

Universidade Federal de Minas Gerais

Igor Costa de Freitas

**Aplicação de pacotes estatísticos e computacionais na  
previsão de séries temporais financeiras**

Belo Horizonte

2022

Igor Costa de Freitas

**Aplicação de pacotes estatísticos e computacionais na  
previsão de séries temporais financeiras**

Trabalho de conclusão de curso  
apresentado à Escola de Engenharia da  
Universidade Federal de Minas Gerais,  
como exigência para aprovação em  
Trabalho de Conclusão de Curso II

Orientador: Marcelo Azevedo Costa

Belo Horizonte

2022

# **Aplicação de pacotes estatísticos e computacionais na previsão de séries temporais financeiras e análise do desempenho de modelos**

**Igor Costa de Freitas**

Graduando em Engenharia de Produção - UFMG

E-mail: [igorfreitascp@gmail.com](mailto:igorfreitascp@gmail.com)

**Marcelo Azevedo Costa**

Professor Titular do Departamento de

Engenharia de Produção/Universidade Federal de Minas Gerais - UFMG.

E-mail: [macosta@ufmg.br](mailto:macosta@ufmg.br)

## **Resumo:**

As competições de previsão de séries temporais têm ganho maior popularidade no campo da ciência dos dados, sobretudo as competições M, que existem há mais de 40 anos. Nessas disputas muitos métodos são testados em busca do melhor ajuste. Os modelos ARIMA, Suavização Exponencial, Naïve-Bayes, XGboost e Redes Neurais são bastante comuns na literatura da previsão de séries temporais, por isso serão utilizados em uma aplicação prática de testagem e seleção de métodos. Para selecionar qual o melhor método é necessário escolher uma fonte de dados coesa, tratar esses dados para a formatação requerida pelo tipo de análise e selecionar uma medida de acurácia adequada. A base de dados escolhida conta com informações de abertura, fechamento, máxima e mínima do preço das ações da *Microsoft Corporation* na bolsa dos Estados Unidos referente a 723 dias, distribuídos em 150 semanas. Essa escolha foi feita por se tratar de dados de alta disponibilidade e por ser uma companhia de grande notoriedade mundial, para reforçar o caráter prático do estudo. Neste estudo são exploradas algumas dessas técnicas comuns nas competições de previsão de séries, o objetivo é demonstrar todo o processo que envolve a análise e teste de desempenho dos métodos preditores, o que irá embasar a investigação e seleção de quais são os melhores modelos para previsão dos retornos históricos.

**Palavras chave:** séries temporais; modelo de previsão; retornos históricos; XGboost; ARIMA; Redes Neurais; Naïve -Bayes; Suavização Exponencial.

## Abstract:

The time series competitions are gaining increasingly popularity among the data science community, especially the M competition, that has existed for more than 40 years. In these contests many methods are tested chasing the best fit. Models such as ARIMA, Exponential Smoothing, Naïve-Bayes, XGboost and Neural Nets often appear in the séries forecasting literature, this paper will use those models in a practical application and method choosing. To select the best one of these it is necessary to pick up a cohesive data source, treat the data for the fit in analysis and define a proper accuracy estimator. The database selected contains information of opening, closing, high and low daily values of *Microsoft Corporation* stock price in the United States stock exchange of 723 days, spreaded in 150 weeks. This choice was made because of the high availability of this kind of data and the great notoriety of the company around the world, which reinforces the practical purpose of this study. In this paper these techniques are exploited with the purpose of demonstrating all the stages of the process of analysis and performance testing of the forecast methods, which will launch a proper knowledge base for selecting the best of those models for returns forecasts.

**Keywords:** Time series; Forecasting models; Returns; XGboost; ARIMA; Neural nets; Naïve -Bayes; Exponential smoothing.

## Sumário

1. Introdução.....	5
2. Materiais e Métodos.....	7
2.1. Modelos.....	7
2.2. Bases de dados.....	10
2.3. Medidas de desempenho.....	14
2.4. Plataformas computacionais.....	14
2.5. Validação.....	15
3. Resultados.....	16
4. Discussão.....	20
5. Conclusão.....	24
6. Referências.....	25
7. Anexos.....	31

## Lista de ilustrações

Figura 1 - Gráfico <i>Candlestick</i> .....	11
Figura 2 - Preços de fechamentos x retornos históricos.....	12
Figura 3 - Teste Dickey-Fuller.....	12
Figura 4 - Esquema de validação.....	13
Figura 5 - <i>Heatmap</i> do RMSE diário de cada modelo.....	18

Figura 6 - Curvas RMSE.....	19
Figura 7 - <i>Heatmap</i> do XGboost e Redes Neurais.....	21
Figura 8 - Box e Violino do XGboost e Redes Neurais.....	21
Figura 9 - Box e Violino do ARIMA e Suavização Exponencial.....	22
Figura 10 - Barras de ganhos dos modelos.....	23

### Lista de Tabelas

Tabela 1 - Modelos.....	10
Tabela 2 - Recorte da base de dados.....	13
Tabela 3 - Configuração dos modelos.....	16
Tabela 4 - Recorte da base de dados resposta.....	17
Tabela 5 - Erros médios quadráticos por modelo.....	17

### Lista de equações

Equação 1 - Raiz quadrada da média dos quadrados dos erros.....	14
Equação 2 - Função de ganho considerada para avaliar os métodos de previsão em relação ao método <i>benchmark</i> .....	23

## 1. Introdução

Previsões fazem parte do dia-a-dia de todas as pessoas e empresas, desde problemas como chegar em determinado local em certo horário à decisões estratégicas da cadeia de suprimentos, como explica MAKRIDAKIS et al. 2021. Existem diversas aplicações para previsões de valores e comportamento ao longo do tempo, como em FILDES et al. 2019, que estuda o cenário de vendas de grandes entidades do varejo e seus métodos para previsão de demanda. MA e FILDES, 2022, apresentam um método de fluxo de demanda baseado nas vendas que pode ter aplicações bem úteis para grandes redes de fornecedores. Existem modelos de inventário, bastante utilizados na cadeia de suprimentos intermediária, que consideram flutuações de vendas e de estoque, um destes modelos é apresentado em PRAK e ROGETZER, 2022, baseando-se em dados da competição M5.

No recorte do mercado de ações a informação mais valiosa é aquela que determina qual será o preço futuro de determinados ativos, pois com esse dado é possível programar uma sequência de negociações financeiras no sentido de multiplicar o patrimônio investido inicialmente. Porém é extremamente complexo prever qual será o preço exato da ação no futuro, sobretudo por um longo período, é neste ponto que aproximações e/ou intervalos de predição ganham seu valor neste campo, pois é possível, utilizando os dados do passado, tentar definir um comportamento padrão e/ou tendências no valor. A Ciência dos Dados sempre se mostrou bastante presente no mercado financeiro com propósito de fornecer previsões de possíveis cenários, encontrar padrões e, conseqüentemente, multiplicar patrimônio. Muitas empresas sofreram e sofrem as conseqüências do mercado de ações, algumas perdem e outras ganham valor. Para sobreviver neste território competitivo e caminhar rumo ao crescimento, inúmeras estratégias são adotadas pelas companhias e investidores, prospectando os prováveis cenários que só podem ser fornecidos ao contemplar dados do passado. Quanto mais dados e de melhor qualidade mais valor informativo possui, é como se os dados fossem ou matéria-prima ou recurso, alguns estudos apresentam projetos abrangentes já com esse *benchmarking* (VOS, 2014) e (KIM, 2017).

Segundo HUNG et al. (2021), investidores mais experientes utilizam análises que vão além do crescimento e decaimento aparente no preço de ações. Por exemplo, mesmo

em um dia negativo(preço de fechamento inferior ao preço de abertura) pode ser observado um comportamento que sinaliza que a compra pode ser um bom negócio, devido à fatores como preço de abertura predominantemente crescente, dilatação do bloco do gráfico (se relaciona com volume de negociações) ou até uma alta e baixa do dia próximas uma da outra. Todos esses detalhes influenciam na tomada de decisão dos profissionais do mercado, são conceitos que estão mais próximos de um conhecimento tácito do que de fato uma predição matemática, embora ainda haja um paralelo com a análise quantitativa em termos de avaliar uma autocorrelação ou um padrão de repetição. Em ambas abordagens são utilizados os dados históricos, para definir padrões futuros. Em um levantamento bibliográfico realizado pelo autor previamente (Freitas, 2022), o campo da previsão de séries foi explorado através das competições históricas de previsão, sobretudo as competições M, a pioneira e a mais famosa.

Em análises estatísticas/computacionais trabalha-se em busca de minimizar o erro. Por isso as competições M adotaram em sua última edição duas modalidades de previsão: acurácia e incerteza (MAKRIDAKIS et al. 2022). Essa mudança indica que cada vez mais a área de previsão de dados tem caminhado no sentido da construção de um modelo que tenha como maior discriminante uma medida de erro. Esta tendência no ajuste de modelo foi dominante nas últimas competições M e dado o levantamento bibliográfico feito anteriormente o presente estudo buscou colocar em prática os mesmo modelos utilizados na disputa de previsões, foi usada uma base de dados histórica das séries de retornos das ações da companhia *Microsoft Corporation*. Esses dados são fontes de outros estudos que buscam avaliar o comportamento do preços, como VU et al. 2021, que utiliza a mesma base de dados que o presente estudo e algumas outras de mesmo formato mas de diferentes instituições para determinar os prognósticos dos valores da ações através de árvores de classificação e validação cruzada. TICKNOR, 2013 e KAZEM et al. 2013 utilizam informações de empresas como a *Microsoft* para ajuste de modelos e análise do mercado de ações.

O objetivo deste trabalho é aprofundar o conhecimento teórico e prático do campo da Ciência dos Dados, avaliando o desempenho dos modelos ARIMA, XGboost, Exponencial, Naïve-bayes e Redes Neurais para dizer quais os mais adequados para este conjunto de dados temporais. Este *mix* de métodos visa contemplar modelos

clássicos e modelos computacionais em constante evolução. HEWAMALAGE et al (2021) demonstra que técnicas mais tradicionais para a previsão de séries temporais podem fazer frente à técnicas mais atuais e também podem servir como referência para outros modelos. Enfim, não existe bala de prata no campo de predição de dados.

Foi dado enfoque mais específico aos recursos disponíveis para previsão de séries. As ferramentas e recursos empregados foram selecionados por familiaridade e forte presença nas competições M. Para isso, as metodologias de Levantamento bibliográfico e pesquisa experimental foram aplicadas no sentido de formar uma base robusta acerca da aplicação de modelos e tendências nas competições, o estudo foi desenvolvido voltado para implementação.

## **2. Materiais e Métodos**

Um estudo prático de séries temporais voltado para o desempenho de modelos requer pelo menos quatro parâmetros. Primeiramente, a definição de quais modelos deseja-se comparar as previsões e performance. Segundo, a escolha de uma série histórica característica. Em seguida é preciso optar por uma medida de desempenho de métodos indicada para a classe de dados escolhida e quais seus aspectos, tudo isso é feito em ambientes de desenvolvimento que envolvem linguagem computacionais, softwares e arquivos com dados.

### **2.1 Modelos**

As metodologias *Auto Regressive Integrated Moving Average* (ARIMA), Naïve Bayes e Suavização Exponencial são métodos utilizados desde o início das competições de séries e no mercado em geral. Essas metodologias já foram amplamente avaliadas em inúmeros estudos (ARMSTRONG e COLLOPY, 1998) com citações encontradas desde o fim da década de 1970 e 80. Neste estudo serão utilizados as metodologias ARIMA, Suavização Exponencial e Naïve Bayes na previsão da série temporal em estudo. Em contrapartida os modelos de *Artificial Neural Networks* (ANN ou simplesmente Redes Neurais) e *Extreme Gradient Boosting* (XGboost) são técnicas mais atuais requerem uma maior capacidade computacional e uma customização e *tunning* mais específica



com o intuito de desenvolver um modelo com erro mínimo. Estes modelos também serão ajustados para previsão da série temporal em estudo.

