Stability Classification of Stock Market Using Temporal Rule-Based Classification

Polla Fattah
Salahaddin University-Erbil
Uwe Aickelin
University of Nottingham Ningbo China
Christian Wagner
University of Nottingham
Tarik A. Rashid (tarik.ahmed@ukh.edu.krd)
University of Kurdistan Hewler
Jafar
Majidpour
Bryar A.
Charmo University

Research Article

Keywords: rule-based temporal classification, stock market data, time series analysis

Posted Date: August 21st, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2124662/v2

Abstract

Economists have sought to predict stock market prices for decades with varying degrees of success. This study classifies stocks according to their stability in two consequent financial quarters (depending on whether the majority of stocks remain in the same stability group, which can indicate forecastability). However, classifying temporal information like stock market data using available methods produces complicated rules that cannot be easily interpreted by human experts. To reduce this complication a new approach of rule-based temporal classification is used. The method combines human-provided rules with machine optimisation to produce classes that can be easily interpreted by experts, enabling them to comprehend the complicated temporal dimension of the data. Rules provided by human experts might be via generalization of the temporal data using statistical functions like standard deviation and averages. Initially, each rule will have a range of possible values that can be reduced to a single one. For this study, we classify the stability of stock market data of the S&P 500 for two constitutive financial quarters to test if stocks have the same stability or not. The results show that the classes for stock markets are different beyond random chance, which might be an indication of the viability of forecasting stock markets depending on their old trends.

1. Introduction

Many algorithms and methods have been developed to predict stock market prices [1], generally falling on either side of a debate between economists. The first group emphasises the essential randomness of the stock market, thus precluding any possibility of future price prediction-based on historical values [2]. The second group claims market prices have an element of predictability [3]. This paper deploys a method to measure the ability to predict stock markets by classifying stocks according to their price stability. It is proposed that the majority (above 50%) of stock markets should follow the same stability class in at least two constitutive periods to be able to predict their future values.

The proposed hypothesis means if we classify stability of shares of the stock market for two consequent quarters, and 50% of the stocks have the same stability class for both quarters, then we can conclude that stock prices follow a pattern, hence methods of predicting their prices might work, otherwise, less than 50% means that prediction methods might deliver random guesses. However, we should emphasise that this method only tests quarterly prediction, not a next-day prediction. To see whether stocks are following the same stability class or not we harvested stock market prices data of S&P 500 from Yahoo Finance of two consecutive quarters (first and second) of the year 2015.[1] We introduced four classes for stock stability classification: very stable, smooth stable, rough stable, and unstable (according to their volatility and price behaviour over time).

As the harvested stock price records do not have any predefined labels for their stability, a modified version of a rule-based classification method[2,3,4] for temporal data was used [4]. This classification method comprises two main stages: creating rules for classification, and optimizing these rules. Classification rules for this method can be provided by human experts or profiling existing data to see possible limits for classes. Rules are generated using aggregative measures of the temporal data, like their standard deviation, mean, and/or sum of differences. Instead of providing crisp splitting values between classes, the rules can be provided in the form of possible ranges of aggregated values between classes; this might result in a big number of crisp rules. To select the best classifier among all provided ranges of classifiers, an optimization stage follows the first stage, which uses provided classifiers to find the best-compacted item sets for each class.

To measure compactness of classes, distance measures like Euclidean distance or statistical measures like standard deviation and variance can be used to measure the distance of items according to their temporal data. Figure 1 shows the flowchart of the classifier. The benefit of this classification method is that it provides simple rules that are easy to understand and modify by experts of the field on the one hand while taking advantage of the temporal dimension of the data to provide the best possible classifier. In this paper, an enhanced version of this method has been used to be able to deal with the requirements of a relatively large data set. This version uses differential evolution to select the best classifier instead of the brute force method used previously.

A profile for each stock of the S&P 500 was created. The profiles represent the market price of stocks for the first quarter of 2015. A pool of classifiers was produced depending on these profiles and selecting a reasonable range of split values between classes. This range was then optimized to find the best possible classifier among these ranges, then the selected classifier model was used to classify the same stocks for the next financial quarter to compare stocks' stability between these two consecutive quarters. The results showed that there is a significant difference between first and second financial quarters for S&P 500 stocks' stability classes, so it can be concluded according to the underlying data and the used method that the stock market values cannot be predicted by entirely relying on their historical data.

