Investment Strategy for most Volatile Signals in the Forex Market

Pedro de Araújo Correia Neto Lopes,
pedro.neto.lopes@técnico.ulisboa.pt,
Instituto Superior Técnico

Abstract—This thesis proposes an adaptive system based on Genetic Algorithms to establish Long positions in the Forex Market at more volatile periods. In this work, the Genetic Algorithm is responsible for optimizing certain parameters, resulting from investment rules based on Technical indicators. Regarding the Technical indicators used, some of these indicators had already been created by other authors while others were developed during the thesis, such as the Size indicator and the Percentage indicator. In this thesis, the Nervous Periods indicator was also created, which aims to limit the periods where transactions can be made in the market, which in this case was intended to be periods where there was great volatility and at the same time it was possible to identify a trend. The experiments were purposely performed on three very volatile currency pairs, USD/ZAR, USD/MXN and USD/TRY, and evaluated according to the Cumulative Return, Cumulative Return per minute and Drawdown evaluation metrics. Data available from 2012 until 2018 for the USD/ZAR, USD/MXN, USD/TRY currency pairs were used to train the Genetic Algorithm as well as to test it. The results were quite promising, having been achieved for the currency pair USD/TRY, a 57.60% Cumulative Return at around 26% of the test period, for the currency pair USD/ZAR, a 13.53% Cumulative Return with near 27% of the test period, and in the case of the USD/MXN currency pair, a Cumulative Return of 17.25% is obtained at about 28% of the test period.

Index Terms—Forex Market, Technical analysis, Technical indicators, Genetic Algorithm, Trading rules.

I. INTRODUCTION

Nowadays, the financial market is quite complex, as it is dependent on many factors that interact with the financial market, such as political events, economic conditions and sometimes even the individual expectations of each investor, thus making the prediction of the price movements quite difficult [1]. However, the financial sector has changed dramatically over time, due to the availability of a large amount of data, and with the increase in computing power and the spread of various Machine Learning (ML) techniques [2], which allows the efficient use of a large data set.

In the Forex market, which is where this thesis focused, there is a great difficulty in manipulating the price as it is very complicated for banks and large funds to have price control, which makes this market more attractive. In addition, the fact that this market is highly liquid and has a large amount of past price data available provides an interesting perspective for scientific and financial research [3]. However, despite this attractiveness for the Forex market, the price structure in this market is quite unstable, which complicates price forecasting, thus leading to the use of optimization methods [4] such as the Genetic Algorithm (GA) that was used in this thesis. This algorithm is the most used of the optimization methods and is based on evolution and genetics concepts, being responsible for generating solutions that evolve over time. Such algorithms can obtain near-optimal solutions if the investor has a reasonable level of computation [4], which is now increasingly available to any investor. This thesis aims to create an adaptive investment strategy capable of establishing Long positions in the Forex market in more volatile currency pairs, such as USD/TRY, USD/MXN and USD/ZAR, to maximize profit and at the same time minimize the risk. The success of this strategy depends on choosing the best periods to be in the market, something that had not yet been done by other authors of other articles and therefore, the main objective of this thesis is to try to limit the periods where the strategy can be in the market, which periods have to be very volatile, and at the same time it has to be possible to spot a trend.

Once these periods have been identified, it is still necessary to decide when to establish a Long position on the asset, and for this purpose, it is necessary to create several investment rules from the Technical indicators or others indicators created.

Finally, the last objective is to exploit the ability of the GA to optimize all investment rules, as well as the weights associated with those same investment rules in the Forex market.

The main contributions made by this thesis were as follows:

• Construction of a system using the GA, able to adapt to the market.
• Signal characterization to understand the best periods to be in the market by creating the Nervous Periods indicator.
• Creation of some Technical indicators and some investment rules related to these indicators (Size indicator and Percentage indicator).
• Conjugation of investment rules that are based on Technical indicators in which information is kept up to the minute with investment rules associated with the created indicators whose information in these indicators is not up to the minute, to buy assets.

This paper is organized as follows: Section II addresses some works that are inside the market inherent in this thesis or using Genetic Algorithms in their works. Section III presents the architecture of the system created with an explanation of each layer of the system. Section IV presents the case studies and the analysis of the results. Section V summarizes the paper’s content and supplies its conclusion and limitations.

II. RELATED WORK

This subchapter consists of two subchapters, which are:

• Works in the Forex Market—In this subchapter, a detailed explanation of some works regarding future price forecasting in the Forex market is presented exclusively through various distinct techniques such as Genetic Algorithms, and Support Vector Machines.
• Works with Genetic Algorithms—As for in this subchapter, the interest focuses on the works performed in financial markets (except for foreign exchange), which resort to the GA.

A. Works in the Forex Market

There is a lot of work done using various ML techniques to try to predict financial markets, in particular, in the Forex market.
using a wide variety of Technical indicators.

A. Hirabayashi in [5] proposed a method of optimizing certain rules for buying or selling currency pairs in the Forex market through a GA. This optimization was made using past historical data and applying them to a set of technical indicators. The goal was for the GA to be able to find the most profitable rule set for a given period, and then apply that same rule set to future data compared to the period where the highest-profit rule set was found. Besides, they also added a model that allowed them to decide the leverage ratio to use, which depended on a correlation coefficient. Finally, they added a buy-and-sell model, which consisted of when a certain level of profit was reached from a buy or sell respectively, were bought or sold more assets, with the prospect that the asset would continue to rise if it were in a Long position or continued to decline in the case of a Short position, thus allowing for higher profits while managing risk. As for the best results, they were able to achieve a profit rate of 80% at the end of the full test period on the EUR/JPY currency pair.

