

**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS MATEMÁTICAS E DE COMPUTAÇÃO**

Pedro Henrique Vieira Pedrosa

**Modelagem preditiva de séries temporais na cotação de
commodities via análise estatística**

São Carlos

2023

Pedro Henrique Vieira Pedrosa

**Modelagem preditiva de séries temporais na cotação de
commodities via análise estatística**

Trabalho de conclusão de curso apresentado ao Centro de Ciências Matemáticas Aplicadas à Indústria do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, como parte dos requisitos para conclusão do MBA em Ciências de Dados.

Área de concentração: Ciências de Dados

Orientador: Prof. Dr. Afonso Paiva

**São Carlos
2023**

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi, ICMC/USP, com os dados fornecidos pelo(a) autor(a)

S856m	<p>Pedrosa, Pedro Henrique Vieira Modelagem preditiva de séries temporais na cotação de commodities via análise estatística / Pedro Henrique Vieira Pedrosa ; orientadora Afonso Paiva. – São Carlos, 2023. 53 p. : il. (algumas color.) ; 30 cm.</p> <p>Monografia (MBA em Ciências de Dados) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2023.</p> <p>1. Redes neurais. 2. Séries temporais. 3. Predição. 4. Commodities. I. Paiva, Afonso, orient. II. Título.</p>
-------	---

Pedro Henrique Vieira Pedrosa

Modelagem preditiva de séries temporais na cotação de commodities via análise estatística

Trabalho de conclusão de curso apresentado ao Centro de Ciências Matemáticas Aplicadas à Indústria do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, como parte dos requisitos para conclusão do MBA em Ciências de Dados.

Data de defesa: 18 de Março de 2023

Comissão Julgadora:

Prof. Dr. Afonso Paiva
Orientador

Professor
Convidado1

Professor
Convidado2

São Carlos
2023

À minha noiva pela compreensão, carinho e apoio incansável

AGRADECIMENTOS

Ao Dr. Afonso Paiva, pela paciência e incentivo.

À minha noiva por todo o suporte que me deu enquanto estava ocupado estudando e não pude me fazer presente.

Aos meus pais por todo carinho e dedicação na minha criação e por me darem o sentimento de pertencimento.

RESUMO

PEDROSA, P. H. V. **Modelagem preditiva de séries temporais na cotação de commodities via análise estatística.** 2023. 53p. Monografia (MBA em Ciências de Dados) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2023.

Esse trabalho tem como objetivo estudar o desempenho de modelos estatísticos e computacionais que sejam capazes de trazer uma previsão assertiva acerca do preço das commodities boi, café, petróleo e ouro analisando as séries temporais extraídas da base dados do *investing.com* no período cinco mil dias entre os anos 2000 e 2019. Foram comparados os modelos ARIMA e os modelos de rede neural GRU e LSTM. Para comparar os resultados dos modelos e validá-los, foram utilizados os métodos estatísticos de erro absoluto médio e erro quadrático médio. Além disso, foi utilizado o método de ADF para verificar a estacionariedade das séries. Foi visto que todas as séries não eram estacionárias, sendo assim, foi necessário aplicar a primeira diferença. Nos modelos de redes neurais foi necessário normalizar os dados e criar novas estruturas de dados para que fosse possível o treinamento do modelo. Nesse estudo o resultado demonstrou uma superioridade dos modelos de redes neurais, tanto o LSTM e GRU tiveram um desempenho muito superior.

Palavras-chave: Séries temporais, Redes Neurais, predição.

ABSTRACT

PEDROSA, P. H. V. . 2023. 53p. Monografia (MBA em Ciências de Dados) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2023.

This work has the objective of studying statistics and computational performance models which can stipulate over the price of certain commodities like cattle, coffee, oil, and gold by analyzing time series extracted from *investing.com* between the years 2000 to 2019. ARIMA models and GRU/LSTM neural networks models were compared for this study. To compare and certify the results it was used the statistics method of Mean Absolute Error and Mean Squared Error. Furthermore, the Adfuller test was used to verify the stationary presented in the time series. Every series were proven to be non-stationary, therefore, it was needed to apply the first difference. On the neural networks models there was the need to normalize the data and create new structures of data so that the training of the model would be possible. In this study the result showed a superiority of the neural networks models, both LSTM and GRU had a better performance.

Keywords: time series, neural networks, prediction

LISTA DE FIGURAS

Figura 1 – Estrutura do neurônio	38
Figura 2 – Estrutura da LSTM	39
Figura 3 – Estrutura da GRU	41
Figura 4 – Decomposição da série temporal de boi	43
Figura 5 – Autocorrelações da série temporal de boi	44
Figura 6 – Decomposição da série temporal de café	45
Figura 7 – Autocorrelações da série temporal de café	45
Figura 8 – Decomposição da série temporal de petróleo	46
Figura 9 – Autocorrelações da série temporal de petróleo	46
Figura 10 – Decomposição da série temporal de ouro	47
Figura 11 – Autocorrelações da série temporal de Ouro	47
Figura 12 – Resumo da Rede LSTM	48
Figura 13 – Resumo da Rede GRU	48

LISTA DE TABELAS

Tabela 1 – Modelos ótimos encontrados pelo auto arima	48
Tabela 2 – Erros encontrados no teste dos Modelos LSTM e GRU	49

LISTA DE ABREVIATURAS E SIGLAS

AR	Lineares Autorregressivos.
ARMA	Autorregressivo de Médias Móveis.
ARIMA	Auto Regressive Integrated Moving Averages.
CME	Chicago Mercantile Exchange.
I	Filtro de integração.
GRU	Gated Recurrent Unit.
LSTM	Long-Short Term Memory.
MA	Médias Móveis.
MAE	Mean Absoluto Erro.
MAPE	Média Absoluta de Erro Percentual.
MBA	Master in Business Administration
ML	Aprendizado de Máquina.
MSE	Mean Squared Error.
PIB	Produto Interno Bruto.
RMSE	Erro Quadrático Médio.
RNA	Redes Neurais Artificiais.
RNN	Redes Neurais Recorrentes.
SARIMA	Modelos Autorregressivos Integrados de Médias Móveis com Sazonalidade.

LISTA DE SÍMBOLOS

ϕ	Letra grega phi
ϵ	Lambda
Θ	Letra grega maiuscula tetha
\in	Pertence
Δ	Delta
\tanh	Delta
R	Conjunto dos reais

SUMÁRIO

1	INTRODUÇÃO	25
1.1	Objetivos	26
1.2	Justificativa	26
1.3	Estrutura do Trabalho	27
2	REVISÃO BIBLIOGRÁFICA	29
2.1	Trabalhos Relacionados	29
3	METODOLOGIA	31
3.1	Descrição do problema	31
3.1.1	Entendimento do Negócio	31
3.2	Entendimento dos dados	32
3.2.1	Séries temporais	33
3.3	Preparação dos dados	34
3.4	Modelagem	34
3.4.1	Modelos estatísticos	34
3.4.2	Aprendizado de Máquina	37
3.4.2.1	Redes neurais recorrentes	38
3.4.2.2	Redes LSTM (Long Short-Term Memory)	39
3.4.2.3	Redes GRU (Gated Recurrent Unit)	41
3.5	Validação	42
4	RESULTADOS	43
4.1	Estudo das series temporais	43
4.2	Modelos Arima e Sarima	46
4.3	Modelos LSTM E GRU	47
5	CONCLUSÃO	51
	REFERÊNCIAS	53

1 INTRODUÇÃO

Entender as tendências dos preços de insumos, matérias primas, ou até produto final é fundamental para que as organizações possam se planejar, definir políticas e planos de ações para alcançar seus objetivos, garantir retorno para o acionista e preservar a existência da instituição. Organizações que operam com produtos de baixo valor agregado, como as commodities, tem esse efeito potencializado e devem estar sempre monitorando notícias e indicadores globais e internos, que possam afetar esses preços.

Devido as commodities serem base de muitos processos produtivos elas acabam sendo amplamente negociadas pelo mundo. Entende-se, commodities como um tipo particular de mercadoria, no estado bruto ou primário com grande relevância comercial, por exemplo o café, o chá, a lã, o algodão, vários tipos de carne, o cobre, minério de ferro, petróleo, etc., como explica(SANDRONI, 2016). As commodities estão inseridas em um contexto complexo onde sua produção é afetada por inúmeros elementos externos que têm grande influência sobre seu processo produtivo e preço, podendo citar o clima, fatores políticos e sociais. Criando uma dificuldade de projeção de valor futuro.

