The Application of Temporal Association Rule Mining in Stock Markets

Matheus H. de Souza[1], **Adriano C. M. Pereira**[1]

[1]Departamento de Ciência da Computação – Universidade Federal de Minas Gerais

{matheushsouza,adrianoc}@dcc.ufmg.br

Abstract. *For a long time it was believed that trying to forecast stock prices from stock markets was an absurd idea given the unpredictable nature of such markets, and, indeed, many studies have suggested that. However with the rise of machine learning and data mining, new approaches for this problem were born and some of them achieved promising results. In the present study we implement and evaluate a predictive model that aims to forecast stock market prices based on a Temporal Association Rule Mining technique. The model is tested with brazilian stock market data and its results are compared to the traditional Buy and Hold strategy. The results here are also promising, since it is shown that the model is capable of overcoming the Buy and Hold baseline despite the fact that it makes a lot of wrong predictions.*

1. Introduction

Algorithmic data analysis applied to stock markets can help new traders and even the seasoned player into making good trading decisions and achieving great results in their investments. However, dealing with stock markets is a tough task, since markets, highly dependant on organic decisions and events, are mostly unstable and unforeseeable.

Several past works have suggested that stock prices cannot be predicted because they are completely random, making it impossible to profit from trying to foretell trading strategies. The most notable studies in this direction are by [Cowles 1933] and [Fama 1965a, 1965b]. However, recent studies [Atsalakis and Valavanis 2009], [Brabazon and O'Neill 2006], [Nair and Mohandas 2015a, 2015b], [Nair et al. 2010], have shown that forecasting stock prices is not an absurd idea and, furthermore, that it can generate much better returns than the traditional and passive Buy and Hold (B&H) strategy.

Association Rule Mining (ARM) is a well stated and solved problem nowadays. To find association rules between items means to identify strong relations between these items, in databases, usually very large, using some measures of interestingness [Agrawal et al. 1993]. The most widespread use of this technique is in market basket analysis, what essentially consists of finding which items are bought together by customers [Trnka 2010]. However, there is a primary issue with ARM algorithms: temporal information is not a considered parameter. This drawback makes the application of temporal association rule mining in stock markets a very challenging task.

The present study takes on the hard duty to implement and evaluate a predictive model to suggest trading strategies based on past stock market data and underlying association rules among this data. The system makes use of Symbolic Aggregate approXimation (SAX) [Lin et al. 2003] to obtain a convenient representation of a time series, that is inputed into the ARM Apriori [Agrawal and Srikant 1994] algorithm in order to find temporal association rules between price ranges and generate trading rules. Also, a way of dealing with the aforementioned problem with ARM algorithms is presented. At the end, the implemented methodology is tested with data from the brazilian stock market and its results are compared with the traditional Buy and Hold strategy.

The remaining of this article is structured as follows: Section 2 briefly describes and cites some related works; Section 3 presents the methodology followed to implement the studied model and Section 4 explains the experiments and shows the achieved results. Finally, in the Section 5, the results are analysed and future works are discussed.

2. Related Works

Studies regarding *Temporal Association Rule Mining* have become very popular since the discovery of the technique [Bettini et al. 1998], [Mannila et al. 1997], [Ale and Rossi 2000], [Atsalakis and Valavanis 2009] and research has been made in various fields of study. The work by [Nguyen et al. 2018] utilises Temporal Association Rule Mining to understand toxicities in a routine cancer treatment; a research by [Wen et al. 2019] makes use of the technique to predict traffic congestion; in [Chok and Gruenwald 2009] the approach is used to estimate missing or corrupted data in sensor network applications; and [Ng et al. 2007] applies it to support the discoveries of crime patterns in a district of Hong Kong.

However, Association Rule Mining algorithms have not been very popular in studies aiming to generate stock trading recommendations. Very little research has been made on the application of ARM for stock markets. Some examples are by [Srisawat 2011], in which ARM was used to identify relationships between individual stocks; [Ting et al. 2006], that made use of ARM to study frequently occurring patterns in stock time series and relationships between price movements of pair of stocks; and [Kumar and Kalia 2011], that used the technique to find similarities between stocks on Indian stock markets. No study that applies the technique to brazilian stock markets was found.

The efforts in trying to predict stock markets tendencies seems to be concentrated in works that make use of Sentiment Analysis (SA) [Medhat et al. 2014] techniques [Mao et al. 2012], [Labiad et al. 2018], [Schumaker and Chen 2009] on comments of users from platforms like Twitter [Twitter 2019], investment forums like Investing [Investing 2019], and financial news websites. This trend, to the detriment of Association Rule Mining and other techniques, may have been due to the great results achieved by the use of SA [Wang et al. 2015] and the difficulties in incorporating temporal information into ARM and finding a proper way to represent time series as a set of items, which is the usual input for ARM algorithms. The methodology proposed in this study addresses both of these problems.

A methodology that uses most of the algorithms and techniques implemented by this study has been proposed in two different papers: first by [Canelas et al. 2012] and later by [Nair et al. 2015]. In these research the authors introduce genetic algorithm optimised *Symbolic Aggregate approXimation (SAX)–Apriori*-based stock trading recommender systems, which can mine temporal association rules from a stock price dataset in order to generate stock trading recommendations. The results found in these two works show that, unlike suggested by [Cowles 1944] and [Fama 1965b], trying to predict stock market prices' behaviour based on past data was actually more profitable than passive strategies.

In fact, the present study was mainly inspired by the two aforecited works: [Canelas et al. 2012] and [Nair et al. 2015] — specially the latter —, and it differs from them since it tries to simplify the technique by sticking only to the fundamental steps, leaving away aspects that do not lead to significant gains in the results [Hamilton 2018], like the use of Hodrick-Prescott Filter [Hodrick and Prescott 1997]. Another difference between the present study and the aforementioned research is the data: here we evaluate the technique applied to the brazilian stock market, whilst the two works focus on the US and Indian markets.

