Universidade Federal do Ceará
Centro de Tecnologia
Departamento de Engenharia de Teleinformática
Curso de Graduação em Engenharia de Computação

COVERED CALL STRATEGIES TO ASSIST STOCK PORTFOLIO MANAGEMENT

Luiky Magno LUZ DE VASCONCELOS

Fortaleza - Ceará
2023

Universidade Federal do Ceará
Centro de Tecnologia
Departamento de Engenharia de Teleinformática
Curso de Graduação em Engenharia de Computação

COVERED CALL STRATEGIES TO ASSIST STOCK PORTFOLIO MANAGEMENT

Autor
Luiky Magno LUZ DE VASCONCELOS

Orientador
Prof. Dr. João Paulo Pordeus Gomes
Coorientador
Prof. Dr. Paulo Antonio Leal Rego

Trabalho de Conclusão de Curso apresentado ao Curso de Gradução em Engenharia de Computação do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de ***Bacharel em Engenharia de Computação****.*

Fortaleza - Ceará
2023

Luiky Magno LUZ DE VASCONCELOS

COVERED CALL STRATEGIES TO ASSIST STOCK PORTFOLIO MANAGEMENT

Este trabalho foi julgado adequado para a obtenção do título de Bacharel em Engenharia de Computação e aprovado em sua forma final pelo Departamento de Engenharia de Teleinformática da Universidade Federal do Ceará.

Luiky Magno LUZ DE VASCONCELOS

Banca examinadora:

Prof. Dr. João Paulo Pordeus Gomes (Orientador)
Universidade Federal do Ceará

Prof. Dr. Paulo Antonio Leal Rego (Coorientador)
Universidade Federal do Ceará

Prof. Me. José Moraes Feitosa
Universidade Federal do Norte do Tocantins

Prof. Dr. Ajalmar Rêgo da Rocha Neto
Instituto Federal do Ceará

Fortaleza, 05 de Julho de 2023

"An investment in knowledge pays the best interest."

Benjamin Franklin.

Resumo

O principal objetivo de um investidor no mercado de ações é realizar lucro seja comprando ações quando acredita que vão se valorizar, seja vendendo quando pensa que o seus preços vão cair. Neste contexto, um instrumento eficaz para melhorar a rentabilidade, mas menos conhecido pelos investidores em geral, é opções sobre ações, que podem ser de dois tipos: call, o direito de comprar uma ação, e put, o direito de a vender. Tendo isso em vista, um investidor pode executar a estratégia de call coberta, em que detém uma ação que espera que tenha certa valorização enquanto vende calls para aumentar os seus rendimentos. A dificuldade está em decidir por quais configurações vender a call, sendo que o parâmetro mais importante é o preço de exercício. O preço de exercício correto permite ao investidor ter mais ganhos, correndo menos riscos de ser obrigado a cumprir os termos do contrato. Dito isso, o objetivo deste trabalho é utilizar dados históricos do mercado de ações brasileiro para investigar diferentes estratégias que um investidor pode aplicar para decidir quais seriam os preços de exercício favoráveis para as ações em sua carteira. Para tal fim, indicadores comuns do mercado financeiro, bem como modelos de Machine Learning, foram empregados para simular diferentes estratégias durante o ano de 2022. Ao comparar as estratégias implementadas entre si e contra a abordagem padrão de Buy and Hold, concluímos que as estratégias com melhor desempenho foram as mais simples que utilizavam indicadores financeiros básicos. Por conseguinte, estas estratégias tiveram um desempenho superior às baseadas em Machine Learning. Além disso, apesar de apresentarem resultados aceitáveis para o período analisado, estas estratégias destinam-se apenas para auxílio geral na gestão de portfólio de ações.
Palavras-Chave: Call coberta, opções de ações, mercado de ações, retorno de ações, machine learning, dados financeiros, Brasil.

Abstract

THE main objective of an investor in the stock market is to make profit by buying stocks when he believes their price will rise and selling them when he thinks their price will fall. In this context, an effective instrument to improve returns, but less known to investors in general, is stock options which can be of two types, call, the right to buy a stock, and put, the right to sell it. With this in mind, an investor can perform the covered call strategy where he holds a stock he expects to have a moderate appreciation while he sells call options in order to enhance his returns. The difficulty lies in deciding which configurations to sell the call option for, where the most important parameter is the strike price. The right strike price allows the investor to have more earnings running less risk of being forced to comply with the terms of the call contract. That being said, the purpose of this work is to use historical Brazilian stock market data to investigate different strategies an investor can apply to decide what would be the favorable strike prices for the stocks in his portfolio. To do so, common trading indicators as well as Machine Learning models were employed to simulate different approaches during the year of 2022. By comparing the implemented strategies among themselves and against the standard Buy and Hold approach, we concluded that the most performative strategies were the simplest ones which used basic trading indicators. Therefore, these strategies outperformed the ones based on Machine Learning. Also, although exhibiting acceptable results for the period analyzed, these strategies are intended only for general assistance in portfolio management.
Keywords: Covered call, stock options, stock market, stock return, machine learning, financial data, Brazil.

Contents

List of Figures

List of Tables

Acronyms

AVG Average.

CDF Cumulative Distribution Function.

ma Moving Average.

PDF Probability Density Function.

STD Standard Deviation.

OvR One-vs-Rest.

vol Volatility.

Chapter 1

Introduction

To begin with, in this chapter, we will introduce some of the background on which the current work was built as well as the problem we decided to approach and the main motivations.

1.1 The Stock Market

When a company needs investments to expand its operations, usually the board of directors legally divides the organization into shares and puts them up for sale in the stock market. These shares, depending on type, give to those who purchase them rights over the company such as being able to take part in decisions and receiving dividends based on the company's profit. The term stock market, in turn, refers to the set of exchanges that allow investors to trade shares of companies securely.

The stock market has several important roles in the economy. Other than the evident one of allowing companies to earn investments, thus creating employment and promoting innovation, the stock market acts as a thermometer of a country's financial well being. If the prices rise, the economy is going well and if the prices fall, we have a strong indicator of a recession cycle. Moreover, since it is accessible to the public in general, it stimulates ordinary people to become investors, therefore improving financial education and the chances of achieving a wealthier future as the shares they hold gain value over time.

In Brazil, one of the most important companies that provides infrastructure to the national financial market is B3. In 2018, the number of investors registered was about 700 thousand and at the end of 2022 this number reached over 5 million. This expressive increase clearly shows that the Brazilian interest in the stock market is growing.

1.2 Improving Return

The main goal of an investor in the stock market is to make profit by choosing stocks that will gain value. A typical investor would buy a stock when he believes the company will grow over time, which would make its share price and the dividends paid rise. Contrarily, he would sell a stock when he thinks the company is not worth its current price and it will lose value as time goes by. This process of holding stocks when we believe they will gain value and releasing them when not is a long term strategy. However, there are investors so-called traders that seek to make profit in short time windows by buying when they predict the price will go up and selling when they predict the price will go down.

The investors that perform trading seek patterns in the price time-series of a stock to forecast future movements. The issue is that predicting in which direction the price will move is extremely difficult, seeing that the price seems to behave in a random walk, i.e. the past variations in the price have no influence in the future. The random walk assumption is even used in the Black-Scholes equation which is a well-known model of the fair price of derivatives [The Pricing of Options and Corporate Liabilities, 1973]. However, traders make use of techniques such as price pattern matching and, more sophisticatedly, machine learning to forecast price changes. Even so, it is estimated that up to 95% of traders end up losing money by day trading, [Is Day Trading Profitable, 2023].

Many researches were conducted to understand how we can increase returns in the stock market using machine learning to predict stock movements, but many of them conclude that the task is exceptionally arduous. For instance, after analyzing several artificial neural network architectures to forecast the behavior of some Brazilian stocks, a study showed that the simpler models were the most reliable and, all the same, they should only be used as support for further analysis [Comparing Artificial Neural Network Architectures for Brazilian Stock Market Prediction, 2020].

In addition to holding good stocks and attempting to predict price behavior, another tool of the stock market which is less known by investors in general, but can be used to help improve return, is stock options. Briefly, stock options, also called derivatives or futures, are a type of asset that depend on an underlying stock. They act as an insurance for great price changes and, depending on type, give to who owns them the right to buy or sell a stock for a predetermined price. With stock options, we can implement strategies that aim to either buy or sell them with the right configurations to increase profit over time.

1.3 Objectives

This work aims at investigating different strategies that combine stock and call options to provide assistance managing a given portfolio of Brazilian stocks in the long term. That being said, the main objectives are:

1. Process and analyze financial data from the Brazilian stock market.
2. Evaluate the performance of traditional covered call strategies.
3. Understand the utility of Machine Learning regarding financial data by applying it to different covered call strategies.
4. Develop new strategies to improve the return of a portfolio of stocks.
5. Evaluate and compare all the strategies analyzed to verify which ones performed the best and what is the improvement towards the standard buy and hold approach.
6. Develop tools that will allow people with interest in financial data to perform their own simulations.

In order to accomplish the objectives, we collected the data from the B3's official website and we opted to use the standard covered call strategy as the base of the work, since we can hold the stocks in the portfolio while we trade stock options. Therefore, the strategies analyzed differ only by the method implemented to pick call options. The methods used were based on traditional trading indicators as well as in machine learning models. Also, the analysis was conducted using Python.

The target public of this work is people with interest in Brazilian financial data and investors that seek to increase the performance of their stock portfolios in the long term.

1.4 Organization

This work is separated in six chapters.

1. **Introduction**. Contextualization of the subject, main objectives and the approach used.
2. **Background**. Background on the main concepts used throughout the work. This chapter covers the theory on the stock market and machine learning used to perform the analysis.
3. **Related Works**. Some works that approach similar problems and how they compare to what is proposed here.
4. **Methodology**. Deep dive on how the study was performed such as from where the data was collected and how the processing was made, which strategies were simulated and how they work, and so on.

5. **Results**. The results of the analysis. This chapter goes over the implications of the simulation outcomes and how the chosen strategies compare.

6. **Conclusion**. An overview of what the work proposed and what was achieved.

Chapter 2

Background

The goal of this chapter is to establish the knowledge used throughout the entire work. We will go over some important concepts on the stock market and machine learning.

2.1 The Stock Market

The stock market allows companies to earn investments by publicly selling shares which are commonly called stocks. Then, investors can trade these shares among them having in view earning dividends and the appreciation of the asset.

In order to decide when to buy or sell, various indicators are used to analyze stocks. Most of them are generated by only taking into account the price time-series. Moreover, by speculating on price movements, many investors also trade derivatives, mainly stock options. They do so to attempt to beat the market and improve their earnings.

2.1.1 Dividends

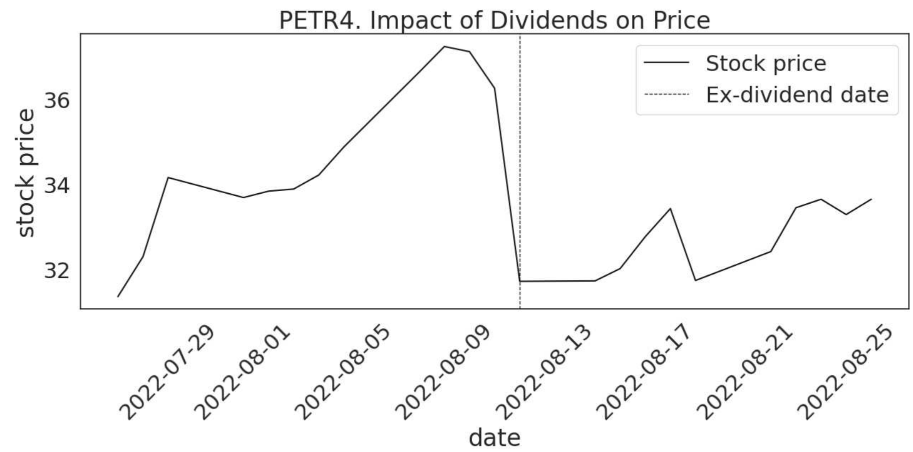
One of the main reasons investors decide to buy a particular stock is because of the dividends paid by the associated company. Dividends are part of a company's profit that is distributed among the shareholders. Some organizations across the world, however, do not pay dividends directly to investors, but reinvest them in the company. In that case, investors only earn with the appreciation of the stock which indirectly contains the company's profit. In Brazil, this is not the case, since all companies must pay dividends to their investors.

When working with dividends, an important date that comes into discussion and depends completely on the company is the ex-dividend date.

Definition 1. ***Ex-Dividend Date****. From this date on, investors that buy the stock will not receive the dividends announced by the company. After the announcement, the value of the stock increases by the amount of the dividends that will be paid, so that the exact value discounts from stock price at the ex-dividend date.*

To illustrate how this date works, let us see an example. In 2022, the company Petrobrás announced the payment of dividends for the stock PETR4 and it defined the ex-dividend date at 12/08/2022. That means investors who held the stock until 11/08/2022 received the dividends and, at 12/08/2022, the stock price started to be traded with dividends discounted.

Figure 2.1: The impact of dividends in the price of PETR4.



Source: created by the author.

In Brazil, companies not only can pay dividends to the shareholders, but JCP [Juros sobre Capital Próprio] which is a type of interest. The practical difference from ordinary dividends is that JCP is taxed by 15% before being discounted from the stock price and deposited to the investors.

2.1.2 Trading Indicators

Traders use a wide variety of indicators to evaluate stocks. Here, we will go over the ones used during this study to perform our analysis.

Historical Volatility

The historical volatility measures the magnitude of the stock price movements, where the movements are represented by the return of the stock. The return, in turn, is calculated based on the difference of the closing prices in a time period such as daily or annually. More precisely, the historical volatility is defined as follows.

Definition 2. ***Historical Volatility***. *The historical stock price volatility is the standard deviation of the natural log of the returns of a stock calculated based on the closing prices.*

$$DailyReturn_i = ln(\frac{StockPrice_i}{StockPrice_{i-1}}) \tag{2.1}$$

$$AnnualHistoricalVolatility = \sqrt{252} * \sqrt{\sum_{i=1}^{n} \frac{(DailyReturn_i - \overline{DailyReturn})^2}{n-1}} \tag{2.2}$$

In this work, the returns were calculated on a daily basis and then converted to an annual distribution to comply with the measurement units of the Black-Scholes model. The conversion from daily to annual values was made considering the year has 252 business days.

Moving Average

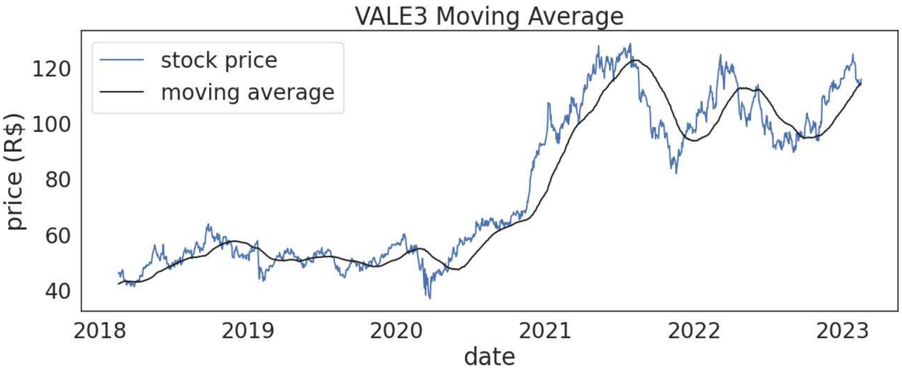
When working with financial time series, it is useful to measure the average of the n last steps. This information provides an indicator about the changes in the trend of a time series.

Definition 3. ***Moving Average***. *Considering a time span of n steps, the moving average of a stock price at any step is the average of the n last steps.*

$$MovingAverage_t = \frac{\sum_{i=1}^{n}(StockPrice_{t-i})}{n} \tag{2.3}$$

The following plot shows the behavior of the moving average in relation to the price. Notice how the moving average is a smoother curve than the price.

Figure 2.2: Comparison between the price of VALE3 (dividends included) and its last 100 days moving average.



Source: created by the author.

RSI

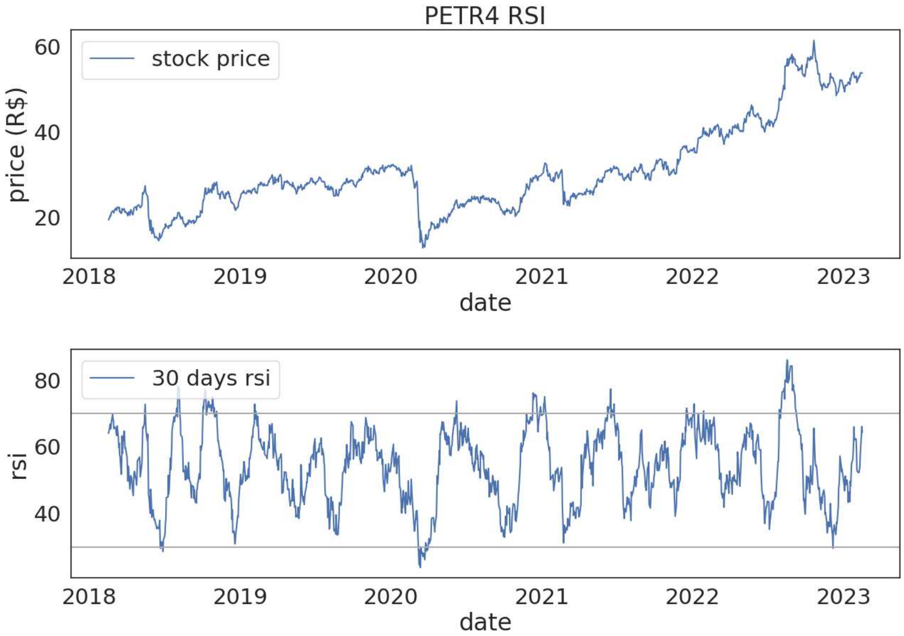
RSI stands for Relative Strength Index and it measures the trend of the price and is used to determine when a stock is either oversold or overbought. This indicator ranges from 0 to 100 where a high value such as 70 indicates the stock is overbought, whereas a low one such as 30 indicates it is oversold. Traders usually consider over 70 an opportunity to sell, since the price tends to fall, while a value of under 30 means exactly the contrary.

Definition 4. ***Relative Strength Index***. *RSI indicates the tendency of the price to return to average values. It is calculated based on the ratio of positive and negative average returns in a time span of n days in the past.*

$$RSI_t = 100 - \frac{100}{1 + \frac{U_t}{D_t}} \tag{2.4}$$

Where U_t is the average of the positive returns in the last n days and D_t is the average of the negative returns in the last n days.

Figure 2.3: Price of PETR4 (dividends included) alongside the historical value of the RSI indicator.

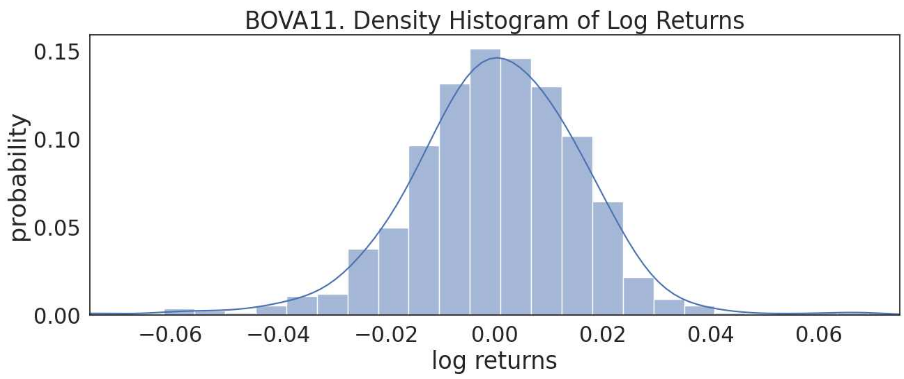


Source: created by the author.