O Brasil integra o atual cenário como sendo um grande produtor de commodities e um importante ator internacional, sendo o segundo maior exportador de grãos do mundo, por exemplo. Essa característica possibilitou o desenvolvimento do mercado de capitais de commodities, melhorando o ambiente de negociação desses ativos. Gerando maior transparência nos preços, acesso e liquidez para a negociação e utilização de instrumentos mais complexos como contratos futuros e outros derivativos, (GOMES, 2002). Graças a esse mercado temos acesso a um grande volume de informação. E hoje quem produz e quem compra, tem a oportunidade de adotar estratégias que eliminem riscos de variações nos preços. Outra característica da precificação de commodity é que elas são cotadas em paridade internacional, então variações externas afetam o preço internamente.

A formação dos preços é um tópico tão importante para a ciência econômica (GOMES, 2002), que dentro da microeconomia, se aperfeiçoou esse ramo através da economia financeira de microstructure theory. Que coloca o papel da informação como de grande importância para a teoria econômica. E sugere uma maior integração entre os modelos econômicos e a hoje conhecida ciência de dados.

Com as características próprias do mercado de commodities, como contratos padronizados, entrega dos produtos fisicamente, garantias físicas de contratos financeiro. Isso da possibilidade para que os produtores possam tomar decisões muito antes da própria produção começar. Um fator levado em consideração nessas decisões é o preço futuro, que dirá se a produção terá um bom retorno. Assim, no presente trabalho pretendemos estudar

e analisar modelos de previsões de preços, pelos métodos estatísticos e computacionais, que deem suporte à tomada de decisão daqueles que operam ou trabalham no mercado de commodities. Esses modelos serão baseados nos modelos estatísticos do tipo SARIMA (Auto-Regressivo Integrado de Médias Móveis com Sazonalidade) e modelos computacionais de redes neurais recorrentes, tais como LSTM (Long-Short Term Memory) e GRU (Gated Recurrent Unit). Utilizando-se da base de dados pública que fornece a cotação do preço do contrato futuro de commodities, *investing.com*¹, que é um agregador de informações sobre o mercado financeiro, tendo diversos tipos de séries temporais. Como é uma plataforma internacional os preços estão todos considerados em dólar. Utilizaremos neste trabalho, a cotação do açúcar, café e boi gordo, que são negociados em contratos de quantidades diversas, sendo para o açúcar sacas de 50 kg, sacas de 60 kg para o café e por arroba de 15 kg no caso do boi. O período de apuração será diário, padronizada pela própria plataforma, tendo assim certeza de que teremos o mesmo número de períodos por ano.

1.1 Objetivos

Tendo em mente que as commodities são influenciadas por diferentes acontecimentos, como variações climáticas, conflitos políticos, variações na demanda, entre outros. O objetivo principal é avaliar quais modelos têm o melhor desempenho mesmo com um cenário que pareça caótico e com determinados padrões e com mudanças abruptas. Os objetivos específicos deste trabalho são:

- Estudar os modelos de predição mais comumente empregados séries temporais, com enfoque em classificá-los de acordo com o seu desempenho de predição;
- Implementar e comparar a performance dos modelos de predição de séries temporais com técnicas estatísticas de predição;
- implementar e comparar a performance dos modelos redes neurais;
- Aplicar os modelos estudados em um conjunto de dados real.

1.2 Justificativa

O estudo de previsão de preços de ativos financeiros desperta muito interesse, talvez falsa impressão de que se prevendo para onde os preços vão, haverá retornos excepcionais para aqueles que conseguirem saber o próximo movimento do preço. O que corrobora para essa visão é que se pesquisado no Google Scholar o termo “previsões de series financeiras”, somente no ano de 2022 serão retornados aproximadamente 15 mil resultados. Para nós é que estamos tentando desenvolver um conhecimento é ótimo, pois é nosso objeto de

¹ <https://br.investing.com/commodities/>

pesquisa e nós traz um vasto material de consulta. Os artigos utilizados por nós nesse estudo, contribuíram para que tentássemos nós aproximar do estado da arte do assunto. Mas dada a abrangência do tema é impossível que tenhamos esgotado o assunto. Assim esperamos contribuir minimamente com o tema, auxiliando na melhora da tomada de decisão de gestores.

Por fim, espera-se que esse trabalho jogue luz a tomadores de decisão que muitas vezes estão sujeitos ao aleatório. Além de submeter teorias tão consolidadas para o escrutínio de dados reais e não homogêneos. Fazendo assim o encontro da academia com o mundo prático.

1.3 Estrutura do Trabalho

O presente trabalho está estruturado da seguinte forma: o Capítulo 2 apresenta trabalhos relacionados ao tema que serviram de base para a construção do conhecimento empregado no presente trabalho. No Capítulo 3 é apresentado o referencial teórico. O Capítulo 4 contém os resultados obtidos. Por fim, no Capítulo 5 é apresentada a conclusão e sugestões para trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

Este capítulo contempla uma revisão da bibliografia sobre séries temporais, redes neurais e outros trabalhos que propuseram fazer uma análise e prever preços de ativos financeiros utilizando ferramentas estatísticas e de Ciência de Dados, além de trazer estudos sobre commodities. Trazendo uma pequena análise do proposto e suas contribuições para o assunto.

2.1 Trabalhos Relacionados

As Redes Neurais Artificiais (RNA) segundo (??) (Gaio, Júnior e Oliveira (2007)) são uma área da computação Natural que trata da solução de problemas a partir de uma mimetização simplificada do cérebro humano em sistemas computacionais. Sendo elas capazes de se adaptar e de aprender a realizar uma tarefa, ou um comportamento, a partir de um conjunto de exemplos de dados.

Ja (??) (Ribeiro e Correa (2021)) utilizaram das redes neurais LSTM e GRU para identificar quais conjuntos de variáveis possuem maior influência no preço. Para avaliação do desempenho dessas redes Neurais utilizaram o MAPE (média absoluta de erro percentual), onde se calcula a média do erro relativo a cada evento na série temporal, confrontado o valor previsto com o dado observado. Utilizaram a RMSE (erro quadrático médio), que calcula a raiz quadrada da média da distância entre os erros da previsão e o valor conhecido. Utilizam também a U de Theil, onde utilizando de uma predição ingênua, que consiste em repetir o ultimo valor da serie, avalia-se a performance da predição.

É comum observar em trabalhos que estão trabalhando com previsões de séries temporais, a comparação entre o método puramente estatístico e o método que se utiliza de computação através das redes neurais. Trabalhos como (??), (??), (??)(Costa (2012), Souza (2008) e Sergio (2017),). É de grande valor, comparar os modelos pois os modelos lineares probabilísticos estão sobre estudo e teste a muito tempo isso faz com que tenhamos um parâmetro confiável sobre a utilidade da rede neural e o ganho de eficácia, uma vez que temos um trade-off entre o custo computacional e a qualidade dos resultados gerados por esses modelos.

Segundo (??) (Miranda, Coronel e Vieira (2013)), que utilizaram de modelos econométricos e de redes neurais, para a revisão de mercado futuro de café arábica, o modelo de redes neurais teve desempenho superior nos três métodos de avaliação. Os modelos comparados foram, médias móveis (MA), lineares Autorregressivos (AR), autorregressivo de médias móveis (ARMA) e uma rede neural. Avaliados com erro médio quadrático, erro médio global e R^2 ajustado com a rede neural performando melhor em

todos esses testes.

Outro estudo que propõe um estudo sobre previsão de preços em commodities é o da (??) (Ribeiro, Sosnoski e Oliveira (2010)), neste trabalho, em vez de uma comparação, o estudo optou por combinar as abordagens tradicionais de séries temporais com modelos de redes neurais. Assim permitiu uma melhora de desempenho do modelo a partir da inclusão de variáveis exógenas. O método de avaliação de desempenho dos resultados obtidos foram através de Erro médio absoluto, o Erro quadrático médio, o erro percentual absoluto médio e a variância do erro.

A principal diferença entre uma rede neural e os métodos estatísticos é que as RNs não estabelecem hipóteses ou suposições sobre a distribuição dos dados, que as tornam mais úteis em situações práticas. Isso é feito através de unidades de processamento que chamamos de neurônios, que têm a função de armazenar conhecimento e experimental e torná-lo disponível para uso. Assim, conseguindo executar funções básicas análogas ao comportamento humano, como o aprendizado, generalização e associação.

(MIGLIATO, 2021) une os dois modelos e utiliza da redes neurais recorrentes (RNN) para a detecção de outliers. Conseguindo assim contribui com a melhor performance do modelo estatístico SARIMA, com isso o modelo teve um melhor desempenho no geral. Utilizando da LTSM e GRU com series uni-variadas reais e públicas, com o período de observação mensal. Os resultados obtidos pelo seu estudo foram validados pela métrica do erro médio absoluto.

A discussão proposta pelo presente trabalho se diferencia dos outros ao comparar métodos estatísticos com métodos de aprendizado de máquina. E converge ao utilizar séries temporais publicas de ativos financeiros.