O modelo ARIMA é uma combinação dos modelos Autoregressive e Moving Average, no qual suas representações estão na forma de um conjunto  $(p,d,q)$ , onde  $p$  representa o coeficiente do modelo auto regressivo,  $q$  denota o modelo de média móvel e  $d$  diz respeito a integração dos dados(PARMAR e BHARDWAJ, 2014). Para um bom ajuste deste método existe uma medida bastante utilizada na estatística, o *Akaike Information Criteria(AIC)*, que mensura a qualidade do ajuste, selecionando a combinação  $(p,d,q)$  ótima, ou seja, com melhor ajuste.

Na Suavização Exponencial o algoritmo trabalha no sentido de continuamente usar dados mais recentes na previsão seguinte(KALEKAR et al. 2004). Isto significa que a previsão do retorno em certo dia será mais influenciada pelos retornos imediatamente anteriores. É comum utilizar Suavização Exponencial em camadas, as quais são indicadas para cada tipo de série histórica. No caso de séries que apresentam tendência, duas camadas são indicadas para o ajuste e em cenários de tendência e sazonalidade o modelo em três camadas é o mais usado. Existem também fatores de sazonalidade multiplicativa que podem ser considerados, porém no presente estudo será usado o método em apenas uma camada, por se tratar da série de retornos.

Os métodos bayesianos de previsão foram criados a fim de descrever modelos de classificação(MUTHUKUMAR e ZHONG, 2021). Porém, esses modelos vem sendo cada vez mais utilizados no campo do Machine Learning. O modelo Naïve Bayes assume que a série temporal é composta por uma única componente independente, sendo a melhor estimativa de predição o valor anterior observado. Por essa razão, o método é conhecido como Naïve (em inglês, ingênuo).

Ao final da análise pretende-se avaliar se em um horizonte de uma semana os modelos mais simples performam tão bem quanto os modelos mais complexos, e entre esses qual apresenta melhor/menor erro quadrático médio em cada dia considerando dados de teste.

Alguns métodos são puramente estatísticos e utilizam de médias móveis e/ou ponderadas dos valores observados, definindo pesos ou coeficientes para cada unidade de previsão. É o caso dos três primeiros métodos descritos acima, que possuem definições matemáticas bem estabelecidas e amplamente aplicadas em diversos problemas de séries temporais. De certa forma, esses métodos definem o núcleo de outros métodos mais complexos que se utilizam de combinações como árvores e florestas de previsão. Atualmente é comum o surgimento de novas formas de prever valores futuros. Alguns métodos vêm se sobressaindo e cada vez mais estão presentes em problemas do tipo previsão de séries. Dois desses métodos serão usados no presente estudo, o XGboost e as Redes Neurais artificiais. Ambos são modelos computacionais de aprendizado de máquina, isto é, foram criados a partir de outros métodos e idealizados a partir de estruturas conhecidas pela ciência. O XGboost por exemplo é definido a partir de árvores/florestas de decisão, nas quais os dados são constantemente recortados na fase de treino. Essas árvores formam um algoritmo de aprendizado em conjunto e combinam vários modelos de Machine Learning em busca do melhor ajuste para previsão (ALIM et al. 2020). Reutilizar uma arquitetura computacional tem sido cada vez mais uma realidade no mundo de códigos de programação. As Redes Neurais Artificiais foram inspiradas a partir do sistema nervoso central biológico, no qual cada neurônio se comunica com outro através de sinapses em uma grande rede ramificada (KOLARIK e RUDORFER, 1994). No caso do algoritmo os nós são valores que se conectam com outros nós e geram camadas de combinações que são repetidamente criadas em busca de melhores valores de saída. É possível definir quantidades máxima e mínima de nós, números de camadas e número de neurônios em cada camada, sempre de acordo com a complexidade do problema.

**Tabela 1** - Breve descrição de cada modelo que será utilizado.

Modelo	Descrição
<i>ARIMA</i>	Modelo autoregressivo integrado de médias móveis, é um método que combina AR(autoregressivo) e MA(moving average). A notação $(p,d,q)$ informa as ordens de cada parte do modelo, sendo $p$ a ordem do modelo AR, $d$ o grau de diferenciação(integração dos métodos) e $q$ é a ordem do modelo de média móveis.
<i>Exponential Smoothing (Suavização Exponencial)</i>	Método que oferece maior peso aos valores passados mais recentes. São utilizadas camadas para lidar com tendência e sazonalidade.
<i>Naïve-Bayes</i>	A princípio um modelo de classificação mas com adaptação para séries temporais. Semelhante ao algoritmo exponencial, com preferência para valores imediatamente anteriores, por isso é conhecido como Naïve (em inglês, ingênuo).
<i>Xgboost</i>	Modelo computacional que utiliza de árvores de decisão na previsão e procura internamente por combinações que oferecem o melhor ajuste.
<i>Neural Nets (Redes Neurais)</i>	Utiliza nós de uma rede em combinações de camadas que procuram por melhores valores da saída. Sua rede é altamente personalizável em aspectos de ramificação e interações.

Fonte: Elaborado pelo autor.

## 2.2 Bases de dados

A fonte de dados utilizada neste trabalho é composta pelos retornos diários da cotação em ações da empresa *Microsoft Corporation* na bolsa de valores dos Estados Unidos. Esta escolha resulta do interesse em testar os modelos em uma de suas principais aplicações, utilizando os registros diários do mercado financeiro.

Para analisar uma base de dados dos preços históricos da ação de uma determinada empresa é preciso entender os conceitos que envolvem a dinâmica de seus preços. Na plataforma R está disponível o pacote *quantmod*, para leitura e processamento de séries temporais financeiras. Com o pacote é possível obter e analisar séries históricas

para diferentes períodos de registro, este recurso foi a fonte principal de dados financeiros no estudo.

**Figura 1** - Gráfico *Candlestick*, característico no mercado de ações; Conjunto *OHLC* em verde representa dias de fechamento maior que abertura. Em laranja, são indicados os dias de fechamento menor que abertura. Extremos dos *boxplots* indicam máximo e mínimo diário. Os limites das caixas indicam abertura e fechamento também diário.

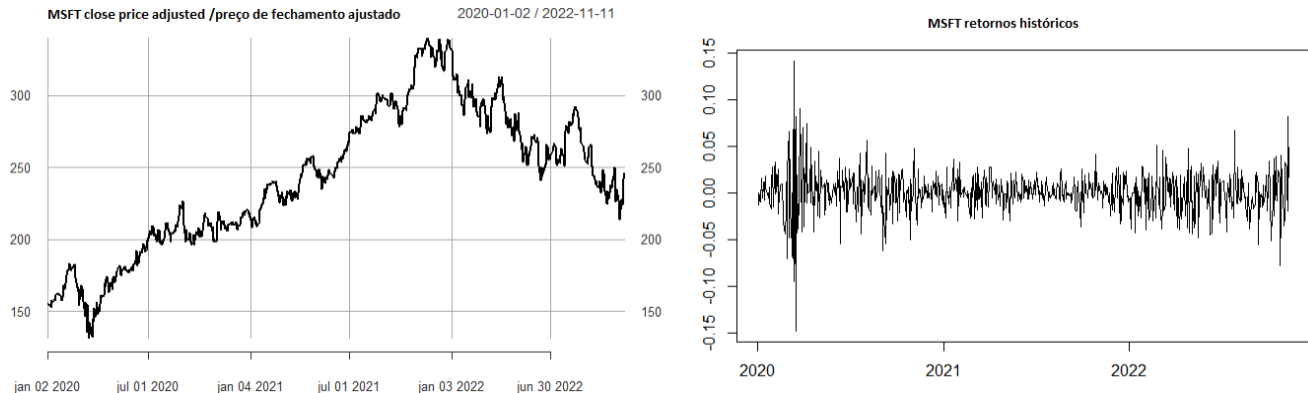


Fonte: Elaborado pelo autor.

É comum ao expressar graficamente as séries de preços de ações a utilização do gráfico *Candlestick*(Figura 1). Trata-se de um tipo de ilustração que contém várias informações(DO PRADO et al 2013). A notação característica do preço da ação envolve o preço de abertura do dia, o preço máximo e mínimo e o preço de fechamento. Essa combinação é conhecida como conjunto *OHLC*(*Open, High, Low, Close*). Há também o volume diário que diz respeito à quantidade de negociações realizadas no dia e o preço de fechamento ajustado, pois trata-se do preço de fechamento após o cálculo de dividendos. Estes valores são hierárquicos, isto é, a sua sequência é tão fundamental quanto seus valores, e sua estrutura deve ser respeitada em qualquer tipo de análise.

São inúmeros os estudos que utilizam bases de dados do mercado financeiro. Alguns se interessam pelo preço de fechamento da ação no dia, porém é comum optar pela análises do retornos(o preço final de hoje subtraído do preço final de ontem, dividido pelo preço final de ontem), para eliminar não-estacionariedade presente nos dados como discutido em ZIEL (2016) e AHMED et al (2020). Dessa forma, as metodologias são mais precisas em suas respostas, pois não precisam considerar a tendência dos erros da série em função do tempo.

**Figura 2** - Preço de fechamento X retornos históricos. O primeiro com não-estacionariedade e o segundo seguindo regime aparentemente estacionário, sem uma tendência ao longo do tempo.



Fonte: Elaborado pelo autor.

Para saber precisamente se a série de retornos é estacionária ou não estacionária, é necessário testar essa condição, para isso foi escolhido o teste Dickey-Fuller Aumentado, apresentado em detalhes por MUSHTAQ, 2011. Este teste frequentemente é utilizado em estudos estatísticos para avaliar tendência, variância e autocorrelação da amostra. Trata-se de um teste estatístico clássico, com hipótese nula e hipótese alternativa, essas:

- **H<sub>0</sub>** : A série apresenta estrutura dependente do tempo e variância não constante ao longo do tempo, isto é, não é estacionária;
- **H<sub>a</sub>** : A série temporal é estacionária.

O teste forneceu o seguinte resultado, deve-se rejeitar a hipótese nula e aceitar a hipótese alternativa pois o *p-valor* do teste(<0,01) foi inferior ao nível de significância(0,05). Abaixo a saída do teste.

**Figura 3** - Resultados do teste Dickey-Fuller Aumentado.

```
> adf.test(MSFT$retornos)
warning: p-value smaller than printed p-value
Augmented Dickey-Fuller Test

data: MSFT$retornos
Dickey-Fuller = -8.7919, Lag order = 8, p-value = 0.01
alternative hypothesis: stationary
```

Fonte: Elaborado pelo autor.

Os dados extraídos pelo *quantmod* necessitam de tratamento, pois estão organizados no formato *OHLC*, apresentando também volume comercializado e valores ajustados. Em um primeiro momento, não são apresentados os retornos. Neste trabalho, também

foram utilizados os pacotes *dplyr* para manipulação de dados e o pacote *lubridate* para manipulação de datas. Utilizando o pacote *lubridate* é possível definir o dia da semana e semana do ano. Um recorte dos dados é mostrado na tabela 2, a seguir.

**Tabela 2** - Recorte da base de dados adequada para análise.

retornos	Date	wday	wdayname	month_day	week	wk_total
0,000000000000000000	02/01/2020	5	qui	2	0	1
-0,01245174985841580	03/01/2020	6	sex	3	0	1
0,00258481914590925	06/01/2020	2	seg	6	1	2
-0,00911775771312173	07/01/2020	3	ter	7	1	2
0,01592837903378120	08/01/2020	4	qua	8	1	2

Fonte: Elaborado pelo autor.