2.1.3 The Gaussian Distribution of Log Returns

In many contexts across various fields of study, the Gaussian distribution (also known as Normal distribution) manifests itself and the stock market is no exception. If we take the daily returns of any stock and plot them in a histogram, we will see they are log normally distributed. That means if we take instead the daily log returns 2.1 and plot a density histogram, we will end up with a bell-shaped curve centralized at the value 0.

Figure 2.4: Density histogram of the daily log returns for BOVA11 considering the period of 2020 to 2022.



Source: created by the author.

This indicates we can assume $\mathcal{X} \sim \mathcal{N}(0, \sigma^2)$ where the random variable $\mathcal{X}$ represents the daily log returns of a stock. That is to say $\mathcal{X}$ follows a Gaussian distribution $\mathcal{N}(0, \sigma^2)$ with mean 0 and variance σ^2 where σ is the daily volatility of the stock.

The Gaussian distribution is endowed with the equalities described below.

Definition 5. ***Gaussian Distribution***. *A Gaussian probability distribution $\mathcal{N}(\mu, \sigma^2)$ has probability density function (PDF) defined as:*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \tag{2.5}$$

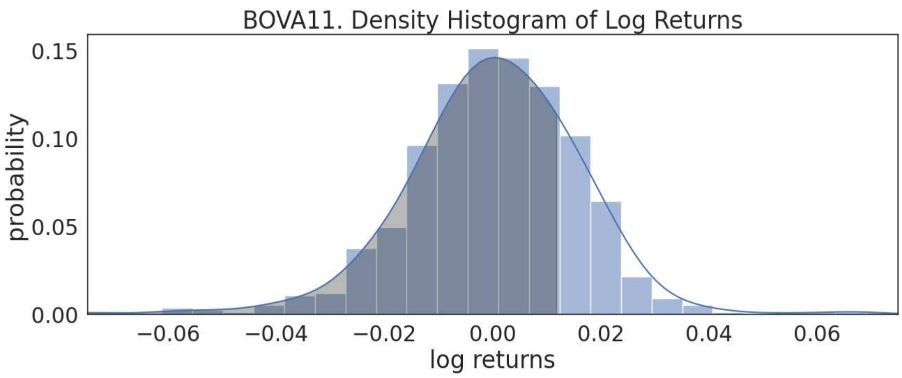
Given the density function above for a value x, the cumulative distribution function (CDF) $\Phi : \mathbb{R} \to [0, 1]$ it is its integral from $-\infty$ to x:

$$\Phi(x) = P(z \leq x) = \int_{-\infty}^{x} f(z)dz \tag{2.6}$$

Where $P(z \leq x)$ is the probability of z being less or equal than a given x.

As a reminder, for a given x, the CDF is the area under the PDF curve from $-\infty$ to x. The next plot illustrates this concept.

Figure 2.5: Graph 2.4 with area under curve highlighted to illustrate the calculation of the CDF.



Source: created by the author.

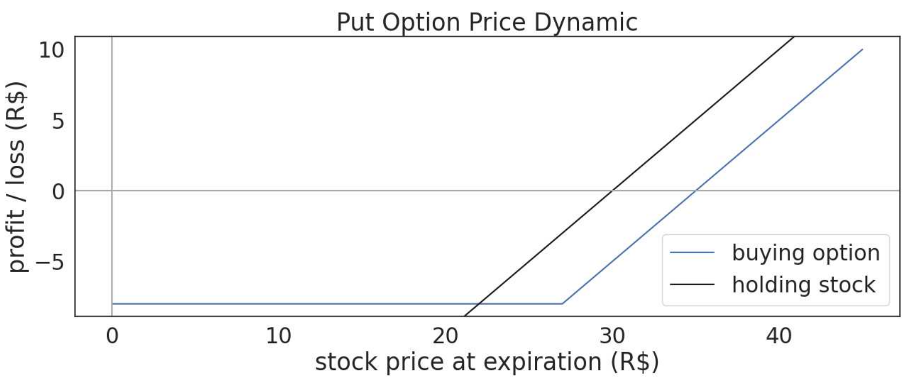
2.1.4 Stock Options

When you buy a car or a house, it is usual to buy insurance as well. You do so, because if something bad happens to your property before the insurance's expiration, you can go to the insurance company and enforce your right to be compensated for the loss. Yet, if you take care, the insurance will never be used and the money you spent with it will make the insurance company richer. What we see in this scenario is that the insurance company is selling you protection against the risk of an event happening while hoping that the event will never happen.

Although with some differences, the stock market has a resource similar to insurances, and it is called stock options. There are two types of option contracts which are calls and puts. Whereas calls give the buyer the right to buy a stock in the future, puts give the buyer the right to sell a stock. It is important to note that they give rights and not obligations to the buyer. Also, both of them are defined by an underlying stock, a premium, a strike price and an expiration date. Considering a put option, the buyer has the right of selling his stock to the seller if the stock price at the expiration date is lower than the strike price. Throughout this work, we only consider European style options. That means options can only be exercised at expiration.

Let us see an example to clarify things. Imagine you hold the stock ITUB4 currently priced at R$ 30. You think the stock will gain value in the long term, but you fear that, in the next few weeks, the price may fall due to politics. Since you do not want to release the stock, you decide to buy for R$ 5 a put option with a strike price at R$ 27 and expiration two weeks from now. In this case, if ITUB4 values R$ 20 at the expiration date, you have the right to sell your stock at R$ 27 to the person you bought the put from. By doing so, you would minimize your loss to R$ 8, since you spent R$ 5 for option premium and R$ 30 for the stock and ended up selling it for R$ 27. If you had not bought the option instead, you would lose R$ 10, since the stock price dropped by this amount. The graphic below illustrates this dynamic of price. As we can see, the loss is limited to R$ 8 when buying the put, but the profit is impacted as we have to invest more.

Figure 2.6: Considering a stock that was bought at R$ 30, the graphic shows the profit and loss dynamic holding the stock vs. buying a put option with strike at R$ 27.



Source: created by the author.

Call options work similarly to puts, but, instead of giving the right to sell, they give the right to buy a stock. They are exercised if the stock price goes over the strike price once the expiration date arrives.

Concerning the price of a stock option which is commonly called premium, it consists of two parts: Intrinsic and extrinsic values.

Definition 6. ***Intrinsic and Extrinsic values***. *On the one hand, the intrinsic value of an option is the difference between the price of the underlying asset and the strike price of the option. On the other hand, the extrinsic value is the difference between the premium and the intrinsic value and decays with time.*

$$IV_{call} = max(S - K, 0) \tag{2.7}$$

$$IV_{put} = max(K - S, 0) \tag{2.8}$$

$$EV_{call} = P_{call} - IV_{call} \tag{2.9}$$

$$EV_{put} = P_{put} - IV_{put} \tag{2.10}$$

Where IV, EV, K, S and P stands for intrinsic value, extrinsic value, strike price, stock price and premium respectively.

The intrinsic value reflects the real value of an option and doesn't change with time. It is worth noticing that the intrinsic value of a call is zero if the current stock price is less than the strike price. In this case, it is better to buy the stock directly from the market instead of using the right given by the call (the buy option), therefore making the call is useless. A similar logic applies to put options, since a put is worthless if the current stock price is higher than the strike price. Whereas the extrinsic value, which is the difference between the premium and the intrinsic value, is generated by the uncertainty the time remaining on the contract and the implied volatility of the stock causes. More on implied volatility in the next topic.

2.1.5 The Black-Scholes Model

Developed in 1973 by Fischer Black, Robert Merton, and Myron Scholes, the Black-Scholes equation is one of the most important concepts in financial theory.

Definition 7. ***Black-Scholes Model***. *The Black-Scholes model is a partial differential equation (PDE) that governs the evolution of the fair price of a European option.*

$$\frac{\partial p}{\partial t} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 p}{\partial^2 s} + rs\frac{\partial p}{\partial s} - rp = 0 \tag{2.11}$$

Where p, s, r, σ are the option price, the stock price, the risk-free interest rate and the stock volatility respectively.

As a partial equation, the Black-Scholes model is demanding to understand. Fortunately, a derivation of this equation exists and is easier to manipulate. For a European call option, we have the following.

$$d_1 = \frac{ln(\frac{S}{K} + T(r + \frac{\sigma^2}{2}))}{\sigma\sqrt{T}} \tag{2.12}$$

$$d_2 = d_1 - \sigma\sqrt{T} \tag{2.13}$$

$$P_c = SN(d_1) - N(d_2)Ke^{-rT} \tag{2.14}$$

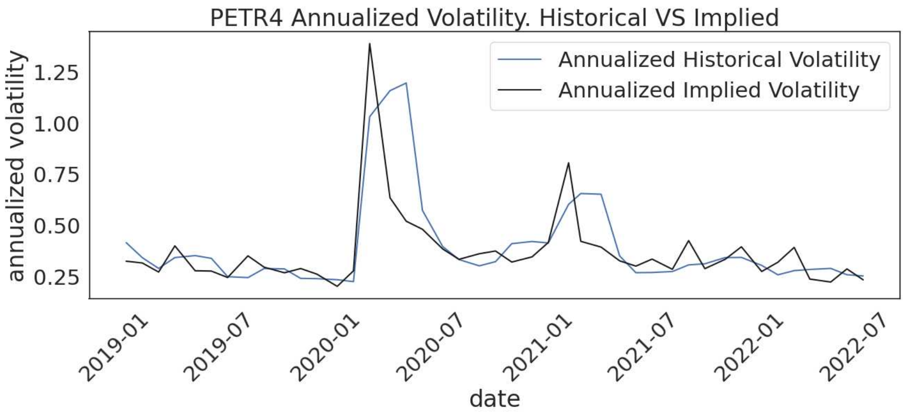
Where P_c, S, K, r, T, σ are the call option price, the current stock price, the strike price, the risk-free interest rate, the call's time to expire and the annualized stock volatility. In Brazil, the risk-free interest rate would be either the Selic or the CDI interest rate. Additionally, the function $N(x)$ is the cumulative distribution function for the standard normal distribution 2.6.

Given S, K, r, T and a call price that was observed in the market, therefore a fair price, since the market is efficient, we can solve the Black-Scholes formula for σ. This value σ is called the implied volatility.

Definition 8. ***Implied Volatility***. *The implied volatility of a stock option represents the forecast the market has about the volatility that the underlying stock is supposed to maintain until the expiration date.*

The implied volatility is used by some traders to make comparisons with the historical volatility in order to understand if the expectations of the market are consistent with the asset's history. Considering the graphic below, the annualized historical volatility was calculated day by day based on the last 60 days and the implied volatility by picking calls where the strike price was equal to the current price, i.e. calls with no intrinsic value.

Figure 2.7: Comparison between the historical and implied volatilities



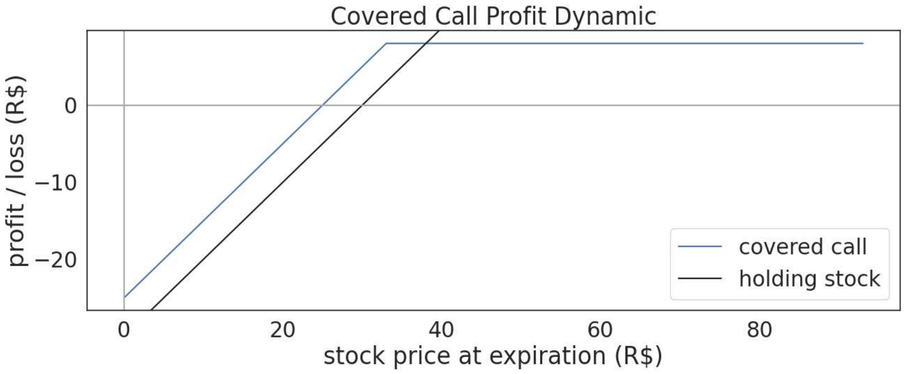
Source: created by the author.

2.1.6 Covered Call Strategy

There are several strategies that use stock options to leverage returns. One of the simplest and most popular ones is the covered call strategy. In this strategy, the investor sells, for each stock he buys, a corresponding call option with strike price usually above the current price of the stock. He does so hoping the stock price will rise, but not sufficiently to surpass the strike price. If he is right, at expiration date, the call option will not be exercised and he will have earned the premium. If he is wrong, he will be forced to sell his stock at the value of the strike price, thus losing the asset's appreciation. By using the covered call strategy, you try to improve return selling to someone that intends to buy a stock protection against the risk of a great rise in the price.

Let us see an example to make it clear. Imagine you bought ITUB4 currently priced at R$ 30 as well as a call with premium of R$ 5, strike price at R$ 33 and expires one month from now. When the call expires, if the stock price is over R$ 33 you have to sell your stock for R$ 33 no matter how above the price is. If the stock price is below R$ 33, you can keep the stock and the premium. The graphic below shows how profit behaves as a function of the stock price at expiration.

Figure 2.8: Considering a stock that was bought at R$ 30, the graphic shows the profit and loss dynamic holding the stock vs. selling a call option with strike at R$ 33.



Source: created by the author.

As we can see, selling the call option limits the profit to R$ 8, although it protects the investor from a fall in the price of up to R$ 5. Therefore, the success of the covered call strategy depends on choosing the right strike price. Considering the price of the stock, a strike margin too high implies low risk, but a low premium as well, whereas a low margin provides a high premium in exchange for a high risk.

2.2 Machine Learning

In a nutshell, Machine Learning is a broad field of Artificial Intelligence where algorithms and statistical models are used to make computer systems learn patterns in data without being given explicit instructions. Besides, by inferring patterns in the data, Machine Learning models are capable of making predictions about events they learned. Generally, we would use Machine Learning when we have a sufficient amount of data available and finding a model analytically is too complicated. In other words, we use Machine Learning when we cannot find the exact model, but we have data to look for patterns in.

By and large when working with Machine Learning, we separate the data we have in features of an event and their respective target values. Then, the models we create will attempt to understand how the behavior of the features, the input, influences the target, the output. Depending on the type of target in analysis, we can differentiate problems into two classes: regression and classification.

2.2.1 Regression vs Classification

Two of the main classes of problems when working with Machine Learning models are regression and classification. On the one hand, regression problems seek to model continuous target variables such as temperature and price. On the other hand, classification problems try to model discrete values, for instance, the type of object in a picture, e.g. a car, a boat. Also, if necessary, a regression problem can become a classification one by discretizing the continuous target variable into classes.

Knowing the type of problem we are approaching, regression or classification, is essential, since the models we have available are different for these two situations. In this work, to implement option-based strategies, we opted for a classification problem and the models we used to analyze the data were the Logistic Regression and the Random Forest Classifier.

2.2.2 Logistic Regression

Logistic Regression is one of the simplest classification models, because it is built on top of a linear model, i.e. a model that assumes the features and the target are linearly related. Considering x the vector of features with size p and y the one-dimension target, here below we have an example of a linear model.

$$y = \beta x = \beta_0 + \beta_1 x_1 + ... + \beta_p x_p \tag{2.15}$$

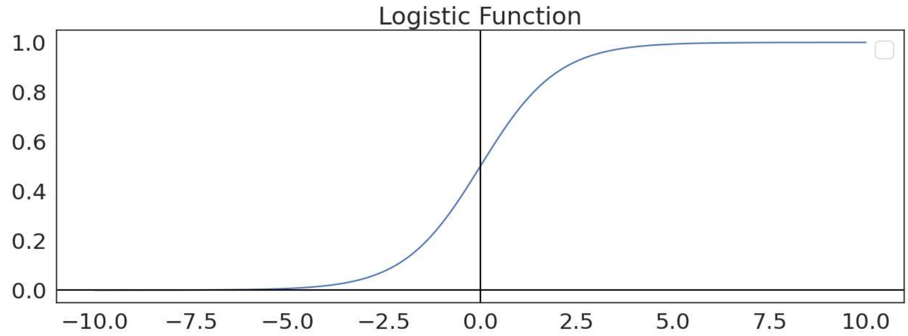
Additionally, the Logistic Regression model has as output the probability of a target class being right, hence a value between 0 and 1. Since the target y is a probability, its relation towards the features x is non-linear. In fact, features and target are related by the logistic function.

Definition 9. ***Logistic Function***. $g : \mathbb{R} \rightarrow (0, 1)$

$$g(z) = \frac{1}{1 + e^{-z}} \tag{2.16}$$

The Logistic Function has the following appearance.

Figure 2.9: The Logistic function.



Source: created by the author.

Given that h is the Logistic Regression hypothesis, we have the following for a binary classification.

Definition 10. ***Logistic Regression.*** $h : \mathbb{R}^{p+1} \to (0, 1)$

$$h(x) = g(\beta x) = \frac{1}{1 + e^{-\beta x}} \tag{2.17}$$

Where β is the set of parameters, also called weights, of the model with size p + 1.

Considering a binary distribution, in order to optimize β for a set of m data points, we have to maximize the logarithmic likelihood described here below. To do so, we can apply gradient descent.

$$l(\beta) = \sum_{i=1}^{m} y^i log(h(x^i)) + (1 - y^i) log(1 - h(x^i)) \tag{2.18}$$

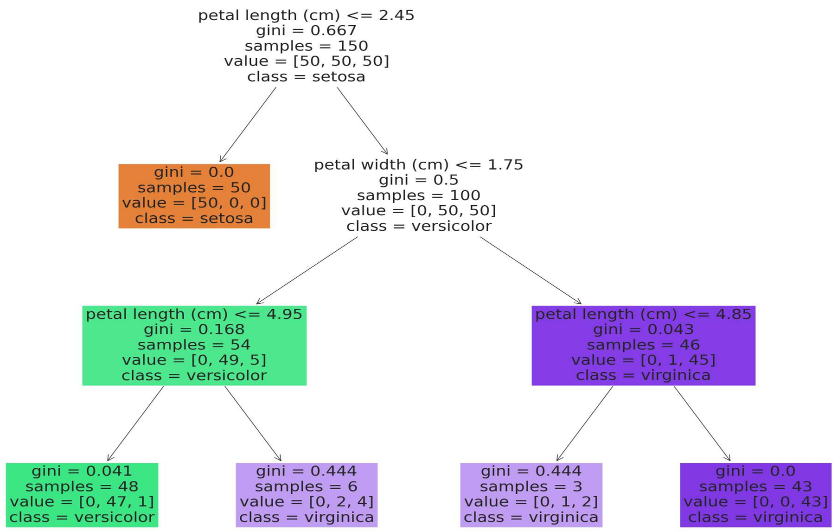
If the target y has k possible values instead of two, thus a multinomial distribution, we can use the Softmax Regression which is the generalization of the Logistic Regression for multiple classes. However, another efficient way to approach classification with multiple classes is by using the One-vs-Rest (OvR) strategy, where, for each class, we train a binary model considering the class in question the positive one the and all the remaining ones the negative class. Then, the prediction is given by the model which indicates the highest probability for its positive class.

2.2.3 Random Forest

Random Forest is one of the most used nonlinear Machine Learning models. It is an ensemble method which means it combines the outputs of other models to come up with its own result. In its case, Random Forests aggregate several Decision Trees as the underlying models to output a result. In turn, a Decision Tree is an elementary Machine Learning algorithm that starts with a simple question about the data and, from the answer, it asks another and another until it reaches the max depth of the tree. The idea is to segment the data with classification questions one after the other up till the moment the model can make a prediction on the input.

The plot below illustrates how a Decision Tree works. The iris sklearn data set was used to train the model.

Figure 2.10: Decision Tree visualization.



Source: created by the author.

A Random Forest model would, then, take the outputs of its Decision Trees and predict the most popular class. Generally, this class would be the most voted one.

In order to train this model, we could minimize the Gini Index. This metric is a measure of impurity, essentially the variance, of our predictions. A high Gini value implies a high level of misclassification.