In [6], an investment solution within the Forex market, in the EUR/USD currency pair, was proposed from investment rules based on Technical indicators and historical past price data. This solution was developed by combining a Support Vector Machine (SVM), used to identify and classify the market into three market types, up trend, down trend or sideways, with a dynamic Genetic Algorithm that optimized investment rules depending on the identified market by SVM. In other words, depending on the classification that SVM assigned to the market, the GA trained in different types of markets. This work also used leverage such as [5], however, a different strategy was applied to decide when to use leverage and what proportion of leverage to use. Their use depended on whether the investment rules and SVM were in agreement. This agreement was obtained when the SVM rated the market as an up trend and the investment rules decided to buy the asset or in case the SVM rated the market as down trend and the investment rules decided to sell the asset, as well as price fluctuations, identified through two Simple Moving Average (SMA) with different periods. Regarding the ratio used in leverage, that was decided based on the values optimized by the GA. As for the results obtained in this article, these were very promising, reaching a Return on Investment (ROI) of 83.5% in the test period from January 2, 2015, to March 2, 2016.

M. Ozturk in [4] proposed a system of buying and selling currency pairs in the Forex market, using various types of investment rules based on a diverse set of Technical indicators. This system was divided into several phases. Initially, the best parameters for each investment rule were chosen individually, based on NET Profit, through GA. Subsequently, based on a set of evaluation metrics, it was decided which rules to discard. Finally, at a later stage, it was decided through the GA, what was the best combination of rules that were not been previously discarded, based on NET Profit and Average Profit per trade (APPT), and it was also decided what weight to give to each investment rule. These researchers also used a greedy search heuristic to select the best combination of rules. To test and compare the performance of the developed system, the rules with the respective weights resulting from both GA and greedy search heuristic were used in the EUR/USD and GBP/USD currency pair test at three different time intervals. Regarding the results obtained when the selection method was the GA, it was concluded that the choice of the best combination of rules using the NET Profit evaluation metric instead of the APPT metric had better results in the 2 out of 3 tests. In this paper, they also demonstrated that using individual investment rules, as well as using the greedy search heuristic to select the best combination of rules, resulted in negative profits while using GA to combine investment rules resulted in positive gains in all experiments performed.

Finally, B.J. de Almeida presented in [7] a method whose objective was to maximize profit in the 9 most traded currency pairs in the foreign exchange market by optimizing the parameters of certain Technical indicators during the period of international economic crisis. In this article, the authors created 15 buy and sell rules based on 4 optimized Technical indicators. In order to evaluate the performance of the implemented GA model, this method was compared with a hybrid system that included a Support Vector Regression (SVR) and a Growing hierarchical self-organizing maps (GHSOM), both in terms of profit using the ROI metric and risk using the Maximum Drawdown (MDD) metric. It was then possible to conclude from the results demonstrated in this article, that the implemented GA outperformed the hybrid system mentioned above in the test period, which included the period of international economic crisis, reaching a ROI of 143.78% and a Drawdown of 39.28%.

B. Works with Genetic Algorithms

In this subchapter, the focus was on works where the GA was used, whose algorithm was central in the development of the thesis, but outside the Forex market, since some works had already been presented in the Forex market, which used the GA in subchapter A of this chapter.

M. Radeerom in [8], developed a stock market system based on investment rules, which were the result of one or more Technical indicators. This system included a GA, which was responsible for the optimization of investment rules, to maximize the profit generated. Before the system was put to the test, a set of shares were first eliminated, which had a negative price to earning (P/E) or negative equity value of the shareholder, thus preventing the system from investing in unacceptable shares for the shareholder. This system was tested on stock indices at the Stock Exchange of Thailand (SET) index. The developed system was also compared with the Buy and Hold (Buy&Hold) strategy, being possible to conclude that this system allowed to achieve higher profits, with more stable results, having reached 11.7% of Cumulative Return during the test period.

In [9], a hybrid system in futures markets, Korea Composite Stock Price Index 200 (KOSPI 200), whose purpose was to discover rules that enabled the investor to understand when to buy and sell a given stock. These rules were based on Technical indicators and a GA was used to optimize them. To evaluate the system’s performance using the GA, tests were made where the number of investment rules and the size of training period were modified. It was also tested the influence that the use of validation data had on the developed system. All of these experiments were then evaluated based on both profit and risk using the Annualized Return, Standard deviation, MDD, and Sharpe ratio metrics. In this paper, it was found that the highest Annualized Return, around 10.59, was achieved when 50 investment rules were used, with a 6 month training period, implying that the combination of an appropriate number of investment rules as well as the length of the training period, had an influence on the performance of the developed system.

Finally, in [10], a system was developed, which the purpose was to manage a financial portfolio through a GA and Technical indicators. To combine all the Technical indicators in the final decision to establish a Long or a Short position within the portfolio, a subsystem was also created, which was responsible for normalizing each indicator in 4 different levels, according to certain rules defined by the author, and assign to each indicator a
This subsystem was very interesting because it allows the combination of various Technical indicators in the final decision to buy or sell a given asset, and so a weighting system based on the one developed by [10] was used. To validate the developed system, it was tested from 2003 to 2009, which included a major financial crash in 2008, and was then compared to various strategies such as Buy&Hold and the random strategy. In terms of the ROI assessment metric, this developed system was able to reach a value of 62.95% over the 6 years tested.

III. SYSTEM ARCHITECTURE

The developed system is divided into four essential layers, the data file module, the Technical indicators and the choice of Nervous Periods module, the Genetic Algorithm module and, finally, the validation of the work performed by the Genetic Algorithm.

In Figure 1, the various layers of the developed solution are represented, as well as the interaction that each layer had with each other, represented by the arrows.

A. Data file Module

In this first layer, minute data from the currency pairs USD/TRY, USD/MXN and USD/ZAR have been taken from 1 January 2012 until 8 January 2019 for USD/TRY, and from January 1, 2012 through April 24, 2019, for USD/MXN and USD/ZAR currency pairs, via the Dukascopy platform. This data contains information about the opening price, the highest price, the lowest price, the closing price, and the volume in each minute, as shown in Table 1.

Table 1 - Dukascopy minute data from USD/MXN currency pair.