A desagregação da data em suas componentes de dia da semana e semana do ano permite que os algoritmos sejam avaliados por características de semana utilizando um procedimento de validação cruzada. É objetivo deste trabalho avaliar a capacidade preditiva dos modelos para cada dia da semana em um conjunto de teste definido para as últimas semanas da série temporal. Com relação ao procedimento de validação cruzada, a figura 3 ilustra os dados para treino(anos de 2020 e 2021) e teste(ano de 2022 até 11 de novembro, sexta-feira da semana 45). Será aplicada uma abordagem dinâmica para a validação cruzada. Em suma, os conjuntos de treino e teste serão dinâmicos. Os conjuntos de teste serão definidos como os dias que compõem a semana seguinte ao conjunto de treino

**Figura 4** - Esquema de validação utilizado.

iter	Semanas de 2020	Semanas de 2021	Semanas de 2022
1	TREINO		TESTE1
2	TREINO		TESTE2
3	TREINO		TESTE3
	.		
	.		
	.		
n	TREINO		TESTEn

Fonte: Elaborado pelo autor.

A estruturação dos dados conforme a tabela 1 permite estruturar uma tabela de previsões diárias para cada semana dos conjuntos de teste. Sucintamente, os conjuntos de treino e de teste podem ser definidos apenas utilizando um contador de semanas. Assim a tabela inicial de treino compreende todas as semanas dos anos de 2020 e 2021 e mais uma semana do ano de 2022. A cada iteração, o conjunto de teste é

simplesmente a próxima semana. Para implementação do procedimento de validação cruzada, é suficiente o tipo *data.frame* para alguns modelos. Para outros modelos é necessária a transformação dos dados para o tipo *timeSeries*.

### 2.3 Medida de desempenho

Neste estudo será aplicada a raiz quadrada da média dos quadrados dos erros de previsão (*RMSE*), mostrada a seguir. Essa medida de erro foi escolhida por representar uma medida na mesma unidade dos valores ajustados (CHAI e DRAXLER, 2014). Por outro lado, é uma medida sensível a valores discrepantes dos erros.

**Equação 1** - Raiz quadrada da média dos quadrados dos erros.

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

O RMSE tende a ser uma medida superior para a avaliação de previsão de séries temporais pois lida melhor com distribuições estatísticas dos erros do que o erro médio absoluto (MAE). É comum utilizar mais de uma métrica em pesquisas sobre séries temporais. Neste estudo, será utilizada apenas uma medida de performance.

### 2.4 Plataformas computacionais

No desenvolvimento deste estudo foram utilizadas as linguagens de programação R, na IDE Rstudio, e Python, utilizado no Visual Studio Code. Na linguagem R, como citado, existem pacotes de funções dedicados a certos tipos de análise. Por se tratar de um software voltado à manipulação e análise de dados existem diversas maneiras de atingir um objetivo seja em ajustes de modelos, análises exploratórias ou plotagem de gráficos.

Um exemplo é o pacote *quantmod* que foi usado para extração de dados, porém o *BatchGetSymbols* é capaz de fornecer os mesmos dados. Os pacotes listados a seguir foram utilizados:

- **quantmod**: Extração dos dados dos preços de ações *Microsoft Corporation*;
- **dplyr**: Manipulação de dados;
- **lubridate**: Manipulação de datas;
- **timeSeries**: Conversão de dados de séries em objeto *Time-Series* do R;
- **forecast**: Ajuste e previsão dos modelos ARIMA e Suavização Exponencial;
- **forecastxgb**: Ajuste e previsão do modelo XGboost;
- **bayesforecast**: Ajuste e previsão do modelo Naïve-Bayes-Stan;
- **nnfor**: Ajuste e previsão do modelo de Redes Neurais;
- **ggplot2**: Plotagem de gráficos e visualização de dados;
- **magrittr**: Utilização de *pipelines* juntamente com *dplyr*;
- **readr**: Manipulação de arquivos de texto;
- **xlsx**: Manipulação de arquivos .csv, .xls, .xlsx;

Para linguagem Python, sobretudo para plotagem e análise final dos dados, foram utilizadas os módulos:

- **pandas**: Busca e manipulação de dados;
- **matplotlib** e **seaborn**: Plotagem de gráficos;
- **datetime**: Manipulação de datas;
- **numpy**: Manipulação de listas e arranjos numéricos;
- **functools**: Utilização do método *reduce*;

Recursos do *Microsoft Excel* também foram empregados para a construção de tabelas e para processamento dos dados em arquivos .csv e .xlsx. A documentação completa do estudo pode ser acessada via GitHub(<https://github.com/igxchl/tg-igor-freitas-time-series>)

## 2.5 Validação

O código computacional implementado computa os erros diários de cada semana para o conjunto dinâmico de teste e, além disso, realiza a previsão semanal considerando a possibilidade de *missing data*, ou dias que representam feriados. Para esses dias, não são estimadas as previsões da série temporal. Uma outra necessidade para comparação de modelos é saber medir o tempo para execução do modelo. Essa é uma informação essencial para avaliar a viabilidade do método.



Com intuito de tornar a análise comparável em mais de uma linguagem foi desenvolvido um pseudocódigo utilizando a linguagem R. O código encontra-se nos anexos.

### 3. Resultados

Uma vez que a estacionariedade foi atingida, como exposto no teste, o modelo pode ser ajustado aos dados. Cada ajuste conta com uma configuração específica do método, essas definem o quanto o modelo irá se ajustar/interpretar os dados. Algumas funções encontram automaticamente as configurações mais indicadas para realizar a previsão, o que facilita um bom ajuste.

Por se tratar do método mais simplório entre os escolhidos, o algoritmo Naïve-Bayes será utilizado como benchmarking em algumas comparações da medida de desempenho para o teste.

**Tabela 3** - Configuração dos métodos ajustados.

Modelo	Especificação
<i>ARIMA</i>	ARIMA(1,0,0) e média zero
<i>Exponential Smoothing</i> ( <i>Suavização Exponencial</i> )	Alpha = 0.83 e Frequência MAX = 5
<i>Naïve-Bayes</i>	Cadeias = 4, Iterações = 2000 e média zero
<i>Xgboost</i>	Lag Max = 8, lambda = 1 e sem sazonalidade
<i>Neural Nets</i> ( <i>Redes Neurais</i> )	Nº nós = 100, replicações = 40 e camadas = 1

Fonte: Elaborado pelo autor.

A partir dos ajustes dos modelos, é formada uma tabela de erros onde as linhas representam as semanas e as colunas são os dias da semana. A tabela 4 ilustra um recorte da tabela de erros. Em dias de feriados, os erros são considerados como dados ausentes (NA) em virtude da ausência de pregão. As tabelas completas de todos os modelos se encontram no Anexo do texto.

**Tabela 4** - Estrutura de um *data frame* resposta gerado em todos os ajustes de modelos, cinco primeiras linhas de uma das tabelas.

<b>Seg</b>	<b>Ter</b>	<b>Qua</b>	<b>Qui</b>	<b>Sex</b>
0,006322005	0,019220906	0,040482841	0,015222159	0,001369312
0,00322461	-0,002260085	0,000501547	0,048571811	-0,015493601
NA	0,011698966	0,014442357	0,00987637	0,025258715
0,008502134	0,037189532	-0,02182714	-0,010255034	-0,027877035
-0,008223061	0,003789748	-0,012585933	0,044127614	-0,015508393

Fonte: Elaborado pelo autor.

A partir da tabela dos erros são estimados os erros de previsão para cada um dos modelos. Este resultado é apresentado na tabela 5. Optou-se por apresentar o RMSE para cada dia da semana e para cada método.

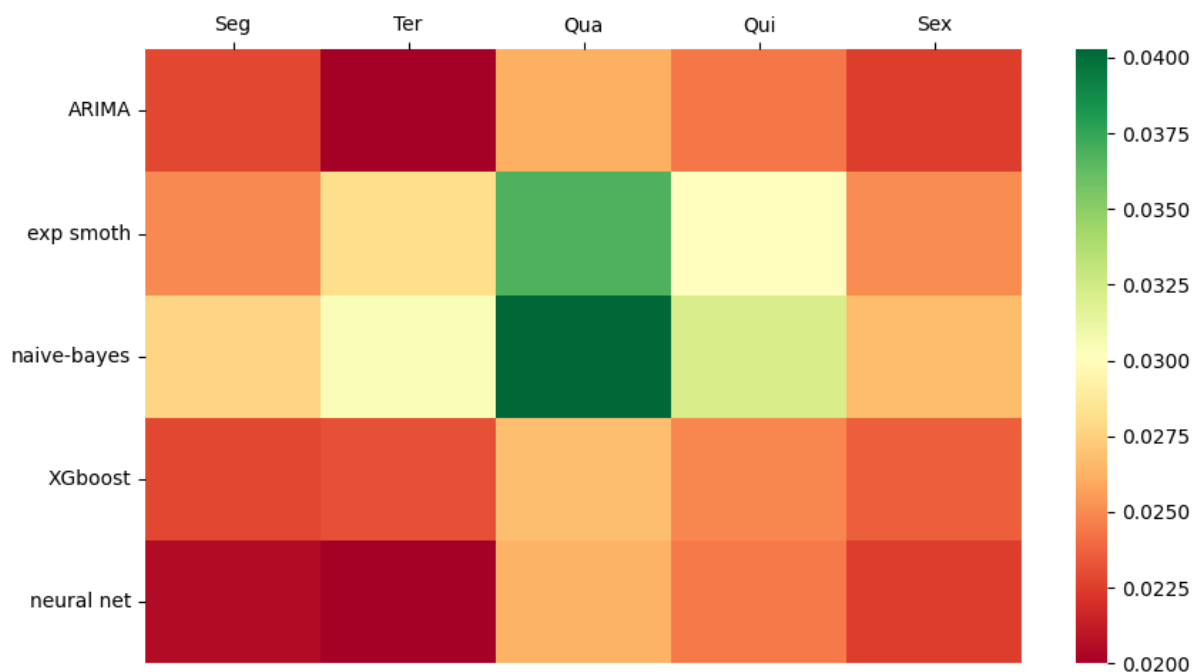
**Tabela 5** - Erros médios quadráticos para cada dia da semana para cada modelo.

<b>Método</b>	<b>RMSE Seg</b>	<b>RMSE Ter</b>	<b>RMSE Qua</b>	<b>RMSE Qui</b>	<b>RMSE Sex</b>
ARIMA	0,022842	0,020037	0,026114	0,024416	0,022512
Exp Smoth	0,026285	0,029245	0,038475	0,030972	0,025816
Naive-Bayes	0,027689	0,030442	0,040272	0,032193	0,026707
XGboost	0,022806	0,023108	0,02678	0,024887	0,023566
Neural Net	0,020521	0,019991	0,026323	0,024434	0,022468

Fonte: Elaborado pelo autor.

Esta tabela pode ser plotada em um mapa de calor para compreensão inicial através de uma visualização gráfica.