Definition 11. ***Gini Index.***

$$Gini = 1 - \sum_{i=1}^{n} p_i^2 \tag{2.19}$$

Where p_i is the relative frequency of the i_{th} class.

Chapter 3

Related Works

In this chapter, we will go over some works that shared a similar subject to the one we discuss in this study in order to understand what have been done so far and how they compare to what we propose here.

3.1 Comparing Neural Network Architectures for the Brazilian Stock Market

The first work we will discuss is entitled Comparing Artificial Neural Network Architectures for Brazilian Stock Market Prediction and was conducted by Suellen Teixeira Zavadzki de Pauli, Mariana Kleina and Wagner Hugo Bonat in 2020. The main objective of this study was to compare the prediction performance of different Artificial Neural Networks to predict some of the most traded stocks, e.g. PETR4, in Brazil from 2019 to 2020. They concluded that the simplest architectures provided the most reasonable results to predict stock prices with satisfactory confidence intervals. Additionally, due to the complexity of the problem approached and the fact that models rely on past data, they suggest using the results only as a support tool in decision making.

In the current study, we also attempted to predict prices, but, instead of opting for a regression problem, we transformed the continuous prices in price ranges so that we could approach a classification problem which is generally simpler. Moreover, contrary to using Neural Networks, we used the Logistic Regression with the One-vs-Rest (OvR) approach and the Random Forest models to perform predictions which rely on less complex learning algorithms. Another difference is that the study made in 2020 only used the Machine Learning models to predict prices and, therefore, improve returns. Here, in addition to the models, we complemented their predictions by trading stock options, namely call options, as a secondary source of profit.

3.2 Covered Call as a Management Strategy for an Investment Club

The next study titled Covered Call as a Management Strategy for an Investment Club was made by Marcelo Notomi Kanazawa, a student from the Federal University of Santa Catarina, in 2011. His work aimed at developing a covered call strategy to improve the return of an investor. To do so, the work used the difference between two moving averages, one longer than the other, to decide for which strike price to trade calls. When the shorter moving average was lower than the longer, the price was considered to be in a movement downwards, so an opportunity to perform the covered call strategy. Also, the results were based on simulations using stocks and stock options from Petrobrás (PETR3) from 2008 to 2010. In his study, he compared the strategy developed with the results generated by only holding the stock during the respective period and concluded that the strategy developed showed an efficient behavior.

Here, we also propose a study of the covered call strategy, but we analyzed a larger portfolio of Brazilian stocks and several strategies for choosing the strike price for which to sell call options during 2022. Among the strategies developed here, there are the ones using popular trading indicators and the ones based on Machine Learning. Additionally, we also study a strategy based on the moving average to decide the strike price of the call contract, but this strategy compares the last 100 days moving average with the stock price at the moment.

3.3 Evaluation of Strategies in the Brazilian Stock Option Market

The last work we will mention in this chapter was conducted by Gabriel Franco Pereira and Pietrangelo Ventura De Biase in 2012 and was named Evaluation of Strategies in the Brazilian Stock Option Market. Their work had as objectives analyzing the Black-Scholes model for stock option pricing and discussed some option-based strategies to improve return such as the bear spread with call contracts strategy. They concluded the risk incurred in the strategies was proportional to their return and that some strategies depend on the market tendency to be successful. In other words, using a strategy that expects the market to go up when it, in fact, goes down resulted in poor results. However, knowing how to apply the strategies is a useful tool to improve returns.

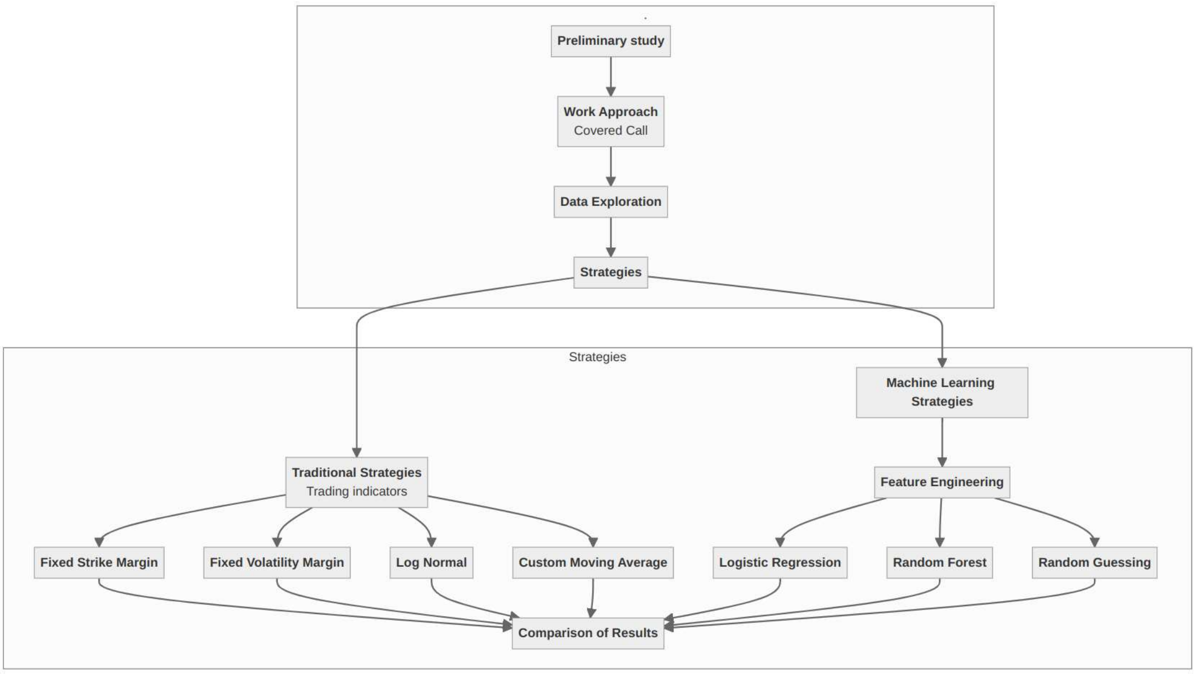
Considering their study, the work presented here differs by focusing on one strategy which is the covered call strategy and different ways we can approach it. Also, a larger stock portfolio was considered, since their study was conducted using only PETR4 and the Ibovespa index.

Chapter 4

Methodology

Now, we will describe how the current study was conducted as well as the decisions taken and which tools were used. This chapter is divided basically in two parts: data manipulation and strategy definitions. The flowcharts here below demonstrate the steps taken in each one of these parts. First, we performed an initial study about the stock market and stock options, then we dealt with the data and, finally, we explored the traditional and machine learning strategies to decide which strike price to sell call options for.

Figure 4.1: Steps taken to define and analyze the option based strategies.



Source: created by the author.

In the next sections, the steps mentioned above are explored in details.

4.1 Data

Since B3 is the largest stock exchange in the Brazilian stock market, it has the main source of financial data we can make use of. The data provided by B3 contains the historical daily information of all stocks available in the market. For instance, we have access to the maximum, minimum, opening and closing prices of all stocks day by day. In addition to it, the data contains information such as premiums, strike prices and expiration dates of the respective stock options. We should note as well that there are other sources in the market where the same data can be found such as MetaTrader, some of them even provide it in real time.

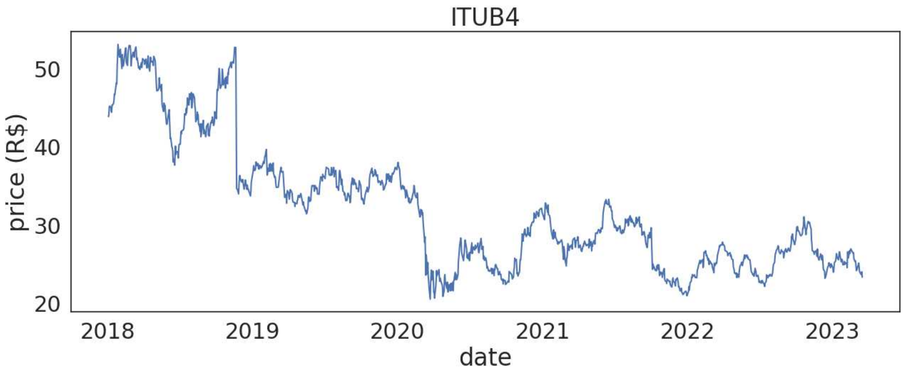
Besides data concerning stock and stock options, we also used data about the dividends distributed by the companies and the occurrences of stock splits in our analysis. These data are available at the websites of the owner companies of the stocks chosen for the study.

In this work, the date range for the data retrieved from B3 goes from January, 2019 to May, 2023.

4.1.1 Data Processing

B3 provides historical data by year in text files formatted accordingly to a particular layout specified in its website. In order to have this data available to be processed by our Python scripts, we had to download it from the website and parse it using Pandas. After having parsed B3's data, we had it available as the following graph exemplifies.

Figure 4.2: ITUB4 price time series.



Source: created by the author.

4.1.2 The Portfolio

In order to perform our analysis, we decided to take stocks from large companies and with high liquidity, in other words, stocks highly traded in the market. Moreover, equally important, we opted for stocks whose stock options were also liquid to guarantee we would not be impacted by lack of trades during the simulations.

That being said, here below we have the list of stocks chosen to compose our portfolio for the study. These stocks belong to large Brazilian companies and have high volumes of negotiations.

- Petrobrás. PETR4.
- Vale. VALE3.
- Itaú. ITUB4.
- Banco do Brasil. BBAS3.
- Ambev. ABEV3.
- B3. B3SA3
- Gerdau. GGBR4
- BOVA11.[1]

4.1.3 Stock Splits and Dividends

If we go back to the graph showing the stock price of ITUB4 4.2, we notice an abrupt fall near the end of 2019. This fall is due to the stock split performed by Itaú back then. To handle stock splits in our analysis, we decided to multiply the price prior to the split by the split factor. We do so to make the stock price curve become consistent as if no split was performed. For ITUB4 where at 21/11/2019 each stock was split in three, this means that the price prior to this date was multiplied by $\frac{1}{3}$.

Along with stock splits, dividends also impact the price of stocks and the calculation of a stock's return. In order to handle dividends, we considered that they were never paid, hence never discounted from the stock price. The reason behind it is to concentrate the source of return in one place, the stock price, which simplifies the simulations and keeps the price more consistent. Also, the non-payment of dividends is one of the assumptions of the Black-Scholes pricing model 7, thus we are in accordance with its specifications.

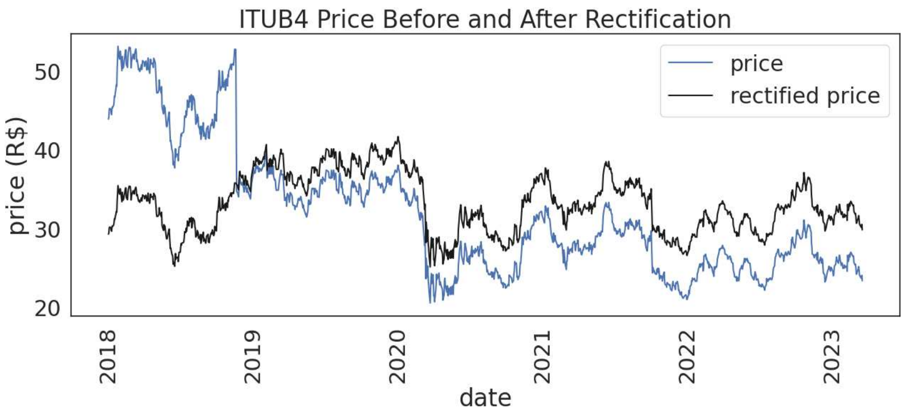
To perform the operation of adding dividends back to the stock price, for each stock, we took the ex-dividend date 1 of each dividend paid and, from this date forward, we increased the stock price by the value of the dividend.

[1]BOVA11 is an Exchange Traded Fund (ETF), therefore it doesn't belong to any particular company.

Moreover, not only the stock price is impacted by stock splits and dividends, but also the strike prices of the options associated with the involved stock. Consequently, we performed the same operations to the strike price of all options concerned.

It is important to notice that the order of the operations was to add back dividends to the stock price and, then, multiply it by the split factor, since the dividends needed to be normalized as well. The graph below illustrates the price of ITUB4 before and after the operations.

Figure 4.3: ITUB4 price before and after rectification.



Source: created by the author.

4.2 Strategies

After manipulating the data to adapt it to our needs, we can start developing and simulating strategies. In this work, we decided to use the covered call structure as the base of the strategies analyzed as well as ignore all transaction costs associated to it. By using the covered call strategy, the usual investor can trade call options while holding the stocks in his portfolio. Since the base of the strategies is the same, they only differ in the way the call options are chosen. Moreover, the decision of which call to sell is basically a decision of which strike margin is favorable given the option premiums available.

We decided to run the simulations selling call options once a month all along the year of 2022. Also, the call options sold had always one month until expiration, thus the sale of new options happened when the ones sold in the previous month expired. To make it clear, for a given strategy, in January, 2022, we sold a call for each stock in the portfolio. In February, the contracts sold in January expired, so we evaluated whether a strike took place or not. If a strike happened, we were forced to sell all the stocks for the strike price and, right after, we took the entire amount available to buy new stocks in order to continue performing operations. If no strike happened, we just bought new calls. The same logic applied until November included. At last, in December, we just evaluated if the calls sold the previous month generated a strike, so no new calls were sold and we consolidated the simulation results. Besides, for all the strategies analyzed, we considered an investment amount of R$ 100,000.00 available at the beginning of each simulation to buy stocks and sell calls so that the return at the end was given by (final amount − initial amount)/initial amount.

A caveat important to state is that finding a call option with exactly the strike price we defined is not always an easy task. For instance, sometimes for a particular month, we wanted to sell a call with a strike price 5% higher than the current stock price, but the closest one we found was with a strike price 7.5% higher. Therefore, for those cases, we opted for selling the call with the closest strike price during the simulations in order to avoid having months without trades.

We opted to simulate the strategies in 2022, because we could analyze performance in a considerable time span where no exceptional events impacting the stock market happened in Brazil, which was not the case, for instance, in 2020 with the Coronavirus pandemic. Additionally, the data available prior to 2022, therefore from 2019 to 2021, was used to train the machine learning models.

This section is divided in two. The first part goes over traditional strategies, that is to say strategies that apply simple logic and are based on standard trading indicators, whereas the second part discusses strategies that use machine learning.

4.2.1 Traditional Strategies

The strategies here analyzed make use of simple approaches and trading indicators to make the decision of what call option to sell.

Buy and Hold

Renowned investors such as Warren Buffet in the US and Luiz Barsi in Brazil consider buy and hold the best way to invest in the stock market. In this strategy, an investor composes a portfolio of stock from companies he believes will grow in the long term and hold these stocks hoping that their price will rise and the dividends will be generous.

Since buy and hold is such an adopted strategy for most investors, it is reasonable to use it as the benchmark we will try to outperform with the following call-based strategies.

Fixed Strike Margin

To recap, in the context of this work, the covered call strategy consists of, for each stock in the portfolio, selling a call option with a strike price we believe will be higher than the stock price at expiration. Therefore, a simple way of deciding which strike price to choose every month is by using a fixed percentage margin in relation to the stock price at the moment of sale. For instance, if we chose a strike margin of 3%, every month we would sell a call with a strike price 3% higher than the current stock price in the hope the stock would rise at most 3% in order to avoid a strike.

In this study, we decided to evaluate the following strike margins: -7%, -5%, -3%, 0%, 3%, 5% and 7%.

Fixed Volatility Margin

When we sell a call option, what we are theoretically doing is selling protection against the risk of losing a great rise in the asset's price. In the stock market, a common measure of risk is the stock volatility. That being said, an interesting strategy we implemented was to decide which call option to sell by comparing the historical volatility to the volatility expected by the market for that particular option, in other words, its implied volatility 8.

Therefore, the same way we did for the fixed strike margin strategy, here, we will also use a fixed margin, but it is a fixed margin of volatility. For instance, a volatility margin of 5% means we chose the call whose implied volatility was the closest to 5% greater than the historical volatility.

On the one hand, to calculate the historical volatility each month, we considered the price of the last 60 days to compute a daily volatility and, then, we annualized this value. On the other hand, the implied volatility was calculated using the equations derived from the Black-Scholes model 7 and we used the Newton Raphson Algorithm. Also, we only considered calls with strike margin between -10% and 10% to avoid taking values too extreme and we used the price of the last 180 days to calculate the historical volatility.

The volatility margins evaluated were the following: 0%, 30%, 50% and 100%.

Log Normal

As mentioned in chapter 1, the log returns of a stock follow a Gaussian distribution. Having this in mind, we can come up with a strategy that uses this information.

Since our goal is to find a strike margin that will cover the stock price at expiration, we can find this value by defining the risk of strike we accept. This value of risk translates to the probability the stock price will go above the strike price at expiration, which is always one month in the future. For instance, a strike risk of 25% means we accept to sell a call with a strike price that has 25% of probability of being exercised at expiration.

For the purpose of choosing a strike margin given a strike probability, we can use the inverse CDF of the 30-day log returns. This is due to the fact the CDF takes as input a value x and outputs the probability p of a value $y \in (-\infty, x]$ happening. Therefore, its inverse must take a probability p and outputs a value x where the range $(-\infty, x]$ has probability p of happening.

Let us visualize this with an example. Say a given month we defined the strike risk as 25%, in other words, we want a strike price k so that we have 75% probability that the stock price will be in the range $[0, x]$ in one month from now. Hence, we calculate the inverse CDF of the value 75% which will yield a log return, and, then, we can convert this value to a strike margin and find a strike price.

For this strategy, we defined strike risks 10%, 30% and 50% to be simulated, since a strike risk of more than 50% means we have more chances of having a strike than not. Also, we used the last 90 days to compute the volatility that shapes the distribution.

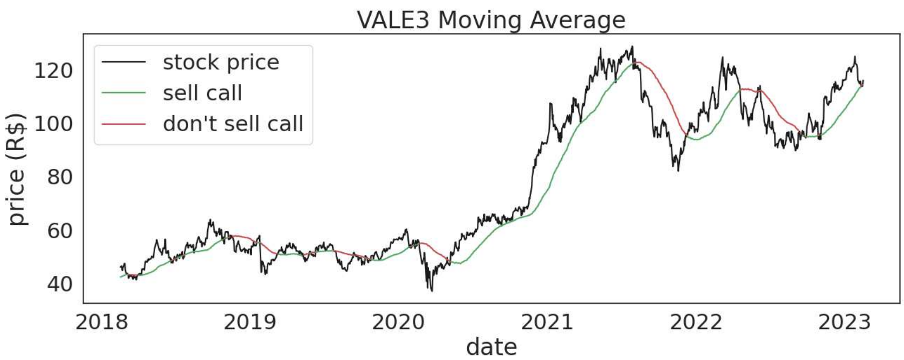
Custom Moving Average

In the previous strategies, a call was sold for each stock in the portfolio every month. Now, to take a different approach, we avoided selling calls when we predicted the stock price would go up by any amount. To do so, we used the last 100 days moving average 2.1.2 as an indicator of the direction the stock would move.

For this strategy, assuming prices tend to return to an average value, when the current stock price was below the moving average, we did not sell any call, because we considered there would be an increase in price, thus a strike would happen. However, if the price was above, we predicted it would fall until the next iteration, then we sold a call option with a strike price 3% lower than the current price for each stock.

In a nutshell, by using the moving average, we sold call options when we believed the price would fall and we just held the portfolio when we predicted it would rise in value. The graph below shows this behavior for VALE3. The green sections would be opportunities and the red ones would be moments to be conservative.

Figure 4.4: VALE3 moving average with highlighted sections based on the ratio between stock price and moving average.