<table>
<thead>
<tr>
<th>Date</th>
<th>Close</th>
<th>Open</th>
<th>High</th>
<th>Low</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>01/01/12 22:00</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>0</td>
</tr>
<tr>
<td>01/01/12 22:01</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>0</td>
</tr>
<tr>
<td>01/01/12 22:02</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>0</td>
</tr>
<tr>
<td>01/01/12 22:03</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>0</td>
</tr>
<tr>
<td>01/01/12 22:04</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>0</td>
</tr>
<tr>
<td>01/01/12 22:05</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>0</td>
</tr>
<tr>
<td>01/01/12 22:06</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>0</td>
</tr>
<tr>
<td>01/01/12 22:07</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>0</td>
</tr>
<tr>
<td>01/01/12 22:08</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>0</td>
</tr>
</tbody>
</table>

Then, for each currency pair, data was created which, instead of being separated by a minute, as stated above data, was separated by a change in the closing price as shown in Table 2, which was entitled as Size candles.

Table 2 - Data separated by 0.0892 variations in closing price in USD/MXN currency pair (Size candles).

<table>
<thead>
<tr>
<th>Date</th>
<th>Close</th>
<th>Open</th>
<th>High</th>
<th>Low</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>01/01/12 22:00</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>0</td>
</tr>
<tr>
<td>03/01/12 09:31</td>
<td>13.839</td>
<td>13.929</td>
<td>13.992</td>
<td>13.838</td>
<td>15341.5</td>
</tr>
<tr>
<td>03/01/12 13:43</td>
<td>13.748</td>
<td>13.839</td>
<td>13.840</td>
<td>13.717</td>
<td>15187.3</td>
</tr>
<tr>
<td>04/01/12 17:55</td>
<td>13.658</td>
<td>13.749</td>
<td>13.759</td>
<td>13.658</td>
<td>31026.8</td>
</tr>
<tr>
<td>05/01/12 10:47</td>
<td>13.748</td>
<td>13.658</td>
<td>13.749</td>
<td>13.648</td>
<td>30572.6</td>
</tr>
<tr>
<td>10/01/12 03:55</td>
<td>13.659</td>
<td>13.748</td>
<td>13.794</td>
<td>13.647</td>
<td>219766.0</td>
</tr>
<tr>
<td>10/01/12 12:32</td>
<td>13.569</td>
<td>13.671</td>
<td>13.681</td>
<td>13.569</td>
<td>15782.0</td>
</tr>
</tbody>
</table>

It is also important to note that the closing price variation over the previous minute was obtained every minute. This variation was then added to the variation obtained in the previous minutes. If the sum of the variations in each minute reached a value higher than the desired value, or in other words, if the difference between the maximum closing price and the minimum closing price in a given period was greater than a certain value, which in the case of Table 2 is 0.0892, a new row of data was created. Then, the value of the closing price where this variation had been reached was saved and the process described above was repeated.

In Figure 2 it is possible to see the Size candles created and presented in Table 2.

Figure 2 – Size candles, with variations of 0.0892 in closing price in USD/MXN currency pair.

Next, new data sets were created that were separated by percentage changes in closing price as shown in Table 3, which were called as Percentage candles.

Table 3 - Data separated by 0.1% change in closing price in USD/MXN currency pair (Percent candles).

<table>
<thead>
<tr>
<th>Date</th>
<th>Close</th>
<th>Open</th>
<th>High</th>
<th>Low</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>01/01/12 22:00</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>13.929</td>
<td>0</td>
</tr>
<tr>
<td>01/01/12 23:20</td>
<td>13.911</td>
<td>13.929</td>
<td>13.930</td>
<td>13.911</td>
<td>11.0</td>
</tr>
<tr>
<td>01/01/12 23:22</td>
<td>13.927</td>
<td>13.911</td>
<td>13.927</td>
<td>13.911</td>
<td>3.0</td>
</tr>
<tr>
<td>01/01/12 03:49</td>
<td>13.941</td>
<td>13.927</td>
<td>13.946</td>
<td>13.909</td>
<td>120.0</td>
</tr>
<tr>
<td>01/01/12 03:52</td>
<td>13.965</td>
<td>13.963</td>
<td>13.972</td>
<td>13.952</td>
<td>29.0</td>
</tr>
<tr>
<td>02/01/12 07:30</td>
<td>13.990</td>
<td>13.965</td>
<td>13.992</td>
<td>13.953</td>
<td>424.18</td>
</tr>
<tr>
<td>02/01/12 08:05</td>
<td>13.975</td>
<td>13.990</td>
<td>13.992</td>
<td>13.975</td>
<td>67.63</td>
</tr>
<tr>
<td>02/01/12 08:17</td>
<td>13.991</td>
<td>13.974</td>
<td>13.991</td>
<td>13.974</td>
<td>40.9</td>
</tr>
<tr>
<td>02/01/12 08:18</td>
<td>13.973</td>
<td>13.973</td>
<td>13.980</td>
<td>13.965</td>
<td>14.0</td>
</tr>
</tbody>
</table>

In this dataset, there had to be a certain percentage change (whether this change was positive or negative) of the closing price. At each minute, the change in percentage of the closing price was saved, and then this change was added to the change in the percentage of the closing price in the following minute. This
process was repeated until the sum had reached the desired closing price change percentage value, which in the case of Table 3 is 0.1%. When this value was reached a new row of data was created, repeating the whole process previously described.

In Figure 3 it is possible to see the Percentage candles, which are arranged in Table 3.

![Percentage candles, with 0.1% change in closing price, in USD/MXN currency pair.](image)

In this layer, 50 different Size candle data files were created for each currency pair, where the row separation of this data was done with different variations between closing prices, and 50 different Percentage candle data files, separated by different percentage changes in the closing price, also for each currency pair.

Regarding the Size candles, these were created with closing price variations from 0.00180 to 0.01 with increments of 0.00018 for USD/TRY, from 0.01 to 0.0982 with increments of 0.001180 for USD/MXN and for USD/ZAR, from 0.007 to 0.07854 with 0.00146 increments. The values of the closing price differences generated were different for each currency pair since each currency pair has a different quotation.