2. Background And Related Works

Classifying stock market data using rule-based temporal classification overlaps with many disciplines, including temporal clustering and classification, rule-based classification, optimization (differential evolution in our case), and stock market forecasting. In the sub-sections below we will briefly explain each of these subjects and critically review related studies.

2.1. Temporal Classification

Classification is a type of supervised machine learning concerned with predicting one of the predefined finite classes for items subject to classification [5]. The temporal and sequence classification is an automatic system that assigns one of the predefined classes to the time series or sequence input [6]. Many temporal classifications have been introduced that reuse traditional classification algorithms using criteria and measurements crafted for temporal data.

Many temporal supervised and unsupervised algorithms use dynamic time warping (DTW) [7] to align between two sequences or time series and find the distance between them. This method was originally used in speech recognition to find human speech patterns [8]. DTW uses a local cost function to compare between two time-series. The operation of time matching between two series using DTW is shown in Figure 2.

Douzal-Chouakria et al. [10] used classification trees to classify time series data by introducing new splits for the tree nodes using time series proximities relying on adaptive metrics considering behaviours and values. The distance-based K-nearest neighbour classification method (KNN) is used with temporal and sequential data with Euclidean distance measure [11]. However, for complex time series, Euclidean distance is sensitive to the time fluctuation, thus DTW has been used [12]. Other methods use Support Vector Machine (SVM) as a temporal data classifier using different kernels [13]. SVM classifies items by separating each class using optimal hyperplanes between them [5].

Model-based classifiers can also be used for temporal and sequential classifications like the Naive Bayes sequence classifier [14] and Hidden Markov Model [15]. In the training step, the parameters of the model are created and trained depending on some assumptions, and a set of parameters describing probability distributions. In the classification step, a new sequence is assigned to the class with the best possible similarity [16].

2.2. Temporal Clustering

Clustering is an unsupervised machine-learning method whose goal is to find natural groupings (clusters) of instances in data sets. All clustering methods strive to detect compacted clusters by maximizing the total sum of inter-cluster distance and minimizing the total sum of the intra-cluster distance be-tween instances [17]. The distance can be measured using Euclidean distance, DTW distance, or any other similarity measures. Jebara et al. [18] used a hidden Markov model (HMM) to cluster time series data, while Oates et al. [15] compared two methods for clustering time-series data sets, first using HMM alone and then using DTW with HMM. DTW returns the minimized area between two time-series variables, which can be used as a similarity measure between the variables. They concluded that using DTW enhances the results of the clusterings of the time series data set.

Rodrigues, Gama, and Pedroso [19] used hierarchical clustering to cluster time series data sets. A hierarchical clustering method works by grouping items into a tree of clusters. The tree can be generated in two ways, either by starting from single items then agglomerating them into a higher structure, or starting from the entire data set and dividing it until ends up with single items in each branch of the tree [20]. Another method used a scaled-up version of DTW [21] with hierarchical clustering, which calculates the distance between temporal variables efficiently and shows the advantage of using DTW with hierarchical clustering. Soheily-Khah et al. [22] proposed k-means-based clustering for temporal data sets using DTW, the Dynamic Temporal Alignment Kernel, and the Global Alignment kernel. Items of a data set are partitioned by K-means clustering, minimizing the total distance of items to a centre of the clusters chosen randomly at the initial stage, but later recalculated iteratively, and items are allocated to the nearest centroid to form clusters with minimum intra-cluster distance [5].

2.3. Rule-Based Classification

Most rule-based classification uses if..else.. form to classify underlying data, which is amenable to easy human comprehension [23]. The rules can be learned through examples or provided by an expert [24]. Many different data mining and analysis methods use rule-based systems for classification, as explained below.

Rule-based classification in fuzzy systems is used, for example, Cordon et al. [25] proposed a new Fussy Reasoning Method (FRM) with better optimization for the system, whereby the rules do not lose their comprehensibility. Ishibuchi [26] compared two kinds of voting schemes for fuzzy rule-based classification. Experts use common sense and vague terms to solve problems and classify situations/items, while an expert system that tries to simulate human experts uses logic to conclude decisions instead of hard programmed solutions [23]. Several expert systems that rely on rule-based logic have been introduced [27]. Many other methods have been introduced that use rule-based systems for classification, like [28], which proposed a generic classifier construction algorithm (ICCA). [29] proposed an algorithm for a rule-based classifier that can extract rules from uncertain data, and [30] used probability estimation for rule learning, inspired by the use of probabilities to construct decision trees.