Source: created by the author.

4.2.2 Machine Learning Strategies

Machine Learning is a tool widely adopted to analyze financial data. In this work, we used it to decide for which strike margin to sell call options during the simulations. To do so, we defined the target and features as well as the models to train and test the data on. Also, we implemented a model that classifies the inputs randomly to serve as a point of comparison.

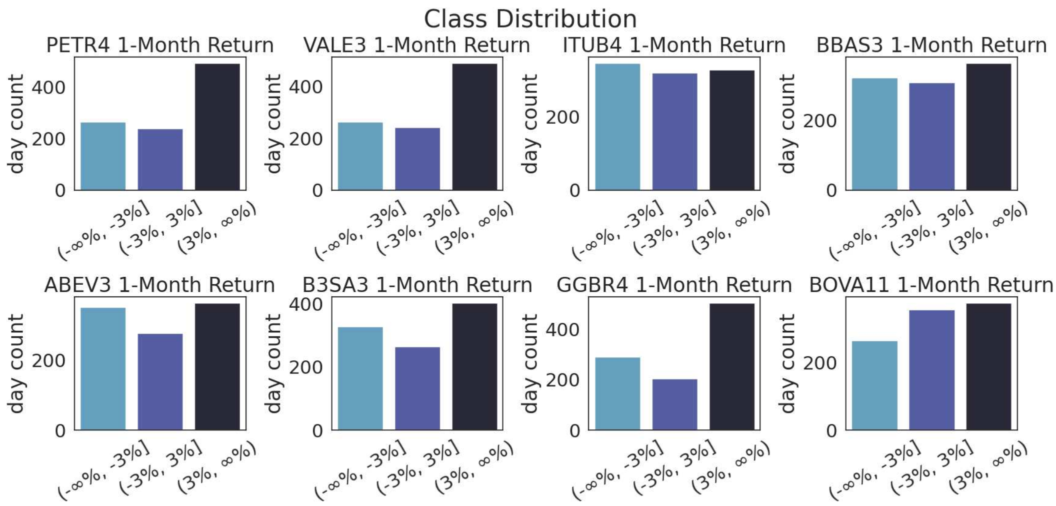
Prediction Target

The goal is to find the right strike margin for which to sell call options. In this scenario, a good strike margin is a value higher than what the stock return will be at expiration, because if we chose a strike margin higher than the future stock return, then no strike would take place and we wouldn't be forced to sell the portfolio at the expiration date.

With this in mind, a reasonable target for the Machine Learning models is the stock return 30 days in the future, as, in the context of this study, the calls always had one month until expiration. Also, since classification problems are in general easier than regression ones, the target was chosen to be classes of stock return, namely the ranges $(-\infty\%, -3\%]$, $(-3\%, 3\%]$ and $(3\%, \infty\%)$.

The graphs below show how the classes are distributed for each stock in the portfolio.

Figure 4.5: Class distribution for each stock from 2019 to 2022.



Source: created by the author.

As for the behavior during the simulation, every month, the models would predict one of the classes listed above. If the classes $(-\infty\%, -3\%]$ and $(-3\%, 3\%]$ were the output, then, the strike margin was chosen to be -3% and 3% respectively, whereas, if the class $(3\%, \infty\%)$ was predicted, strike margin was chosen to be 8% to avoid a strike at expiration.

Features

Having defined the target to be the stock return 30 days in the future, we needed to select features plausible of being capable of predicting the target. Therefore, we started with some simple ones to form a first group of features.

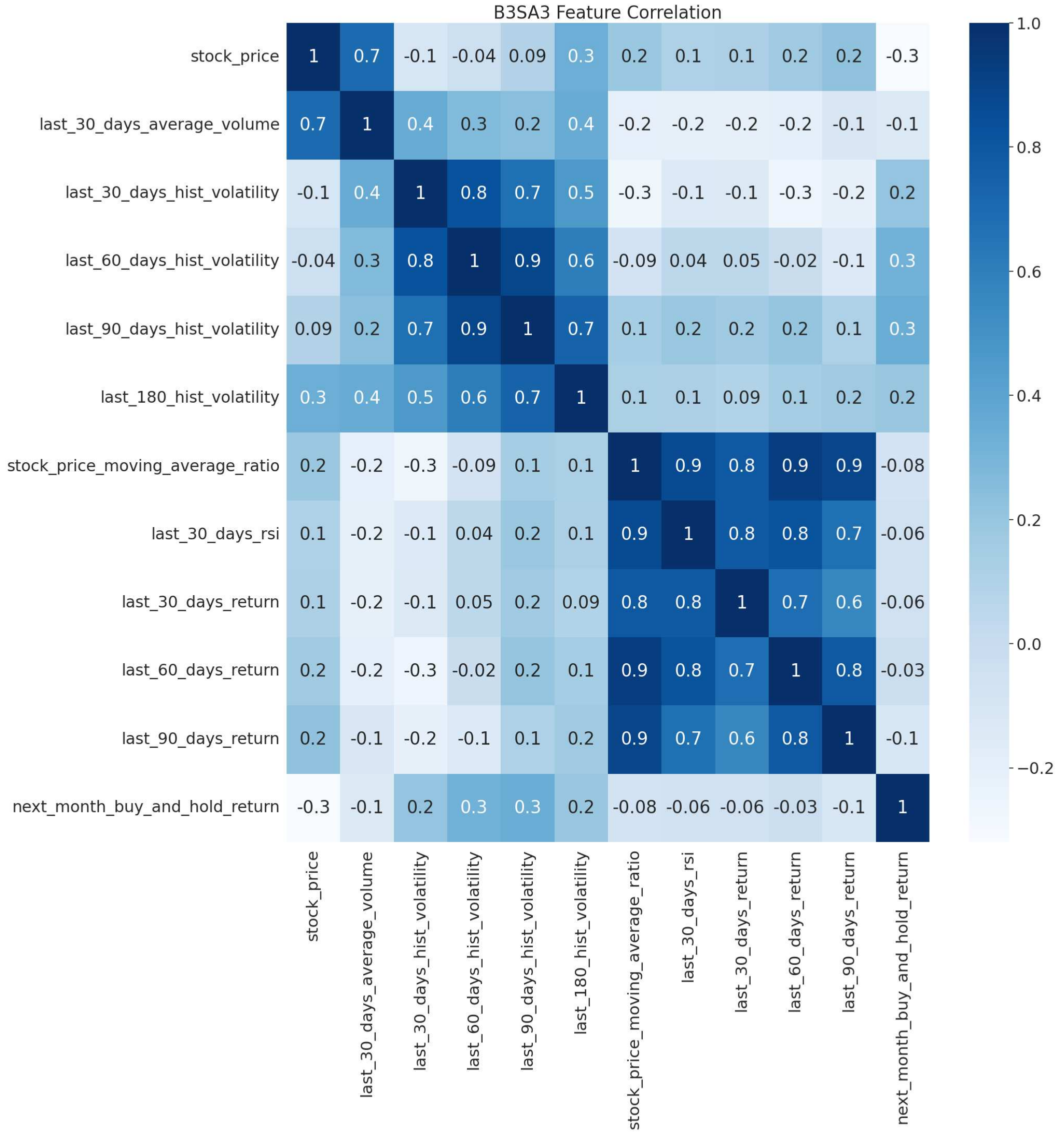
- The current stock price.
- The historical volatility considering the last 180 days.
- The current stock price divided by the last 100 days moving average
- The stock return in the last 30 days.
- The stock return in the last 60 days.
- The stock return in the last 90 days.

Then, we added some more features to the previous group to compose a second one so we could verify if the training and test accuracies would increase. That being said, the second group of features was:

- The current stock price.
- The historical volatility considering the last 180 days.
- The current stock price divided by the last 100 days moving average
- The stock return in the last 30 days.
- The stock return in the last 60 days.
- The stock return in the last 90 days.
- The historical volatility considering the last 30 days.
- The historical volatility considering the last 60 days.
- The historical volatility considering the last 90 days.
- The last 30 days RSI 4.
- The last 30 days average volume traded.

The plot below shows the correlation of the features described above for the stock B3SA3.

Figure 4.6: B3SA3 feature correlation.



Source: created by the author.

Models

Seeing that target value chosen for analysis is made of classes, we only considered classification Machine Learning models to predict stock return. Here, the models chosen were the Logistic Regression 2.2.2 with the One-vs-Rest approach, which is the simplest classification model, and the Random Forest Classifier 2.2.3, which consists of a more elaborate approach.

For each stock, the models mentioned above were trained using data from 2019 to 2021 and tested using data from 2022, the same data used during the simulations. Moreover, the parameter maximum depth of the Random Forest models was set to 5.

Random Guessing Approach

In order to verify the validity of the results generated by the Logistic Regression and the Random Forest, a random guessing model was implemented. This model outputs the target classes with equal probability. By doing so, we can compare the models and confirm whether or not the real one brought any improvement over a random process.

Chapter 5

Results

Now that we understand how the study was conducted, in this chapter, we will go through the outcomes of the simulations of each one of the strategies discussed previously so that we can evaluate how they performed and explore the main results.

5.1 Traditional Strategies

Let us start by verifying how the traditional strategies performed.

5.1.1 Buy and Hold

To begin with, we have the results generated by the simulations of the Buy and Hold strategy during 2022.

Table 5.1: Buy and Hold results in percentage.

Strategy	Stock Return								AVG	STD
	PETR4	VALE3	ITUB4	BBAS3	ABEV3	B3SA3	GGBR4	BOVA11		
buy and hold	26.31	7.72	4.26	13.35	0.97	-12.12	19.75	-5.06	6.9	11.89

Source: created by the author.

As we can see some stocks resulted in profit by only holding them such as PETR4, while others resulted in loss namely B3SA3. However, the average result considering the 8 assets was positive with value 6.9%. Therefore, this is the benchmark we expect the other strategies to outperform and we will plot it alongside the next results to simplify comparisons.

5.1.2 Fixed Strike Margin

Moving forward, the next strategy was the one using fixed strike margins.

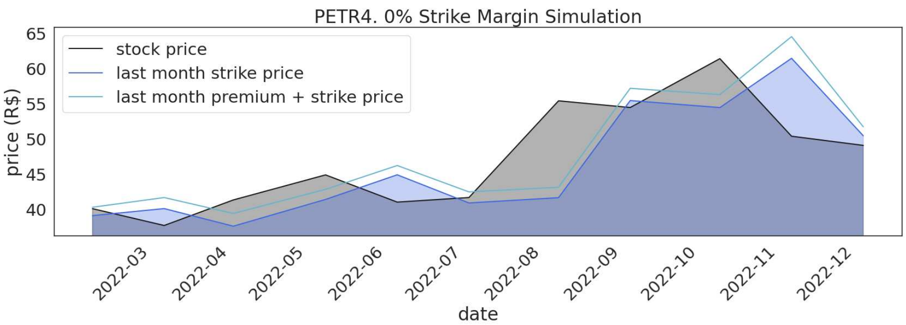
Table 5.2: Fixed strike margin strategy results in percentage.

Strategy	Stock Return								AVG	STD
	PETR4	VALE3	ITUB4	BBAS3	ABEV3	B3SA3	GGBR4	BOVA11		
buy and hold	26.31	7.72	4.26	13.35	0.97	-12.12	19.75	-5.06	6.9	11.89
-7%	1.93	11.19	8.35	2.11	15.98	1.04	12.97	9.04	7.83	4.92
-5%	0.19	11.03	5.38	3.12	19.76	1.15	16.97	8.89	8.32	6.77
-3%	1.62	13.82	2.27	6.0	20.88	2.53	20.85	5.47	9.18	7.52
0%	1.44	16.06	2.08	4.03	16.88	-1.25	16.57	0.17	7.0	7.5
3%	3.19	15.28	-4.52	4.69	12.68	-3.83	18.4	-4.56	5.17	8.71
5%	3.4	15.95	-5.35	4.74	9.21	-8.18	16.98	-7.11	3.7	9.34
7%	5.04	15.28	-4.29	5.75	7.12	-12.25	17.69	-8.15	3.27	10.03

Source: created by the author.

The plot above has some interesting results. First, selling call options with no matter which fixed strike margin was a strategy that performed poorly for PETR4. This was mainly due to the fact the price of PETR4 exhibited a huge spike next to August, 2022 which caused our options to be struck and, also, we ended up losing the respective rise in price. The graph below shows this behavior. Second, the fixed strategies performed the best when the result holding the stock was small or even negative which was the case for ABEV3 and B3SA3.

Figure 5.1: Simulation results for PETR4 with a fixed strike margin of 0% in 2022.

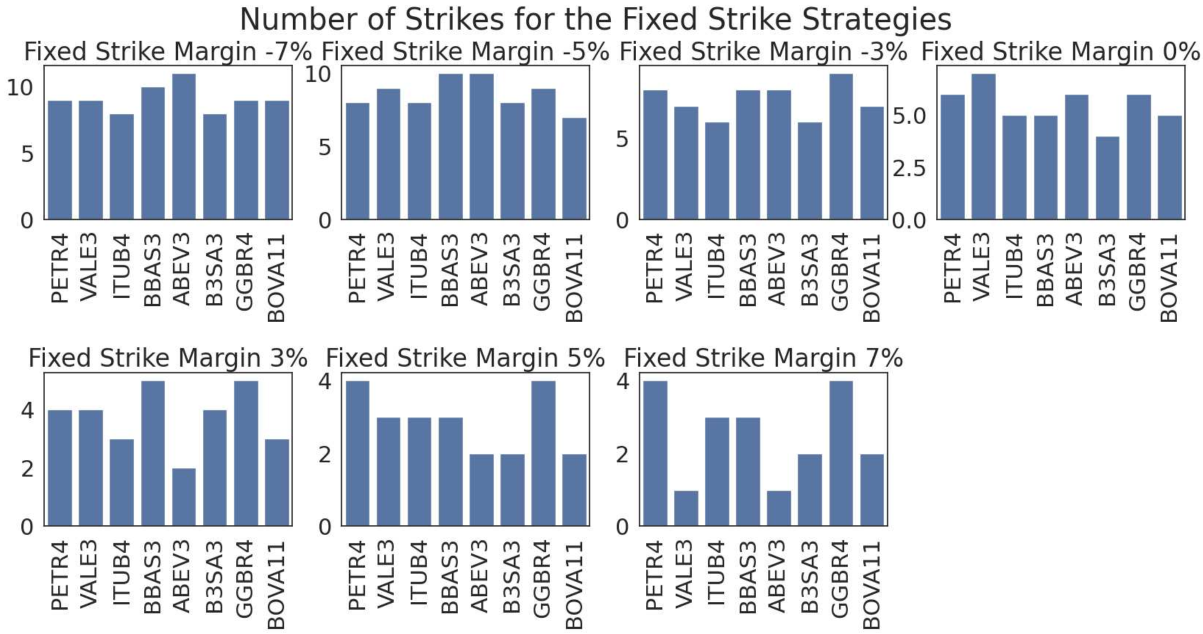


Source: created by the author.

Nevertheless, considering the entire portfolio, the average result using some of the strike margins surpassed the Buy and Hold approach. This was the case for the non positive margins -7%, -5%, -3% and 0% where -3% was the best with 9.18% of profit in 2022, an increase of 2.28% in relation to Buy and Hold.

Additionally, the plot below shows the amount of strikes for each strike margin configuration and each stock in the portfolio. We verify the lower the strike margin was, the more strikes happened. All the same, lower strike margins mean higher premiums and, in the case of the strike margin -3%, the strikes were compensated.

Figure 5.2: Number of strikes for each fixed margin strategy for each stock.



Source: created by the author.

5.1.3 Fixed Volatility Margin

For this strategy, we used how above the implied volatility of the available calls was in regards to the historical volatility to decide which contracts to sell. The results are in the following table.

Table 5.3: Fixed volatility margin strategy results in percentage.

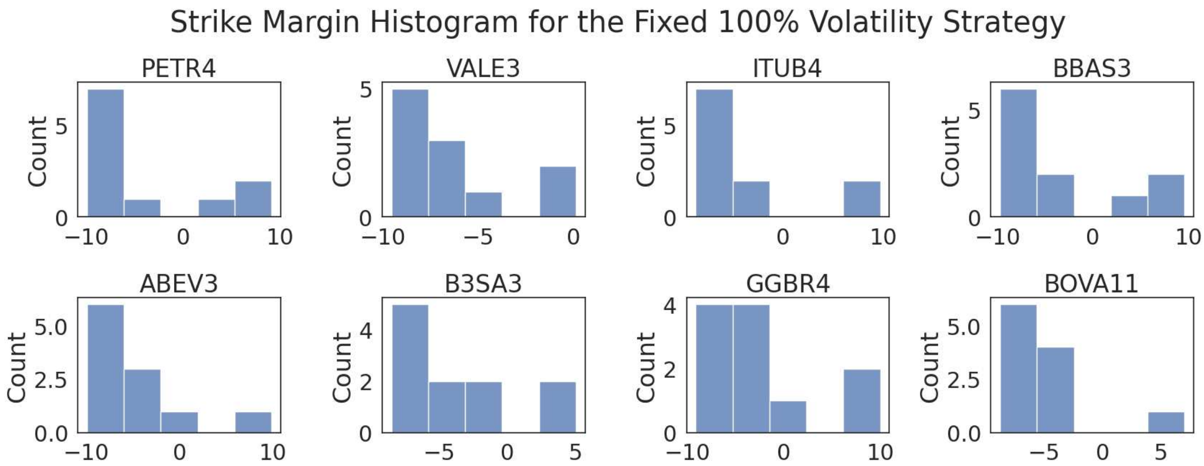
Strategy	Stock Return								AVG	STD
	PETR4	VALE3	ITUB4	BBAS3	ABEV3	B3SA3	GGBR4	BOVA11		
buy and hold	26.31	7.72	4.26	13.35	0.97	-12.12	19.75	-5.06	6.9	11.89
0%	1.82	20.04	3.34	7.42	5.82	2.11	20.63	1.23	7.8	7.5
10%	4.35	11.42	1.38	11.87	22.34	-3.67	21.57	-2.44	8.35	9.47
30%	-5.49	13.40	-10.7	8.11	19.3	-6.46	23.68	10.02	6.48	11.88
50%	-9.56	19.94	1.88	10.16	19.95	0.59	24.7	16.71	10.55	11.14
100%	-0.26	19.94	4.03	1.02	19.27	0.59	24.7	17.76	10.88	9.78
150%	-0.26	19.94	4.03	1.02	19.27	0.59	24.7	17.76	10.88	9.78

Source: created by the author.

Although the results above seem promising, especially considering the average, we must be careful analyzing them. The difference between the implied volatility and the historical volatility has a large amount of noise contained, thus it is difficult to say the margins chosen were the only responsible for the results achieved. With this in mind, a volatility margin of 10%

and 20% appeared to have no practical difference other than what could be just randomness. Moreover, from volatility margins above 40%, it was rare to find matching call options and the simulations always took the calls with the highest implied volatilities which tended to be the ones with lowest strike prices as we see in the graphs below. Therefore, since 2022 was a year of general fall in prices considering the portfolio studied, it is understandable that the simulations parameterized with these values seemed to perform well.

Figure 5.3: Histograms of the strike margins chosen during the simulation of the fixed 100% volatility margin strategy.



Source: created by the author.

5.1.4 Log Normal

The log normal strategy consisted of using the Gaussian distribution of the log returns to decide for which strike margin to sell call options.

Table 5.4: Log normal strategy results in percentage.