As for the Percentage candles, these were created with percentage price changes ranging from 0.05% to 2.5% with 0.05% increments for USD/MXN and USD/ZAR currency pairs, while for USD/TRY pair, were created with percentage changes in closing price, from 0.05% to 3% with 0.0602% increments.

**B. The Technical Indicators Module and the Choice of Nervous Periods**

To predict future asset price trends and to create investment strategies, investors often resort to fundamental and technical analysis, and in this case, only technical analysis will be used.

As for the technical indicators already created, there are many varieties, having been used as Trend indicators, SMA and Double Crossover (DCM), and as Momentum indicators, Relative Strength Index (RSI) and Rate of Change (ROC). Both Trend indicators and Momentum indicators used minute data, which are shown in Table 1. Two indicators were also created, one indicator that uses the Size candles, called the Size indicator, and another indicator that uses Percentage candles referred to as Percentage indicator. Also, since it was quite important to consider what type of market you were investing in, another indicator was created that fits the Technical Volatility indicators, called the Nervous Periods indicator. This indicator aimed to select the most volatile periods, which were called nervous periods, and at the same identify the main trend, something that was not possible with the Average True Range (ATR) indicator, which is only responsible for measuring market volatility, giving no guarantees regarding the recognition of a trend. The Nervous Periods indicator, unlike the above, was not used to establish a Long position or a position to stay out of the market, but to limit the periods in which such positions could be established. Then a specific investment rule was created for each indicator.

In order to combine all of these investment rules associated with all of the above indicators (with the exception of the Nervous Periods indicator), to decide when to establish a Long position or when to exit the position, it was necessary to normalize each investment rule and then a system that assigned a weight to each investment rule.

First, since this thesis was only focused on making Long positions, normalization consisted only of the following classifications:

- When one of the investment rules advised the investor to establish a Long position, that investment rule was assigned a classification of 1.
- When one of the investment rules no longer advised the investor to establish a Long position, that investment rule was assigned a classification of 0.

Then, after normalization, it was necessary to build a system capable of assigning a weight to each of the investment rules regarding the decision to establish a Long position on the market or to remain out of the market, with the objective that all indicators had an influence on the final decision. For this purpose, a system was constructed in which the final position at a given instant is equal to the sum of the classification assigned to each investment rule at the same instant multiplied by its weight, as represented in equation 1. This system was based on the weight system created by [10].

\[
\text{Position}_{\text{final}(m)} = \sum_{r=0}^{\infty} W_r \times \text{Classification}(r,m)
\]  

(1)

Where Classification \((r, m)\) corresponds to the classification given to investment rule \(r\) in minute \(m\), \(W_r\) corresponds to the weight assigned to each investment rule \(r\), while Position_{final}(m) corresponds to the final position in minute \(m\). Subchapter C of this chapter provides an explanation of how each weight was assigned to each investment rule.

It is also important to note that the sum of all weights was always equal to 1, and each of the weights individually assigned to each investment rule could not be greater than 1 nor less than 0. There was also a weight called required weight, which had values in the case of this thesis, ranging from 0.70 to 1, being responsible for deciding when Position_{final}(m) was large enough to affect the Long position or not. When the Position_{final}(m) was greater than or equal to the required weight, a Long position was established if it was not already established, whereas when the Position_{final}(m) becomes less than the required weight, the asset was sold if the Long position was previously established or remains out of the market if a Long position was not established yet.

It will then be explained in detail, how the investment rule associated with the nervous periods indicator chose the periods where there would be greater interest in performing transactions and how the classification of the investment rules of the Percentage Indicator and Size indicator were made. The classification of the investment rules associated to the SMA, DCM, RSI, ROC indicators will not be shown in this extended summary.

**Investment Rules - Nervous Periods Indicator**

To establish in which periods transactions could be made, the Nervous Periods indicator was created. These periods were established by analysing the value of the SMA indicator slope that corresponded to the Nervous Periods indicator output value. Then an investment rule was created from this indicator. This rule
consisted that if the slope of the SMA indicator was greater than a certain value, it was considered that transactions could be made, while if the value of the slope was lower or equal to a certain period, it was considered that it was not outside the intended period, and no transaction could then be made, as represented in equation 2.

\[
\text{Periods} = \begin{cases} 
\text{Intended periods, } \text{slope}_{\text{sma}}(x, m) > H \\
\text{Outside the intended periods, } \text{slope}_{\text{sma}}(x, m) \leq H
\end{cases}
\] (2)

In equation 2, \(\text{slope}_{\text{sma}}(x, m)\) corresponds to the value of the SMA slope with the period \(x\) in the minute \(m\), while \(H\) corresponds to the value that defines the boundary between the period being considered a period in which transactions can be made and the period in which transactions cannot be performed. Both the constant and the period value of the SMA will be explained in this chapter on subchapter C as they were obtained.

In Figure 4, you can see the green and red dots that correspond, respectively, to the beginning and end of the periods stipulated by the investment rule associated with the Nervous Periods indicator.

**Investment Rule - Percentage Indicator**

Finally, the second indicator created to establish Long positions was the Percentage indicator. First, an analysis was made of the Percentage candles, analysing in each row whether the closing price was higher than the opening price. Next, a vector was created, which assigned the value 1 when the closing price was higher than the opening price, and a value 0 when it was lower than the opening price. Subsequently, a SMA of period 12 was made for the particular case of the USD/TRY and USD/ZAR currency pairs, and period 10 was made for the case of the USD/MXN currency pair of this vector, called percentage of green candles. This percentage of green candles corresponded to the value assigned each minute by the Percentage indicator. Based on this indicator, an investment rule was created that assigned a classification of 1 when the Percentage indicator or the percentage of green candles exceeded 0.55 if that moment was within the established periods. This Classification remained at 1 until the moment when the period set by the Nervous Periods indicator ended.

In Figure 6, it is represented with the red signal, the classification assigned overtime to the investment rule for the Percentage indicator.