3. Methodology

The methodology implemented by this study can be summarized into three main steps: (1) the time series data is split into training and testing subsets; (2) the time series is converted into a symbol representation, with the help of a genetic algorithm optimization; and lastly, (3) association rules are extracted from this symbolic representation to create trading rules that will result in trading loss or profit.

Figure 1 below illustrates the whole methodology, and its details are discussed in the underneath subsections.

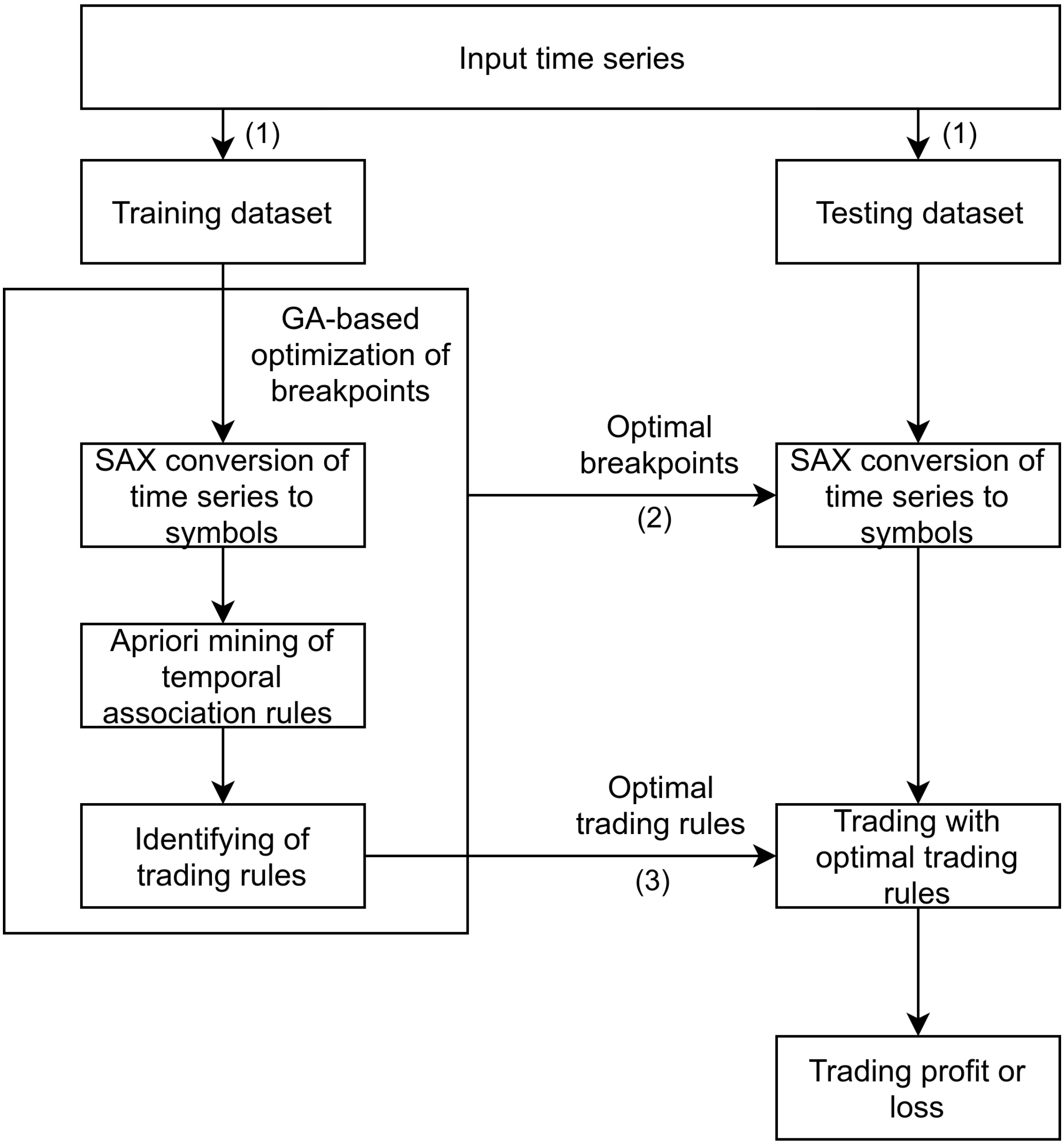

Figure 1. Block diagram of the implemented model.

3.1. Training and testing data

The model's input data is a time series of the closing prices of a share in a stock market. The training data was set to be 70% of the dataset, while the other 30% represent the testing data. Since this study deals with stock markets and aims to be a predictive model, the training data is the 70% data at the beginning of the data set, and the testing data is the reminder 30%. A *sliding window* or a *k-folds* approach was not implemented since the objective is to determine if the model can forecast future events based on past data, causing the need to respect time order in the training-test split.

Figure 2. Train-test split of an input time series example.

3.2. SAX conversion of a time series to symbols

Once the data is split, it is then converted into symbols with the help of SAX algorithm [Lin et al. 2003]. The SAX algorithm works by first converting the time series consisting of N samples into a piecewise aggregate aproximation (PAA) representation, which is essentially a discretization of the time series into M intervals, such that $N \geq M$, and then converting this representation into a string, by assigning a symbol to each of these intervals.

The algorithm takes two parameters: the window size D, which is the number of samples that are considered together and converted into a single symbol, and the number of symbols to be used, S. It takes as input a time series $\boldsymbol{Y} = \{y_1, y_2, \cdots, y_N\}$ consisting of N samples. As a first step, the data is normalized, what of course does not affect the time series original shape. Then, with the use of PAA, the dimensionality of the data set is reduced. In PAA, the total number of samples is divided into M equal sized bins (windows) with each bin consisting of D samples, such that $M \times D = N$. A series $\boldsymbol{B} = \{b_1, b_2, \cdots, b_M\}$ with M number of samples is generated, such that each b_i is the average value inside the bin i, i.e.:

$$b_i = \frac{M}{N} \sum_{j=\frac{N}{M}(i-1)+1}^{\frac{N}{M}i} y_i, \qquad i = 1, 2, \cdots, M \tag{1}$$

To convert the resulting series to a symbolic representation, the amplitude of the resulting series is divided into S intervals, such that each interval is assigned a unique symbol. Since the data is normalized, it is assumed that the samples follow a normal distribution $N(0, 1)$. Therefore, $S - 1$ breakpoints on the normal distribution curve is selected in a way that equiprobable intervals are produced. The whole SAX conversion of time series to symbols is illustrated by the following figure.