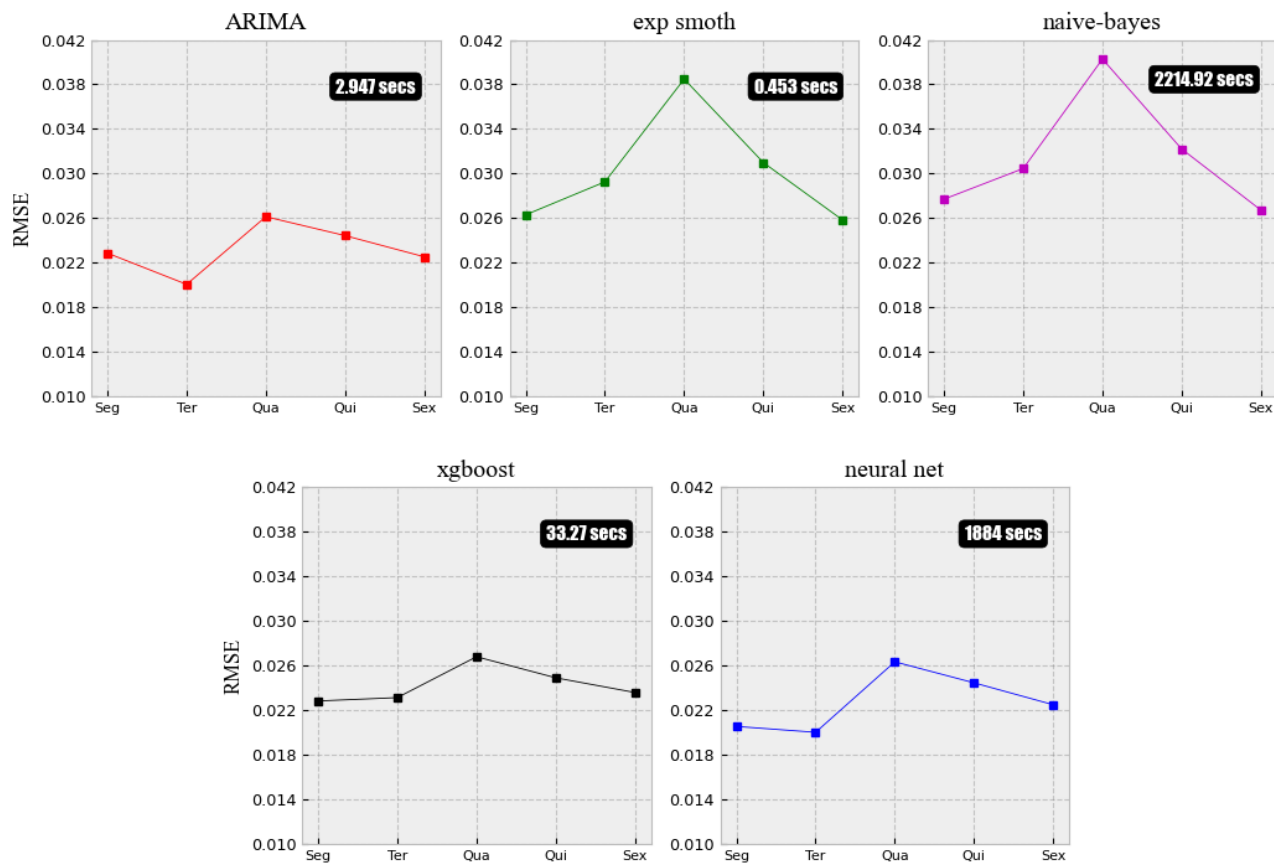
**Figura 5** - Mapa de calor do RMSE de cada dia para os modelos.



Fonte: Elaborado pelo autor.

Essa ilustração mostra que será preciso uma investigação maior para descobrir qual o melhor modelo. Chama a atenção a disparidade e elevados RMSEs para os modelos Naive-Bayes e Exponencial. Os demais modelos apresentam valores de RMSE similares. Na Figura 4 é possível observar os erros quadráticos em forma de curva para cada modelo para cada dia da semana.

**Figura 6** - Curva RMSE de cada modelo com tempos de ajuste.



Fonte: Elaborado pelo autor.

A partir da figura 4, é possível notar uma consistência do algoritmo XGboost, que obteve desempenho RMSE ligeiramente menor que os concorrentes da semana. O modelo ARIMA permaneceu em uma faixa de erro pequena e com desvio semelhante aos dois modelos mais sofisticados, o que a princípio pode indicar uma não superioridade por parte destes sobre métodos estatísticos clássicos. Ambos os casos eram esperados, pois tratam-se de algoritmos bastante robustos por levarem em consideração um bom número de cenários (ALIM, 2020) e serem capazes de encontrar sua melhor configuração.

Para previsão de séries temporais o melhor modelo seria aquele que apresenta o menor erro de previsão (FANG et al. 2022). Outro ponto importante é o tempo

computacional requerido por cada método. Este resultado também é apresentado na figura 4. Em geral, a Suavização Exponencial é o método mais rápido, seguido pelo modelo ARIMA e XGboost. O método Naïve Bayes e de Redes Neurais apresentam o maior tempo computacional.

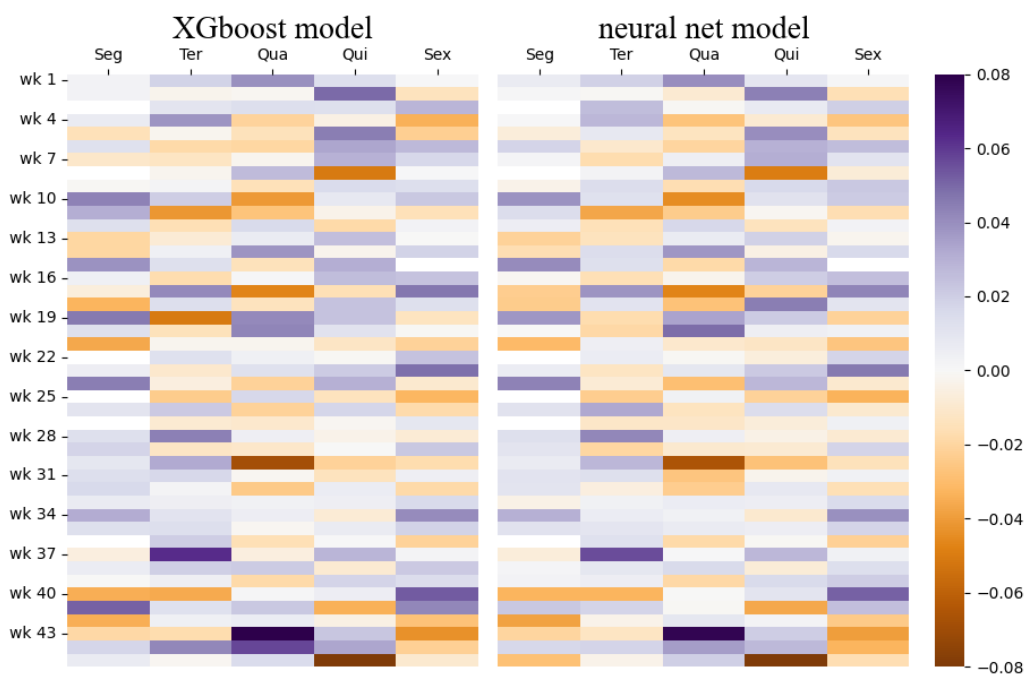
#### 4. Discussão

Considerando a base de dados estudada, qual seria o melhor método para o ajuste? Se for considerado o erro quadrático médio (RMSE) ao longo da semana, a escolha seria feita pelo modelo de Redes Neurais. Este modelo apresentou em valores absolutos as menores médias. Porém, caso o critério chave seja redução do tempo computacional do ajuste, o modelo de Suavização Exponencial seria a opção. Estes modelos são as “melhores” técnicas de previsão considerando a base de dados em estudo, as metodologias empregadas e as métricas de desempenho avaliadas(STOCKINGER e DUTTER, 1987).

Considerando os resultados numéricos alcançados é possível dizer que o modelo Naïve Bayes ficou aquém das expectativas, com um tempo de ajuste excessivamente grande e erros bem esparsos. O ajuste do método durou mais de 35 minutos para uma base de dados de treino com 505 observações até 723 observações considerando o conjunto de teste, base relativamente pequena. As Redes Neurais por sua vez mostraram bons valores de RMSE, mas apresentaram tempo computacional bastante elevado, ficando acima dos 31 minutos. É importante destacar que não foi utilizada nenhuma opção de processamento paralelo no ajuste dos modelos, o que é bastante usual na área(WELLENS et al. 2022).

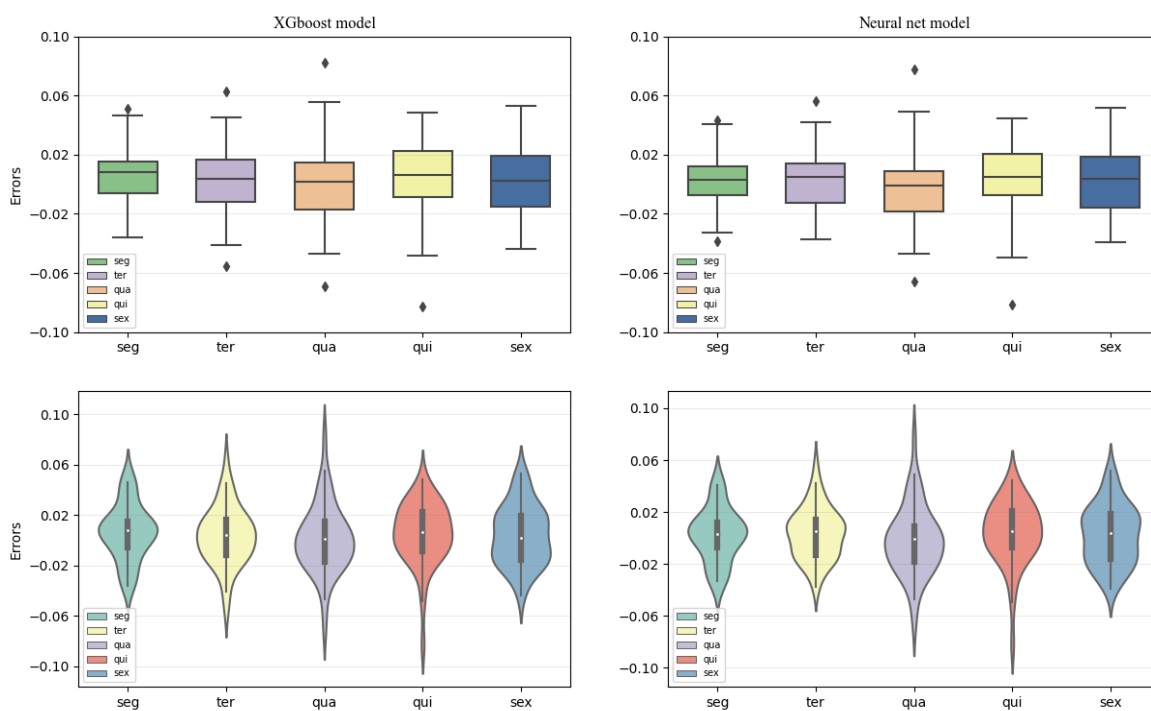
Na figura 5, a seguir, são exibidos os erros diários entre XGboost e Redes Neurais em escala de cores. Em geral, o método *Extreme Gradient Boosting* não demonstra uma aparente superioridade quando comparado com as Redes Neurais. A figura 6 apresenta *Box-plots/Violin plots* para a distribuição dos erros diários. Esses gráficos são comuns em análises da distribuição empírica(DRAGO e SCEPI, 2011). É possível observar que no *Box-plot* as diferenças são quase nulas entre os dois modelos, mas o *Violin plot* consegue expor algumas pequenas diferenças, sobretudo na assimetria dos dados.

**Figura 7** - *Heatmap* dos erros quadráticos médios dos modelos XGboost e Redes Neurais, valores mais claros indicam erros mais próximos de zero.



Fonte: Elaborado pelo autor.