**Investment Rule - Size Indicator**

The first of the two indicators created to make Long positions was the Size indicator. This indicator had as its value, the number of Size candles in a row with the closing price higher than the opening price. From this indicator, an investment rule was created that assigned a classification of 1 from the next minute to be identified three rows in a row of Size candles data, where the closing price was higher than the opening price in the case of next minute be within the Nervous Periods until the end of the previously established periods. If three rows in a row with the closing price above the opening price were not found within the established periods by the investment rule associated with the Nervous Periods indicator, this investment rule was assigned a classification of 0.

In Figure 5, you can see the purple sign, which corresponds classification of the investment rule associated with the Size indicator overtime.

**C. Genetic Algorithm module layer**

1) **Choice of Initial Population**

The first phase of the GA corresponds to the choice of the initial population, which represents the parameters that the GA is intended to optimize.

Regarding the Nervous Periods indicator, the GA was responsible for optimizing the parameter SMA Period and constant \(H\), being that the parameter SMA Period corresponds to the period of Simple Moving Average used to calculate the slope, while the constant \(H\), corresponded to the value of slope from which transactions could be made. As for the Size indicator, it was only necessary to optimize the closing price variation parameter, as this indicator used the Size candles and they were created from the closing price variation. Regarding the SMA indicator, it only had one parameter that needed optimization which was the SMA Period. This parameter corresponded to the Simple Moving Average period used by this indicator. Regarding the DCM indicator, it had two parameters that needed optimization, the Fast SMA Period and the Slow SMA Period. These periods corresponded respectively to the period of two SMA, one with a shorter period and one with a longer period than the other. Regarding the RSI indicator, it only needed the RSI Period optimization, while the ROC indicator needed the optimization of the ROC Period parameter. The RSI Period and the ROC Period as the name implies corresponded respectively to the RSI indicator Period and the ROC indicator Period. About the
Percentage indicator, it only needed to optimize of the percentage variation of the closing price, since it used the Percentage candles, and these were obtained through a specific value of percentage change of the closing price.

Regarding the parameters, Size weight, Percentage weight, RSI weight, ROC weight, DCM weight and SMA weight, these respectively corresponded to the weight given by the investment rules regarding the Size indicator, Percentage indicator, RSI indicator, ROC indicator, DCM indicator and SMA indicator, which were also optimized by the GA. These weights refer to the weight that each investment rule had individually, in the decision to establish a Long position on the market. Finally, the last parameter that the GA was responsible for optimizing was the required weight. This weight was responsible for imposing the barrier between establishing a Long position or staying out of the market.

Then, after choosing the parameters to be optimized, a previous study was done, to understand what would be the best limits for each parameter, aiming to restrict the values that the GA could initially generate. In this particular work, 50 initial chromosomes were created, with random values in each gene of each chromosome (except for the closing price variation and percentage closing price variation genes), and each gene was limited to the values shown in Table 4.

<table>
<thead>
<tr>
<th>Genes</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Increments</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMA Period slope</td>
<td>12000</td>
<td>18000</td>
<td></td>
</tr>
<tr>
<td>Constant H</td>
<td>2*10E-6</td>
<td>8*10E-6</td>
<td></td>
</tr>
<tr>
<td>Closing Price Variation</td>
<td>USD/TRY - 0.0018</td>
<td>USD/TRY - 0.01</td>
<td>USD/TRY - 0.00018</td>
</tr>
<tr>
<td></td>
<td>USD/MXN - 0.092</td>
<td>USD/MXN - 0.01</td>
<td>USD/MXN - 0.00118</td>
</tr>
<tr>
<td></td>
<td>USD/ZAR - 0.007</td>
<td>USD/ZAR - 0.07854</td>
<td>USD/ZAR - 0.00146</td>
</tr>
<tr>
<td>SMA Period</td>
<td>12000</td>
<td>18000</td>
<td></td>
</tr>
<tr>
<td>Fast SMA Period</td>
<td>800</td>
<td>1300</td>
<td></td>
</tr>
<tr>
<td>Slow SMA Period</td>
<td>2400</td>
<td>3400</td>
<td></td>
</tr>
<tr>
<td>RSI Period</td>
<td>5000</td>
<td>15000</td>
<td></td>
</tr>
<tr>
<td>ROC Period</td>
<td>5000</td>
<td>15000</td>
<td></td>
</tr>
<tr>
<td>Percentage change of the closing price</td>
<td>USD/TRY - 0.05%</td>
<td>USD/TRY - 0.03%</td>
<td>USD/TRY - 0.0062%</td>
</tr>
<tr>
<td></td>
<td>USD/MXN - 2.5%</td>
<td>USD/MXN - 0.05%</td>
<td>USD/MXN - 0.05%</td>
</tr>
<tr>
<td></td>
<td>USD/ZAR - 2.5%</td>
<td>USD/ZAR - 0.05%</td>
<td>USD/ZAR - 0.05%</td>
</tr>
<tr>
<td>Weight Size</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Weight DCM</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Weight RSI</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Weight ROC</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Weight SMA</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Weight Percentage</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Required Weight</td>
<td>0.7</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

In the case of the closing price change and closing price change percentage genes, which correspond to the values required to create respectively the Size candles and the Percentage candles, the Genetic Algorithm was limited to generating a value from 50 values previously selected in subchapter A of this chapter.

In Figure 7, shows the structure of a chromosome whose genes/parameters will be optimized in the GA following phases.