3 METODOLOGIA

Essa seção é destinada a apresentar os aspectos metodológicos utilizados para que consigamos alcançar os objetivos do presente trabalho. Como iremos trabalhar com séries temporais com a intenção de prever valores futuros, para auxiliar na tomada de decisão, Utilizando de bases de dados de commodities distintas com o objetivo de testar os métodos utilizados no trabalho, assim construindo um conhecimento prático para estarmos preparados com qualquer tipo de commodities que venham trabalhar na vida profissional. Para guiar este projeto montaremos um “pipeline” de trabalho da seguinte forma:

- Descrição do problema;
- Entendimento dos dados;
- Preparação dos dados;
- Modelagem;
- Avaliação.

3.1 Descrição do problema

O problema base desse trabalho em termos gerais é entender qual método de predição de valores futuros de preços de commodities é melhor para auxiliar a tomada de decisão de um gestor. De uma forma específica, o problema é comparar o desempenho do modelo SARIMA com as redes neurais recorrentes LSTM e GRU. O primeiro passo desse problema é a obtenção das séries históricas dos ativos analisados. Passando pela limpeza e estruturação dos dados, foi feita uma análise das características das séries temporais e avaliada a sua necessidade de diferenciação. O particionamento dos dados ocorreu quando essas características estavam elencadas.

A partir disso se modelou em python na plataforma da anaconda com o Jupyter notebook, para a escrita e execução do código, tanto o SARIMA como as redes neurais LSTM e GRU. Todas as bibliotecas utilizadas são públicas e ensinadas durante o curso, como por exemplo, os, pandas, matplotlib, numpy, stats models, pmdarima e tensor flow.

3.1.1 Entendimento do Negócio

Como discutido na introdução a necessidade da predição para a tomada de decisão de qualquer organização. Neste trabalho a predição que será feita é sobre as commodities agrícolas, através de seus contratos futuros. Contratos futuros são acordos celebrados entre

dois agentes em que o proprietário de um ativo compromete-se a entregá-lo a outro agente, em uma data futura.

Como o objetivo é validar modelos neste trabalho iremos trabalhar com algumas commodities para solidificar os conhecimentos do curso de MBA. Essas commodities são ouro, café, petróleo, boi. Não há motivo específico para escolhermos essas commodities específicas o mais importante e avaliar o desempenho dos modelos, assim quanto mais diversas forem as commodities, melhores serão as avaliações sobre o modelo. O recorte temporal que escolhemos, será do dia primeiro de janeiro de 2000 até 21 de dezembro de 2019, estado periodizada em dias. Os dados serão divididos em um conjunto treino e um conjunto de teste. Isso foi feito para que o modelo seja treinando sem influência do conjunto de teste, esse fica desconhecido para o modelo. Após o treinamento e então validamos com o conjunto de teste, para garantir a melhor precisão. Assim ficara possível avaliar qual dado tem uma melhor aderência ao modelos utilizados.

3.2 Entendimento dos dados

No livro “Análise de Séries Temporais” (MORETTIN; TOLOI, 2004), os autores são bem diretos ao generalizar que normalmente séries financeiras são não-estacionárias. O que quer dizer que a série apresenta média e variância dependentes do tempo. Para detectar essa característica podemos aplicar o teste de raiz unitária, neste trabalho utilizaremos o de Dickey-Fuller. Não sendo possível presumir a estacionariedade dos dados, será necessário transformá-los. segundo (MORETTIN; TOLOI, 2004) a transformação mais usual consiste em tomar as diferenças sucessivas da série original até obter uma série estacionária. Isto é, subtrair o valor t da série por $t + 1$, onde t representa um instante de tempo na série. Considerando que nossos dados são diários, a diferença será o valor da cotação de um dia pelo seu sucessor.

Observada a estacionariedade, a escolha do modelo está condicionada à análise da tendência e da sazonalidade dos dados. Faremos isso após o tratamento de dados faltantes ou de dados duplicados. Utilizaremos a biblioteca *statsmodels*, com a função “seasonal decompose”. Essa função gera uma imagem com 3 gráficos que demonstram essas características.

As bases utilizadas nesse trabalho têm o período de coleta das amostras diariamente e foram extraídas do site *investing.com*. Esses dados são obtidos em arquivos CSV, e foram acessados utilizando da biblioteca Pandas. Foi utilizado a cotação dos contratos futuros operados nos mercados norte americanos, uma vez que é onde o mercado financeiro se concentra ali, temos o maior volume de negociações e de acesso mundial. Complementar a essas características os preços já são negociados em dólar, fator muito importante para a pauta de exportação.

A base é constituída de cinco colunas sendo elas: Data, último, abertura, máxima, mínima, volume e variação. Divididos em dados, quantidade monetária, valores absolutos e porcentagem. A data definida será utilizada como índice da série e é marcada com os dias de trabalho norte-americanos, uma vez que é onde os contratos são negociados. A coluna última é referente ao último preço que o ativo teve no dia, ele que temos interesse em prever. A abertura é o primeiro preço do dia. A coluna mínima e máxima são referentes a valores marcantes no dia. O volume versa sobre a quantidade de contratos negociados no dia. E a variação resume o comportamento do preço do ativo no dia em porcentagem.

3.2.1 Séries temporais

A maneira mais direta que podemos definir uma série temporal é como qualquer conjunto de observações ordenadas no tempo, (MORETTIN; TOLOI, 2004). O resultado destas observações nós dão um série de dados que se relacionam uns com os outros de maneira dependente. As séries temporais são compostas por quatro componentes: tendência, sazonalidade, cíclica e erro. Componentes que ajudam a entender as particularidades da série que estamos observando, além de serem consideradas na hora de aplicar-se qualquer modelo de previsão. Assim cabe entender o que cada uma contribui com o modelo de séries temporais.

As componentes de tendência causam frequentes mudanças graduais em um longo espaço de tempo, como o efeito do crescimento populacional constante no produto interno bruto, esses fatores falham na tentativa de produzir mudanças repentinas, mas produzem variações graduais e perenes ao longo do tempo. As componentes cíclicas marcam oscilações subidas e de quedas, de maneira suave e repetitiva, ao longo da componente de tendência.

As componentes sazonais na série são aquelas variações são bem marcadas em um determinado período do ano, do mês, da semana, ocorrendo sempre em um determinado período, independente do sentido da oscilação. Morettin e Tolo (2004) diferencia componentes sazonais e cíclicas, essencialmente pela previsibilidade inerente à primeira. Já os movimentos cíclicos tendem a ser irregulares, ocorrendo sobre um período de muitos anos, por exemplo o ciclo de baixa ou alta do mercado acionário. A última componente é o ruído aleatório ou erro, que compreende a variabilidade intrínseca aos dados e não pode ser modelado.

Com as componentes bem marcadas podemos cumprir melhor os objetivos da análise de séries temporais que são, de investigar o evento gerador da série temporal, fazer previsões de valores futuros da série de curto ou longo prazo. Podemos descrever o comportamento da série utilizando de gráficos e procurar padrões relevantes no dados, segundo (MORETTIN; TOLOI, 2004). Os autores ainda alertam que os modelos seja eles probabilístico ou estocásticos devem ser simples e objetivo em relação aos seus parâmetros.

Percebe-se na literatura estudada que o estudo da série temporal é desenvolvido

através modelos, que são definidos por (MORETTIN; TOLOI, 2004) como uma descrição probabilística de uma érie temporal. Na econometria um modelo é uma forma de descrever, a relação não exata das variáveis econômicas que tentam explicar algum fenômeno prático. Sempre considerando a componente de erro citada anteriormente.

3.3 Preparção dos dados

A preparação dos dados se deu no Google Collab, plataforma do google, que permite trabalharmos com Python em ambientes com o processamento descentralizado. Foi optado pela plataforma uma vez que modelos de redes neurais são melhores processados por unidades de processamento gráfico. Ao abrirmos as bases de dados que foram baixadas manualmente do *investing.com*, indexamos o data frame pela coluna de datas, ordenamos por em ordem crescente o data frame e por último excluimos 5 colunas que não utilizamos no processamento, “Abertura”, “Máxima”, “Mínima”, “Vol.” e “Var”.

Tendo um dataframe com somente os dados que iramos utilizar, dividiu-se a base em treino e teste seguindo a proporção 80% e 20%, respectivamente. Normalizamos os valores através dos valores mínimos e máximos. Com os dados normalizados criou um novo array com o número de elementos passados vamos considerar, agrupando os numero de elementos, nesse caso foram de dois em dois e o número de fatures, isso foi feito para adequar os dados para modelos de aprendizado de máquina recorrente. No nosso estudo utilizamos todos os time steps igual a dois em todas a séries temporais.