2.4. Differential Evolution

Differential Evolution (DE) is a heuristic search algorithm introduced by Storn et al. [31], who described it as simple and efficient. DE is a special type of genetic algorithm that uses crossover and mutation while producing the next generation, as it happens according to the nature of DNA and derives natural evolution from creating solutions (species) that are optimized for the environment. This algorithm proved its success and it has been used in many different areas [32]. In this study, we used DE to optimize provided rules by human classifiers. The optimization focuses on minimizing the distance between items within classes using their temporal attributes.

3. Methodology

To test our hypothesis, stock market data for two quarters were collected and pre-processed, then the stocks were profiled to find initial limits for the rule-based classification. We used the first quarter of the stock price to build the final classifier rules, and then these rules were used to classify the first and second quarters. The classification results of these two quarters were compared to find similarities between them. To confirm classification results we clustered the data using temporal and non-temporal clustering and compared between the results of two quarters and results of all of the above methods.

A sample of stock market price data for two consecutive financial quarters was collected. The data were cleaned to contain only continued stocks for all time points (financial days), then the data is normalized to remove the price gap between different stocks, as the focus of the tests is the stability of the prices and not the price magnitude itself. The profile for each symbol is then created to help with the classification process. Profiling stock symbols yields proper statistics about them so that the human

classifier can make the initial decision about class boundaries. The stocks are classified into four classes which can be defined as:

1. Very stable: the price of this class of stocks has a relatively small change over time.
2. Unstable: the price of this class of stocks has a relatively big change over time.
3. Rough stable: the price of this class of stocks is larger than that of the *very stable* class, with a large difference in price between two constitutive days on average.
4. Smooth stable: the price of this class of stocks is larger than a *very stable* class, with a small difference in price between two constitutive days on average.

Two points should be noted from the rules above. First, the rules are prioritised, whereby if any item can fulfill a higher rule then it will follow that classification, even though its variable's values can be true for lower priority classes. Second, terms like *large* and *small* can be vague, however, they are only for definition purposes here, to give a general idea about each term. In practice, more specific values will be provided for each of these terms. These specific values might change from one data set to another, depending on the context and origin of the data.

The first quarter of data is used to optimize the initial rules and find the best crisp rules for classifying the data set, then this classifier is used to classify both first and second quarters to compare stocks' stability classes between these two quarters. The optimization process uses differential evolution to find the most compacted classes that can be produced using one classifier among the pool of provided classifiers in the form of range values separator between classes. Multiple criteria are used to calculate the compactness of the classes, like the Euclidean distance between items and the standard deviation of each class.

The classification results are compared with temporal and non-temporal clustering. For the temporal clustering, we use a hierarchical clustering method with DTW as a measure of dissimilarity to calculate the distance matrix between items. For this test only closed price temporal attribute is used, as it is a single varied method. For non-temporal clustering, some aggregated statistics like standard deviation and the mean of each stock is used with hierarchical clustering using Euclidean distance between items.

4. Tests And Results

To test the predictability of the stock market prices, stock market sample data was harvested, classified, and clustered to find their classification of stability. Detailed steps are explained in the subsections below.

4.1. Data Gathering and Pre-Processing

Sample data of the S&P 500 was collected from Yahoo Finance using an R script to automate downloading data of all stocks from 1-1-2015 to 1-7-2015, which represents half years. The data has been normalized then multiplied by 1000 to create a scale of integer numbers from 1 to 1000 to simplify rules for classification, without losing the precision of the data. The data is split into two financial quarters, the first from 1-1-2015 to 31-3-2015 and the second from 1-4-2015 to 1-7-2015, so that the comparison can be made between these two quarters. A small number of stocks did not have the complete prices set, so that we removed them, which concluded 497 symbols. The available temporal variables of the data for each stock are Open, High, Low, Close, Volume, and Adj. Close. The most used variable for our tests was Close values of symbols. Table 1 shows a sample of the data with its headers after the pre-processing stage.

4.2. Profiling Stocks

Profiling aims to find initial classification rules with a range of overlapping areas to separate classes from each other. To profile stocks, their mean, standard deviation, and difference of price between every constitutive two days are calculated, along with their linear regression. A sample of the visualised profiles is shown in Figure 3. The initial rules for classification were produced by a human classifier with ranges in the forms of [min, max] values compared with an aggregated temporal variable. The used temporal variable in our tests is the Close price of every symbol and their difference between two constitutive days. While the used aggregative function is the standard deviation, the mentioned temporal variables for each symbol separately. Table 2 shows the exact rules which are provided for the classifier to optimize.