![Figure 7 - Chromosome representation.](image)

2) Fitness function

Then, in order to evaluate each chromosome individually, a fitness function was created that depended exclusively on the percentage of gains metric and the Maximum Drawdown (MDD) as a percentage. With regard to MDD, this is an indicator used to assess the relative risk of one strategy against another, thus focusing on capital preservation. This indicator allows us to measure the largest loss in terms of capital [11], however, it does not give us information about what has been gained or lost over time. To obtain this information, the percentage of gains was used to understand if the strategy used was profitable. With these two evaluation metrics, it was possible to try to obtain the chromosomes that gave greater guarantees in terms of earnings, but at the same time did not give the investor a great risk. Percentage of gains and percentage of Maximum Drawdown metrics are given respectively by equation 3 and equation 4.

\[
p \cdot \text{of gains}(t) = p \cdot \text{of gains}(t - 1) + \frac{\text{price}_{\text{Sell}}(t) - \text{price}_{\text{purchase}}(t)}{\text{price}_{\text{purchase}}(t)} \times 100 \quad (3)
\]

\[
\text{MaximumDrawdown}(t) = \text{Max}((\text{Drawdown}(t), \text{Drawdown}(t - 1))) \times 100 \quad (4)
\]

In equation 3, \(\text{price}_{\text{purchase}}(t)\) corresponds to the price at which it entered the market, while \(\text{price}_{\text{sale}}(t)\) corresponds to the price at which the asset was sold. After a transaction, it is added to \(p \cdot \text{of gains}(t)\), which corresponds to the current earnings percentage, the \(p \cdot \text{of gains}(t - 1)\), which represents the previous earnings percentage, plus the earnings percentage, earned on the current transaction, \(t\) which corresponds to the split \(\frac{\text{price}_{\text{sale}}(t) - \text{price}_{\text{purchase}}(t)}{\text{price}_{\text{purchase}}(t)}\).

With respect to equation 4, \(\text{Drawdown}(t)\) and \(\text{Drawdown}(t - 1)\) respectively represent, the Drawdown at time \(t\) and \(t - 1\). Regarding \(\text{MaximumDrawdown}(t)\), corresponds to the Maximum Drawdown at the moment \(t\), which is equal to the maximum of Drawdown up to the instant \(t\).

In conclusion, to give more weight to the risk in comparison to the profit obtained, the Maximum Drawdown metric was twice penalized when building the fitness function, to obtain final chromosomes that were not at great risk, while sacrificing, sometimes part of the possible profit in the face of risk reduction. The chosen fitness function is then represented by equation 5.

\[
\text{fitness function}(X) = p \cdot \text{of earnings}(X) - 2 \times \text{Maximum Drawdown}(X) \quad (5)
\]

Where \(p \cdot \text{of gains}(X)\) and \(\text{Maximum Drawdown}(X)\) correspond to the percentage of gains and MDD as a percentage of chromosome \(X\) respectively, while \(\text{fitness function}(X)\), corresponds to the fitness function of chromosome \(X\).
3) Selection

Then follows the selection phase. There is no general selection to use for all cases, as each selection has its advantages and disadvantages, always depending on the problem at hand [12]. In this problem, it was intended that in this selection the best chromosomes were chosen according to the fitness function and that these same characteristics would be propagated for the following generations. There are several types of possible selections to use, but only Elitist selection ensures the guarantees stated above and so it was the selection chosen in this problem.

At this stage were chosen from the initial 50 chromosomes, the top 25, according to the fitness function explained earlier. Then these top 25 were used to generate another 25, while the 25 that were not chosen were discarded.

4) Crossover

Following the selection phase, the crossover phase follows. In this phase, the 25 best chromosomes chosen in the selection phase, gave rise to 25 more (50% crossover percentage compared to the initial population), in the particular case of this thesis, through the two-point crossover technique. This crossover was chosen as it allows to increase the research space of the GA [13], thus preventing the GA from converging too early. This method uses two cutoff points, having been chosen the cutoff points 3 and 8. That is, the first child early. This method uses two cutoff points, having been chosen the cutoff points 3 and 8. That is, the first child was the result of chromosomes 0 to 2 inclusive and 9 to 15 inclusive, of the first parent and chromosomes 3 to 8 inclusive of the second parent.

5) Mutation

Finally, after the crossover phase follows the mutation phase. This phase is quite important as, it is also responsible for ensuring that the GA does not converge too early for solutions that are not optimal, and for ensuring some variability in the population.

The mutation performed for each gene is shown in Table 5.

<table>
<thead>
<tr>
<th>Genes</th>
<th>Lower Mutation</th>
<th>Upper Mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMA Period slope</td>
<td>-50</td>
<td>50</td>
</tr>
<tr>
<td>Constant H</td>
<td>-0.0000001</td>
<td>0.000001</td>
</tr>
<tr>
<td>SMA Period</td>
<td>-50</td>
<td>50</td>
</tr>
<tr>
<td>Fast SMA Period</td>
<td>-5</td>
<td>5</td>
</tr>
<tr>
<td>Slow SMA Period</td>
<td>-10</td>
<td>10</td>
</tr>
<tr>
<td>RSI Period</td>
<td>-30</td>
<td>30</td>
</tr>
<tr>
<td>ROC Period</td>
<td>-30</td>
<td>30</td>
</tr>
<tr>
<td>Weight Size</td>
<td>-0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Weight DCM</td>
<td>-0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Weight RSI</td>
<td>-0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Weight ROC</td>
<td>-0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Weight SMA</td>
<td>-0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Weight Percentage</td>
<td>*1</td>
<td>*1</td>
</tr>
<tr>
<td>Required Weight</td>
<td>-0.02</td>
<td>0.02</td>
</tr>
</tbody>
</table>

*1 To ensure that after the mutation phase, the sum of the weights remained equal to 1, the Percentage weight mutation was always equal to the symmetrical of the sum of the weighting mutations for the investment rules, Size, DCM, RSI, ROC and SMA.

After the mutation phase, the 25 chromosomes resulting from crossover and mutation were evaluated in the selection phase, as well as the 25 best initially selected in the selection, which corresponded to the 50 chromosomes of the first generation. This process was repeated 20 times in the particular case of this thesis. At the end of the 20 times the best twentieth generation chromosome with the optimized genes was obtained, according to the fitness function previously explained.

D. Genetic Algorithm Validation

Finally, in this last phase, the genes optimized by the GA were tested in the test period, to understand the viability of the GA. In Figure 8, a possible chromosome with the optimized genes is represented.

IV. RESULTS

A. Evaluation Metrics

To evaluate the various case studies three distinct assessment metrics were used, Cumulative return, Cumulative return per minute and Drawdown, all of which were in percentage.