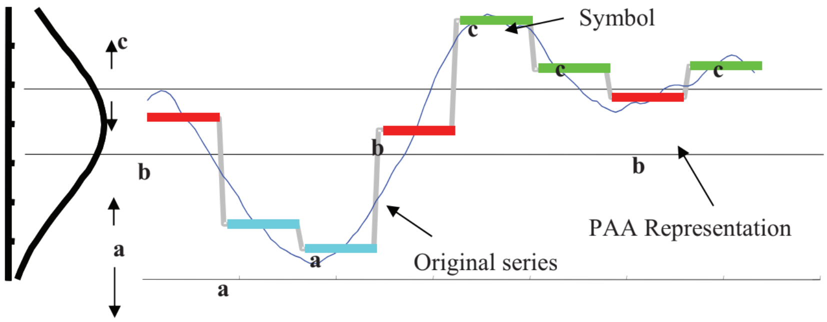

Figure 3. SAX representation of a series adapted from [Lin et al. 2007]

In Figure 3, the original series was $N = 128$ samples long, $D = 16$ and $S = 3$. The set of symbols is $\boldsymbol{\alpha} = \{a, b, c\}$. The series then got reduced to the sequence of $M = 8$ symbols *baabccbc*, consisting of three unique symbols.

In this study, since the objective is to make a one day ahead forecast of the stock market and the input time series is the daily closing stock prices, it is always the case that $D = 1$ and $M = N$, i.e., we want to preserve the input time series size and only assign a symbol to each price range determined by S.

3.2.1. GA-based optimization of the SAX breakpoints

The section above described the original SAX conversion of a time series to symbols process. In this work this process is implemented with a slight modification in order to optimize the final profit. This modification consists of a GA-based optimization of the aforementioned breakpoints. This approach is described below.

The SAX breakpoints are the bounds $\boldsymbol{B} = \{b_1, b_2, \cdots, b_{S-1}\}$ on the normal distribution curve, such that every produced interval is equiprobable. For the methodology implemented by this study, instead of using equiprobable intervals, we select, using a genetic algorithm, optimized breakpoints within the upper and lower bounds given by the set of tuples $\{(-\infty, b_1), (b_1, b_2), (b_2, b_3), \cdots (b_{S-2}, b_{S-1}), (b_{S-1}, \infty)\}$.

The GA selects the best number of breakpoints, and the breakpoints itself, considering 100 generations and 30 solutions per generation. At each generation 15 individual solutions are kept, and the other 15 are killed to be replaced by combinations of the top 15 ones, using crossover and mutation.

The objective function optimized by the GA is simply the total profit in the training step:

$$\text{current money} - \text{invested money} \tag{2}$$

3.3. Apriori temporal association rule mining

After the conversion of the input time series to symbols, the association rules between the prices that will generate the trading rules can be identified with the help of Apriori [Agrawal and Srikant 1994] algorithm. The objective of *Association Rule Mining* algorithms like Apriori is to find association rules between items in a trasactional database. A transactional database is simply a database in which each instance is a set of items selected together for any purpose. Let $\boldsymbol{I} = \{i_1, i_2, \cdots, i_k\}$ be the k unique items in a transactional database. An association rule over these items will be of the type

$$I_A \Rightarrow I_B, \tag{3}$$

(meaning I_A "implies" I_B) where $I_A, I_B \subset I$ and $I_A \cap I_B = \{\}$. I_A is called *"antecedent"*, and I_B is called *"consequent"*.

The support for an association rule is given by

$$support(I_A \Rightarrow I_B) = P(I_A \cup I_B) \tag{4}$$

And the confidence of an association rule is

$$confidence(I_A \Rightarrow I_B) = P(I_A|I_B) \tag{5}$$

3.3.1. Incorporating temporal information into item sets

Temporal information is not a considered parameter when mining association rules from a transactional database. However, since we are dealing with time series and stock prices, time information is of fundamental importance. The incorporation of time information to ARM has been attempted, for instance, by [Lianga et al. 2005] and [Winarko and Roddick 2007]. The same methodology is implemented in this study and it works as follows:

After the time series is converted to a symbolic representation using SAX with the help of GA, we transform this symbolic representation into a transaction set such that each item is a symbol: Let $\boldsymbol{S} = \{S_1, S_2, \cdots, S_\beta\}$ be the set of unique symbols in the symbolic representation. Then the symbolic representation of the time series is $\boldsymbol{I} = \{I_1, I_2, \cdots, I_M\}, I_i \in \boldsymbol{S}$, which we use to form the transaction database, taking two symbols at a time, i.e., we make a transaction database consisting of two-item sets: $d_1 = \{I_1, I_2\}$, $d_2 = \{I_2, I_3\}$, $\cdots$, $d_{M-1} = \{I_{M-1}, I_M\}$. The transaction database is then $\boldsymbol{D} = \{d_1, d_2, \cdots, d_{M-1}\}$.

By making a transaction database consisting of two-item sets only, we ensure that the generated rules will associate only two single items. It is trivial to see that, in such rules, the rule's antecedent temporally precedes the rule's consequent in exactly one time unit, since, in the generated rules, the order of items is preserved according to the order in the transaction dataset and the dataset is formed from a temporal series.

3.4. Generating trading rules

With the help of the Apriori [Agrawal and Srikant 1994] algorithm, the frequent two-item sets are obtained from the transaction database mentioned in the section above. These frequent two-item sets will finally generate the trading rules to be used to try forecast the best action when trading.