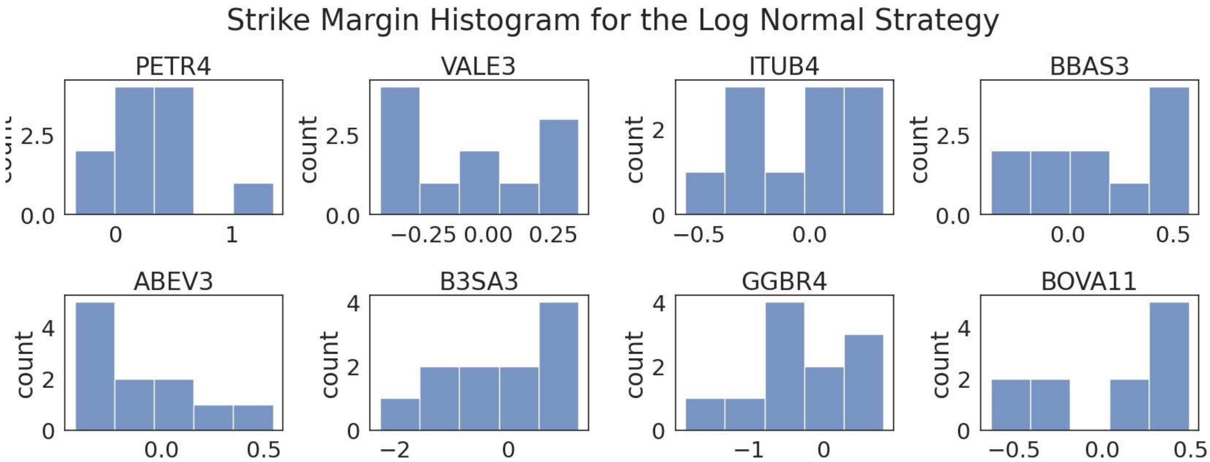
Strategy	Stock Return								AVG	STD
	PETR4	VALE3	ITUB4	BBAS3	ABEV3	B3SA3	GGBR4	BOVA11		
buy and hold	26.31	7.72	4.26	13.35	0.97	-12.12	19.75	-5.06	6.9	11.89
10%	7.05	8.49	-3.05	10.79	8.87	-14.91	9.71	-5.34	2.7	8.75
30%	1.67	12.33	-4.24	4.1	11.8	-11.26	13.77	-4.07	3.01	8.58
50%	2.67	14.38	2.08	5.73	16.88	-0.48	18.43	0.82	7.56	7.24

Source: created by the author.

Considering the portfolio average return, the only strike probability that performed better than Buy and Hold was 50%. Even so, the improvement was only 0.66% which is a small amount to consider it an alternative option against Buy and Hold.

For the strike probability with the best result, the strike margins chosen during the simulation were close to 0% as the following histogram shows.

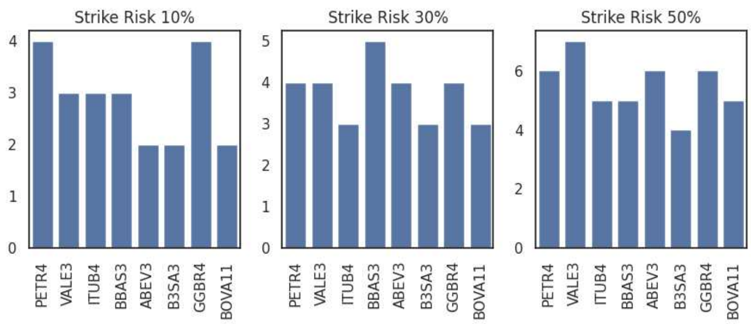
Figure 5.4: Histograms of the strike margins chosen during the simulation of the log normal strategy for a strike probability of 50%.



Source: created by the author.

Moreover, a considerable number of strikes happened for each stock, especially when using a strike probability of 50% which means the strike margins found with the distribution of log returns is not reliable to predict the stock movement.

Figure 5.5: Number of strikes for each log normal strategy for each stock.



Source: created by the author.

It was expected a poor performance for this strategy, seeing that the log normal distribution of returns needs a considerable amount of past data to be calculated and doesn't capture recent trends in the stock price movement.

5.1.5 Custom Moving Average

This strategy differs from the rest, since calls were not sold every month, only when the stock price was above the last 100 days moving average. We considered this to be an indicator of movement downwards that would prevent strikes.

Table 5.5: Custom moving average strategy results in percentage.

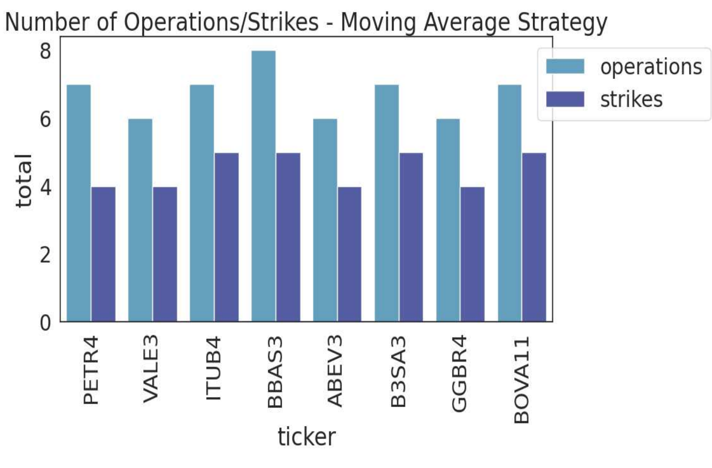
Strategy	Ticker								AVG	STD
	PETR4	VALE3	ITUB4	BBAS3	ABEV3	B3SA3	GGBR4	BOVA11		
buy and hold	26.31	7.72	4.26	13.35	0.97	-12.12	19.75	-5.06	6.9	11.89
custom ma	33.0	4.0	0.01	14.51	14.6	-3.34	38.58	1.15	12.9	14.58

Source: created by the author.

So far, this strategy was the one with the best performance. The average return was 12.31%, an increase of 5.91% considering the Buy and Hold strategy for the same period. Also, this time, we were able to outperform the return holding the stock for PETR4.

Now, if we look at the next plot, we see that the number of strikes was reduced when we compare it to the fixed strike strategies, after all we don't sell call options for some months. Also, for most stocks, the number of strikes was about half the number of sold calls.

Figure 5.6: Number of operations and strikes for the custom moving average strategy for each stock.



Source: created by the author.

5.2 Machine Learning Strategies

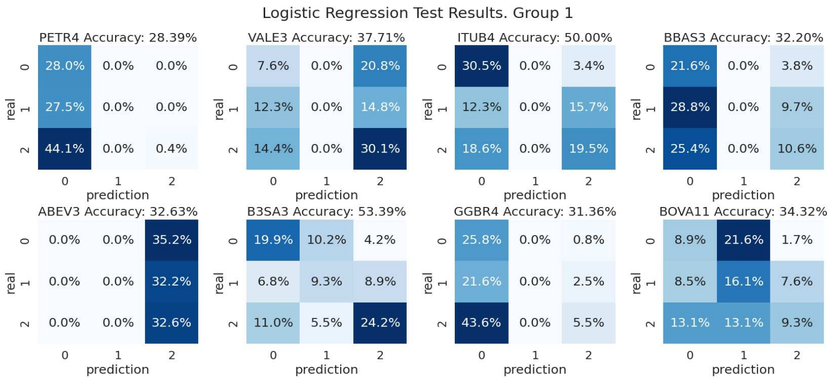
Now that we have seen how the traditional strategies did, let us inspect the Machine Learning strategies.

5.2.1 Logistic Regression

This strategy consisted of using the simplest classification model to predict the return range of a stock 30 days in the future in order to choose the right strike margin.

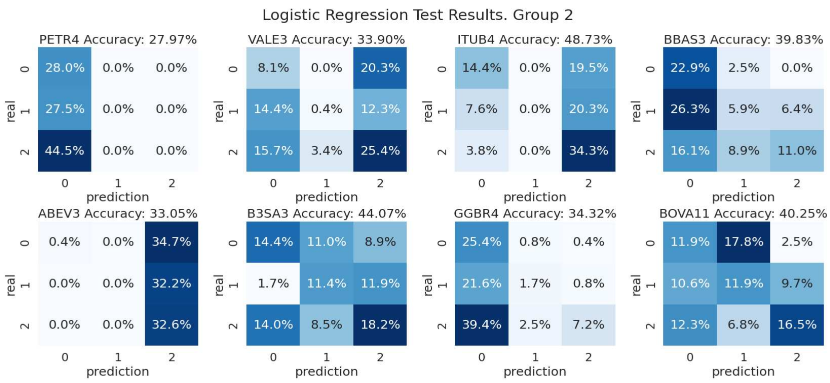
After training the models with data from 2019 to 2021, we tested them with data from 2022. The test results show that, in general, the models made most of their predictions to just one class, thus having a high rate of misclassification. Also, even after adding more features, the accuracies didn't improve, therefore we decided to run the simulations only with the first group of features 4.2.2.

Figure 5.7: Logistic Regression test confusion matrices using the first group of features 4.2.2.



Source: created by the author.

Figure 5.8: Logistic Regression test confusion matrices using the second group of features 4.2.2.



Source: created by the author.

The simulation results for the first group of features are shown in the next table.

Table 5.6: Logistic Regression strategy results in percentage.

Strategy	Stock Return								AVG	STD
	PETR4	VALE3	ITUB4	BBAS3	ABEV3	B3SA3	GGBR4	BOVA11		
buy and hold	26.31	7.72	4.26	13.35	0.97	-12.12	19.75	-5.06	6.9	11.89
logistic regression	3.34	8.56	17.26	7.4	6.98	5.78	15.32	-2.7	7.74	5.93

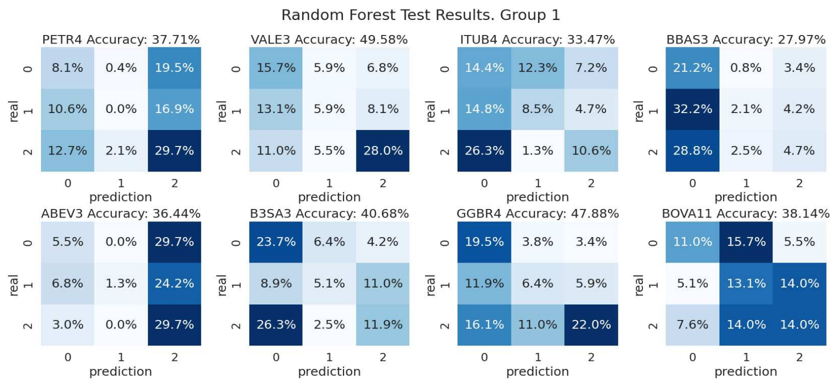
Source: created by the author.

As we can see, the Logistic Regression models didn't perform well having only an increase of 0.84% compared to the Buy and Hold strategy in average.

5.2.2 Random Forest

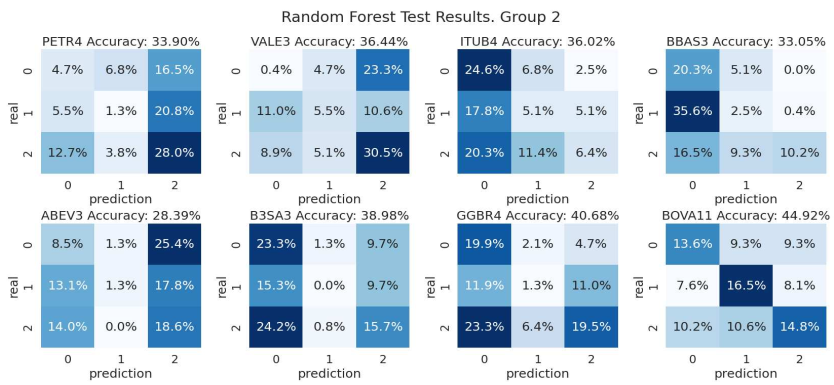
Here, we used a more sophisticated model which was the Random Forest Classifier to predict the return ranges. The test results for both feature groups 1 and 2 are shown below.

Figure 5.9: Random Forest test confusion matrices using the first group of features 4.2.2.



Source: created by the author.

Figure 5.10: Random Forest test confusion matrices using the second group of features 4.2.2.



Source: created by the author.

The Random Forest models tended to choose one target class and direct most of their predictions to it. Moreover, the models exhibited a poor performance for both feature groups where the group containing more features performed worse. That being said, we only kept the models trained with the first feature group for simulation.

Table 5.7: Random Forest strategy results in percentage.

Strategy	Stock Return								AVG	STD
	PETR4	VALE3	ITUB4	BBAS3	ABEV3	B3SA3	GGBR4	BOVA11		
buy and hold	26.31	7.72	4.26	13.35	0.97	-12.12	19.75	-5.06	6.9	11.89
random forest	7.5	20.35	-1.09	13.06	8.9	1.51	21.91	-0.63	8.94	8.4

Source: created by the author.

Even though the models had unsatisfactory accuracies, we notice a considerable improvement of 2.04% over the average result of the Buy and Hold strategy. However, this seems to be due to the fact the covered call per se improves the return for periods of general fall in the market. In this sense, we also calculate the correlation between the test accuracies and the simulation results and the value found was 0.55. This value is relatively low considering the accuracies and the simulation results were calculated based on the same data.

5.2.3 Random Guessing

With the goal of verifying that the Machine Learning models were responsible for any improvement over the Buy and Hold strategy, we implemented a model that guesses randomly the target classes. The results of this approach are shown in the next table.

Table 5.8: Random guessing strategy results in percentage.

Strategy	Stock Return								AVG	STD
	PETR4	VALE3	ITUB4	BBAS3	ABEV3	B3SA3	GGBR4	BOVA11		
buy and hold	26.31	7.72	4.26	13.35	0.97	-12.12	19.75	-5.06	6.9	11.89
random guessing	0.81	9.71	4.32	7.51	24.99	-2.67	16.53	6.09	8.41	8.26

Source: created by the author.

Despite being completely random, the model achieved 1.51% of improvement over the benchmark in average.

5.3 Discussion

Beforehand, the table below contains the assembled results of the Buy and Hold strategy, the best fixed margin strategy, the best volatility margin strategy, the best log normal strategy, the custom moving average strategy, the Logistic Regression strategy for the first group of features, the Random Forest strategy for the first group of features and the random guessing strategy.

Table 5.9: All simulation results in percentage.

Strategy	Stock Return								AVG
	PETR4	VALE3	ITUB4	BBAS3	ABEV3	B3SA3	GGBR4	BOVA11	
buy and hold	26.31	7.72	4.26	13.35	0.97	-12.12	19.75	-5.06	6.9
strike margin -3%	1.62	13.82	2.27	6.0	20.88	2.53	20.85	5.47	9.18
vol margin 100%	-0.26	19.94	4.03	1.02	19.27	0.59	24.7	17.76	10.88
log normal 50%	2.67	14.38	2.08	5.73	16.88	-0.48	18.43	0.82	7.56
custom ma	33.0	4.0	0.01	14.51	14.6	-3.34	38.58	1.15	**12.9**
logistic regression	3.34	8.56	17.26	7.4	6.98	5.78	15.32	-2.7	7.74
random forest	7.5	20.35	-1.09	13.06	8.9	1.51	21.91	-0.63	8.94
random guessing	0.81	9.71	4.32	7.51	24.99	-2.67	16.53	6.09	8.41

Source: created by the author.

As to the next table, we took the improvement each strategy produced for each stock in relation to the Buy and Hold strategy and, then, we calculated the average and standard deviation per strategy.

Table 5.10: Relative improvements towards the Buy and Hold strategy in percentage.

Strategy	AVG	STD
strike margin -3%	2.52	13.03
vol margin 100%	3.99	15.50
log normal 50%	0.67	11.66
custom ma	**6.0**	**7.42**
logistic regression	0.85	11.77
random forest	2.04	9.88
random guessing	1.52	13.56

Source: created by the author.

Besides, another interesting way to visualize the simulation results is by running them in different time windows and considering just the average results. Therefore, we ran the simulations on different six-month intervals shifted by one month.

Table 5.11: Average simulation results in percentage in different six-month time windows shifted by one month during 2022.

Strategy	Period						
	Jan-Jul	Feb-Aug	Mar-Sep	Apr-Oct	May-Nov	Jun-Dec	Jan-Dec
buy and hold	-5.01	4.94	2.34	10.52	4.68	10.8	6.9
strike margin -3%	4.60	5.44	5.56	6.09	2.22	5.45	9.18
vol margin 100%	2.89	4.15	3.16	7.93	3.65	8.26	10.88
log normal 50%	2.33	3.40	3.87	5.24	1.69	6.13	7.56
custom ma	0.28	12.06	10.58	11.55	6.85	10.41	12.9
logistic regression	0.76	3.4	1.68	3.78	1.61	6.57	7.36
random forest	2.83	4.61	3.14	5.61	3.37	6.58	8.94

Source: created by the author.

On the face of the tables above, the best average result was promoted by the custom moving average strategy which also held the least standard deviation of improvements over Buy and Hold considering different time windows and the whole interval of 2022. This strategy had a differential in the way it chose the call options which was by only selling for months it predicted the prices would go down. This was the main responsible for its performance. Additionally, the Machine Learning models didn't exhibit satisfactory performances since the accuracies of the models were low and the simulation results didn't outperform the Buy and Hold strategy the way we expected. Moreover, the random guessing strategy was able to match the Machine Learning models which only calls into question the efficiency of such techniques to predict stock prices. Also, the fixed strike margin strategy proved to be a reasonable benchmark, whereas the volatility margin strategy had its results mainly due to randomness.

5.4 Contributions

Among the contributions of this work, there is a study on how call options can be leveraged in the Brazilian stock market to improve returns over time. Besides, the Python scripts and data used throughout this work to simulate and plot strategies are available on GitHub at `https://github.com/luikymagno/stock-options`.

The only difference between the strategies analyzed here is the way we choose the strike price and, therefore, the call to be traded in order to perform the covered call strategy. To do so, we create a base simulation script and implemented the strategy pattern so that we could seamlessly input different functions to pick call options and evaluate how they performed in 2022. With this in mind, developers and investors interested in performing their own simulations as well as expanding the work done until here can have public access to the tools they need.

Chapter 6

Conclusion

The work presented here consisted of a study on how stock options can be exploited to improve returns and better manage risk in the long term for a given portfolio of stocks from companies with good growth perspectives. To do so, a broad research on some of the main concepts of the stock market was made. Then, we developed option-based strategies that could reach our needs. In this context, we decided to only use variations of the covered call strategy, since, with this approach, stock options can be traded while we hold the respective stocks. Namely, these variations were, first, what called traditional approaches: sell calls with fixed strike margins 4.2.1, use the implied and historical volatility ratio to choose calls 4.2.1, use the log normal distribution of stock returns to find a safe strike margin 4.2.1, calculate the last 100 days moving average of the stocks in order to know when to sell calls 4.2.1. Subsequently, we developed Machine Learning based strategies that aimed at predicting the stock price range one month in the future so that we could decide on which strike margin to sell calls. The models chosen were the Logistic Regression and the Random Forest classifier 4.2.2. Also, we established as a benchmark the standard Buy and Hold approach 4.2.1 where we buy a stock and hold it hoping the associated company will thrive.

The strategies analyzed here were simulated from January to December, 2022. During this period, the best result was generated by one of the simplest strategies which was the one using the moving average indicator. The fixed strike margin procedure also performed well when considering the non-positive margins, whereas the volatility-based strategy seemed to improve returns, but it turned out to be due mainly to randomness and the fact the covered call model tends to increase profit in periods of fall. Additionally, the log normal approach exhibited the worst performance. Now, regarding the Machine Learning strategies, we expected better results, however, this was not the case. Due to the complexity of the problem of, basically, predicting stock prices, both Logistic Regression and Random Forest were not able to express satisfactory accuracies. A proof of this was that a completely random guessing approach met the performances of the Machine Learning models.