The first evaluation metric, Cumulative return (CR), is the percentage measure of how much investment has increased or decreased over a specific period, in other words, how much wealth has been accumulated or lost due to investments over time [14]. This metric can be given by equation 6.

\[
\text{Cumulative Return}(t) (\%) = \frac{\text{Profit}_t - \text{Profit}_{\text{purchase}}}{\text{Profit}_{\text{purchase}}} \times 100
\]  

(6)

The CR metric is important as it allows us to realize how profitable each strategy is, however, it does not allow us to take into account the time each strategy was in the market, and therefore it was not possible to quantify the risk of each strategy through this metric, since the longer the investor is in the market, the greater the risk that the investor runs.

With this in mind, a second evaluation metric called Cumulative Return per minute (CRPM) was created, which was the result of CR divided by the minutes each strategy had been in the market. With this metric, it was already possible to take into account both profit and risk, being this metric given by equation 7.

\[
\text{Cumulative Return per minute}(t) = \frac{\text{Cumulative Return}(t)}{X}
\]  

(7)

In equation 7, the constant \(X\), corresponds to the number of minutes the strategy was in the market.

Finally, also to quantify the risk of each strategy, the Drawdown evaluation metric was used, which consists of a persistent decrease in the price of a given investment, due to a sequence of decreases in the investment price [15,16]. This metric is therefore responsible for directly measuring the accumulated losses an investment incurs, and in the case of this thesis, its value was equal to the current CR minus the CR maximum value to date [17], as represented in equation 8. Since Cumulative Return is in percentage, the Drawdown evaluation metric is also in percentage.

\[
\text{Drawdown}(t) = \text{Cumulative Return}_{\text{max}}(t) - \text{Cumulative Return}(t) (\%)
\]  

(8)
B. Case Study - Inclusion of a variety of investment rules in different strategies

Before starting this case study, a previous study was done to understand which training set had better results for each currency pair, between a large training set and a small training set in terms of profit and risk. The large training set consisted of starting training the GA with 1800000 rows of data, which corresponds to about 4 years and 7 months and test it on the next 200000 rows. Then the next training used the 1800000 lines used in the previous training plus the previous test set and tested on the next 200000 lines. This process was repeated until there were no longer rows of file data to test the GA. The small training is almost the same as the large training, with the difference that the first test started to train the GA with 400000 lines, instead of 1800000. After this study, it was concluded that the USD/ZAR currency pair had better results with a large training and the USD/MXN and USD/TRY currency pairs had better results with a small training. With this in mind, the following strategies used the most favourable training set.

Initially, 10 Genetic Algorithms were tested where there were no limitations by the investment rule associated with the Nervous Periods indicator and where neither the Size and Percentage investment rules were used to decide when to make Long positions. It was therefore only used for Long positions, the investment rules relative to indicators Trend, SMA and DCM and Momentum indicators, RSI and ROC. Next, the Genetic Algorithm test which had the percentage of gains values closest to the average among the 10 Genetic Algorithms tested was chosen. This test was called as Outside Nervous Periods.

Subsequently, 10 Genetic Algorithms were also tested, and in these tests, the investment rule regarding the Size indicator was used to establish Long positions on the market as well as the investment rules associated with the SMA, DCM, RSI and ROC indicators. Also in this strategy, the investment rule referring to the Nervous Periods indicator was used, to limit the periods where market positions could be established. As was done in the previous test, the Genetic Algorithm test which had percentage gains values closer to the average among the 10 Genetic Algorithms tested was chosen. This test was called Within Nervous Periods.

Then, the investment rule associated with the Percentage indicator was added to the strategy Within Nervous Periods, which now has an influence on the decision to take Long positions on the market, as well as like the other rules already used in the strategy within Nervous Periods, as explained earlier. Once more, 10 Genetic Algorithms were tested, and then the test which had the percentage of gains values closer to the average among the 10 Genetic Algorithms tested was chosen. The nearest test was then termed as Complete within Nervous Periods.

Finally, the Buy&Hold strategy, which consists of an investment strategy, in which the investor buys assets and holds them for a long period, irrespective of market fluctuations, has also been used [18]. This strategy was used to understand the benefits that the other strategies had on simply buying when the test period begins and selling when it ends, which is the decision made by investors that use the Buy&Hold strategy.

Subsequently, the four strategies previously referred to in this case study were compared according to the CR, CRPM, and Drawdown metrics to understand which strategy is the most profitable as well as which has the least risk.

In Figures 9, 10, and 11, it is possible to see the CR, CRPM, and Drawdown of the four previously mentioned strategies for the USD/TRY currency pair.

From Figure 9, it was possible to see that, the strategy Complete Within Nervous Periods with small training is the worst in terms of the metric of CR compared to the remaining strategies. Since the strategy Within Nervous Periods and the strategy Complete within Nervous Periods were less time on the market, the four strategies were again compared, but this time they were evaluated according to CRPM metric to understand how much return they allowed, given how long they were in the market. After this evaluation that is represented in Figure 10, it was possible to verify that the strategy Complete within the Nervous Periods is the strategy that allows the highest CRPM. This result was predictable, as this strategy is less time-consuming in the market, thus allowing for less risk compared to the remaining three strategies, although the strategy Within Nervous Periods is also very short time in the market.

For the Drawdown, shown in Figure 11, it was not possible to draw any insights into the least risky strategy over time, as strategies Outside Nervous Periods, Within Nervous Periods and
Complete within Nervous Periods take advantage of each other at different test periods, however, it can be noted that the three strategies mentioned above contain less risk than the Buy&Hold strategy.

It was therefore possible to conclude that, in the case of the USD/TRY currency pair, the inclusion of the investment rule associated with the Percentage indicator together with the investment rules of the Size and Nervous Periods indicators have a positive risk intervention as they allow for a higher profit given the time they have been in the market, and thus a higher opportunity cost.