Suppose that from the transaction database $\boldsymbol{D} = \{d_1, d_2, \cdots, d_{M-1}\}$, only the two-item set $d_1 = \{I_S, I_T\}$ was considered to be a frequent set by Apriori. Then the following trading rule would be generated: $I_S \Rightarrow I_T$. I_S is a symbol that represents a certain price range for a share, as well as I_T. This rule can be of three types:

1. $I_S > I_T$: when the trader is faced with a price range in the interval represented by I_S, they should sell their shares, since this rule says the the price in the next day is likely to fall.
2. $I_S < I_T$: the trader should buy shares, since this rule says the the price in the next day is likely to rise.
3. $I_S = I_T$: the trader should not do anything, since the price is likely to stay in the same price range.

Only rules of types 1 and 2 are considered and stored.

4. Experiments and Results

To put the proposed methodology to test, the implemented model was evaluated in comparison to the results of a Buy and Hold strategy using data from the brazilian stock market. The subsections below are dedicated to explain the experiments and present its results with more details.

4.1. The data used in the experiments

The proposed model was tested with closing prices of 57 shares from the main brazilian stock market index (Ibovespa), considering 4 time frames: the entire years of 2016, 2017 and 2018, individually, and the whole period between 2016 and 2018. The criteria used to evaluate the model was the return over investment provided by it, in comparison to the Buy and Hold strategy, that is a fundamental baseline for stock market investments.

Since the Ibovespa index changes every twelve months, the selected shares were the ones that remained in the index from 2016 to 2018. The table below lists the studied shares:

ABEV3	CMIG4	FLRY3	MGLU3	SBSP3
BBAS3	CSAN3	GGBR4	MRFG3	TAEE11
BBDC3	CSNA3	GOAU4	MRVE3	TIMP3
BBDC4	CVCB3	GOLL4	MULT3	UGPA3
BBSE3	CYRE3	HYPE3	NATU3	USIM5
BRAP4	ECOR3	IGTA3	PCAR4	VALE3
BRFS3	EGIE3	ITSA4	PETR3	VIVT4
BRKM5	ELET3	ITUB4	PETR4	VVAR3
BRML3	ELET6	JBSS3	QUAL3	WEGE3
BTOW3	EMBR3	KLBN1	RADL3	
CCRO3	ENBR3	LAME4	RENT3	
CIEL3	EQTL3	LREN3	SANB11	

Table 1. Studied shares.

4.2. The experiment

For each of the aforementioned considered time frames and for each of the selected shares, the experiment made to test the model was conducted as follows:

1. Two trading agents are initialized, each owning 0 shares and 1000 money – the currency here is not relevant.
2. The model is trained with the initial 70% of the input data, as described in Section 3.
3. A simulation is executed as if the first agent from the first step trades for the remaining 30% of the days using the trading rules provided by the model, while the second agent uses a Buy and Hold (B&H) strategy — i.e., buys as many shares as her money lets her and doesn't sell it for the whole period — for the same remaining days.
4. The Return Over Investment (Equation 6) for each agent is computed and compared. Also, the number of correct and incorrect predictions are counted accross the simulation.

$$\text{ROI} = \frac{\text{current value of investment} - \text{initial investment}}{\text{initial investment}} \tag{6}$$

4.3. Results for the first time frame: the year of 2016

The following graphs show the ROI obtained by each strategy in trading each of the shares listed in the Table 1, for the data from the year of 2016. Figure 4 shows, for each model, the Return On Investment for each individual share, and Figure 5 shows the aggregate ROI, i.e., the sum of all ROIs. Figure 5 shows the total correct and wrong predictions made by the model accross all testing cases in this time frame.

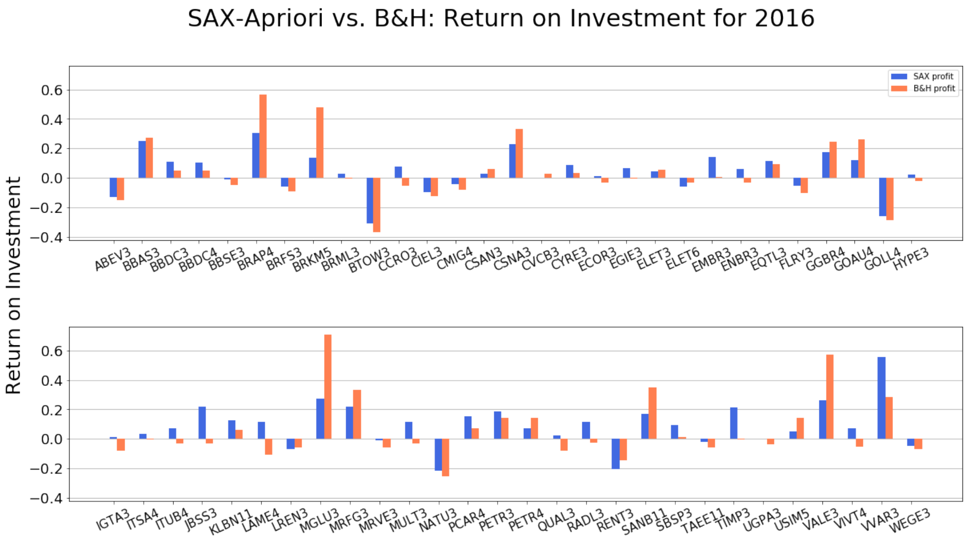

Figure 4. SAX-Apriori vs. B&H: ROI for 2016.

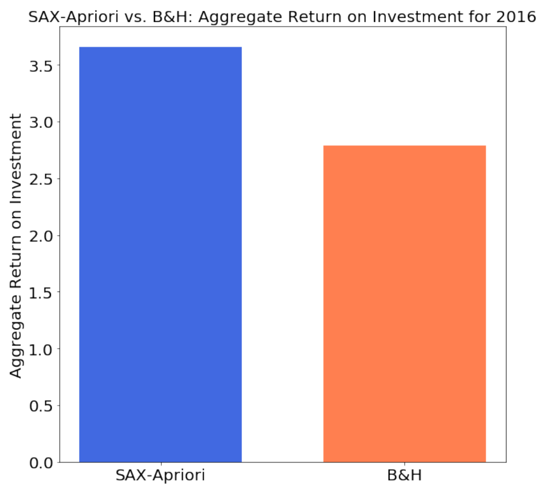

Figure 5. SAX-Apriori vs. B&H: Aggregate ROI for 2016.