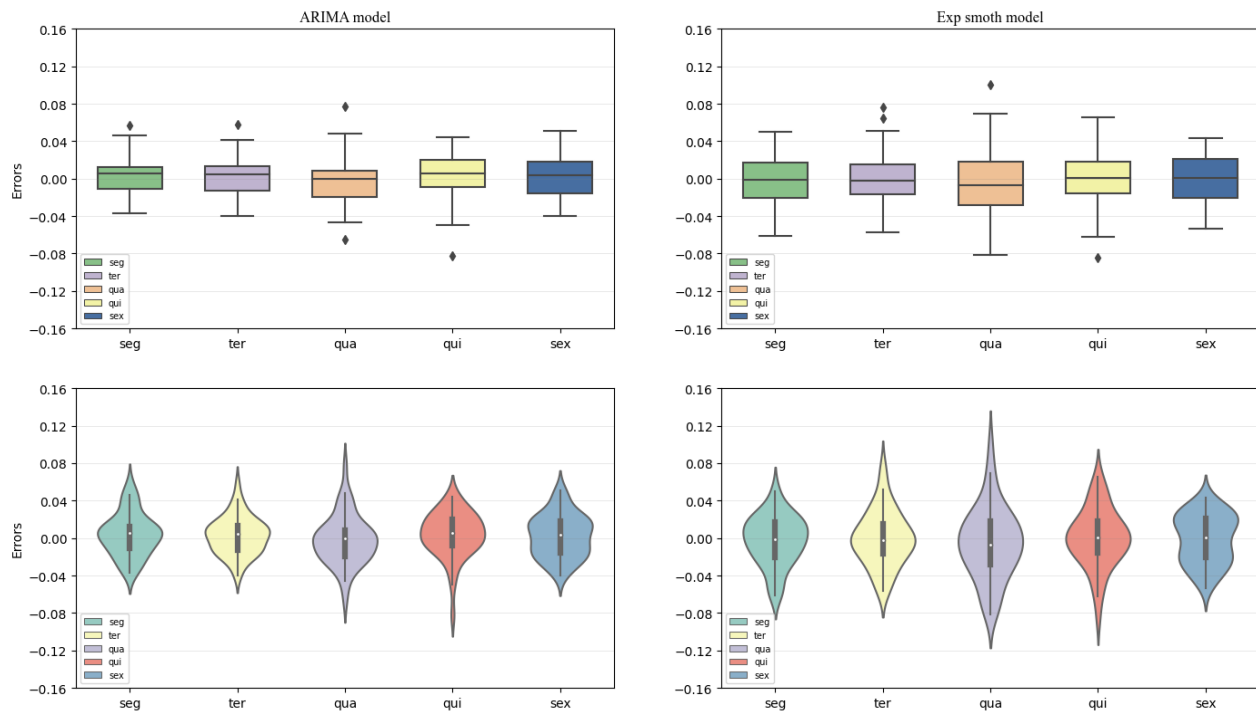
**Figura 8** - *Box e Violin plot* dos modelos XGboost e Redes Neurais, os modelos apresentam desempenho bastante parecidos.



Fonte: Elaborado pelo autor.

Em suma , o modelo ARIMA demonstra performance superior até considerando outros modelos estudados. Segundo SHARMA e YADAV (2020), trata-se de fato de um ótimo modelo para previsão de séries temporais

**Figura 9** - Box e Violin plot dos modelos ARIMA e Exponencial, o primeiro se mostra superior ao segundo.



Fonte: Elaborado pelo autor.

Em aspectos gerais os métodos ARIMA e XGboost são as melhores escolhas para este estudo, Ambos apresentam valores baixos para o RMSE. Porém é interessante investigar mais isso de maneira numérica, uma vez que os gráficos dos modelos não concluem definitivamente. Uma forma encontrada para uma abordagem mais quantitativa e que confronta modelos uns contra os outros é a alternativa de se assumir os ganhos em porcentagem para o RMSE para cada modelo em relação ao modelo *benchmarking* definido para o estudo, no caso o método Naïve-Bayes. Este ganho é simplesmente o valor do RMSE de cada dia de cada um dos outros quatro modelos dividido pelo RMSE de cada dia do modelo Naïve-Bayes, isto sendo subtraído de 1, ou seja, o resultado será descritivamente: para cada modelo há uma porcentagem de

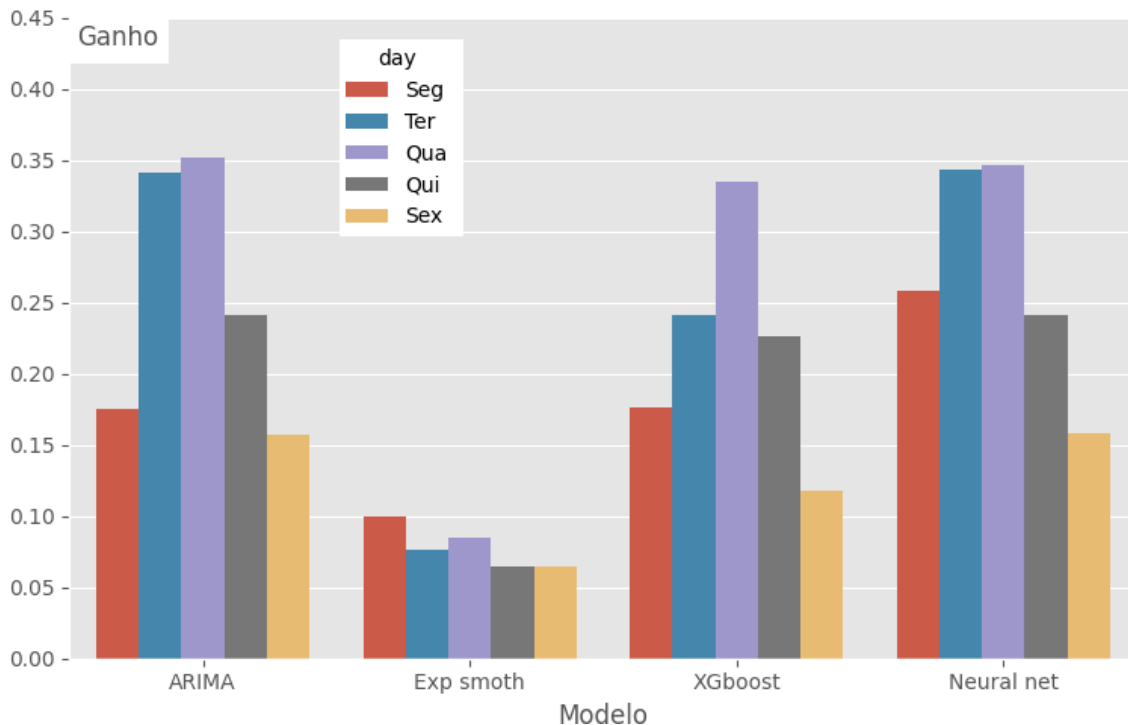
diminuição(ganho) que o RMSE teve para cada dia nas previsões em relação ao método base, Naïve-Bayes.

**Equação 2** - Função de ganho considerada para avaliar os métodos de previsão em relação ao método *benchmark*.

$$\alpha_{ganho} = 1 - \frac{x_i^j}{y_{Naïve-Bayes}^j} \quad \forall \quad i = 1, \dots, 4; j = 1, \dots, 5 \quad (2)$$

Esses valores podem ser observados na figura 8 abaixo, no qual são plotados os ganhos em relação ao modelo Naïve-Bayes. É possível observar que o modelo Exponencial evolui pouco e ao longo dos dias da semana e não se destaca, já os modelos ARIMA, XGboost e Redes Neurais demonstram forte superioridade em suas previsões.

**Figura 10** - Gráfico de barras sobre a porcentagem que cada modelo reduziu seus RMSE's em relação aos RMSE's do modelo Naïve-Bayes.



Fonte: Elaborado pelo autor.

A hipótese da superioridade se confirma com essa análise, os modelos ARIMA e XGboost mostram melhor desempenho em todos quesitos, o primeiro especialmente,



este gerou ganhos maiores que o segundo em todos os dias, exceto na segunda-feira. O fato do tempo de aplicação do método de Redes Neurais ser muito elevado é o único fator que pesa contra seu desempenho, pois seus ganhos foram expressivos.

## **5. Conclusão**

É interessante frisar que por mais que este processo seja considerado simples, existe um conjunto de opções disponíveis em software que oferecem perspectivas diferentes sobre um mesmo problema de séries temporais, bem como há campo para criação de métricas específicas para o problema. Aliado ao conhecimento teórico, as linguagens computacionais e seus pacotes têm muito a oferecer, passam por manutenção constante e possuem forte participação da comunidade. Em uma cadeia de suprimentos e mundo dos negócios cada vez mais acirrados, as atenções estão voltadas para cada uma das vantagens que diferentes abordagens (quantitativas e qualitativas) podem gerar.

Aplicações de sucesso de modelos estatísticos e computacionais são demandas atuais da indústria e em setores fundamentais da economia. Porém cada conjunto de dados requer um tratamento prévio, particularmente com o uso de metodologias de validação cruzada. Neste estudo, os modelos foram selecionados baseado na literatura das competições M. Preços e retornos de ações são informações bastante valiosas para investidores e administradores, e representam um grande desafio em termos de modelagem.

As técnicas utilizadas procuram expor uma gama de possibilidades para construção de métodos preditores. Em geral, o melhor método não é apenas aquele que possui um indicador ótimo em um tipo de avaliação e de baixo/médio desempenho em outros, e sim aquele que se mostra capaz de lidar com variações da amostra e com medidas de desempenho regulares considerando diversas características. Neste estudo, os métodos XGboost e ARIMA foram selecionados como a melhor escolha para previsão de retornos financeiros em curto prazo. O modelo ARIMA demonstrou robustez durante todas as etapas do estudo, e ao final superou outros modelos sofisticados na questão de ganhos. Pode-se dizer que é um modelo de previsão poderoso e que pode fazer

parte do conjunto de algoritmos de uma análise de dados mais profunda, mesmo que para efeitos de comparação. O XGboost abre possibilidades de *tunning* interessantes, pois apresenta vários parâmetros a serem ajustados (GUPTA et al., 2020 e SHOUWEN et al, 2020), se bem explorados, podem oferecer um ajuste robusto capaz de lidar com variações sem apresentar *overfitting*, que é quando o modelo desempenha muito bem em seu conjunto de dados, mas não apresenta precisão adequada quando prevê novos dados.

É possível também fazer este tipo de análise para o conjunto *OHLC*, considerando seus diferentes componentes. Como por exemplo prever o preço de fechamento ou a alta do dia, o que funciona bem se o foco for operações durante o pregão da bolsa, prática conhecida como *day-trade*.

## 6. Referências

AHMED, Shabbir et al. **Data-driven state awareness for fly-by-feel aerial vehicles via adaptive time series and gaussian process regression models**. In: International Conference on Dynamic Data Driven Application Systems. Springer, Cham, p. 57-65, 2020.

ALIM, Mirxat et al. **Comparison of ARIMA model and XGBoost model for prediction of human brucellosis in mainland China: a time-series study**. BMJ open, v. 10, n. 12, p. e039676, 2020.

ARMSTRONG, J. Scott; COLLOPY, Fred. **Integration of statistical methods and judgment for time series forecasting: Principles from empirical research**. 1998.

ASSIMAKOPOULOS, Vassilis; NIKOLOPOULOS, Konstantinos. **The theta model: a decomposition approach to forecasting**. International journal of forecasting, v. 16, n. 4, p.521-530, 2000.

BOJER, Casper Solheim; MELDGAARD, Jens Peder. **Kaggle forecasting competitions: An overlooked learning opportunity**. *International Journal of Forecasting*, v. 37, n. 2, p. 587-603, 2021.

BOX, George EP. Science and statistics. *Journal of the American Statistical Association*, v. 71, n. 356, p. 791-799, 1976.

CHAI, Tianfeng; DRAXLER, Roland R. **Root mean square error (RMSE) or mean absolute error (MAE)**. *Geoscientific Model Development Discussions*, v. 7, n. 1, p. 1525-1534, 2014.

CHENG, Changqing et al. **Time series forecasting for nonlinear and non-stationary processes: a review and comparative study**. *lie Transactions*, v. 47, n. 10, p. 1053-1071, 2015.

DO PRADO, Hércules A. et al. **On the effectiveness of candlestick chart analysis for the Brazilian stock market**. *Procedia Computer Science*, v. 22, p. 1136-1145, 2013.