Concerning the USD/ZAR currency pair, the respective tests performed for the CR, CRPM and Drawdown evaluation metric are shown respectively in Figures 12, 13 and 14.

Complete within Nervous Periods, followed by strategies Within Nervous Periods, Outside Nervous Periods and Buy&Hold. It is also important to note that the strategy Complete within Nervous Periods has a much greater advantage over other strategies when evaluated by the CRPPM evaluation metric compared to when assessed by the CR evaluation metric. This major advantage comes from not only being the most profitable strategy but also the strategy that allows the investor to be less time in the market and therefore incur less risk.

This conclusion concerning risk is further supported by verifying from Figure 14 that the strategy Complete Within Nervous Periods has much lower Drawdown throughout the test, as well as a lower MDD compared to the other strategies.

With all the tests taken into account, it was then possible to conclude that the inclusion of the investment rule associated with the Percentage indicator together with the investment rules concerning the Size and Nervous Periods indicators in the strategy Complete within the Nervous Periods has positive results in the USD/ZAR currency pair case. These positive results come from two important factors, higher profit and at the same time lower risk. As for the lower risk, this is due to this strategy being much less time in the market, having a lower Drawdown over practically the entire testing period and having a lower MDD. Concerning profit, it was possible to see that, even though this strategy was less time in the market, it gave rise to a higher CR when compared to the other strategies analysed in this case study.

Finally, the CR, CRPM, and Drawdown evaluation metrics are used again, but this time they are used to evaluate the USD/MXN currency pair, and these tests are represented respectively in Figures 15, 16, and 17.

Through the Figures 12 and 13, it was found that for both evaluation metrics CR and the CRPM, the best strategy is

![Figure 12 – Cumulative Return of the strategies Complete within Nervous Periods with large training, Within Nervous Periods with large training, Outside Nervous Periods with large training, and Buy&Hold, for the USD/ZAR currency pair.](image1)

![Figure 13 – Cumulative Return per minute of the strategies Complete within Nervous Periods with large training, Within Nervous Periods with large training, Outside Nervous Periods with large training, and Buy&Hold, for the USD/ZAR currency pair.](image2)

![Figure 14 - Drawdown of the strategies Complete within Nervous Periods with large training, Within Nervous Periods with large training, Outside Nervous Periods with large training, and Buy&Hold, for the USD/ZAR currency pair.](image3)

![Figure 15 – Cumulative Return of the strategies Complete within Nervous Periods with small training, Within Nervous Periods with small training, Outside Nervous Periods with small training, and Buy&Hold, for the USD/MXN currency pair.](image4)

![Figure 16 – Cumulative Return per minute of the strategies Complete within Nervous Periods with small training, Within Nervous Periods with small training, Outside Nervous Periods with small training, and Buy&Hold, for the USD/MXN currency pair.](image5)
As a result of the analysis of Figures 15 and 16, it became clear that the best strategy is Complete within Nervous Periods, taking into account the CR and CRPM evaluation metrics. As with the USD/ZAR currency pair, it was possible to verify that in the USD/MXN currency pair the difference between the strategy Complete within Nervous Periods compared to the other three strategies evaluated increased when evaluated by the CRPM in comparison to when assessed by the CR metric. This difference is again due to the fact that strategy Complete within Nervous Periods is the one with the least time on the market.

For the Drawdown evaluation metric, it was possible to see from Figure 17 that the strategy Complete within Nervous Periods has almost always less Drawdown than the other strategies during the test period, and it can also be seen that it is the strategy with the lowest MDD.

The conclusion drawn from this USD/MXN currency pair regarding the inclusion of the investment rule associated with the Percentage indicator along with the investment rules for the Size and Nervous Periods indicators is similar to the conclusion drawn for the USD/ZAR currency pair in terms of profit and risk. Regarding the risk, this similarity is due to the fact that, the strategy Complete within Nervous Periods is able to be less time on the market, have a lower Drawdown in practically the entire test and have a smaller MDD, as had already been achieved in the USD/ZAR currency pair. As for profit, this strategy, similarly to what had happened in the USD/ZAR currency pair, was able to achieve higher profit values even with less time on the market, compared to the other strategies tested in this case study.

V. CONCLUSIONS

In this work, it was possible to realize that the Genetic Algorithm played a very important role in the optimization of the necessary parameters that came from the investment rules associated with the various Technical indicators.

In this thesis, it was also possible to measure the positive results, both in terms of profit and risk, that come from choosing specific periods, with greater volatility and where it is possible to simultaneously identify a trend. With regard to the Nervous Periods indicator investment rule, it has had a key influence on the results, as it has allowed profit to remain similar or in some cases even higher compared to strategies that have no limitations in periods when they can do transactions, such as Buy&Hold and the outside Nervous Period strategy. In addition to the profit advantages, it was possible to achieve much lower risk as this strategy allowed the investor to spend much less time in the market and therefore has a higher opportunity cost as well as lower Drawdown in almost all the test set.

As regards to the inclusion of the investment rules associated with the Size and Percentage indicators together with the investment rules already used in other strategies, these allowed for a more advantageous choice in terms of when Long positions would be most advisable.

In conclusion, it was possible to create an adaptive system capable of trading in USD/TRY, USD/MXN and USD/ZAR currency pairs, with various investment rules associated with different Technical indicators, having been able to achieve 57.60% CR in USD/TRY currency pair, with only 26.04% of time in the market over the full test period, which is about 6 months. In the cases of the USD/ZAR and USD/MXN currency pairs, this system also achieved very positive results, with a CR of 13.53% at 26.82% of the time in the market, for the USD/ZAR and 17.25% of CR at only 28.04% of time in the market, for the USD/MXN currency pair.

This work has had quite positive and innovative results as has been said, even though transactional costs were not considered as spread. However, it would be interesting to see if these results remained positive in other currency pairs and even experiment the investment strategies implemented, in other markets.

Finally, I would suggest the implementation of a Principal Component Analysis (PCA) to decide which are the best Technical indicators, as well as a population increase, something that was not done due to computational limitations as well.

REFERENCES