Table 1. Sample of the S&P 500 Data Showing Temporal Variables, Date as Linear Integer and Stock Symbols

Date	Open	High	Low	Close	Vol	Adj.Cls	S
1	587	567	489	482	73	473	A
2	440	406	367	351	137	344	A
3	352	322	243	243	141	239	A
4	303	282	292	333	300	327	A
5	426	504	454	539	146	529	A
6	556	508	474	487	87	479	A
7	489	455	412	404	227	397	A

Table 2. Rules

Class	**Rule**
Very Stable	*[1100, 1300]* < SD*(*CLOSE*)&&* *[500, 750]* < SD*(*CLOSEDIF F *)*
Unstable	*[1600, 2000]* > SD*(*CLOSE*)&&* *[650, 1000]* > SD*(*CLOSEDIF F *)*
Rouged Stable	*[550, 800]* > SD*(*CLOSEDIF F *)*
Smooth Stable	All remains

4.3. Classifying Stocks

The initially provided rules can create a pool of crisp classifiers from which the optimiser can choose the best classifier with the minimum distance between items of the same class. The criteria used to measure the distance between items in each class in this experiment are the complete distance between all items, distance of all items to the class centroid, standard deviation, sum of squared error, and interquartile range. Differential evolution is used to optimise the provided rules according to the first financial quarter of 2015 for S&P 500 stocks. After this process, the rules have specific values as limits between classes instead of the range of values. An example of the generated rules is shown in Table 3. It can be seen that the resultant rules have exact values for comparison rather than a range of values as in Table 2.

Table 3. Generated Rules by Differential Evolution Using SSE As Distance Measure

Class	**Rule**
Very Stable	$1218 < SD(CLOSE)\&\&$ $620 < SD(CLOSEDIFF)$
Unstable	$1939 > SD(CLOSE)\&\&$ $948 > SD(CLOSEDIFF)$
Rouged Stable	$762 > SD(CLOSEDIFF)$
Smooth Stable	All remains

After the optimization stage the produced rule is used to classify first and second quarters of S&P 500 stocks, then items classes are compared between these two quarters to find the percentage of stock symbols maintained in the same classes in both quarters. The comparison and number of items in each class for both quarters produced by compactness measures are shown in Table 4. Each compactness measure imposes different classes on the items. The distinctive measure is the complete distance, as it behaves uniquely and ignores one of the classes altogether, with fewer classes so that it presents the highest percentage of items with the same class between two quarters. However, all of the comparison results between the two quarters are less than 50%, which might mean that stock symbols have different stability behaviour in two constitutive financial quarters.

Table 4. Number of Items in Each Class and Percentage of Compatible Results between Two Quarters for Different Compactness Measures

Compactness Measuring	Quarters	Number of items in classes				Quarters Agr %
		VS	US	RS	SS	
centroidDist	1st Qt	63	291	0	143	45%
	2nd Qt	159	165	0	173	
completeDist	1st Qt	120	123	128	126	34%
	2nd Qt	242	43	83	129	
SD	1st Qt	42	172	99	184	34%
	2nd Qt	112	59	79	247	
SSE	1st Qt	110	141	132	114	33%
	2nd Qt	232	48	90	127	
Quantile	1st Qt	87	245	30	135	37%
	2nd Qt	197	119	19	162	

5. Clustering Stock Market Data

To confirm the results of the rule-based classification, the same temporal and aggregated variables used for classification are reused for clustering, as K-means is used twice to cluster the stocks in both financial quarters into four clusters, first according to the aggregated attributes (non-temporal) and then according to the closed price (temporal). Moreover, hierarchical clustering using DTW for measuring dimensions between items is used with the closed price attribute of the data. To have comparable results with the classification, for the non-temporal K-means clustering standard deviation of close price and standard deviation of the difference between two days price is used, the same non-temporal attributes are used to generate rules for the previously described classification. Moreover, the temporal attribute which is used for both temporal clusterings K-means and hierarchical is closed value, because this variable is used to optimize rules of classification. The closed temporal variable is transposed to create multiple attributes for each stock so that it will be possible to use these clustering algorithms.