We should note as well that, even though some of the covered call strategies performed reasonably well against the Buy and Hold approach, it is really hard to make them outperform investments based on the risk-free interest rates in Brazil. For instance, in Brazil, one of these risk free rates is the Selic rate, which reached over 12% of interest per annum in 2022. That being said, the best strategy which was the one using the moving average was only able to match this more secure type of investment.

With this in mind, we arrived at the conclusion that it is possible to employ stock options, precisely call contracts, to aid the management of stock portfolios. Confirming this, with the strategies proposed, we observed that 1) during fall periods, the losses were softened, 2) during periods of little price change, the return was given by the option premiums simulating a fixed income and 3) during periods where the prices rose, a greater return was exchanged for a smaller, but more regular profit. Nevertheless, the results presented here should be taken carefully and are not intended to be used as the main source of income, but as a source of assistance in decision making. Also, we verified the inefficiency of using simple Machine Learning models to predict stock prices.

Finally, since this work was conducted taking into consideration only the year of 2022, the results may be completely different for other periods with different price behaviors, especially in the presence of exceptional events such as the COVID-19 pandemic in 2020. Furthermore, since only simple Machine Learning models were applied, little can be implied to the results more sophisticated models would produce. In fact, this opens opportunities for future work. Besides, there is a vast diversity of other strategies that not only make use of call contracts, but put contracts which were not explored in this study. Therefore, the current work can also be expanded in this regard.

Appendices

Appendix A

Detailed Results

In chapter 4, we presented the main results and implications of the several strategies analyzed in this study. Now, in this appendix, we will present in details the simulation results achieved in each group of strategies for each stock. Therefore, these strategies are, namely, fixed strike margin strategy 4.2.1, volatility margin strategy 4.2.1, log normal strategy 4.2.1, custom moving average strategy 4.2.1, logistic regression strategy 4.2.2, random forest strategy 4.2.2 and random guessing strategy 4.2.2. The table below shows the average improvement over Buy and Hold for each strategy. Moreover, the graphs shown in the following sections contain all the import information we need to know about the behavior of the simulations such as number of strikes and strike margins chosen by the strategies.

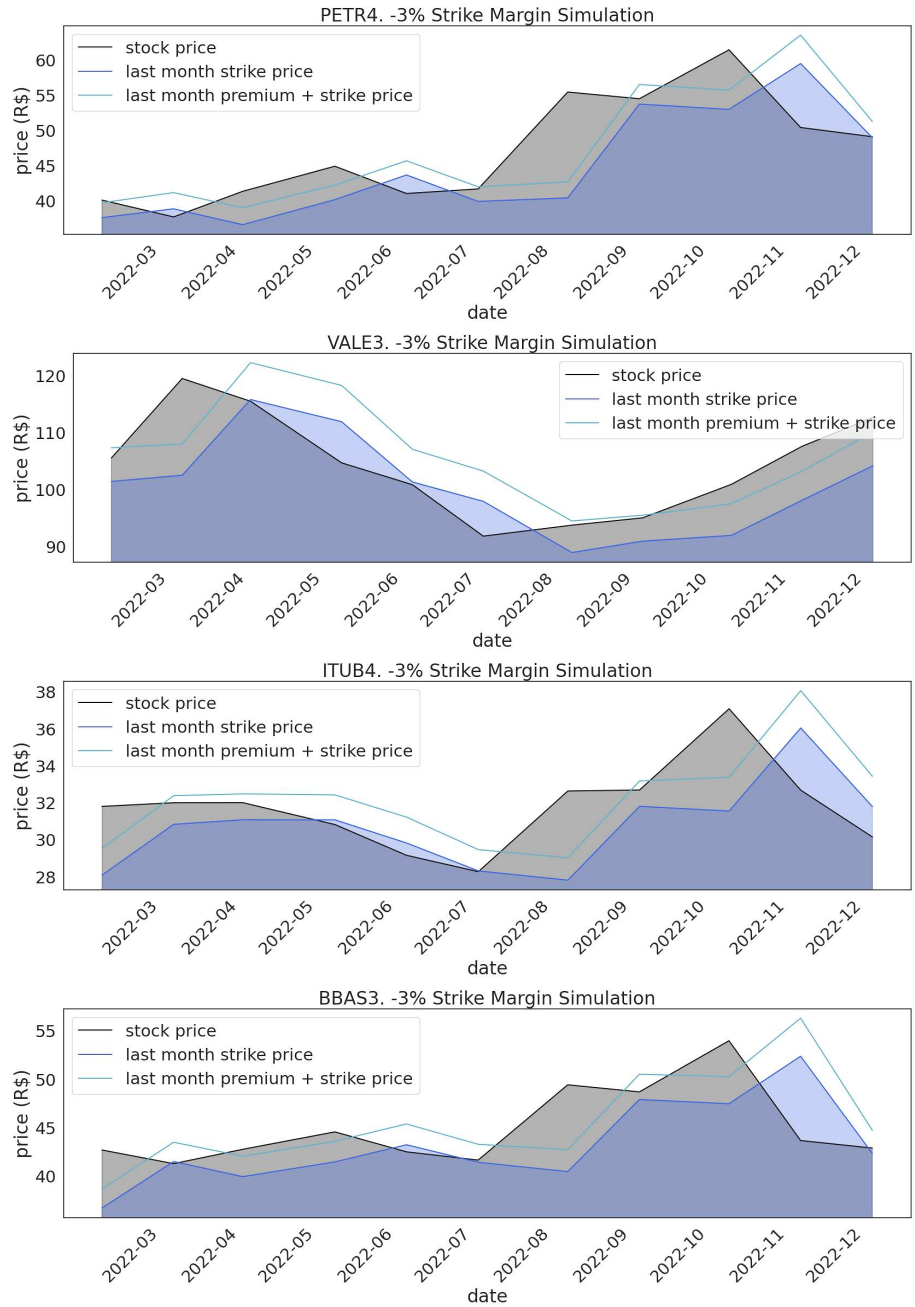
Table A.1: All relative improvements towards the Buy and Hold strategy in percentage.

Strategy	AVG	STD
strike margin -7%	1.82	13.09
strike margin -5%	1.91	13.45
strike margin -3%	2.52	13.03
strike margin 0%	0.1	12.19
strike margin 3%	-1.73	10.77
strike margin 5%	-3.19	9.83
strike margin 7%	-3.62	8.57
vol margin 10%	1.46	11.33
vol margin 30%	-0.41	15.42
vol margin 50%	3.65	17.19
vol margin 100%	3.99	15.50
vol margin 150%	3.99	15.50
log normal 10%	-4.2	7.6
log normal 10%	-3.88	10.13
log normal 50%	0.67	11.66
custom ma	**6.0**	**7.42**
logistic regression	0.85	11.77
random forest	2.04	9.88
random guessing	1.52	13.56

Source: created by the author.

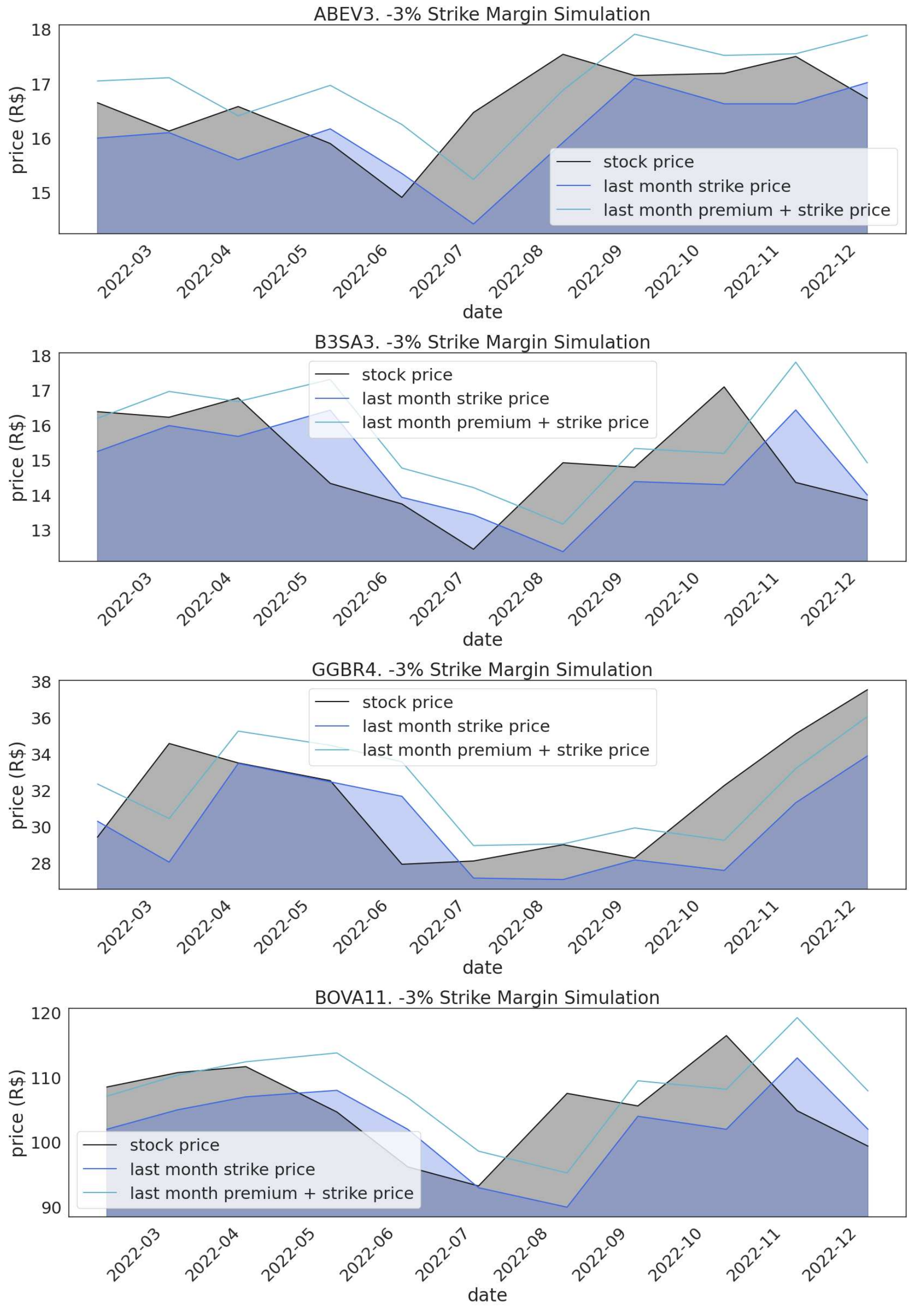
A.1 -3% Fixed Margin Results

Figure A.1: -3% fixed strike margin results.



Source: created by the author.

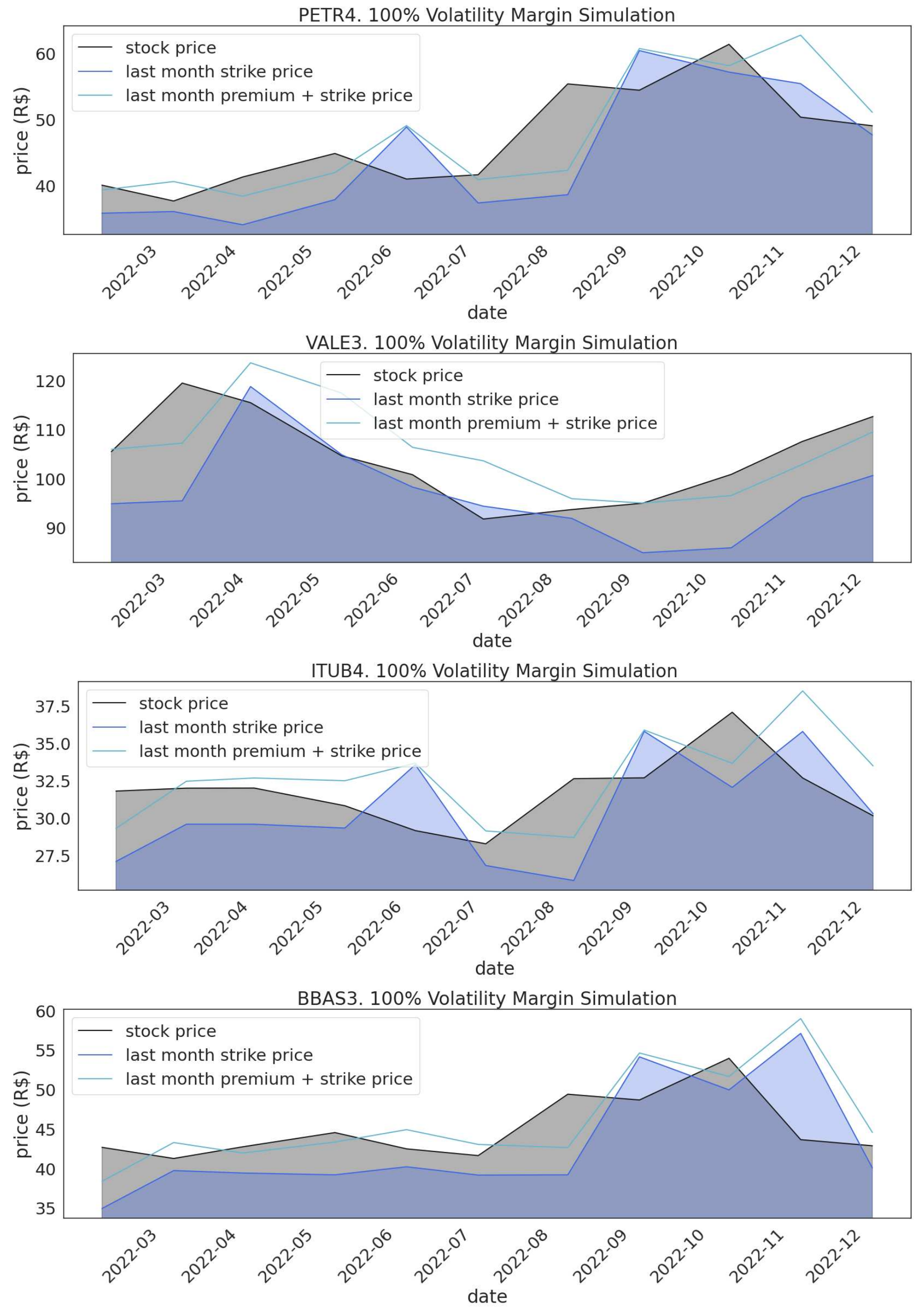
Figure A.1: -3% fixed strike margin results.



Source: created by the author.

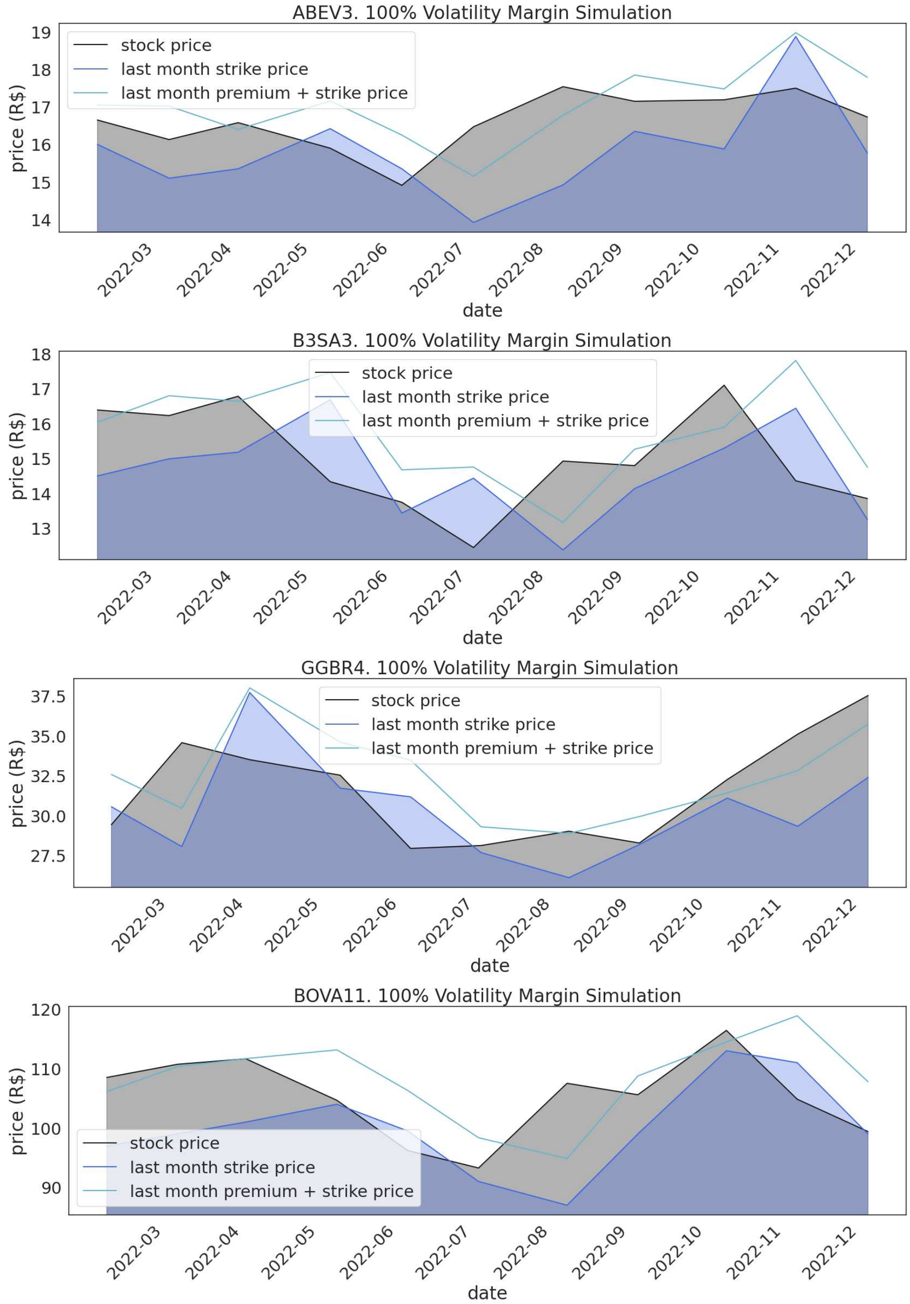
A.2 100% Volatility Margin Results

Figure A.2: 100% fixed volatility margin results.



Source: created by the author.