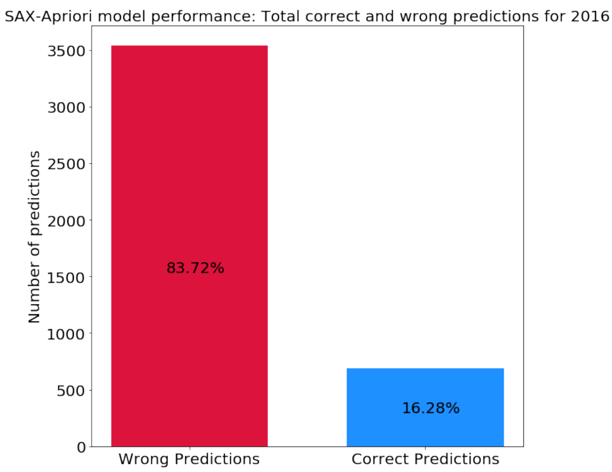

Figure 6. SAX-Apriori performance: Total correct and wrong predictions for 2016.

For this time frame, we can see that aggregate return for the SAX-Apriori based model was slightly better than the B&H return, and for the most cases, the SAX-Apriori model performed better (39 against 18). However, when analysing the number of wrong and right predictions, we see that the model made a lot more mistakes than correct predictions.

The next subsections present the same results for each of the other time frames considered by this study.

4.4. Results for the second time frame: the year of 2017

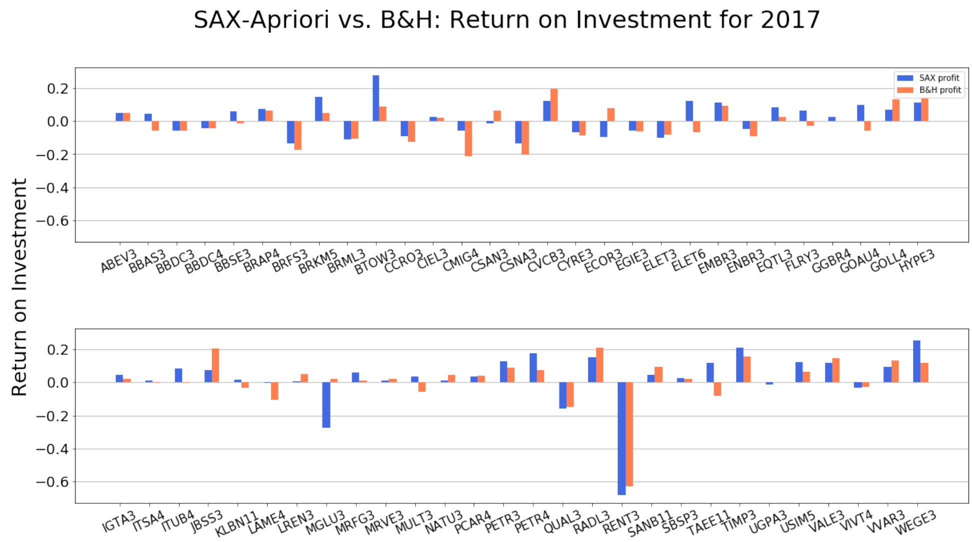

Figure 7. SAX-Apriori vs. B&H: ROI for 2017.

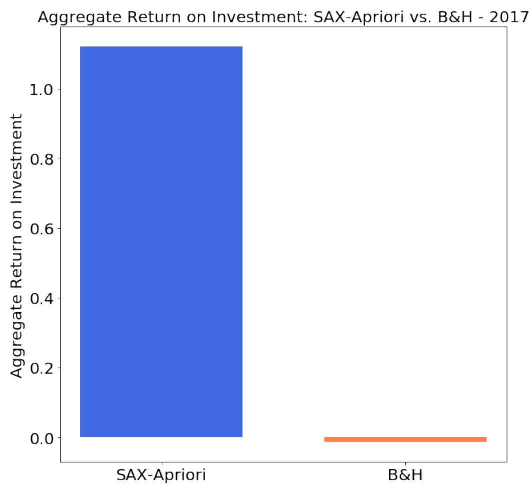

Figure 8. SAX-Apriori vs. B&H: Aggregate ROI for 2017.

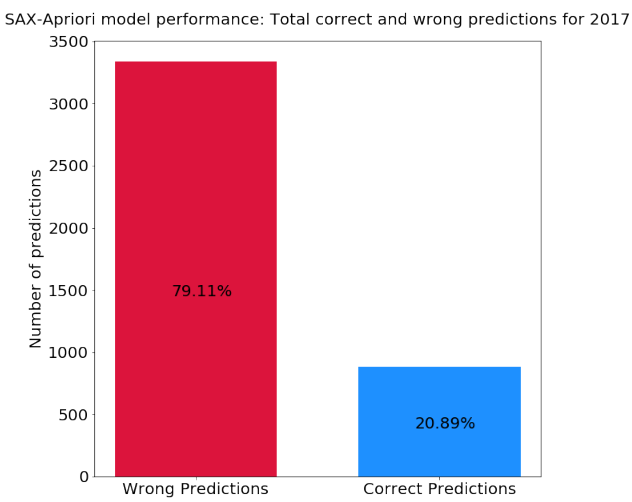

Figure 9. SAX-Apriori performance: Total correct and wrong predictions for 2017.

Once again, for this time frame, we can see that aggregate return for the SAX-Apriori based model was better than the B&H return. Also for the most cases, the SAX-Apriori model, again, performed better (33 agains 24). The high percentage of wrong predictions persisted.