DRAGO, Carlo; SCEPI, Germana. **Visualizing and exploring high frequency financial data: Beanplot time series**. In: **New perspectives in statistical modeling and data analysis**. Springer, Berlin, Heidelberg, p. 283-290, 2011.

EHLERS, Ricardo S. **Análise de séries temporais**. *Laboratório de Estatística e Geoinformação*. Universidade Federal do Paraná, v. 1, p. 1-118, 2007.

FANG, Zheng-gang et al. **Application of a data-driven XGBoost model for the prediction of COVID-19 in the USA: a time-series study**. *BMJ open*, v. 12, n. 7, p. e056685, 2022.

FILDES, Robert; MA, Shaohui; KOLASSA, Stephan. Retail forecasting: Research and practice. *International Journal of Forecasting*, v. 38, n. 4, p. 1283-1318, 2022.

FREITAS, Igor Costa. **Contribuições das competições M para o aprimoramento dos métodos de previsão de séries temporais**. Relatório técnico, 2022.

GUERRERO, Victor M. Time-series analysis supported by power transformations. **Journal of forecasting**, v. 12, n. 1, p. 37-48, 1993.

GUPTA, Aashish et al. Novel xgboost tuned machine learning model for software bug prediction. In: **2020 International Conference on Intelligent Engineering and Management (ICIEM)**. IEEE, 2020. p. 376-380.

HEWAMALAGE, Hansika; BERGMEIR, Christoph; BANDARA, Kasun. **Recurrent neural networks for time series forecasting: Current status and future directions**. **International Journal of Forecasting**, v. 37, n. 1, p. 388-427, 2021.

HIBON, Michele; EVGENIOU, Theodoros. **To combine or not to combine: selecting among forecasts and their combinations**. **International journal of forecasting**, 21.1: 15-24, 2005.

HUNG, Chih-Chieh; CHEN, Ying-Ju. **DPP: Deep predictor for price movement from candlestick charts**. **Plos one**, v. 16, n. 6, p. e0252404, 2021.

Jl, Shouwen et al. An application of a three-stage XGBoost-based model to sales forecasting of a cross-border e-commerce enterprise. **Mathematical Problems in Engineering**, v. 2019, 2019.

KALEKAR, Prajakta S. et al. **Time series forecasting using holt-winters exponential smoothing**. **Kanwal Rekhi school of information Technology**, v. 4329008, n. 13, p. 1-13, 2004.

KAZEM, Ahmad et al. Support vector regression with chaos-based firefly algorithm for stock market price forecasting. **Applied soft computing**, v. 13, n. 2, p. 947-958, 2013.

KIM, Jee-Ae et al. **Towards actualizing the value potential of Korea Health Insurance Review and Assessment (HIRA) data as a resource for health research: strengths, limitations, applications, and strategies for optimal use of HIRA data**. **Journal of Korean medical science**, v. 32, n. 5, p. 718-728, 2017.

KOLARIK, Thomas; RUDORFER, Gottfried. **Time series forecasting using neural networks. ACM Sigapl Apl Quote Quad**, v. 25, n. 1, p. 86-94, 1994.

LIM, Bryan; ZOHREN, Stefan. Time-series forecasting with deep learning: a survey. **Philosophical Transactions of the Royal Society A**, v. 379, n. 2194, p. 20200209, 2021.

LIU, Siyao et al. **Self-similar behaviors in the crude oil market. Energy**, v. 211, p. 118682, 2020.

MA, Shaohui; FILDES, Robert. Forecasting third-party mobile payments with implications for customer flow prediction. **International Journal of Forecasting**, v. 36, n. 3, p. 739-760, 2020.

MAKRIDAKIS, Spyros et al. **The accuracy of extrapolation (time series) methods: Results of a forecasting competition. Journal of Forecasting**, 1(2):111–153, 1982.

MAKRIDAKIS, Spyros et al. **The M2-competition: A real-time judgmentally based forecasting study. International Journal of forecasting**, 9.1: 5-22, 1993.

MAKRIDAKIS, Spyros; HIBON, Michele; and MOSER, Claus. **Accuracy of forecasting: An empirical investigation. Journal of the Royal Statistical Society. Series A (General)**, 142(2):97–145, 1979.

MAKRIDAKIS, Spyros; SPILIOTIS, Evangelos; ASSIMAKOPOULOS, Vassilios. **The M4 Competition: Results, findings, conclusion and way forward. International Journal of Forecasting**, v. 34, n. 4, p. 802-808, 2018.

MAKRIDAKIS, Spyros; SPILIOTIS, Evangelos; ASSIMAKOPOULOS, Vassilios. **M5 accuracy competition: Results, findings, and conclusions. International Journal of Forecasting**, 2022.

MAKRIDAKIS, Spyros; SPILIOTIS, Evangelos; ASSIMAKOPOULOS, Vassilios. **The M5 competition: Background, organization, and implementation.** *International Journal of Forecasting*, v. 38, n. 4, p. 1325-1336, 2022.

MAKRIDAKIS, Spyros; SPILIOTIS, Evangelos; ASSIMAKOPOULOS, Vassilios. **The M5 Uncertainty competition: Results, findings and conclusions.** *International Journal of Forecasting*, 2021.

MUSHTAQ, Rizwan. Augmented dickey fuller test. **Université Paris**. 2011.

MUTHUKUMAR, Pratyush; ZHONG, Jie. **A stochastic time series model for predicting financial trends using nlp.** arXiv preprint arXiv:2102.01290, 2021.

NEWBOLD, Paul; GRANGER, Clive WJ. **Experience with forecasting univariate time series and the combination of forecasts.** *Journal of the Royal Statistical Society: Series A (General)*, 137.2: 131-146, 1974.

PARMAR, Kulwinder Singh; BHARDWAJ, Rashmi. **Water quality management using statistical analysis and time-series prediction model.** *Applied Water Science*, v. 4, n. 4, p. 425-434, 2014.

PAVLYSHENKO, Bohdan M. Linear, machine learning and probabilistic approaches for time series analysis. In: **2016 IEEE First International Conference on Data Stream Mining & Processing (DSMP)**, p. 377-381, 2016.

PRAK, Dennis; ROGETZER, Patricia. Timing intermittent demand with time-varying order-up-to levels. **European Journal of Operational Research**, 2022.

SHARMA, Sudhi et al. **Analyzing the robustness of ARIMA and neural networks as a predictive model of crude oil prices.** *Theoretical and Applied Economics*, v. 27, n. 2, p. 289-300, 2020.

SPILIOTIS, Evangelos et al. **Product sales probabilistic forecasting: An empirical evaluation using the M5 competition data.** International Journal of Production Economics, v. 240, p.108237, 2021.

STOCKINGER, Norbert; DUTTER, Rudolf. **Robust time series analysis: A survey.** Kybernetika, v. 23, n. 7, p. (1), 3-88, 1987.

TICKNOR, Jonathan L. A Bayesian regularized artificial neural network for stock market forecasting. **Expert systems with applications**, v. 40, n. 14, p. 5501-5506, 2013.

VOS, Rutger Aldo et al. **Enriched biodiversity data as a resource and service.** Biodiversity data journal, n. 2, 2014.

VU, Tien Thanh et al. An experiment in integrating sentiment features for tech stock prediction in twitter. In: **Proceedings of the workshop on information extraction and entity analytics on social media data.** p. 23-38, 2012.

WELLENS, Arnoud P.; UDENIO, Maxi; BOUTE, Robert N. **Transfer learning for hierarchical forecasting: Reducing computational efforts of M5 winning methods.** International Journal of Forecasting, v. 38, n. 4, p. 1482-1491, 2022.

ZIEL, Florian. **Iteratively reweighted adaptive lasso for conditional heteroscedastic time series with applications to ar–arch type processes.** Computational Statistics & Data Analysis, v. 100, p. 773-793, 2016.

## 7. Anexo

### Anexo I - Pseudocódigo citado no item 2.5

```

bk = 106                                     # train_set : 2020 - 53 wks ; 2021 - 52 wks ; test_set : 2022 - 45 wks
unq_wks = unique(MSFT[total_weeks])         # wk 1, wk 2, ..., wk 150 ; absolute number of weeks
dt_day_errors = DataFrame(nrow = 0, ncol = 5) # final data frame
start_time = SystemTime()                   # measure time of fitting and debugging

while (bk < length(unq_wks)+1) {             # while its easier for debugging

  train_set = subset(MSFT, total_weeks < bk) # TRAIN SET 2020 & 2021
  test_set = subset(MSFT, total_weeks == bk)  # TEST SET THE NEXT WEEK 2022 YEAR UNTIL PRESENT WEEK
  bk = bk+1                                   # increment the week counter

  fit = fit_model_method(train_set[returns])  ### model FIT code      ###
  pred = prediction(fit, n_of_rows(test_set) ) ### model predict code ###
  test_set[model_pred] = numeric(pred[mean])  # take the model predict
  days = LIST()                               # days in the week been predicted
  day_error = LIST()                          # day-error in the week been predicted

  for j in 1:nrow(test_set) :
    day_error = LIST(day_error, test_set[j,8]-test_set[j,1]) # compute the errors of each day

  test_set[day_error] = day_error             # insert the pred in test set (debugging)
  total_days = LIST(2,3,4,5,6)               # each day in a normal week
  day_error_correct = LIST()                  # empty list that will receive: pred week existing days - final week pred

  for k in 1:5 :
    day_line = subset(test_set, week_day == total_days[k]) # subset by day, 1 row if exists 0 if not
    if(nrow(day_line) < 1)                                # if day doesnt exist compute NA
      day_error_correct[k] = NA
    else                                                    # if day exist compute teh errors
      day_error_correct[k] = day_line[1,9]                 # the error column in test_set by index

  dt_day_errors = combine_rows(dt_day_errors, day_error_correct) # increment the final data with the current week
}

End_time = SystemTime()                                     # debug the end time
Total = numeric(End_time - start_time)                     # total time of fitting SECS

column_names(dt_day_errors) = LIST('Seg', 'Ter', 'Qua', 'Qui', 'Sex')
day_rmse = squared_root(mean(dt_day_errors^2))             # apply the RMSE function (apply it by hand its easier for debugging)
model_name = "your_model_name"
# WRITE the results of dt_day_errors & day_rmse of each the model in .csv files