The results of all clusterings, including the number of items in each cluster, and the percentage of agreements between two quarters using each method are shown in Table 5. However, the percentage of agreement between clusters might not be sufficient for clustering due to the risk of instability in cluster labels, so the Jaccard index and Folkes-Mallows FM index are used to measure the similarities between clusterings of two financial quarters for each clustering algorithm. By comparing the results of all clustering methods using three measures (similarity percentage, Jaccard and FM indices), all the results indicate that the similarity of stock prices' stability between the first and second quarter are less than 50%, which might indicate different stability behaviour for each stock price.

Table 5. Temporal and Non-Temporal Clustering Results of Two-Quarters of Stock Market Data

Method	Qt	Cluster item count				Jac	FM	%
		Cl1	Cl2	Cl3	Cl4			
K-mean Agri	Qt1	89	81	176	151	0.17	0.30	27%
	Qt2	175	34	182	106			
K-means Closed	Qt1	141	155	83	118	0.28	0.43	16%
	Qt2	53	138	164	142			
DTW Closed	Qt1	103	134	191	69	0.31	0.48	47%
	Qt2	102	251	113	31			

6. Conclusion

This paper has proposed a method to contribute to the discussion about the feasibility of predicting stock market prices. The proposed method consists of classifying the price stability of two constitutive periods, then comparing the items' class membership between these two periods to test the consistency of items' class stability in two different periods. A high percentage of items' class similarities between two periods

can be considered to be a predictor of stock market data. To test this method a sample of S&P 500 stock market data was classified with the temporal rule-based classification method after cleaning and profiling stocks for creating rules. To confirm the results of the classification the same data and items aggregated statistics were clustered using K-means and hierarchical clustering with DTW. The results for both classification and clustering conclude that the stability classes of stocks are inconsistent between the two periods, which could suggest a lack of ability to predict stock market prices, however, this test and its results do not suggest the impossibility of predicting the very next time point of the stocks prices; rather it indicates the difficulty of predicting the next period of stock prices and how these stocks will perform for the entire duration of the next period, such as a financial quarter.

Declarations

1. The code to harvest and clean data with the data itself is available at https://goo.gl/U0STqJ
2. The code for profile creator and the classifier is available at https://goo.gl/xiRcF8
3. https://github.com/pollaeng/RuleBased_Classification
4. https://github.com/pollaeng/Calculate_Cluster_Change_Overtime

Funding (The is not applicable)

Conflicts of interest/Competing interests (There is no conflict of interest is found**)**

Availability of data and material (The links for the data are given**)**

Code availability (The link for the software is given**)**

References

[1] G. Atsalakis and Valavanis K., “Surveying stock market forecasting techniques - Part I: Conventional methods,” Zopounidis C., Computa-tion Optimization in Economics and Finance Research Compendium,, no. August, pp. 49–104, 2013.

[2] E. F. Fama, “The Behavior of Stock-Market Prices,” p. 34, 1965.

[3] A. W. Lo and A. C. MacKinlay, “Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test,” Review of financial studies, vol. 1, no. 1, pp. 41–66, 1988.

[4] P. Fattah, U. Aickelin, and C. Wagner, “Optimising Rule-Based Classification in Temporal Data,” in 1st International Conference on Engineering and Innovative Technology, SU-ICEIT 2016, 2016, p. 8.

[5] M. J. Zaki and M. J. Meira, Data Mining and Analysis: Fundamental Concepts and Algorithms. New York: Cambridge University Press, 2013.

[6] S. Laxman and P. Sastry, “A survey of temporal data mining,” Sadhana, vol. 31, no. April, pp. 173–198, 2006.

[7] D. Berndt and J. Clifford, “Using dynamic time warping to find patterns in time series,” Workshop on Knowledge Knowledge Discovery in Databases, vol. 398, pp. 359–370, 1994.

[8] L. Rabiner and B.-H. Juang, “Fundamentals of speech recognition,” 1993.

[9] M. Muller,¨ Information retrieval for music and motion, 2007.

[10] A. Douzal-Chouakria and C. Amblard, “Classification trees for time series,” Pattern Recognition, vol. 45, pp. 1076–1091, 2012.

[11] L. Wei and E. Keogh, “Semi-supervised time series classification,” Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD ’06, p. 748, 2006.