4.5. Results for the third time frame: the year of 2018

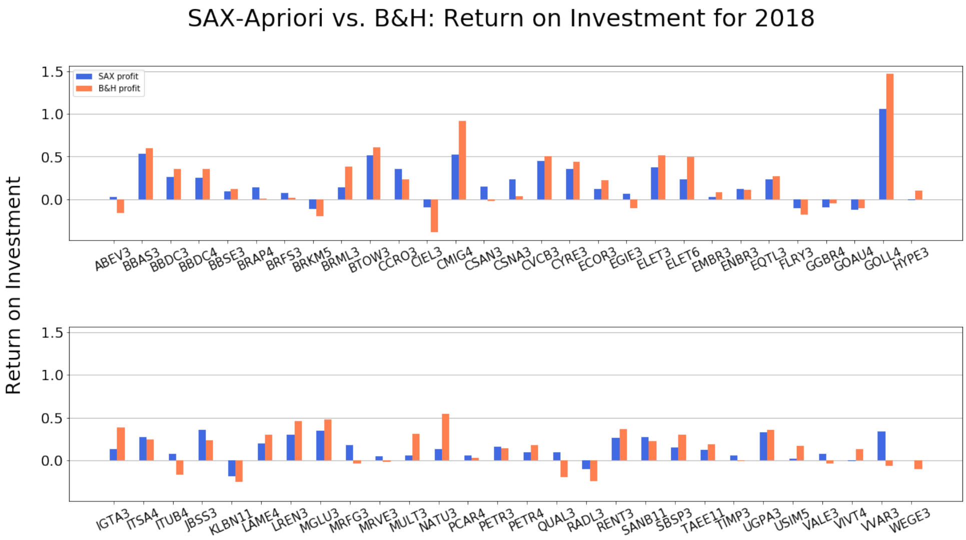

Figure 10. SAX-Apriori vs. B&H: ROI for 2018.

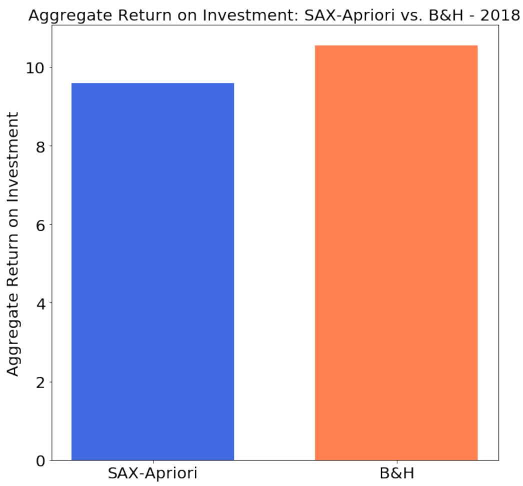

Figure 11. SAX-Apriori vs. B&H: Aggregate ROI for 2018.

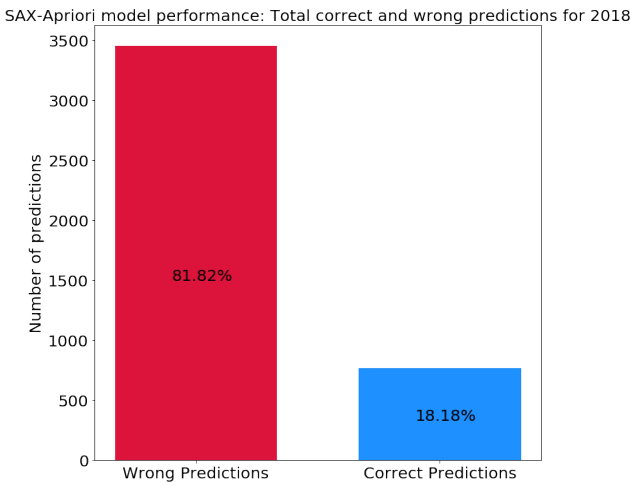

Figure 12. SAX-Apriori performance: Total correct and wrong predictions for 2018.

The results for this time frame were a lot different. We can see that this time the aggregate return for the SAX-Apriori based model was not able to overcome the B&H return. Also, the Buy and Hold strategy performed better for the majority of cases (31 against 26). The high mistake rate for the model persisted nonetheless.

4.6. Results for the fourth time frame: 2016 to 2018

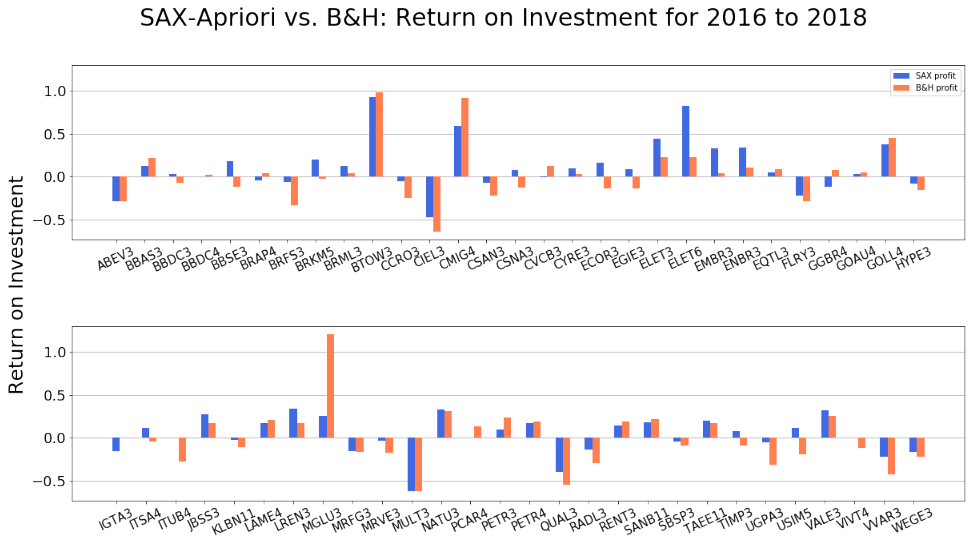

Figure 13. SAX-Apriori vs. B&H: ROI for 2016 to 2018.