3.4 Modelagem

Como esse trabalho pretende comparar métodos de modelagem, vamos neste tópico passar mais sobre os modelos utilizados para treinar o algoritmo de predição. As técnicas de aprendizado são divididas em supervisionado e não supervisionado de acordo com (MORETTIN,). O aprendizado supervisionado nos permitirá considerar uma distinção entre variáveis preditoras e respostas. No caso do aprendizado não-supervisionado, todas as variáveis do conjunto de dados são consideradas sem distinção, o objetivo é descrever associações e padrões entre essas variáveis.

3.4.1 Modelos estatísticos

Nos modelos estatísticos, a primeira parte do treinamento é a separação dos dados em conjuntos de treinamento e teste. O conjunto de treinamento é utilizado para construir o modelo, sem a influência do conjunto de testes. O modelo construído é colocado à prova comparando os dados do conjunto de treinamento e de teste. Assim podendo avaliar a acurácia do modelo.

O modelo estatístico ARIMA (Auto regressive integrated Moving Averages) é um dos principais modelos para previsão de séries temporais, também é conhecido como

modelo Box-Jenkins. Esse método é baseado em um ciclo iterativo, no qual a escolha do modelo é realizada com base nos próprios dados, (SOUZA, 2008). Sendo dividido nas seguintes etapas (WERNER; RIBEIRO, 2003):

- **Identificação:** Consiste em descobrir qual dentre as várias versões dos modelos de Box-Jenkins descreve o comportamento da Série.
- **Estimação:** consiste em estimar o parâmetro ϕ do componente auto-regressivo, o parâmetro θ do componente de médias móveis e a variância de ϵ_t ;
- **Verificação:** consiste em avaliar se o modelo estimado é adequado para descrever o comportamento dos dados. Caso o modelo não seja adequado, o ciclo é repetido, voltando-se à fase de identificação. Um procedimento utilizado é identificar não só um único modelo, mas alguns modelos que são, então, estimados e verificados;;
- **Previsão:** Quando se obtém um modelo satisfatório, passa-se para a última etapa da metodologia de Box-Jenkins, que constitui o objetivo principal da metodologia: realizar previsões.

Modelos matemáticos como o ARIMA, visam capturar o comportamento da correlação seriada ou autocorrelação entre os elementos da série temporal. Utilizando-se de três componentes chamada de “filtros”: o componente auto-regressivo (AR), o filtro de integração (I) e o componente de médias móveis (MA). Cada um desses filtros surge de modelos específicos, que foram combinados no modelo ARIMA.

A equação 3.1 representa a descrição do modelo AR(p), em que Z_t , valor atual da série, é descrito por seus valores passados, representados por Y_{t-1} , Y_{t-2} e pelo ruído aleatório ϵ_t

$$Z_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t \quad (3.1)$$

O ϕ é o parâmetro de ajuste que tem o comprimento p, ordem do modelo, e descreve como Z_t se relaciona com os seus valores passados, Z_{t-1} , Z_{t-2}, \dots, Z_{t-p} . Sendo assim, eles usa de observações passadas para compor a equação do modelo que projeta o próximo valor. Esse modelo precisa suportar a estacionalidade e que as auto covariâncias sejam independentes.

O operador auto regressivo estacionário de ordem p é dado por,

$$\phi(B) = 1 - \phi_1 B - \phi_1 B^2 - \dots - \phi_p B^p \quad (3.2)$$

Sendo assim, substituindo 3.2 em 3.1, podemos escrever o modelo AR(p) como

$$\phi(B)ZT = \epsilon_t \quad (3.3)$$

Em um modelo de médias móveis (do inglês, moving average), a série Z_t resulta da combinação dos ruídos brancos, ϵ , do período atual com aqueles ocorridos em períodos anteriores, $\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-q}$. Assim, um modelo de médias móveis de ordem q ou $MA(q)$ é dado por:

$$Z_t = \epsilon_t + \Theta_1\epsilon_{t-1} + \Theta_2\epsilon_{t-2} + \dots + \Theta_q\epsilon_{t-q} \quad (3.4)$$

Sendo que 3.4 pode ser descrita como

$$Z_t = \Theta(B)\epsilon_t \quad (3.5)$$

Quando uma série temporal apresenta média e variância dependentes do tempo, é porque ela não é estacionária. A não-estacionariedade de uma série implica que: a) há tendência nos dados e eles não permanecem ao redor de uma linha horizontal ao longo do tempo e/ou b) a variação dos dados não permanece constante sobre o tempo, indicando que a variância está se alterando. É nesse aspecto dos dados que o filtro de integração se encarrega. Ele é a transformação pela tomada das diferenças sucessivas da série original até obter uma série estacionária.

$$\Delta Z_t = Z_t - Z_{t-1} \quad (3.6)$$

A partir da definição dos filtros descrita por 3.6 o modelo ARIMA pode ser descrito conforme 3.7

$$\phi(B)\Delta^d Z_t = \Theta(B)\epsilon_t \quad (3.7)$$

Os modelos ARIMA que exibem padrões sazonais devem utilizar os modelos SARIMA (Seasonal ARIMA) que incorpora a sazonalidade da série. Esse modelo é construído com uma parte não sazonal com parâmetros (p,d,q) , e uma sazonal, com parâmetros (P,D,Q) , em que

- p é a ordem do modelo AR não sazonal
- d é a parte de integração não sazonal
- q é a ordem do modelo de médias móveis não sazonal
- P ordem do modelo AR sazonal

- D é a parte de integração sazonal
- Q é a ordem do modelo de médias móveis sazonal

A equação do modelo SARIMA inclui termos para modelar a sazonalidade e é dada por 3.8

$$\phi(B)\Phi(B^m)W_t = \Theta(B)\Theta(B^m)\epsilon_t \quad (3.8)$$

3.4.2 Aprendizado de Máquina

Aprendizado de Máquina (do inglês, Machine Learning) são técnicas de análise de dados que estão no campo da inteligência artificial. Essas técnicas ocorrem o através de redes neurais, que simulam matematicamente o funcionamento do cérebro biológico. Conforme (HAYKIN, 2001), a estrutura básica é representada pelo funcionamento dos neurônios, onde os dados são levados ao seu núcleo e ocorre o processamento desses dados os transformando em informação. Com essa porcesso tanto o cérebro humano e uma rede neural artificial adquirem conhecimento através do estímulo disponível no seu ambiente, fixando desse conhecimento por meio de conexões neurais.

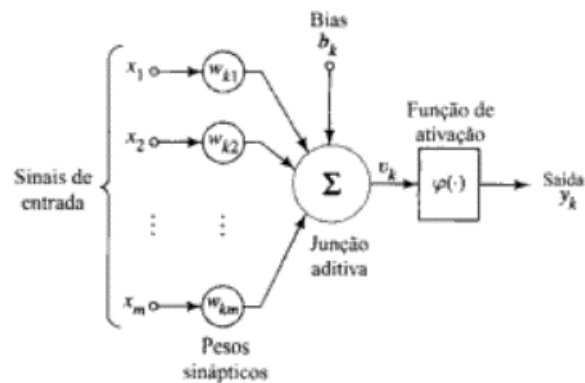
Uma definição mais formal de uma rede neural é de acordo com (HAYKIN, 2001): “um processador maciçamente paralelamente distribuído de unidade de processamento simples, que tem a propenso natural para armazenar conhecimento experimental e torná-lo disponível para o uso.”

O que irá ocorrer internamente na rede neural para realizar o processo de aprendizagem é um algoritmo de aprendizagem, que tem a função de modificar os pesos sinápticos da rede de uma forma que se alcance o objetivo da rede. Esse aprendizado pode ser classificado de três diferentes formas: aprendizado supervisionado, aprendizado não supervisionado e aprendizado por reforço.

A menor parte de uma RNA, é um neurônio que é segundo (HAYKIN, 2001) uma unidade de processamento de informação, sendo ele fundamental para o modelo. Os elementos básicos do modelo neuronal são os pesos sinápticos, que atribui um peso para cada sinal de entrada/estímulo. Depois de somar os sinais de entrada são ponderados pelas respectivas sinapses do neurônio atribuindo peso a cada sinal. Com isso se é ativada a função de ativação para delimitar o intervalo da saída de um neurônio. O neurônio também conta com um bias aplicado de forma externa, que tem o objetivo de aumentar o diminuir a entrada líquida da função de ativação. Na Figura 1, fica bem claro o processo do processamento.