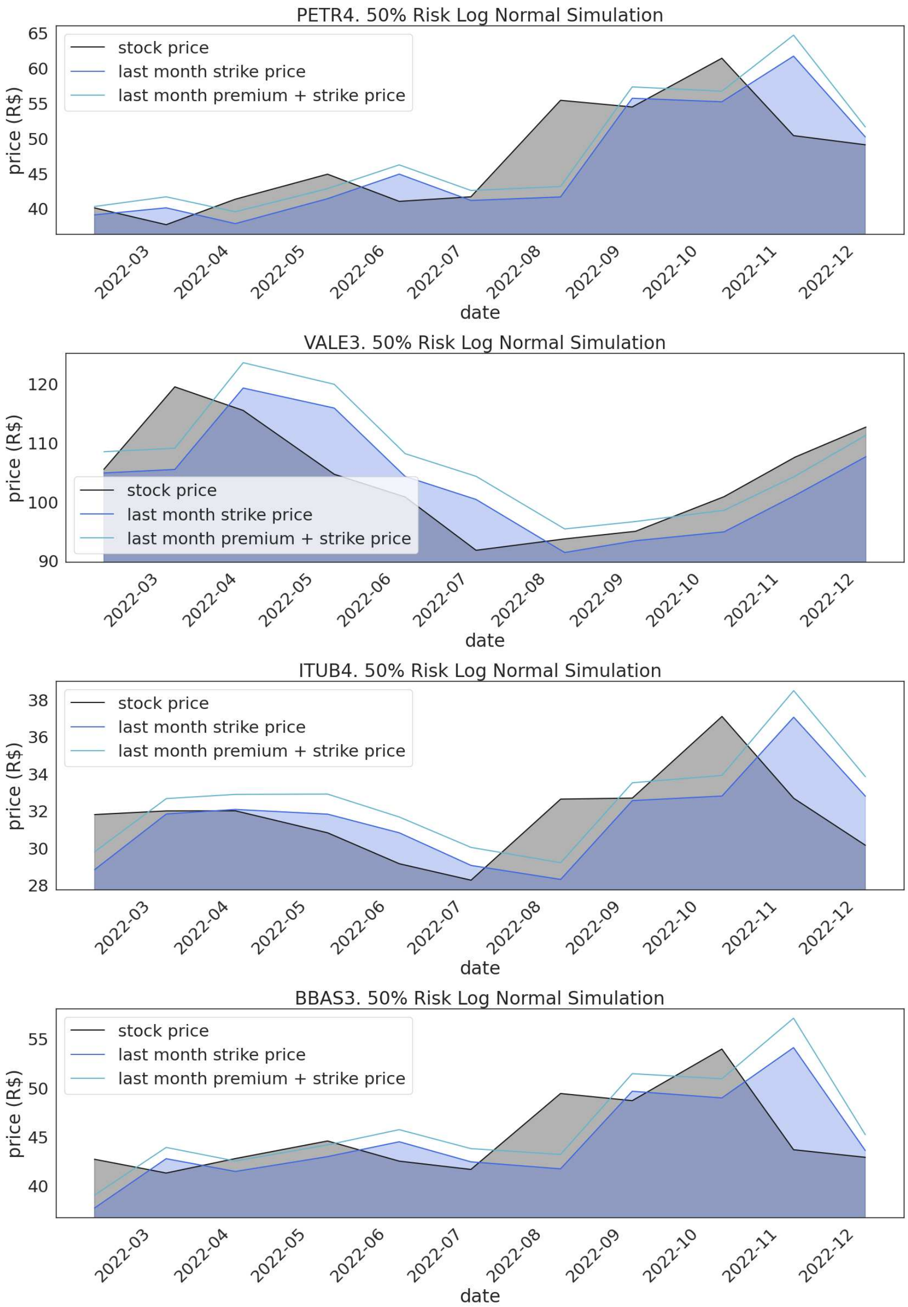
Figure A.2: 100% fixed volatility margin results.



Source: created by the author.

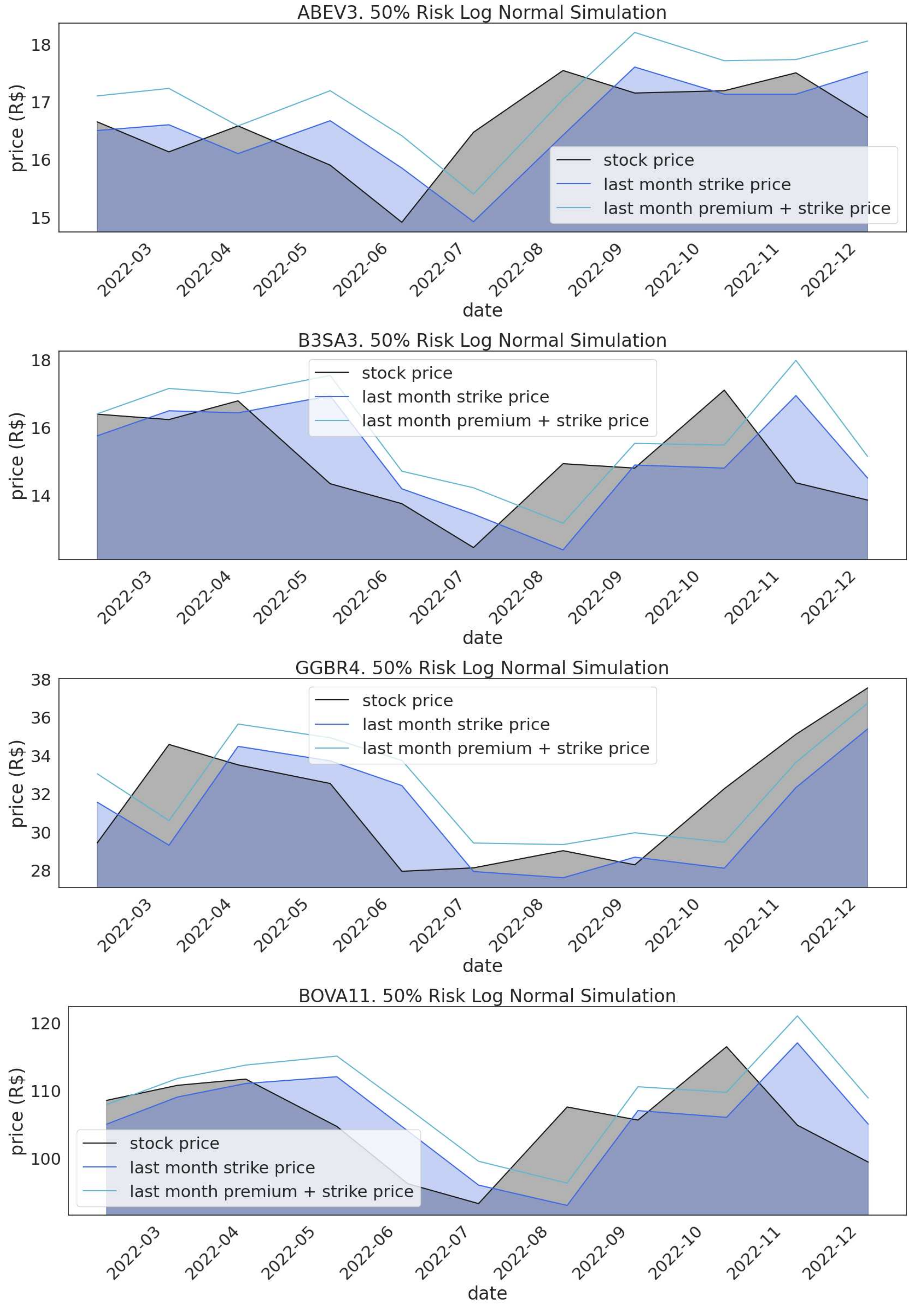
A.3 Log Normal with 50% Strike Risk Results

Figure A.3: 50% strike probability log normal results.



Source: created by the author.

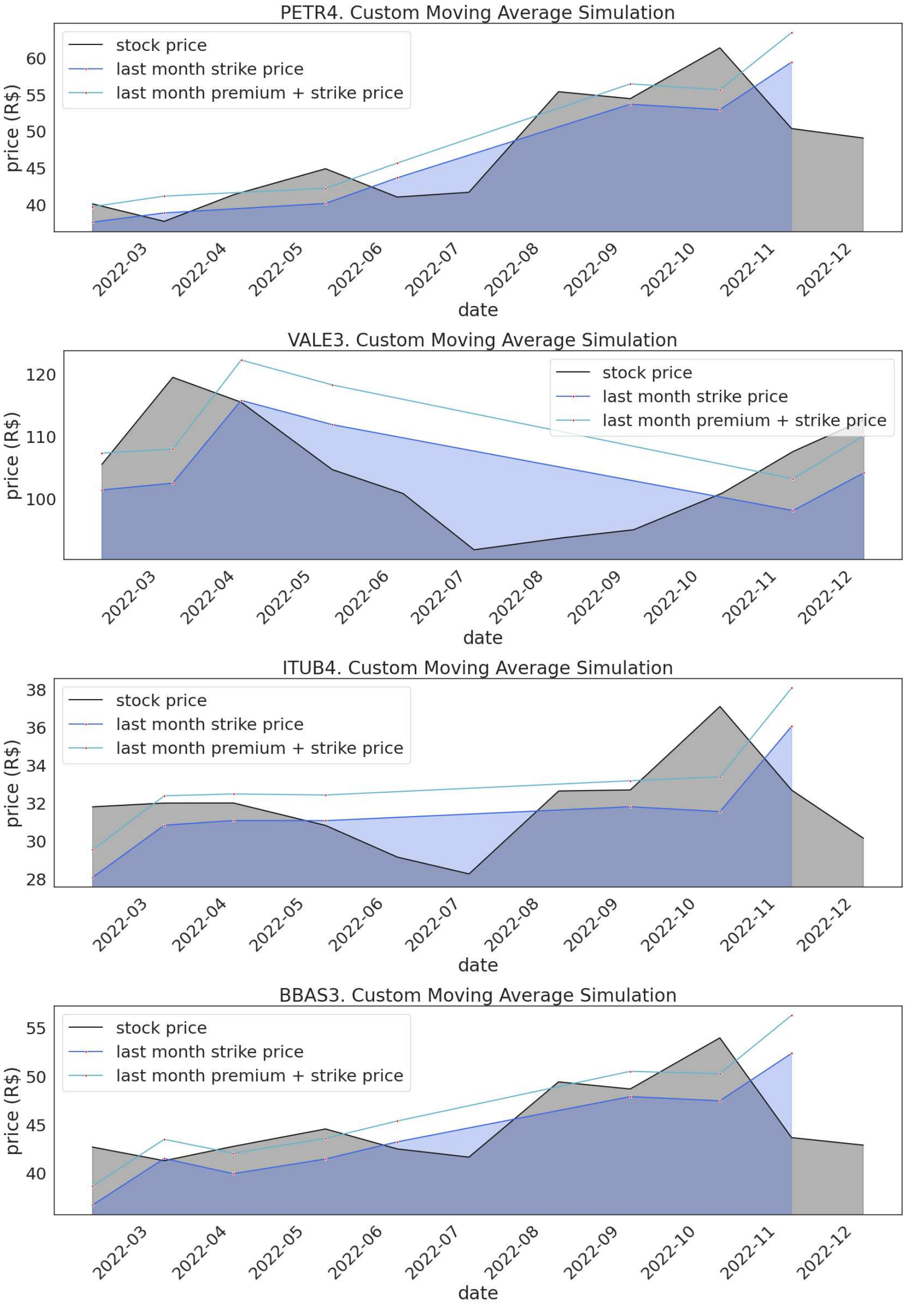
Figure A.3: 50% strike probability log normal results.



Source: created by the author.

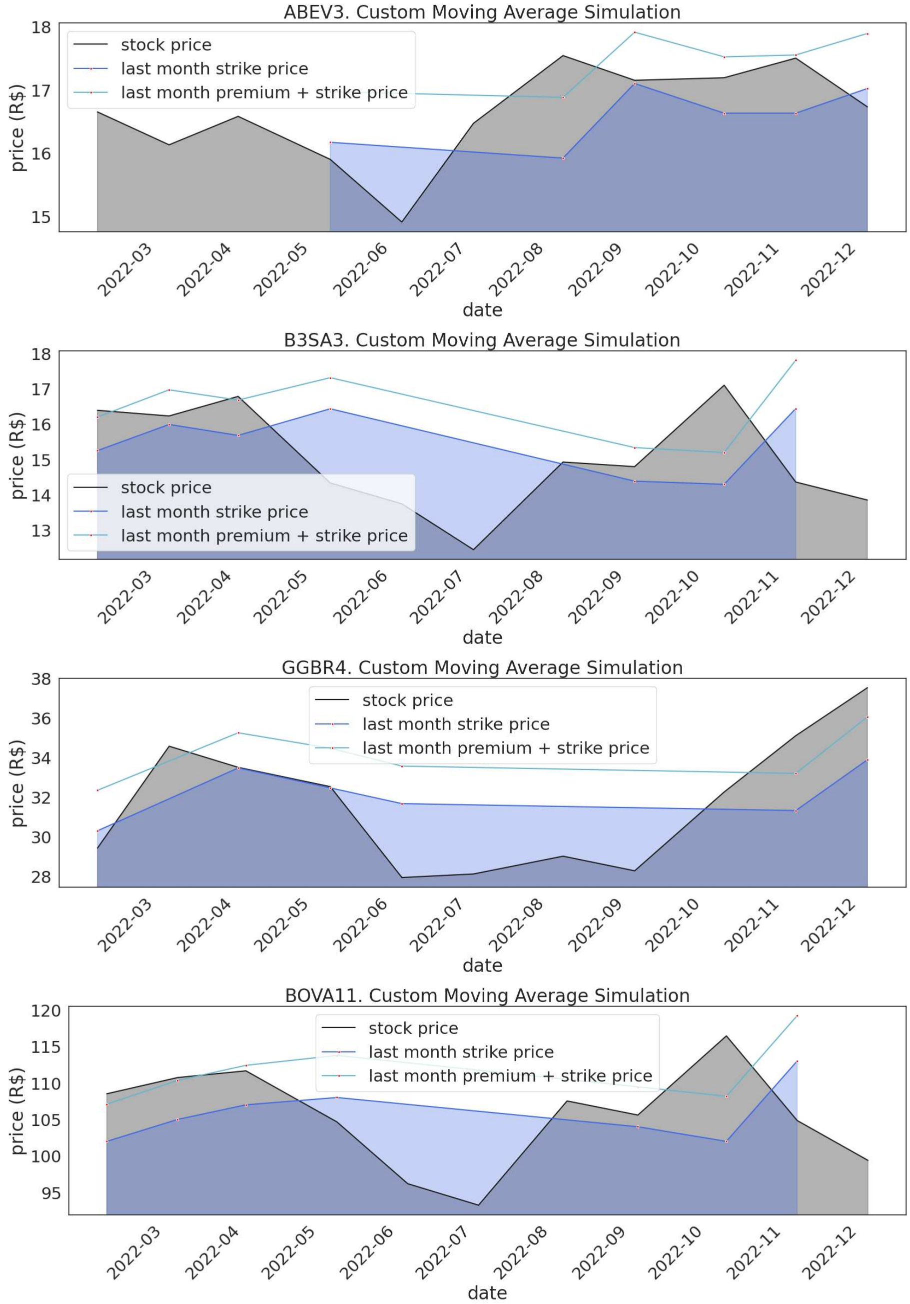
A.4 Custom Moving Average Results

Figure A.4: Custom moving average results. Notice that call options were not sold for some months.



Source: created by the author.

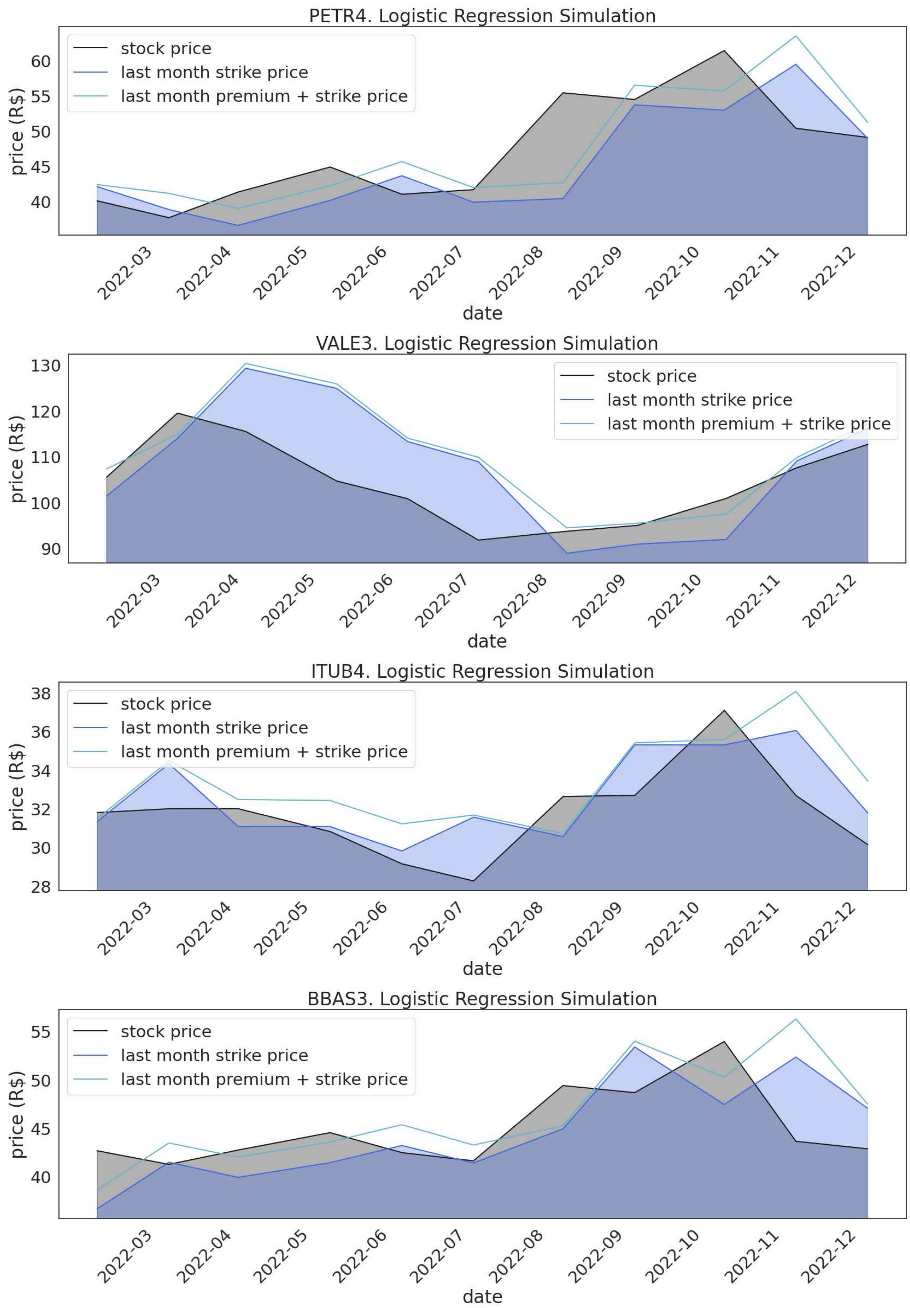
Figure A.4: Custom moving average results. Notice that call options were not sold for some months.



Source: created by the author.