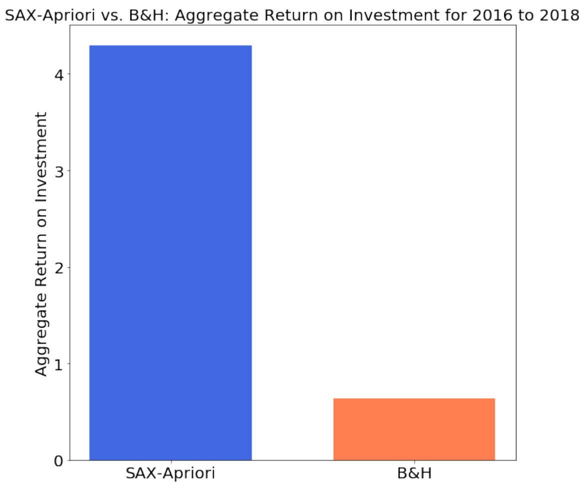

Figure 14. SAX-Apriori vs. B&H: Aggregate ROI for 2016 to 2018.

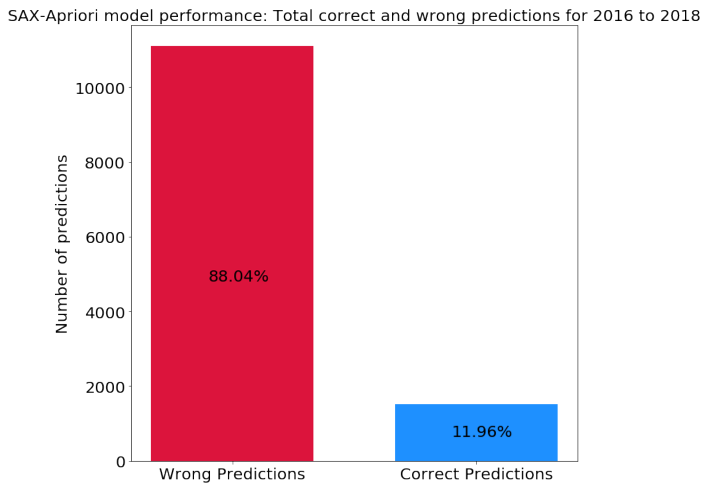

Figure 15. SAX-Apriori performance: Total correct and wrong predictions for 2016 to 2018.

In this time frame we have the most succesfull case for the SAX-Apriori based model over the B&H strategy: the aggregate return was much higher and the model performed better for 73% of the cases (42 agains 15). However, the number of wrong predictions did not fall.

4.7. Best method for each time frame

Finally, the Figure 16, below, shows how many times each strategy performed better accross the considered testing time frames. What is counted here is for how many shares each strategy yielded a greater ROI than the other.

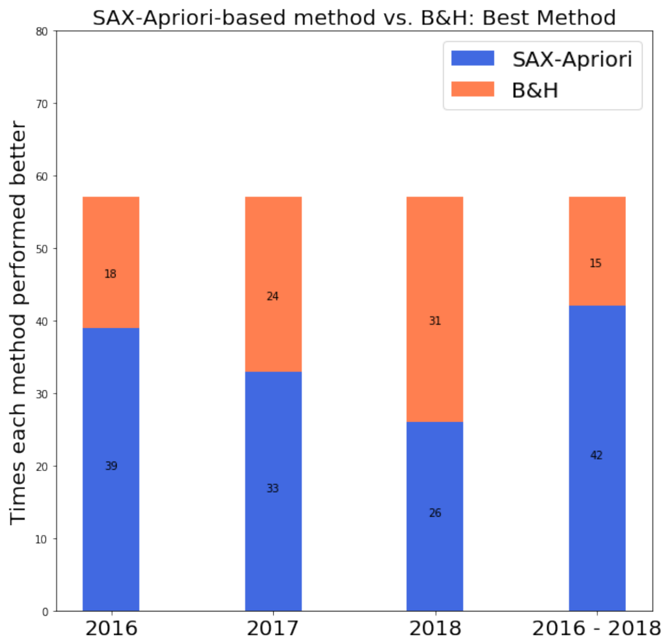

Figure 16. SAX-Apriori vs. B&H: Times each method performed better.

We see that, for 3 out of the 4 studied time frames, trading using the trading rules generated by the implemented model is considered to be the better strategy.

5. Conclusion and Future Works

In the present study we have implemented and evaluated a predictive model, based on Temporal Association Rule Mining, that tries to predict future stock market price ranges with the help of SAX [Lin et al. 2003], Apriori [Agrawal and Srikant 1994] and a Genetic Algorithm optimization. The system was validated using stock prices' temporal series of 57 stocks from the brazilian stock market index Ibovespa, considering 4 different time frames: the years 2016, 2017 and 2018, individually, and also the whole period between 2016 and 2018.

The results showed that the proposed model worked well for many cases and can, indeed, overcome the baseline of the traditional Buy and Hold strategy. However it is also shown that, albeit profitable overall, the model makes a lot of wrong predictions, what is an expected result, given the organic and unpredictable nature of stock prices' time series. Since the model tries to forecast future prices based on past patterns in the data, drastic changes in the market dynamics cannot be previewed, what causes the wrong predictions.

Given the promising results of the implemented model and the drawback of it making so many wrong predictions, what is believed to be due to radical changes in the market's dynamics, for future works it is planned to combine the results of the so far implemented method with a sentiment analysis technique applied to comments from platforms like Twitter and investment forums, with the hopes of improving even further the predictions. It is believed that analysing opinions of a massive number of users from such platforms will help with forecasting the drastic changes in the market dynamics. Thus, a more robust model will be created, providing predictions based on past price patterns and future tendencies, indicated by users, investors, analysts and other relevant people.

References

Agrawal, R., Imielinski, T., and Swami, A. (1993). Mining association rules between sets of items in large databases. pages 207–216.