Existem vários tipos de função de ativação, estão entre elas a Softmax, Relu, Tanh e sigmóide, elas se diferenciam no que entregam para a próxima camada através de uma

Figura 1 – Estrutura do neurônio



Fonte: (HAYKIN, 2001)

transformação não linear. A sigmóide é uma função suave e é continuamente diferenciável, tendo como saída valores entre 0 e 1 em um formato de S. A Tanh é muito semelhante a sigmóide mas tem saídas com valores -1 e 1. A Relu é uma função de ativação que não ativa todos os neurônios ao mesmo tempo, se a entrada for negativa, tornando a rede esparsa e eficiente e fácil para a computação. A função softmax transforma as saídas para cada classe para valores entre 0 e 1 e também divide pela soma das saídas. Isso gera uma probabilidade de a entrada estar em uma determinada classe ajudando com quando o problema múltiplas classes. Dada especificidade entre as funções em uma mesma redes vemos diferentes funções de ativação ao longo da rede.

3.4.2.1 Redes neurais recorrentes

Podemos dizer que a arquitetura mais básica da rede neural é unidirecional não existindo nenhum tipo de ciclo ou seja a informação segue do início até o final, dependendo somente da entrada corrente. Diante da necessidade de que os dados passados influenciem no aprendizado da rede, surge a arquitetura de uma rede neural recorrente (RNN).

Nesta arquitetura surgem os gates que armazenam informações prévias. Essa memória especial da RNN é chamada de estados escondidos recorrentes e possibilitam a rede prever o valor da próxima entrada da serie temporal. Isso ocorre porque a RNN se beneficia das informações sequencias passadas, porém existe uma limitação nesta memória da RNN o comprimento da serie temporal sendo possível utilizar somente alguns períodos atrás.

Segundo Migliato (2021) a RNN pode ser definida como:

$$h_t = \theta(W_x x_t + W_h h_t - 1 + b_t) \quad (3.9)$$

sendo θ e uma função não linear W_x e W_h são matrizes de pesos e b_t é uma constante (viés). Partindo dessa definição se tem os mais diversos tipos de RNNs, variando as quantidades de entradas tanto quantidades de saídas, sendo uma entrada para muitas saídas, muitas entradas para uma saída, entre outras.

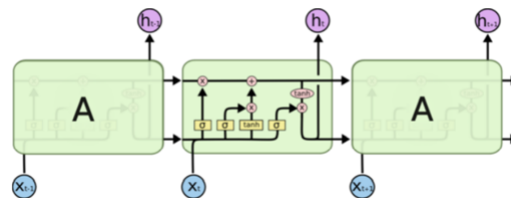
Apesar de existir uma grande variedade de tipos de RNNs, o aprendizado com esse tipo de rede ainda é desafiador, pois de acordo a [Migliato \(2021\)](#), existe uma dificuldade de aprender as dependências de dados muito distantes no tempo, gerando dois principais, a explosão ou o desvanecimento de gradientes. Sendo o desvanecimento, o valor dos gradientes, tendendo a desaparecer ao longo das camadas, ocorrendo quando o algoritmo desta definindo valores cada vez menores para a matriz de pesos. O problema da explosão se trata do gradiente alcançando valores extremamente altos, isso ocorre quando a RNN aloca valores cada vez mais altos para a matriz de pesos.

O desenvolvimento da RNN levou ao surgimento de duas arquitetura bem sucedidas que são as redes LSTM (Long Short-Term Memory) e GRU (Gated Recurrent Unit).

3.4.2.2 Redes LSTM (Long Short-Term Memory)

A principal ideia que trás fanhos na capacidade de aprendizado da Rede reconrrente com memória de longo e curto prazos (LSTM) é uma célula de memória que tem a capacidade de manter seu estado durante do aprendizado e essa células são reguladas através de unidades não lineares, os gates (portões). Esses gates tem a capacidade de ajustar o fluxo de informação que são fornecidas para a célula. Segundo [Migliato \(2021\)](#) um portão é uma unidade sigmoidal que é ativada a partir da entrada atual X_t e da camada escondida anterior h_{t-1} . Esses portões são de três tipos: esquecimento, entrada e saída. Abaixo uma imagem :

Figura 2 – Estrutura da LSTM



Fonte: ([OLAH, 2015](#))

Seguindo a descrição de ([??](#)), o funcionamento da LSTM tem como entrada um vetor X_t no tempo t e h_{t-1} é uma camada escondida anterior, no tempo $t - 1$. As matrizes de pesos são as seguintes:

- Pesos de entrada : $W_f, W_i, W_c, W_o \in R^{N \times M}$;

- Pesos de vieses: $B_f, B_i, B_c, B_o \in R^N$

Temos como primeira etapa na LSTM o portão do esquecimento, responsável por filtrar se as informações passadas ainda são necessárias no estado da célula. Nesse portão as entradas são x_t no tempo T e H_{t-1} , multiplicado por matrizes de pesos W_f e somados a um viés B_f , que serão processados por uma função logística, tendo como saída valores entre 0 e 1. Todos os valores do estado da célula anterior C_{t-1} é multiplicado pela saída desse portão. Assim o modelo entende o quanto da célula anterior sera empregado no estado da célula atual. Essa equação é da por,

$$f_t = \theta(W_f[H_t - 1, X_t] + B_f), \quad (3.10)$$

onde θ é a função de ativação:

$$\theta(x) = \frac{1}{1 + \exp(x - x)}. \quad (3.11)$$

Continuando o processamento, a rede tem que decidir quanto da nova informação será mantida no estado da célula. Esse portão se divide em duas partes. Sendo a primeira o portão de entrada, recebendo a entrada X_t no tempo T e H_{t-1} , multiplicados por matrizes de peso W_i e comados a um viés de B_i , também processados por uma função logística entregando resultados entre 0 e 1. Temos a seguinte equação nesse portação.

$$i_t = \theta(W_i[H_t - 1, X_t] + B_i), \quad (3.12)$$

Na continuação do processamento, chegamos na segunda parte, um novo estado C'_t é computado por meio de uma função tangente hiperbólica e multiplicado pela saída da primeira parte. Então temos um novo estado dado por:

$$C'_t = \tanh(W_c^t[H_t - 1, X_t] + B_c), \quad (3.13)$$

onde \tanh é a função tangente hiperbólica de ativação

$$\tanh(x) = \frac{\exp^2 x - 1}{\exp 2x + 1}. \quad (3.14)$$

O resultado dessas etapas se tem a atualização do estado da célula, Sendo dado por:

$$C_t = F_t * C_{t-1} + I_t \odot C'_t \quad (3.15)$$

onde \odot é a representação de uma multiplicação elemento a elemento.

chegando então no calculo a saída da LSTM. Começando pelo achatamento do estado da célula por uma tangente hiperbolica com saída de valor entre -1 e 1. A seguir esse valor é multiplicado pelo resultado do portão de saída O_t , que define quanto do estado atual compora a saída da LSTM. Esse portão de saída é tido como:

$$O_t = \omega(W_o[H_t - 1, X_t] + B_o), \quad (3.16)$$

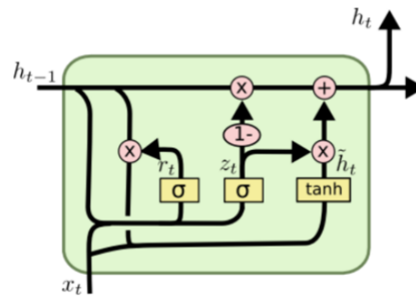
e a saída de H_t é dada por

$$h_t = O_t \odot \tanh(C_t). \quad (3.17)$$

3.4.2.3 Redes GRU (Gated Recurrent Unit)

A GRU é uma rede RNN com certa grua de semelhança da LSTM, porém que tem a arquitetura mais simples. A figura abaixo mostra uma representação do funcionamento da GRU.

Figura 3 – Estrutura da GRU



Fonte:([OLAH, 2015](#))

A principal diferença entra LSTM e GRU, de acordo com ([MIGLIATO, 2021](#)) é que a GRU não faz uma distinção entre as células de memórias. Como fica claro na imagem as ativações de H_t da GRU no tempo t é uma interpolação linear entre a ativação previa h_{t-1} e o candidato à ativação H'_t :

$$h_t = (1 - z_t)H_t - 1 + z_t h'_t \quad (3.18)$$

sendo o portão de atualização Z_t que define quando se deve atualizar sua ativação, ou parâmetros. A atualização calculada por esse portão é da seguinte forma.

$$z_t = \theta(W_z x_t + U_z h_t - 1) \quad (3.19)$$

Apesar do procedimento da soma linear entre o estado atual com o novo estado computado, ser semelhante ao processo da LSTM. Na GRU, não possui um mecanismo de

controle do grau de exposição do estado, assim ele ocorre inteiramente a cada ativação. O candidato a ativação h'_t é calculado da seguinte forma:

$$h'_t = \tanh(W_x \cdot t + U(r_t \cdot h_{t-1})) \quad (3.20)$$

Sendo esse r_t o portão de reset e \odot é uma multiplicação elemento a elemento. O cálculo desse portão de reset R_t tem uma forma parecida ao portão de esquecimento:

$$z_t = \theta(W_r x_t + U_r h_t - 1) \quad (3.21)$$

3.5 Validação

Para comparar as previsões, por diferentes métodos, foram utilizadas duas estatísticas de análise: erro absoluto médio, Erro Quadrático Médio (MAE, Mean Squared Error em inglês). Segue abaixo as fórmulas de cada estatística:

$$MAE = \frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j) \quad (3.22)$$

$$MSE = \frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2 \quad (3.23)$$

4 RESULTADOS

Nesse capítulo iremos discutir e expor os resultados do processamento das séries temporais e seus resultados, junto com o desenvolvimento dos passos metodologias para chegar no resultados objetivados do estudo. Começando pela decomposição das séries e passando pelo resultado da aplicação dos modelos.