```



## Anexo II - Tabela com erros diários do modelo ARIMA.

Seg	Ter	Qua	Qui	Sex
0,0100316835304183	0,0175845678467582	0,0405336602890873	0,0094552339113988	0,0012490232496583
0,0011714769149516	-0,0008267264536863	-0,0088531814966180	0,0438605625762992	-0,0161604077653593
NA	0,0203380494768512	0,0011310283432508	0,0065739950093777	0,0201907286075555
0,0070306586201876	0,0257171005965589	-0,0262906019304923	-0,0093904605322170	-0,0265685255263798
-0,0163502476026524	0,0116444127252845	-0,0147821803182959	0,0407660868784056	-0,0142189231805728
0,0129730194131391	-0,0089278804298461	-0,0209370859490572	0,0300058001411452	0,0256577840661961
0,0101538389235633	-0,0200601818042472	0,0055912324716937	0,0303395543794603	0,0111276187554862
NA	0,0056903790529523	0,0260069947431525	-0,0493696523976071	-0,0080441214852605
-0,0062178971092506	0,0150815796138802	-0,0166854001716206	0,0156853851108146	0,0218135396465139
0,0463131945182060	0,0099052583886450	-0,0437463622480479	0,0111468295562322	0,0207365902595676
0,0208692170521522	-0,0396458357812364	-0,0232660226988133	-0,0017942116929931	-0,0163243450773489
-0,0041295285174821	-0,0074005857470312	0,0105342920954373	-0,0106149016052407	0,0019707781420282
-0,0208726922326353	-0,0140718480850063	0,0063706222862732	0,0190216536790424	-0,0022241327482478
-0,0120019380504817	0,0093552432363132	0,0422402028428069	-0,0077779134745221	0,0173006547717049
0,0457439301398812	0,0109595832084192	-0,0179058983813310	0,0282419171518848	NA
0,0075092847906298	-0,0185483388665407	-0,0016654953534101	0,0203283798175820	0,0254333422378539
-0,0154998050794864	0,0361905859929541	-0,0462199840658751	-0,0216939430012459	0,0430417881092906
-0,0104388864047660	0,0064548325613437	-0,0266079248100508	0,0443330176003319	0,0107598629070762
0,0414262094580289	-0,0184955893467843	0,0346876269823666	0,0210022136799550	-0,0214155567794355
-0,0071505543599814	-0,0171917594869456	0,0481330407202634	0,0049504275981821	0,0032714081649694
-0,0300165690016181	0,0046250386249825	-0,0100785667433445	-0,0119139653151941	-0,0266024747908343
NA	-0,0020909907219333	0,0016245840163655	-0,0076121907831371	0,0179557400819472
0,0112849568118822	-0,0145598506978849	0,0092914651364878	0,0217143257511416	0,0457088100989079
0,0573016278120579	-0,0124505625143050	-0,0274494552424108	0,0275687711391057	-0,0098267853066692
NA	-0,0266258185056056	0,0042171713533703	-0,0219644565308153	-0,0331393888327348
0,0018499037539711	0,0356312736812111	-0,0145237757490535	0,0144672744397760	-0,0097337929922081
NA	-0,0144419107584843	-0,0109532521197710	-0,0074724618291034	0,0038260006668518
0,0139202198857608	0,0417030003544001	0,0048831597583286	-0,0043646846561654	-0,0093382928258482
0,0078882163337978	-0,0190240658951338	-0,0098112656299723	-0,0087670890874622	0,0178659835290841
0,0120021015562172	0,0262945739262240	-0,0654380713668505	-0,0276696436168132	-0,0146384628509509
0,0067608157599591	0,0137005948512438	-0,0270559362282512	-0,0029942092673393	0,0036798128912033
0,0112935434262771	-0,0062466222900222	-0,0231156366625721	0,0084830927008759	-0,0159263515441170
-0,0086313726677510	0,0049615544145623	0,0055447897629577	0,0051823791263045	0,0149667034187422
0,0346172254113611	0,0046726075545814	0,0037801459237103	-0,0100753058810328	0,0397206281169496
0,0226022638574974	0,0065241969962276	0,0075487308707717	0,0048776845953659	0,0177441441490723
NA	0,0166863471533624	-0,0194794714805768	-0,0003819986714463	-0,0221417366463538
-0,0132822841033859	0,0575570389743511	-0,0003876292751681	0,0282006064512531	0,0035389128288680
0,0015918773430380	0,0082814843854377	0,0144848814816080	-0,0085084930998635	0,0127015939881669
0,0053493036165461	0,0034834227362160	-0,0194733559835373	0,0147457232386289	0,0193852600946514
-0,0285256303182656	-0,0351748377952611	-0,0009243827784713	0,0095751265645552	0,0508783226834418
0,0345173004387088	0,0133164524160147	-0,0006160565946547	-0,0378398604261503	0,0243089143095878
-0,0330767209118619	-0,0056531261648150	0,0088688895528743	0,0012938985171400	-0,0252547051942050
-0,0276411537154280	-0,0121444118713822	0,0767357893107918	0,0198635526262044	-0,0402478857543878
0,0058792546177876	0,0195342556255616	0,0347544394211765	0,0267311332458076	-0,0333633231460649
-0,0374085358859636	-0,0024007331472678	0,0185646614487229	-0,0821494964214548	-0,0170262599945573

**Anexo III - Tabela com erros diários do modelo Suavização Exponencial.**

Seg	Ter	Qua	Qui	Sex
-0,003703291702357	0,008775609259783	0,030016419388783	-0,000469595579878	-0,008881247271364
-0,002419339218432	-0,003946286099192	-0,012132049562735	0,040635548777509	-0,019403627018062
NA	0,033332296215534	0,006746306414433	0,014696400072700	0,027461295692164
-0,017319698877245	0,010417223316163	-0,044664373268198	-0,026720218243134	-0,044252870751570
0,016546093690466	0,032508542594975	0,010147699875143	0,064322127707380	0,009801345515107
0,024081756711260	-0,004223899220488	-0,014063353353539	0,036144481938085	0,032045480418415
-0,023452860192960	-0,042130841801658	-0,020360186511696	0,005693634636473	-0,013957474941062
NA	-0,011401378393798	0,013762576581169	-0,063224309871535	-0,021363873005755
0,009260633500195	0,027090464965100	-0,003527048721814	0,028462927511434	0,034717241080550
0,019151719080351	-0,007653801388808	-0,064483308669398	-0,008538406141627	0,000703294492281
-0,003478282176685	-0,055181900224595	-0,041652721969348	-0,019258686704284	-0,034087173836454
0,020024202816976	-0,000582209711985	0,030826919805495	0,000404076668996	0,017178132444313
-0,022385353969649	-0,014428143850755	0,005645446272018	0,018414150305733	-0,002869173643513
-0,017483484723136	0,013438611638103	0,037059040470192	-0,005790462177930	0,015053234390306
0,027276057277294	-0,000937626667130	-0,031867612193396	0,014928870987559	NA
-0,022604472780951	-0,037178463085535	-0,023854059675131	-0,000757528613280	0,004005753188927
-0,047030500101067	0,014786711300793	-0,070726055404087	-0,045249703805873	0,019194914195702
-0,055538625779756	-0,021039152751601	-0,059596100981164	0,013059402082492	-0,021048811968846
0,023631766945017	-0,031909243386792	0,019896008548323	0,006644029976425	-0,035910076128813
0,013735409820751	-0,005153496423937	0,062938261252848	0,018890332549174	0,017481924812737
-0,035514908642547	0,000468723738794	-0,014653080722827	-0,016358160922371	-0,031087279950812
NA	0,030015727996850	0,022978746968325	0,017073001478883	0,041609014304656
-0,007912917735175	-0,026569720868162	-0,004946522404326	0,008167003763193	0,031947397546332
0,002360328987045	-0,049257481134901	-0,069787779278797	-0,013082346302462	-0,050992539785232
NA	-0,018600703203449	0,008394534264134	-0,016646011752117	-0,028159341044509
0,042082807204552	0,063335035278893	0,016848007342426	0,044765210508297	0,020878521667242
NA	-0,005348469402147	-0,005534021758985	-0,000977722638384	0,010005918731440
0,011005114715865	0,040217956765539	0,002981486837797	-0,006144978188348	-0,011153948620829
0,018739544265169	-0,011648794687558	-0,001439344812023	-0,000680926105126	0,026034078556308
-0,006448535923384	0,014449297639682	-0,079176918382276	-0,040865657318438	-0,027990091551307
0,028241907924992	0,029992096528453	-0,009318700712299	0,014340268779138	0,021126492488804
0,008195770457513	-0,008022361918144	-0,025259521076472	0,006441725053482	-0,017996266892907
0,008357729977130	0,016291569129443	0,018450505432053	0,017649360934427	0,027555845082292
0,017180569555893	-0,007458178914424	-0,009823344723560	-0,023270030793848	0,026412445642099
-0,019891432988173	-0,022038492696951	-0,024855363743601	-0,026467161997163	-0,013892784740115
NA	-0,003620062611738	-0,033705579415022	-0,016260103735325	-0,037570999762634
0,011426405559237	0,074685838807593	0,018794723930250	0,046826608370927	0,022315642144938
-0,005286160822250	0,002280515655814	0,008250888335062	-0,014680575094710	0,006513063113176
-0,007747582133728	-0,005343200265878	-0,029434570769118	0,005085964965461	0,009645406028266
-0,051375714836880	-0,051525561665996	-0,018998878673645	-0,008042187439472	0,033139754082734
-0,022079984412539	-0,026632687281978	-0,044891287229078	-0,080990938032234	-0,019134279215854
-0,054118670166282	-0,018956652080415	-0,006403334713394	-0,013477477833962	-0,040153500067317
-0,000563323399080	0,006832813454792	0,097780832578460	0,040380750005282	-0,019595946686252
0,044674659163428	0,045877945408574	0,064186840194077	0,055397363594720	-0,004507040338209
-0,006284511962804	0,018596617604086	0,042035201011251	-0,059282971898412	0,005987780224124



**Anexo IV - Tabela com erros diários do modelo Naïve-Bayes.**

Seg	Ter	Qua	Qui	Sex
-0,003543395484047	0,007941076265950	0,028437805829068	-0,001659685729271	-0,009115554377236
-0,000134102405383	-0,003611447917594	-0,008711102365478	0,043150256025355	-0,016736002662537
NA	0,040641425788769	0,014520895100272	0,024648948331048	0,036280960140323
-0,018421928598053	0,008313966669607	-0,046546903999936	-0,029792633098445	-0,047945684632581
0,019828773013401	0,035644261468166	0,011310775411329	0,066389546728386	0,012118494683711
0,031611359166255	0,003005670783124	-0,008204434332470	0,042546509705476	0,038851358313837
-0,024409654003864	-0,042598821070243	-0,021046203126370	0,004322716631714	-0,016298740352337
NA	-0,009701657152550	0,014226841554460	-0,062328418915338	-0,017183386486992
0,004559389502024	0,021686324984771	-0,007471828225042	0,021544658097486	0,028249785765009
0,016512293682923	-0,010529785977206	-0,067580861659163	-0,010866096435770	-0,000812758300737
-0,007276615388864	-0,057316611011776	-0,044361107678943	-0,020256678373783	-0,037090482107543
0,021813346436592	0,002212963466118	0,031535560771604	0,001935980613011	0,019040932820226
-0,023898100523025	-0,015360091178110	0,001679672574175	0,016236511100155	-0,003675695499333
-0,013516875081827	0,015680684394614	0,041387068163449	-0,003262051631279	0,016304458974139
0,025361709429851	-0,005600939966079	-0,033581913161480	0,012134059811289	NA
-0,029330392186486	-0,043841647154181	-0,028222385041724	-0,008911049974393	-0,000585204457456
-0,048634320524235	0,014472062986842	-0,071856296435323	-0,045525283444618	0,016648833767230
-0,065045063588814	-0,032811986753533	-0,069939459040796	0,001995278882879	-0,031577860450883
0,029120732126865	-0,028279177565155	0,024354269000731	0,010321991206371	-0,033153283276381
0,019270165600069	0,001896059467317	0,071429390535998	0,027581998485379	0,025306622402481
-0,033765695630554	0,002754683191658	-0,013310335939711	-0,016878230236353	-0,031903784565399
NA	0,030176613536389	0,026367489068978	0,019325638332753	0,044632077628985
-0,011801928741591	-0,029854232175881	-0,009079136059019	0,002553643603753	0,027316383826356
-0,001935888466558	-0,054424186461258	-0,074506079874163	-0,020573444256864	-0,052757841095495
NA	-0,012939780197038	0,012653857551624	-0,014180125412288	-0,020294721562018
0,044016578087734	0,066945249090517	0,018365846773396	0,049432065109317	0,024409700582800
NA	-0,003843432596963	-0,001859963625425	0,003727834487391	0,012446658432374
0,008939760536355	0,037524282131125	0,001128937299801	-0,007974042594677	-0,013989541139305
0,019358615946700	-0,010437333875755	-0,000100605147207	0,001913874517770	0,026557795079374
-0,010862232178731	0,009178156230146	-0,084725164989407	-0,046924614981518	-0,031974947304299
0,027757221120983	0,028108343192018	-0,012271133272114	0,013126285293193	0,017576401915532
0,006742215228051	-0,009566432172346	-0,026818445457608	0,004109697523365	-0,019277123569049
0,011110532319008	0,020413181924457	0,022328367769139	0,022237196811140	0,030026309560455
0,017973375557926	-0,007400086175678	-0,010820222098831	-0,026127687642021	0,025216975631505
-0,027259234712256	-0,029371970843986	-0,031652368751369	-0,034583054606407	-0,022628941418611
NA	-0,003635972808416	-0,034829751374246	-0,019262766331750	-0,040001125968571
0,013481142127173	0,076923341727023	0,021691984502640	0,048577115398377	0,026231094146853
-0,000649162161512	0,004555202760252	0,011031591710156	-0,011597267824555	0,009908652907311
-0,010070088624788	-0,006438508476040	-0,034026731433012	0,002744652938460	0,006720498329284
-0,051016279260861	-0,052373075778048	-0,020393064386978	-0,007325778619981	0,032194983293847
-0,030612268845726	-0,033486121687719	-0,052940751089372	-0,088808689851541	-0,025828063889322
-0,063395028877279	-0,029652907909049	-0,017428906887986	-0,022903518687781	-0,048505014128025
0,002656723017734	0,010756552630000	0,102170867961106	0,044341287893787	-0,013162277331564
0,054748978857694	0,057480331430080	0,076018571651127	0,065778535183000	0,006553529814194
0,004362707502095	0,030202064321160	0,050278019071100	-0,050063022648444	0,017275100563139