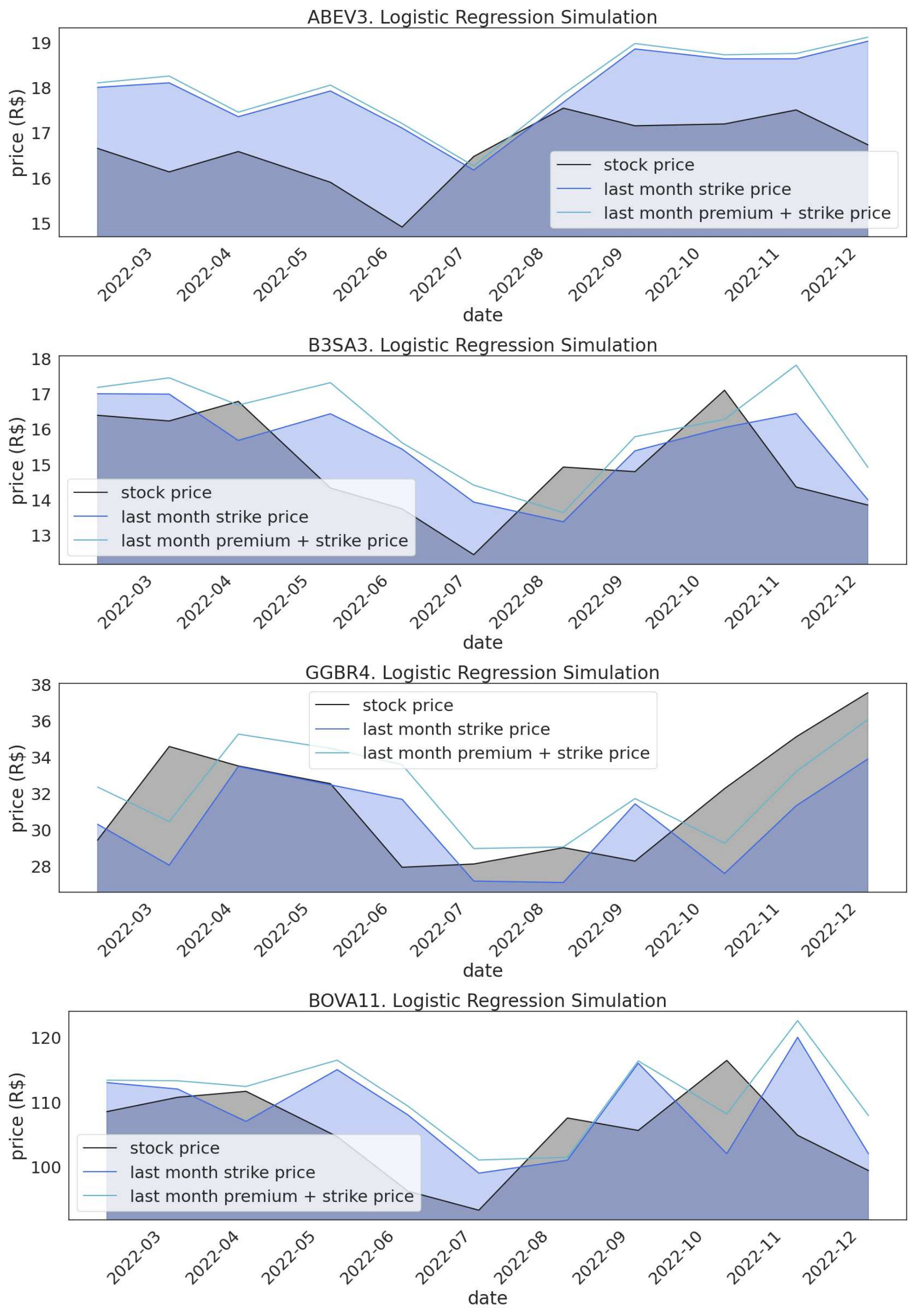
A.5 Logistic Regression Results

Figure A.5: Logistic regression results.



Source: created by the author.

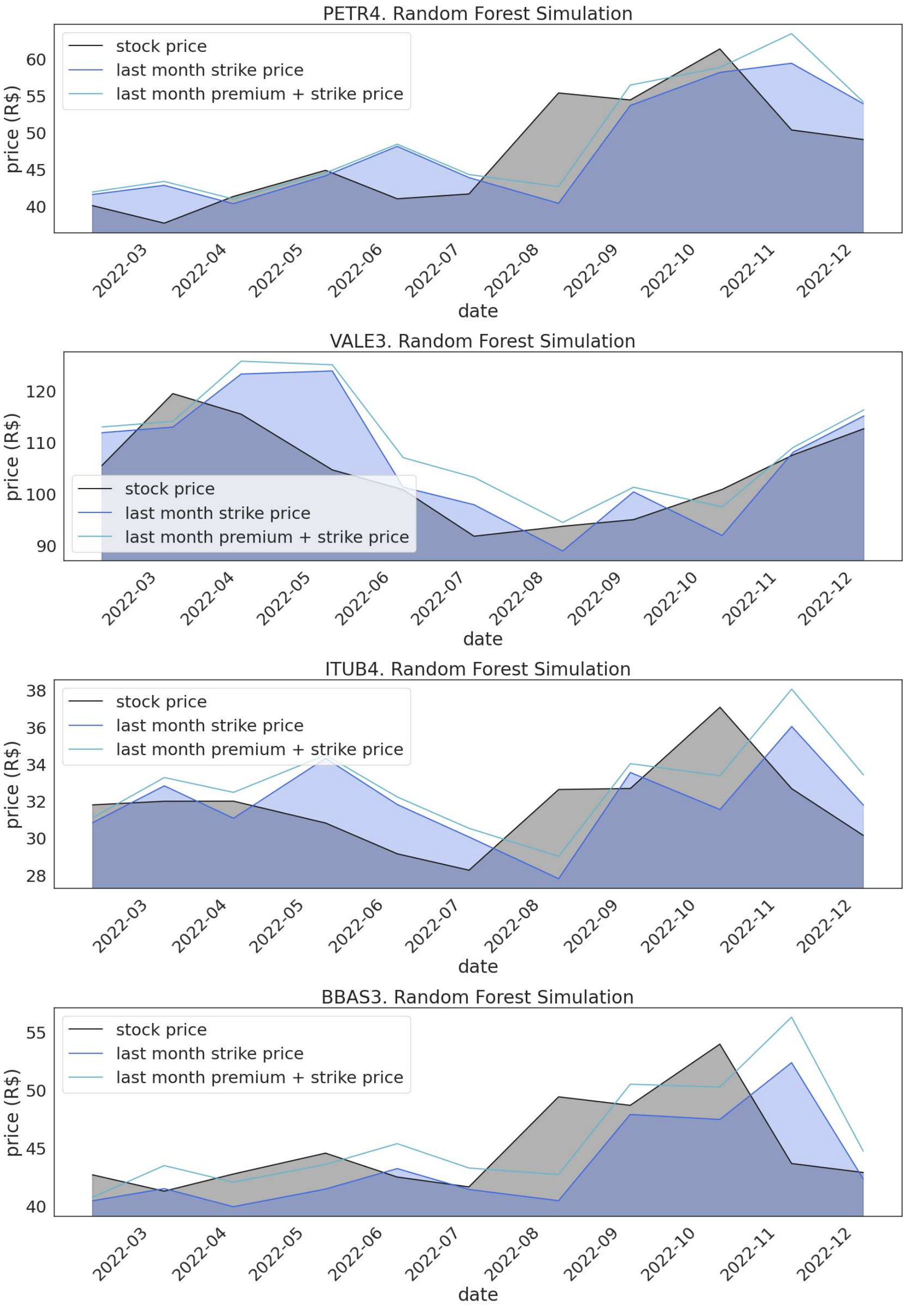
Figure A.5: Logistic regression results.



Source: created by the author.

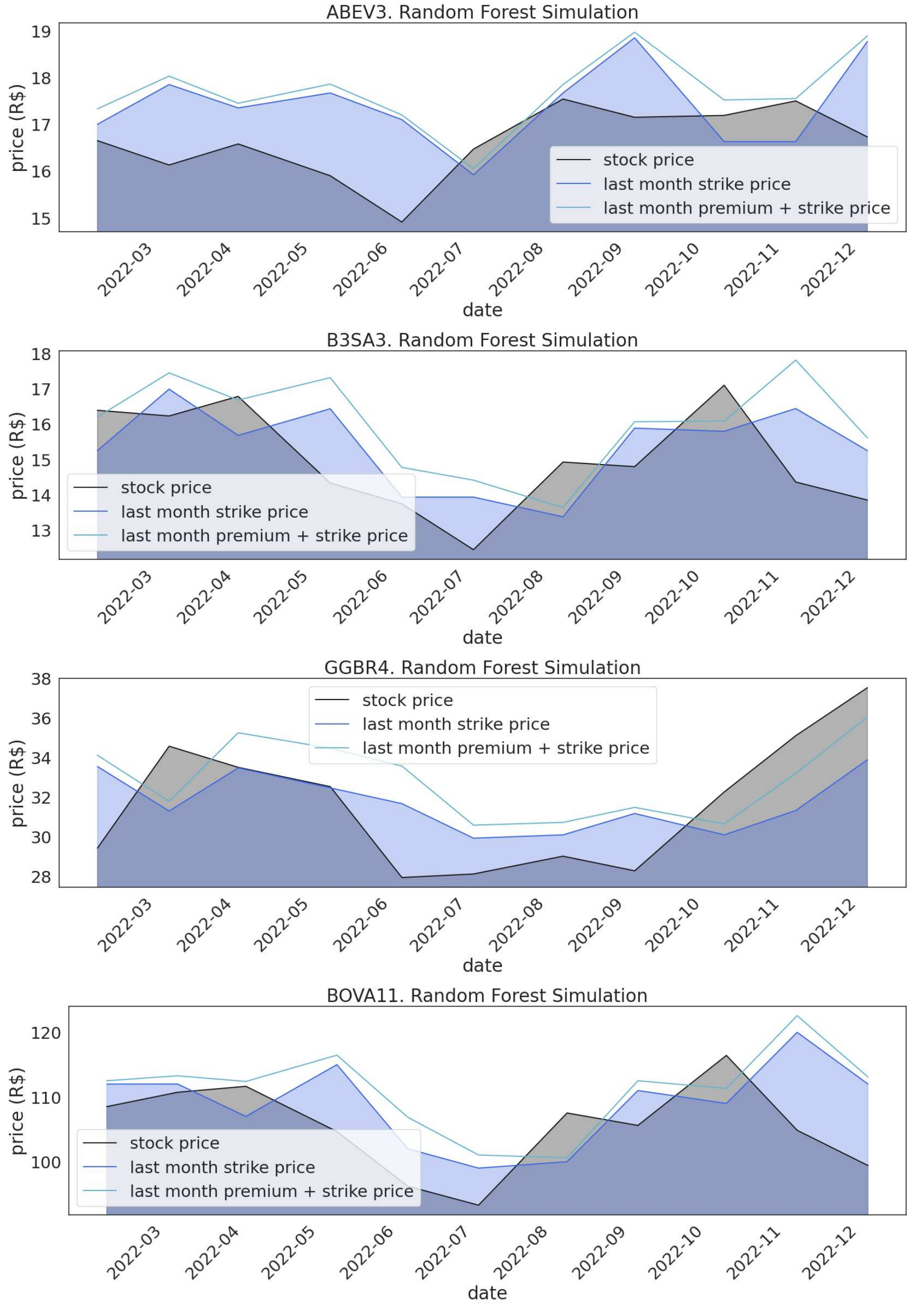
A.6 Random Forest Results

Figure A.6: Random forest results.



Source: created by the author.

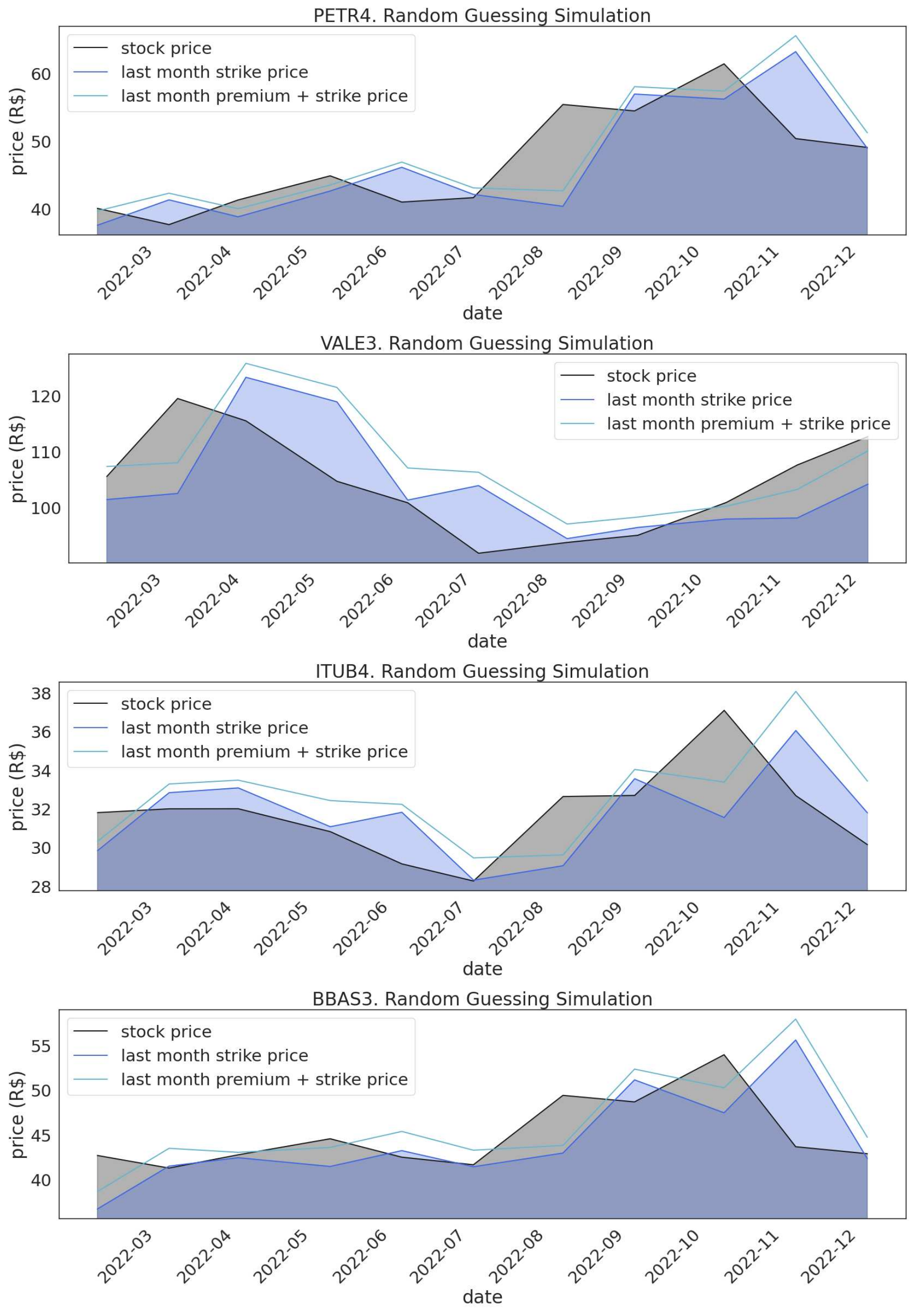
Figure A.6: Random forest results.



Source: created by the author.

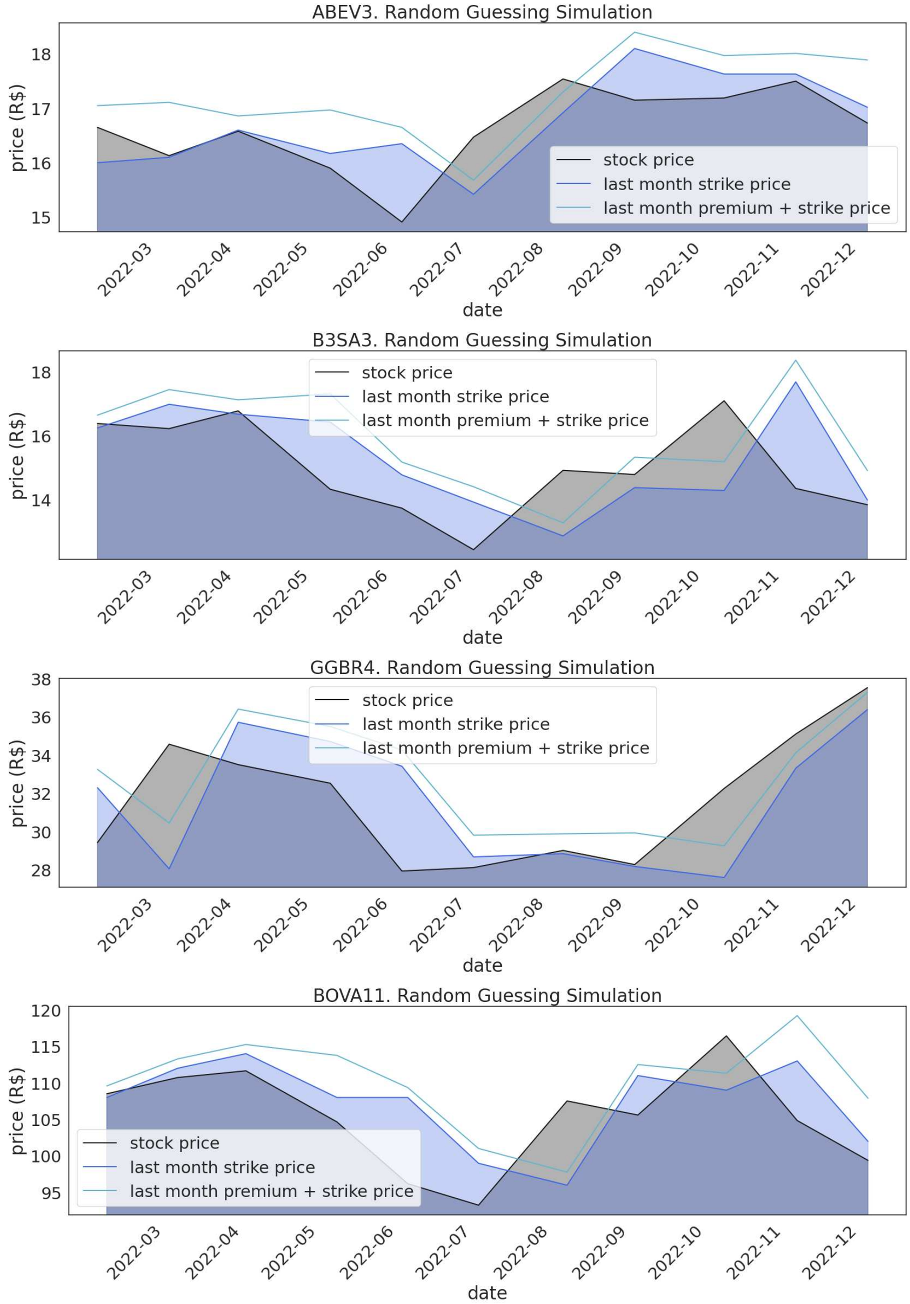
A.7 Random Guessing Results

Figure A.7: Random guessing results.



Source: created by the author.

Figure A.7: Random guessing results.



Source: created by the author.

References

Teixeira Zavadzki de Pauli, S., Kleina, M. & Bonat, W.H. Comparing Artificial Neural Network Architectures for Brazilian Stock Market Prediction. Ann. Data. Sci. 7, 613–628 (2020). https://doi.org/10.1007/s40745-020-00305-w

BLACK, F.; SCHOLES, M. The Pricing of Options and Corporate Liabilities. Journal of Political Economy, v. 81, n. 3, p. 637–654, May 1973.

Tailab, Mohamed. (2016). Fundamentals of Stock Options: With Selected Solutions.

Kanazawa, Marcelo Notomi. Venda coberta de opção como estratégia de gestão para clube de investimentos. 2011. TCC (graduação) - Universidade Federal de Santa Catarina, Centro Sócio Econômico, Curso de Economia.

Pereira, Gabriel Franco & Ventura De Biase, Pietrangelo. Avaliação de Estratégias no Mercado Brasileiro de Opções de Ações. 2012. TCC (graduação) - Escola Politécnica, Universidade Federal do Rio de Janeiro, Curso de Engenharia de Produção.

AURÉLIEN GÉRON. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. [s.l.] "O'Reilly Media, Inc.", 2019.

A Bolsa do Brasil — B3. Available at: https://www.b3.com.br/en_us/. Accessed: 08/06/2023.

BLYSTONE, D. Is Day Trading Profitable? How To Get Started. Available at: https://www.investopedia.com/articles/active-trading/053115/average-rate-return-day-traders.asp-day-traders.asp . Accessed: 08/06/2023.

Juros Sobre Capital Próprio (JCP): o que são e como lucrar. Available at: https://www.infomoney.com.br/guias/juros-sobre-capital-proprio-jcp/. Accessed: 08/06/2023.

John. CodeArmo. Available at: https://www.codearmo.com/blog/implied-volatility-european-call-python. Accessed: 09/06/2023.

Banco Central do Brasil. Available at: https://www.bcb.gov.br/controleinflacao/historicotaxasjuros. Accessed: 09/06/2023.

PANDAS. Python Data Analysis Library — pandas: Python Data Analysis Library. Available at: https://pandas.pydata.org/. Accessed: 09/06/2023.

MetaTrader 5, plataforma de negociação para Ações e Forex. Available at: https://www.metatrader5.com/pt. Accessed: 07/07/2023.