4.1 Estudo das series temporais

Na Figura 4 contém a decomposição referente aos contratos futuros de boi negociados pela bolsa CME (Chicago Mercantile Exchange) negociada em dólar. Como observada na imagem a série não tem uma tendência muito uniforme, a sazonalidade apresentada se remete a períodos anuais.

Após a decomposição da série e gerar a imagem anterior, realizamos o teste ADF na série temporal e obtivemos uma estatística ADF de $-1,785$ e um valor p de $0,388$. Isso indica que não podemos rejeitar a hipótese nula de que a série é não-estacionária. As estatísticas críticas correspondentes para um nível de significância de 1%, 5% e 10% foram $-3,432$, $-2,862$ e $-2,567$, respectivamente. Também se elaborou os testes de autocorrelação, e autocorrelação parcial com 45 lags, esse estudo demonstrado na figura 5 abaixo mostra como é forte a correlação entre os dados. Mostrado a possível necessidade da diferenciação.

Figura 4 – Decomposição da série temporal de boi

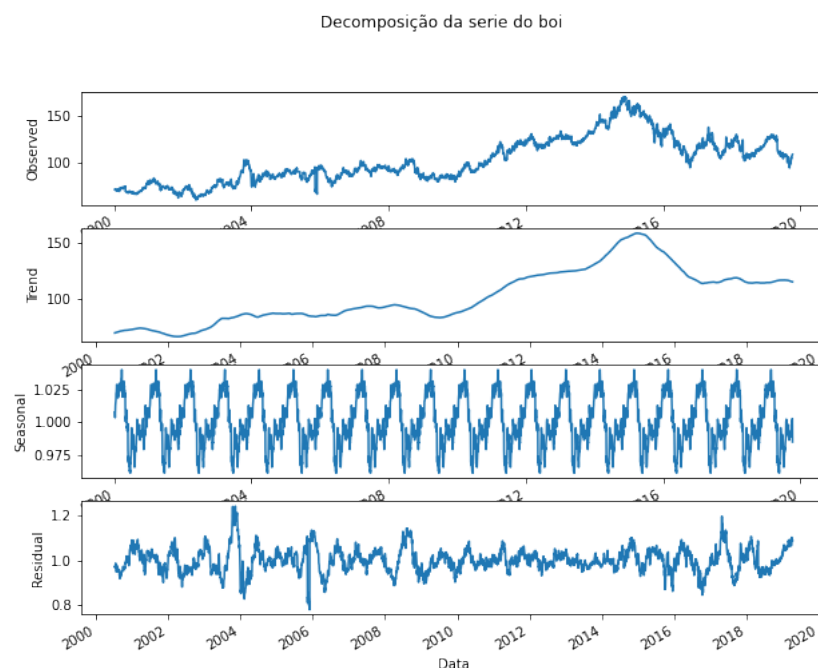
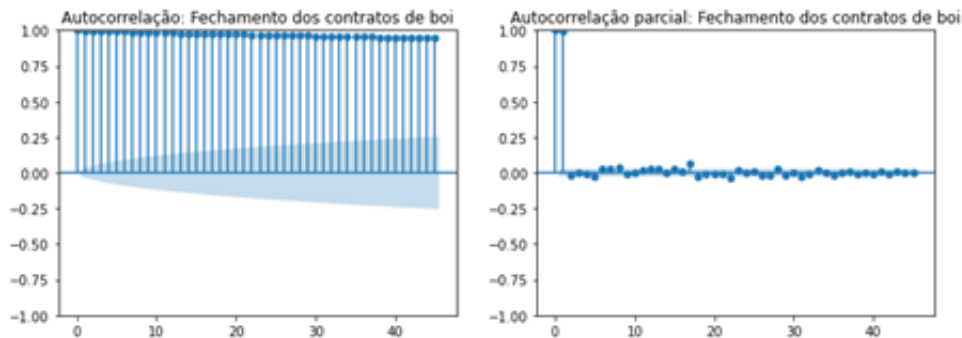


Figura 5 – Autocorrelações da série temporal de boi



Fonte: Autor(2023)

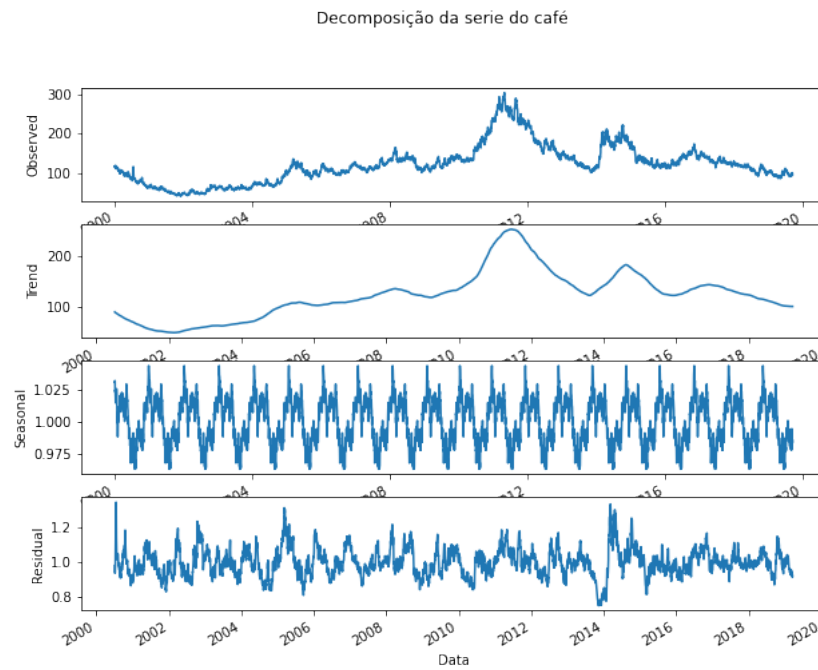
A figura 6 está contendo a decomposição da série temporal do café mostra, uma série sem uma tendência muito clara ao longo do tempo, com uma sazonalidade anual. Com base nos resultados do teste ADF aplicados nos contratos de café, podemos dizer que a série não é estacionária. Isso é indicado pela estatística do teste ADF -1.5184 , que é maior do que o valor crítico no nível de significância de 5%, -2.8621 e de todos os outros níveis de 1%, -3.431 e de 10%, -2.5670 . Além disso, o p-valor de 0.5243 é maior do que o nível de significância de 5%, o que significa que não podemos rejeitar a hipótese nula de que a série possui uma raiz unitária e não é estacionária. Os correlogramas, na figura 7, dessa série também demonstram um forte correlação entre os dados sériais.

A próxima série decomposta foi a do petróleo, sendo os contratos de Brent de petróleo futuros negociados por barril. A decomposição ficou da seguinte forma, vista na figura 8 uma sazonalidade mais bem marcada ao longo dos anos, porém sem uma tendência uniforme. Sendo que o teste ADF também apontou para a não estacionariedade da série, com o ADF de -2.0239 e p-value: 0.2761 e os valores críticos sendo -3.4316 , -2.8621 e -2.5670 , para a significância de 1%, 5% e 10% respectivamente. Nesse caso o teste de ADF foi muito importante pois, visualmente temos a impressão de um resíduo estacionário. Assim não rejeitamos a hipótese nula e utilizamos da diferenciação ao modelar a série. Ainda nos correlogramas desta série, observados na Figura 9, o padrão de uma forte autocorrelação se mantém.