Agrawal, R. and Srikant, R. (1994). Fast algorithms for mining association rules in large databases. In *Proceedings of the 20th International Conference on Very Large Data Bases*, VLDB '94, pages 487–499, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Ale, J. M. and Rossi, G. H. (2000). An approach to discovering temporal association rules. In *Proceedings of the 2000 ACM Symposium on Applied Computing - Volume 1*, SAC '00, pages 294–300, New York, NY, USA. ACM.

Atsalakis, G. S. and Valavanis, K. P. (2009). Surveying stock market forecasting techniques - part ii: Soft computing methods. *Expert Syst. Appl.*, 36(3):5932–5941.

Bettini, C., Sean Wang, X., Jajodia, S., and Lin, J.-L. (1998). Discovering frequent event patterns with multiple granularities in time sequences. *IEEE Trans. on Knowl. and Data Eng.*, 10(2):222–237.

Brabazon, A. and O'Neill, M. (2006). *Biologically Inspired Algorithms for Financial Modelling*.

Canelas, A., Neves, R., and Horta, N. (2012). A new sax-ga methodology applied to investment strategies optimization. In *Proceedings of the 14th Annual Conference on Genetic and Evolutionary Computation*, GECCO '12, pages 1055–1062, New York, NY, USA. ACM.

Chok, H. and Gruenwald, L. (2009). Spatio-temporal association rule mining framework for real-time sensor network applications. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management*, CIKM '09, pages 1761–1764, New York, NY, USA. ACM.

Cowles, A. (1933). Can stock market forecasters forecast? *Econometrica*, 1(3):309–324.

Cowles, A. (1944). Stock market forecasting. *Econometrica*, 12(3/4):206–214.

Hamilton, J. D. (2018). Why you should never use the hodrick-prescott filter. *The Review of Economics and Statistics*, 100(5):831–843.

Hodrick, R. J. and Prescott, E. C. (1997). Postwar u.s. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1):1–16.

Investing (2019). Investing.com.

Kumar, M. and Kalia, A. (2011). Mining of emerging pattern: Discovering frequent itemsets in a stock data 1 2.

Labiad, B., Berrado, A., and Benabbou, L. (2018). Short term prediction framework for moroccan stock market using artificial neural networks. In *Proceedings of the 12th International Conference on Intelligent Systems: Theories and Applications*, SITA'18, pages 20:1–20:6, New York, NY, USA. ACM.

Lianga, Z., Xinmingb, T., Li-na, L., and Wenlianga, J. (2005). Temporal association rule mining based on t-apriori algorithm and its typical application.

Lin, J., Keogh, E., Lonardi, S., and Chiu, B. (2003). A symbolic representation of time series, with implications for streaming algorithms. In *Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery*, DMKD '03, pages 2–11, New York, NY, USA. ACM.

Lin, J., Keogh, E., Wei, L., and Lonardi, S. (2007). Experiencing sax: A novel symbolic representation of time series. *Data Min. Knowl. Discov.*, 15:107–144.

Mannila, H., Toivonen, H., and Inkeri Verkamo, A. (1997). Discovery of frequent episodes in event sequences. *Data Min. Knowl. Discov.*, 1(3):259–289.

Mao, Y., Wei, W., Wang, B., and Liu, B. (2012). Correlating s&p 500 stocks with twitter data. In *Proceedings of the First ACM International Workshop on Hot Topics on Interdisciplinary Social Networks Research*, HotSocial '12, pages 69–72, New York, NY, USA. ACM.

Medhat, W., Hassan, A., and Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4):1093 – 1113.

Nair, B., Mohandas, V., and Sakthivel, N. (2010). A genetic algorithm optimized decision tree-svm based stock market trend prediction system. *International Journal on Computer Science and Engineering*, 2.

Nair, B. B., Mohandas, V. P., Nayanar, N., Teja, E. S. R., Vigneshwari, S., and Teja, K. V. N. S. (2015). A stock trading recommender system based on temporal association rule mining. *SAGE Open*, 5(2):215824401557994.

Ng, V., Chan, S., Lau, D., and Ying, C. M. (2007). Incremental mining for temporal association rules for crime pattern discoveries. In *Proceedings of the Eighteenth Conference on Australasian Database - Volume 63*, ADC '07, pages 123–132, Darlinghurst, Australia, Australia. Australian Computer Society, Inc.

Nguyen, D., Luo, W., Phung, D., and Venkatesh, S. (2018). Ltarm: A novel temporal association rule mining method to understand toxicities in a routine cancer treatment. *Knowledge-Based Systems*, 161:313–328.

Schumaker, R. P. and Chen, H. (2009). Textual analysis of stock market prediction using breaking financial news: The azfin text system. *ACM Trans. Inf. Syst.*, 27(2):12:1–12:19.

Srisawat, A. (2011). An application of association rule mining based on stock market.

Ting, J., Fu, T.-C., and Chung, F.-L. (2006). Mining of stock data: Intra- and inter-stock pattern associative classification. pages 30–36.

Trnka, A. (2010). Market basket analysis with data mining methods. In *2010 International Conference on Networking and Information Technology*, pages 446–450.

Twitter (2019). Twitter.com.

Wang, G., Wang, T., Wang, B., Sambasivan, D., Zhang, Z., Zheng, H., and Zhao, B. Y. (2015). Crowds on wall street: Extracting value from collaborative investing platforms. In *Proceedings of the 18th ACM Conference on Computer Supported Cooper-*

ative Work & Social Computing, CSCW '15, pages 17–30, New York, NY, USA. ACM.

Wen, F., Zhang, G., Sun, L., Wang, X., and Xu, X. (2019). A hybrid temporal association rules mining method for traffic congestion prediction. *Computers and Industrial Engineering*, 130.

Winarko, E. and Roddick, J. (2007). Armada – an algorithm for discovering richer relative temporal association rules from interval-based data. *Data and Knowledge Engineering*, 63:76–90.