## Anexo V - Tabela com erros diários do modelo XGboost.

Seg	Ter	Qua	Qui	Sex
0,0063220051970409600	0,0192209057483934000	0,0404828410673532000	0,0152221594838622000	0,00136931234444469300
0,0032246101600208500	-0,0022600846681641900	0,0005015474500447150	0,0485718113032977000	-0,0154936005048514000
NA	0,0116989664944316000	0,0144423567138297000	0,0098763695045978000	0,0252587150893895000
0,0085021335617516700	0,0371895319098453000	-0,0218271400395231000	-0,0102550341160876000	-0,0278770352676103000
-0,0082230610532947300	0,0037897481727472400	-0,0125859332575120000	0,0441276144907621000	-0,0155083927395363000
0,0124145079709276000	-0,0170967047539590000	-0,0187504586064180000	0,0337015020383785000	0,0270521134621640000
-0,0100578012667184000	-0,0128017770300788000	0,0013638594672101800	0,0305682950595452000	0,0171935482904311000
NA	0,0015215657424944000	0,0269966544020517000	-0,0485718696790287000	-0,0015779019897352900
0,0003140089951458870	0,0031770694534668500	-0,0170874293126093000	0,0156192900965110000	0,0129769019292508000
0,0440102353772033000	0,0193829345207325000	-0,0399157198497266000	0,0096735376594145900	0,0233836642852681000
0,0316309215565238000	-0,0410117695510932000	-0,0204985488264342000	-0,0034024244491186500	-0,0150119091891454000
0,0137767875053497000	-0,0155358295003769000	0,0158479179262769000	-0,0177685006916406000	0,0049532213263949700
-0,0207050903110326000	-0,0118233092405282000	0,0035165079338602700	0,0221819125137206000	-0,0009653685854327240
-0,0198008292264362000	0,0049019189630771200	0,0388310984674316000	-0,0014738751431404600	0,0189557926613252000
0,0374780413986078000	0,0119307742650094000	-0,0168383335733995000	0,0303250201542686000	NA
0,0024356554454200000	-0,0187191056064895000	0,0034028309155613700	0,0315765009862026000	0,0187058897269805000
-0,0063535188298593500	0,0416267220191983000	-0,0470199862438712000	-0,0161480479268521000	0,0463583723129487000
-0,0309577754952268000	0,0123467338448972000	-0,0116558993708546000	0,0299306928481768000	0,0088489854368561900
0,0461998370674926000	-0,0552337527453215000	0,0402459523514505000	0,0215622623386458000	-0,0097370090569024500
0,0129731367962695000	-0,0145713501404108000	0,0431042751384240000	0,0117540896866699000	-0,0009640539923995160
-0,0359522562569476000	-0,0032654159171169300	-0,0028281184358109900	-0,0123596188810471000	-0,0192991328451706000
NA	0,0134437795709053000	0,0051738489911854500	0,0004088213844062330	0,0228614120597168000
0,0058797185963094200	-0,0080562197191771400	0,0146104609625968000	0,0244818187874212000	0,0496009009944171000
0,0443778326284640000	-0,0064256527829640300	-0,0217991575280776000	0,0300642306626384000	-0,0105144597399269000
NA	-0,0241586630879940000	0,0113603412623069000	-0,0120675917763204000	-0,0371925127226103000
0,0121459774886564000	0,0208620702936693000	-0,0187924623110340000	0,0170215793317186000	-0,0224852338028281000
NA	-0,0079539480061248100	-0,0109700857113826000	0,0002793455145022610	-0,0027386750272812000
0,0135549201481433000	0,0454171387159100000	0,0043154476221163300	-0,0038314054046728600	-0,0088403758371542500
0,0176343260663903000	-0,0125401309317856000	-0,0123846739097506000	-0,0017128734316684800	0,0225946349003681000
0,0082544105563336900	0,0328607699130262000	-0,0693004016682304000	-0,0221072441029844000	-0,0185593603038421000
0,0134487106150070000	0,0163154506773038000	-0,0014151212410520300	-0,0140779146988437000	0,0047826500770036300
0,0163907190548905000	-0,0010285414422321900	-0,0239981653210550000	0,0062396725762321500	-0,0177779400341644000
0,0033152695450502400	0,0058308062060559200	0,0044509256906302100	0,0072029599364633600	0,0189586213778082000
0,0318906655173299000	0,0110470202169909000	0,0041418452165943300	-0,0093048408536935400	0,0403776355822539000
0,0135687824561035000	0,0130472627151154000	-0,0087562320657697000	0,0035803442831463500	0,0190373149097706000
NA	0,0208554679996006000	-0,0144177786429187000	0,0023220967862653800	-0,0193959816197380000
-0,0057588321089774300	0,0628214444485424000	-0,0057054136742831200	0,0287476648289621000	0,0027124484634911700
0,0044477221033105700	0,0199356572674849000	0,0161635529693250000	-0,0089559468242521600	0,0180200335793774000
0,0029191326268870300	0,0047110617520930200	-0,0168462078488800000	0,0176743278856996000	0,0207786016400549000
-0,0337500631191598000	-0,0369121286937015000	0,0014709704868305400	0,0048540790233430100	0,0529177290576323000
0,0509432601114079000	0,0123751415399541000	0,0225780188464406000	-0,0354981112858845000	0,0426924581670988000
-0,0311145193505393000	-0,0060783196941875900	0,0068987983387186500	-0,0036099117956045500	-0,0165691120045253000
-0,0197066612602166000	-0,0179520332538190000	0,0822743876022937000	0,0205650483515993000	-0,0435914081059094000
0,0165201994949224000	0,0397776235767442000	0,0554694479108688000	0,0331369135904039000	-0,0211559649596786000
0,0095684358614431800	-0,0040075890135238300	0,0181397607839769000	-0,0826879730594874000	-0,0117726319728924000



## Anexo VI - Tabela com erros diários do modelo Redes Neurais.

Seg	Ter	Qua	Qui	Sex
0,006370746355752	0,018849647317892	0,040090457446892	0,009604442478231	0,001192790786745
0,000820700676145	-0,000706246204615	-0,008892009668158	0,043875588672086	-0,016163587123485
NA	0,025855383429772	-0,000730606371330	0,007219487286938	0,019984382906401
0,000266651013354	0,028003573206762	-0,027078023377599	-0,009133868352535	-0,026666520860971
-0,007342381001189	0,008620067903320	-0,013740774816512	0,040433653015726	-0,014087129176547
0,017765402671465	-0,010540253260284	-0,020379707393335	0,029828127898290	0,025729126378619
0,001511098366750	-0,017166883241949	0,004603772048014	0,030657593196183	0,011006483618647
NA	0,002047909453011	0,027211864427978	-0,049775022024726	-0,007914585158947
-0,003564106191039	0,014265725273867	-0,016351788413048	0,015638187820201	0,021892501389316
0,039089950272782	0,012284429803623	-0,044545077476967	0,011399825050804	0,020641525684712
0,014203801204746	-0,037499816843163	-0,023970638587916	-0,001576603322853	-0,016405090455022
0,005587118403675	-0,015019294125285	0,016389835392194	-0,014033007744305	0,002741048031012
-0,021748908498104	-0,013791698379211	0,006281891743562	0,019050595777277	-0,002232728171969
-0,016547210929858	0,014374885431381	0,037995314263470	-0,004854188384652	0,015989508183584
0,040739001519630	0,012525317575206	-0,018404667951060	0,028391815229895	NA
-0,001268162090254	-0,015842152394839	-0,002517748984434	0,020578782077417	0,025342063879624
-0,023261019173971	0,038556192227889	-0,046956574476991	-0,021480222878777	0,042964395122798
-0,023873543464961	0,010625929563194	-0,027931018666369	0,044724484397287	0,010616270345949
0,038091228331912	-0,017449781999897	0,034355469935219	0,021103491363320	-0,021450614741917
-0,000401205154409	-0,019290111399096	0,048801646277689	0,004753717574015	0,003345309837577
-0,031039805591979	0,004943826789362	-0,010177977672258	-0,011883057871802	-0,026612176900243
NA	0,006129441797799	-0,000907539230726	-0,006813284720168	0,017722728105605
0,005792422656876	-0,012864380476111	0,008758817987726	0,021872344155245	0,045652737938383
0,043389350428194	-0,008228459693753	-0,028758757837649	0,027946675138686	-0,009963518344083
NA	-0,023653964847874	0,003341272619709	-0,021699273396542	-0,033212602688934
0,011554602359649	0,032806830433991	-0,013680197502477	0,014237005663394	-0,009649683177661
NA	-0,011696897318351	-0,011762553275321	-0,007224881185629	0,003758800771523
0,012812763427864	0,042025605477538	0,004789135549796	-0,004337329476349	-0,009346299908830
0,010593493175776	-0,019794845776950	-0,009585395901415	-0,008826977194518	0,017888027466916
0,006863708823178	0,027761542386243	-0,065864673635715	-0,027553412571876	-0,014677846804745
0,010824847296517	0,012575035899978	-0,026735761340774	-0,003076791849337	0,003709431860329
0,010259602015141	-0,005958530360517	-0,023195689518844	0,008505556611109	-0,015932435335279
-0,004199342438888	0,003734496713425	0,005893433016036	0,005092288518409	0,014998772666275
0,030461481864847	0,005822733394530	0,003457567585394	-0,009989118484894	0,039693357951053
0,011670311278492	0,009523251569714	0,006706380523064	0,005094582269502	0,017668959526550
NA	0,011901010515757	-0,018184506287527	-0,000739030607831	-0,022049926635140
-0,007311066279795	0,055948366968561	0,005276385421222	0,028089136531895	0,003578170305906
0,001751055392960	0,009317731871024	0,015288104550272	-0,007643358879500	0,013550279328386
0,002780633010220	0,005185014878070	-0,018906355625170	0,015614180109410	0,020173621172215
-0,032893295547241	-0,033043142376357	-0,000516459384005	0,010440231850167	0,051622173372373
0,022078525829066	0,017525822959628	-0,000732776987472	-0,036832427790628	0,025024231025751
-0,038508701040848	-0,003346682954982	0,009206634412039	0,002132491291471	-0,024543530941884
-0,020372967804041	-0,012976830950169	0,077971188173499	0,020571105600321	-0,039405591091213
0,016634826372895	0,017838112618040	0,036147007403544	0,027357530804187	-0,032546873128742
-0,028582176381389	-0,003701046814498	0,019737536592666	-0,081580636316996	-0,016309884194461