A última série temporal que foi analisada foi o Ouro, com o contrato com a quantidade de uma Onça. Apresenta uma sazonalidade anual, com uma tendência mais clara para mais da metade da série, porém perdendo essa tendência após 2012, como visto na sua decomposição, representada na figura 10

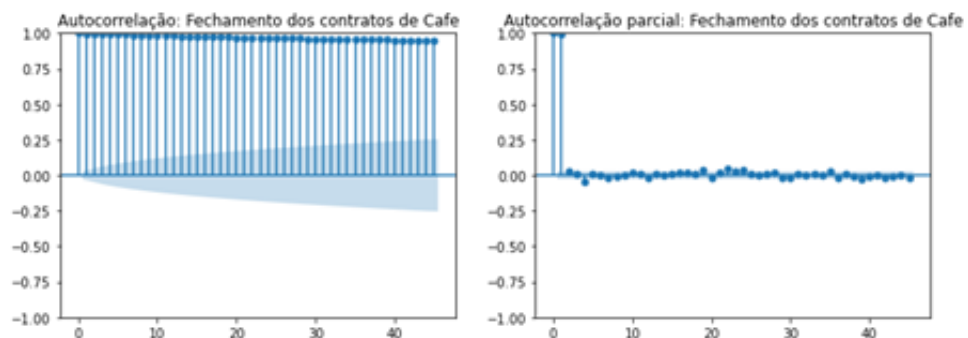
Repetindo o processo anterior após a decomposição já calculamos o ADF da série do ouro e obtivemos o valor de -0.8634 com o p-valor de 0.7997 , com os valores críticos sendo 1% -3.431 , 5% -2.8621 , 10% -2.5673 . Dados esse teste aceitamos a hipótese

Figura 6 – Decomposição da série temporal de café



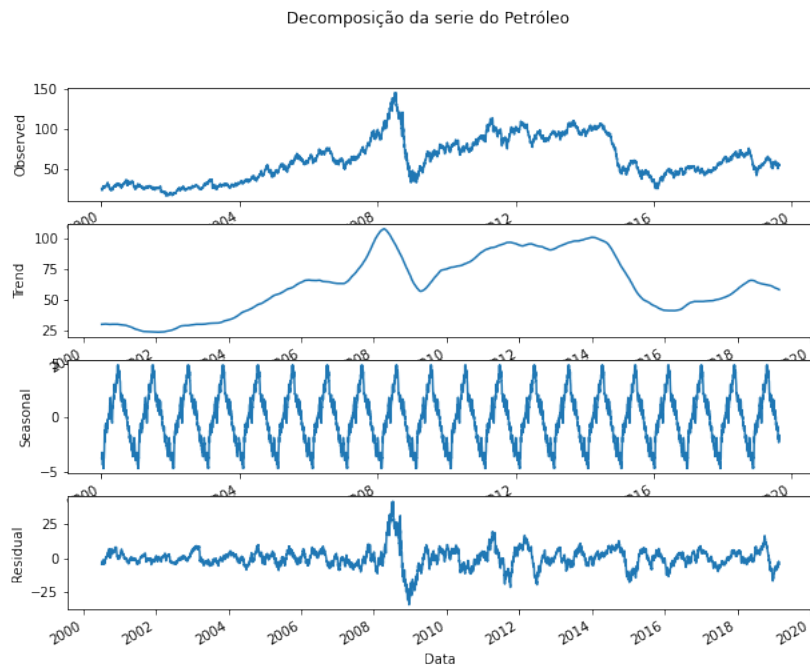
Fonte:Autor(2023)

Figura 7 – Autocorrelações da série temporal de café



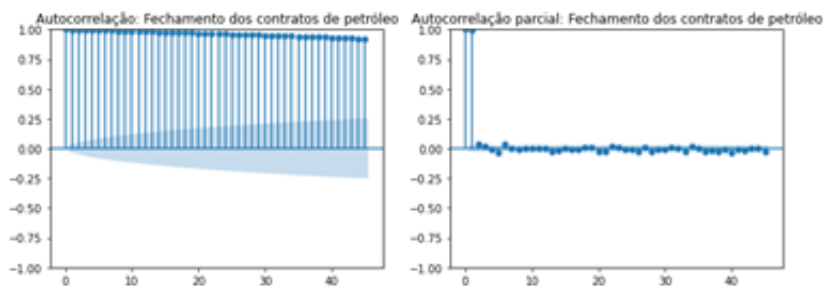
Fonte:Autor(2023)

Figura 8 – Decomposição da série temporal de petróleo



Fonte:Autor(2023)

Figura 9 – Autocorrelações da série temporal de petróleo



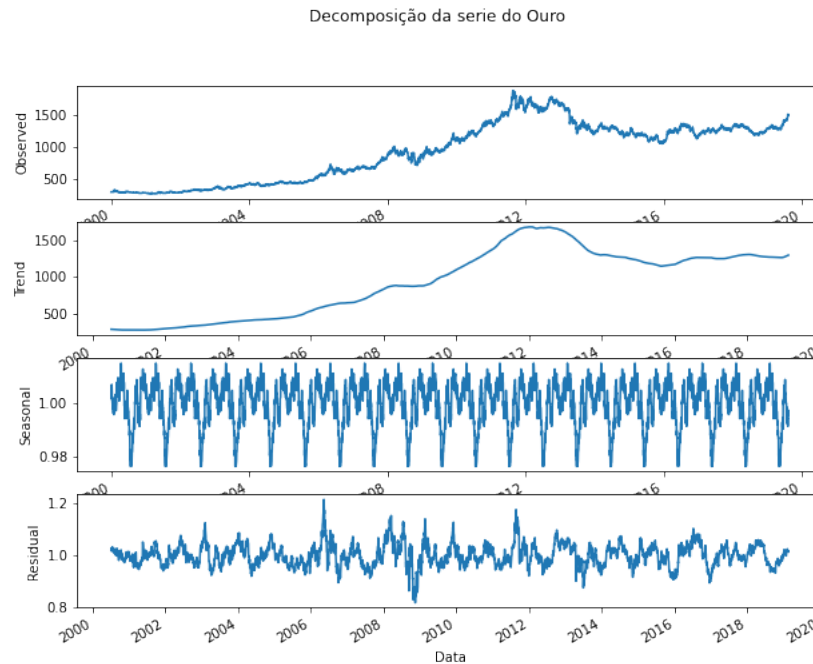
Fonte:Autor(2023)

nula e presumimos a não estacionariedade da série. Os correlogramas apresentados não diferenciam do padrão encontrado na outras séries como podemos observar na figura 11 .

4.2 Modelos Arima e Sarima

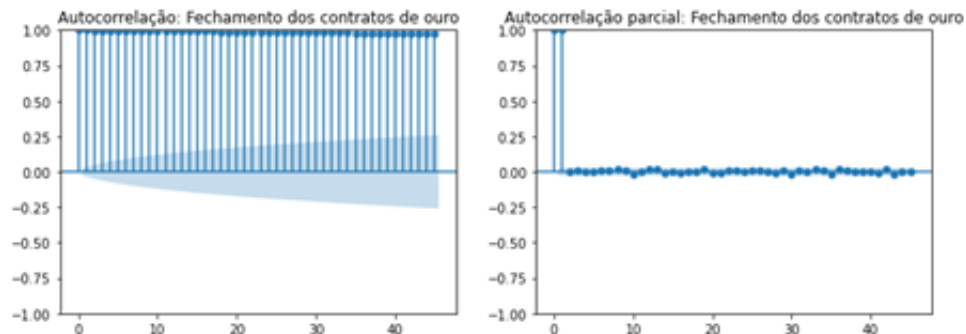
Após os estudos do capítulo anterior se descartou o uso do a utilização do modelo Arima uma vez que ele não considera a sazonalidade. Para a definição modelo individuais para as séries utilizou a biblioteca pmdarima com o pacote auto arima, e estabeleceu como M igual a 5. Utilizando de dias úteis de uma semana. Segue a tabela com o resumo dos atributos gerados pelo auto arima:

Figura 10 – Decomposição da série temporal de ouro



Fonte:Autor(2023)

Figura 11 – Autocorrelações da série temporal de Ouro



4.3 Modelos LSTM E GRU Fonte:Autor(2023)

Os modelos de aplicados a data a seguir fora executado utilizando o pacote keras do tensor flow, para conseguirmos executar esses tipo de modelos, as séries foram normalizadas pelo meétodo “min-max” e tratadas com uma função que separa o treinamento e o teste em dois arrays, que si diferenciam pelo número de elementos passados que consideramos para prever o próximo nesse casa utilizamos dois, para os dois modelos. Os modelos foram construidos utilizados os parâmetros demonstrados durante o curso. Na figura 12 podemos ver o resumo do modelo LSTM, contendo 5 camadas sendo elas dívidas em duas LSTM com 16 e 32 parâmetros, a terceira sendo uma de Dropout de 0,2 e a última uma camada densa, combinando os dados. A LSTM totalizou 7.457 paramentos e camadas que necessitam de função de ativação foram ativadas por uma relu. Na figura 12 mostra o resumo da rede e

Tabela 1 – Modelos ótimos encontrados pelo auto arima

Série Temporal	Arima	SARIMA
Boi	(0,1,0)	(0,0,1,5)
Café	(0,1,0)	(0,0,0,5)
Petróleo	(3,1,1)	(0,0,1,5)
Ouro	(0,1,0)	(0,0,0,5)

Fonte:Autor(2023)

como é o fluxo interno dela.

