005). *Derivativos Agrícolas*. São Paulo: Globo.
- DaMatta, F. M., & Ramalho, J. D. (2006). Impacts of drought and temperature stress on coffee physiology and production: a review. *Brazilian Journal of Plant Physiology*, 18, 55-81.
- Ferrari, S. M. (2006). *Identificação online de motores d e indução através de modelo discreto para sinais senoidais*. Universidade Estadual de Santa Catarina.
- Fracaro, N. (2018). *Estacionariedade das Séries Temporais do Modelo Matemático ARIMAX de Propulsores Eletromecânicos*. Ijuí: Universidade Regional do Noreoeste do Estado d Rio Grande do Sul - UNIJUÍ.
- Gutierrez, C. E., & Almeida, F. M. (2013). MODELAGEM E PREVISÃO DO PREÇO DO CAFÉ BRASILEIRO. *Revista de Economia - Universidade Federal do Paraná*, 39, 7-27.
- Hochreiter, S., & Schmidhuber, J. (1987). Long Short-Term Memory. *Neural Computation*, 9, 1735 - 1780.
- Junior, T. S., & Ikeda, R. H. (2004). Mercados eficientes e arbitragem: um estudo sob o enfoque das finanças comportamentais. *Revista Contabilidade & Finanças*, 97-107.

- Keiel, G., & Bender, F. A. (2018). Modelagem de séries temporais financeiras: uma abordagem estatística para a identificação de modelos de média condicional. *Scientia cum Industria*, 22.
- Kingma, D. P., & Ba, J. L. (2015). ADAM: A Method For Stochastic Optimization. *ICLR*, 1-15.
- Leticia Biagi Vilela, T. H. (2016). APLICAÇÃO DE REDE NEURAL NARX PARA A PREVISÃO DO PREÇO DA SOJA. *XXXVI ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO*. João Pessoa/PB.
- Luciano Ferreira, G. L.-S. (2005). Utilização de Redes Neurais como Estratégia de Previsão de Preços no Agronegócio. *SEGeT – Simpósio de Excelência em Gestão e Tecnologia*.
- Macrotrends. (6 de Maio de 2019). *Macrotrends*. (Macrotrends) Acesso em 6 de Maio de 2019, disponível em <https://www.macrotrends.net/2535/coffee-prices-historical-chart-data>
- Malkiel, B. G., & Fama, E. F. (1970). Efficient Capital Markets; a review of theory and empirical work. *Journal of Finance*.
- Margarido, M. A., & Junior, H. d. (2009). Teste para mais de uma Raíz Unitária: Uso do software sas na elaboração de uma Rotina para o teste Dicey-Pantula. *Pesquisa e Debate*, 149-170.
- McCulloch, W. S., & Pitts, W. H. (1943). A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics*, 115-133.
- Ministério da Agricultura, Pecuária e Abastecimento. (18 de Setembro de 2018). *Ministério da Agricultura, Pecuária e Abastecimento*. (Ministério da Agricultura, Pecuária e Abastecimento) Acesso em 15 de Junho de 2019, disponível em <http://www.agricultura.gov.br/assuntos/politica-agricola/cafe/cafeicultura-brasileira#section-1>
- Morettin, P., & Toloi, C. (2004). *Análise de Séries Temporais*. Blucher.
- Nelson, D. M. (2017). *Uso de Redes Neurais Recorrentes para a previsão de Séries Temporais Financeiras*. Belo Horizonte: UFMG.

Paulo Sergio Ceretta, M. B. (2010). PREVISÃO DO PREÇO DA SOJA: uma comparação entre os modelos ARIMA e redes neurais artificiais. *Informações Econômicas*, 40(9).

S. Hochreiter, J. S. (1987). "Long Short-Term Memory". *Neural Computation*, 1735-1780.

Tarcísio Saraiva Rabelo Júnior, R. H. (2004). Mercados eficientes e arbitragem: um estudo sob o enfoque das finanças comportamentais. *Revista Contabilidade & Finanças*, 97-107.

United States Department of Agriculture. (2019). *Coffee: World Markets and Trade*. Washington: Foreign Agricultural Service.

Universidade de Duke. (2 de Junho de 2019). *ARIMA models for time series forecasting*. Acesso em 05 de Maio de 2019, disponível em Identifying the numbers of AR or MA terms in an ARIMA model: <https://people.duke.edu/~rnau/411arim3.htm>

Valer, L. A. (2019). *Modelo Matemática ARIMAX de um propulsor eletromecânico utilizado em naves do tipo multirrotor*. Ijuí: UNIJUÍ.

White, H. (1982). Maximum Likelihood Estimation of Misspecified Models. *Econometrica*, 1-25.

Zuben, F. J. (14 de Junho de 2005). *Redes Neurais Recorrentes (Parte 1)*. Acesso em 13 de Maio de 2019, disponível em DCA/FEEC/Unicamp: ftp://ftp.dca.fee.unicamp.br/pub/docs/vonzuben/ia353_05/