Figura 12 – Resumo da Rede LSTM

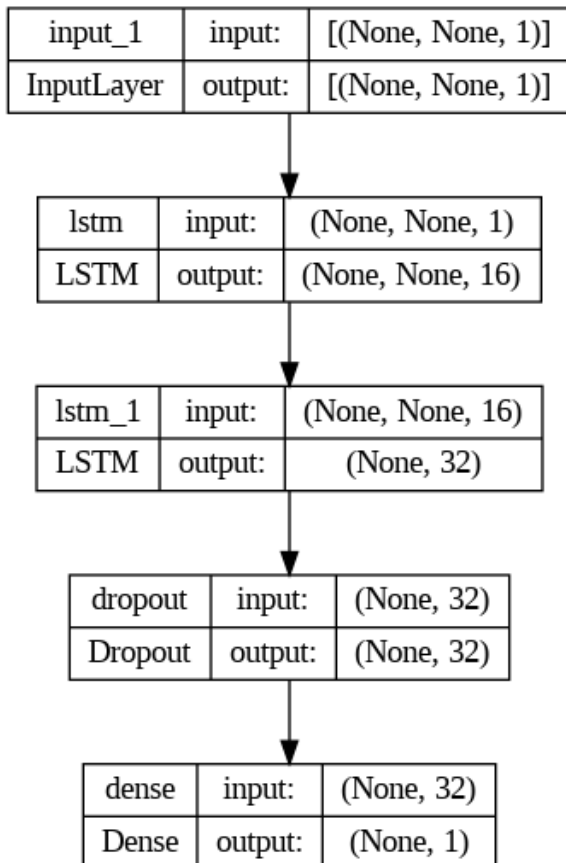
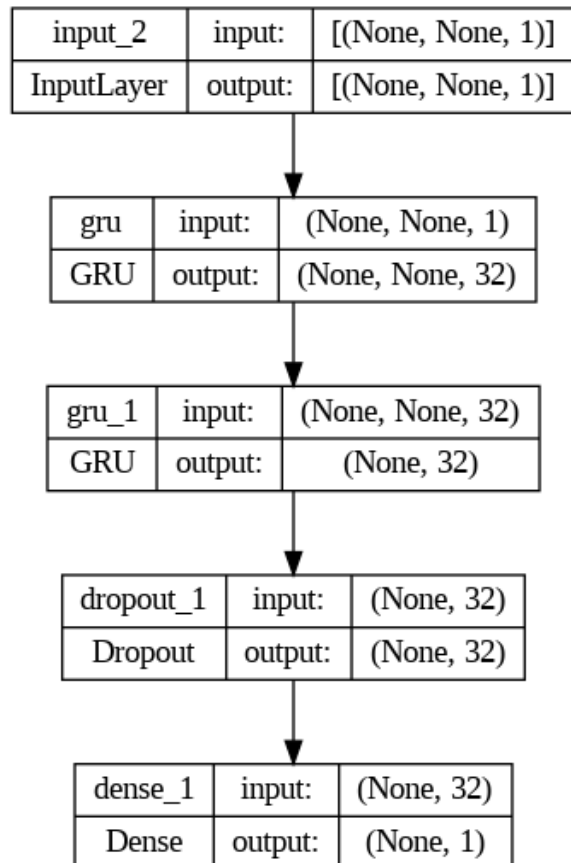


Figura 13 – Resumo da Rede GRU



Fonte: Autor(2023)

Na rede GRU, o número de parâmetros aumentou saltou para 9.729, mas a configuração é quase a mesma, trocamos as camadas LSTM por camadas GRU, como vista na figura 13.

Durante o estudo utilizamos de os dados foram divididos em duas instancias de array já que isso é uma necessidade do modelo, deslocando sempre uma unidade na a mostra utilizada como entrada para o modelo, nesse trabalho utilizou-se de 2 unidades da

amostra para prever a próxima. Todos os modelos foram rodados durante 80 épocas. E para função perda foi considerado o erro quadrático médio, otimizador adam com taxa de aprendizado 0.0001, métricas MAE. As convergências dos treinamentos convergiram por volta sempre da décima época.

Os erros obtidos por cada modelo em todas as séries estão resumidos na tabela a seguir

Tabela 2 – Erros encontrados no teste dos Modelos LSTM e GRU

Série Temporal	Modelo	MAE	MSE
Boi	LSTM	0.0160	0.0004
	GRU	0.0193	0.0006
	SARIMA	23,652	654,3864
Café	LSTM	0.0003	0.0137
	GRU	0.0004	0.0151
	SARIMA	14,0656	312,68
Petróleo	LSTM	0.0186	0.0006
	GRU	0.0168	0.0005
	SARIMA	9,6390	149,1924
Ouro	LSTM	0.0283	0.0014
	GRU	0.0189	0.0006
	SARIMA	125,1458	19349,4323

Fonte: Autor(2023)

5 CONCLUSÃO

É notável como temos cada vez mais acesso a dados com volume e de fontes confiáveis que podem nos auxiliar a tomada de decisão. Do ponto de vista de uma organização saber utilizar de base de dados públicas que tragam alguma percepção a respeito sobre as variáveis que impactam o dia-a-dia da empresa é um ganho competitivo muito importante. Que pode dar suporte a dinâmica competitiva do mundo atual.

O resultado obtido foi satisfatório em consolidar o que antes era uma impressão para um conhecimento empírico, uma vez que o estudo mostrou o que ainda não sabemos e qual o caminho a trilhar para desenvolver as habilidades adquiridas no programa de estudo do MBA, juntando teoria com prática.

Do ponto de vista da predição dos modelos, as redes neurais tiveram uma baixa variabilidade nos erros e um bom desempenho nos testes, porém ainda temos que fazer mais testes saber qual é a frequência ideal para utilização dos dados, utilizar mais elementos ao longo do estudo traz um ganho de performance. Entender em quais casos o ARIMA poderá ter um desempenho relevante para a análise. Saber onde podemos deixar os modelos com o custo computacional melhor e mantendo a confiança.

A pesquisa estudou o desempenho de métodos de predição para séries temporais financeiras, utilizando de dados públicos e volumosos para gerar uma boa fundamentação, na tomada de decisão. Analisando do desempenho dos modelos estudados, mostrou a tamanha complexidade do assunto. E partir desse estudo começou a se construir um caminho para o conhecimento que permita melhorar a tomada de decisões, uma vez que o estudo trouxe uma teoria e prática de como os modelos se comportam diante de um cenário de muita volatilidade, uma falta de padrão nos dados e com dados com diferentes dinâmicas.

Para o emprego dessa ferramenta, ainda se nota necessária, uma maior experiência e experimentação dos modelos para maior confiança. E assim começar a se construir uma cultura de tomadores de decisões pautados por dados.

Nos futuros trabalhos se pretende trabalhar com moedas e ver como elas se comportam na predição de modelos de redes neurais, com o intuito final de trabalhar com dados de alta frequência. Prevendo tendências de movimento.

REFERÊNCIAS

- GOMES, M. F. **Formação de Preços de Commodities no Brasil**. 2002. 52 p. Dissertação (Dissertação de Mestrado) — FGV - Escola de Administração de Empresas de São Paulo, São Paulo, 2002. Área de Concentração: Economia de Empresas.
- HAYKIN, S. **Redes neurais: princípios e prática**. [S.l.]: Bookman Editora, 2001.
- MIGLIATO, A. L. T. **Detecção de Outliers em Dados não Vistos de Séries Temporais por meio de Erros de Predição com SARIMA e Redes Neurais Recorrentes LSTM e GRU**. 2021. Tese (Doutorado) — Universidade de São Paulo, 2021.
- MORETTIN, P.; TOLOI, C. **Análise de Series Temporais, Edgard Blucher**. [S.l.]: São Paulo, SP, 2004.
- MORETTIN, P. A. Mae 5905: Introdução a ciência de dados.
- OLAH, C. **Understanding LSTM Networks**. 2015. <<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>>. Acesso em: 20/02/2023.
- SANDRONI, P. **Dicionário de economia do século XXI**. [S.l.]: Editora Record, 2016.
- SOUZA, R. C. T. D. Previsão de séries temporais utilizando rede neural treinada por filtro de kalman e evolução diferencial. Curitiba: Pontifícia Universidade Católica do Paraná. Tese de doutorado, 2008.
- WERNER, L.; RIBEIRO, J. L. D. Previsão de demanda: uma aplicação dos modelos box-jenkins na área de assistência técnica de computadores pessoais. **Gestão & Produção**, SciELO Brasil, v. 10, p. 47–67, 2003.