[12] L. Kajan, A. Kertesz-Farkas, D. Franklin, N. Ivanova, A. Kocsor, and S. Pongor, “Application of a simple likelihood ratio approximant to protein sequence classification,” Bioinformatics, vol. 22, no. 23, pp. 2865–2869, 2006.

[13] R. Sitaram, H. Zhang, C. Guan, M. Thulasidas, Y. Hoshi, A. Ishikawa, K. Shimizu, and N. Birbaumer, “Temporal classification of mul-tichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface,” NeuroImage, vol. 34, no. 4, pp. 1416–1427, 2007.

[14] V. S. Tseng and C. H. Lee, “Effective temporal data classification by integrating sequential pattern mining and probabilistic induction,” Expert Systems with Applications, vol. 36, no. 5, pp. 9524–9532, 2009.

[15] T. Oates, L. Firoiu, and P. Cohen, “Clustering time series with hidden Markov models and dynamic time warping,” Proceedings of the IJCAI-99 workshop on neural, symbolic and reinforcement learning methods for sequence learning, pp. 17–21, 1999.

[16] Z. Xing, J. Pei, and E. Keogh, “A brief survey on sequence classifi-cation,” ACM SIGKDD Explorations Newsletter, vol. 12, no. 1, p. 40, 2010.

[17] P. Esling and C. Agon, “Time-series data mining,” ACM Computing Surveys (CSUR), vol. 45, no. 1, pp. 1–34, 2012.

[18] T. Jebara, Y. Song, and K. Thadani, “Spectral Clustering and Em-bedding with Hidden Markov Models,” 18th European Conference on Machine Learning, ECML 2007, Proceedings of, vol. 4701, pp. 164–175, 2007.

[19] P. P. Rodrigues, J. Gama, and J. P. Pedroso, “Hierarchical clustering of time-series data streams,” IEEE Transactions on Knowledge and Data Engineering, vol. 20, no. 5, pp. 615–627, 2008.

[20] T. Warren Liao, “Clustering of time series data - A survey,” Pattern Recognition, vol. 38, no. 11, pp. 1857–1874, 2005.

[21] E. J. Keogh and M. J. Pazzani, “Scaling up dynamic time warping for datamining applications,” Knowledge discovery and data mining, vol. In 6th ACM, pp. 285–289, 2000.

[22] S. Soheily-Khah, A. Douzal-Chouakria, and E. Gaussier, “Generalized k-means-based clustering for temporal data under weighted and kernel time warp,” Pattern Recognition Letters, vol. 75, pp. 63–69, 2016.

[23] M. Negnevitsky, Artificial Intelligence, 2002.

[24] L. X. Wang and J. M. Mendel, “Generating fuzzy rules by learning from examples,” IEEE Transactions on Systems, Man and Cybernet-ics, vol. 22, pp. 1414–1427, 1992.

[25] O. Cordon, M. J. del Jesus, and F. Herrera, “A proposal on reasoning methods in fuzzy rule-based classification systems,” International Journal of Approximate Reasoning, vol. 20, no. 1, pp. 21–45, 1999.

[26] H. Ishibuchi, T. Nakashima, and T. Morisawa, “Voting in fuzzy rule-based systems for pattern classification problems,” Fuzzy Sets and Systems, vol. 103, no. 2, pp. 223–238, 1999.

[27] E. Styvaktakis, M. H. J. Bollen, and I. Y. H. Gu, “Expert system for classification and analysis of power system events,” IEEE Transac-tions on Power Delivery, vol. 17, no. 2, pp. 423–428, 2002.

[28] A. Giacometti, E. Miyaneh, P. Marcel, and A. Soulet, “A generic framework for rule-based classification,” Proceedings of LeGo, pp. 37–54, 2008.

[29] B. Qin, Y. Xia, S. Prabhakar, and Y. Tu, “A rule-based classification algorithm for uncertain data,” Proceedings - International Conference on Data Engineering, pp. 1633–1640, 2009.

[30] J.-n. Sulzmann and F. Johannes, “A Study of Probability Estimation Techniques for Rule Learning,” 2009.

[31] R. Storn and K. Price, “Differential evolution a simple and efficient heuristic for global optimization over continuous spaces,” Journal of global optimization, vol. 11, pp. 341–359, 1997.

[32] A. Application, N.-c. Portfolio, D. Ardia, K. Boudt, P. Carl, M. Katharine, and B. G. Peterson, “Differential Evolution with DE-optim,” pp. 27–34, 2011.

Figures

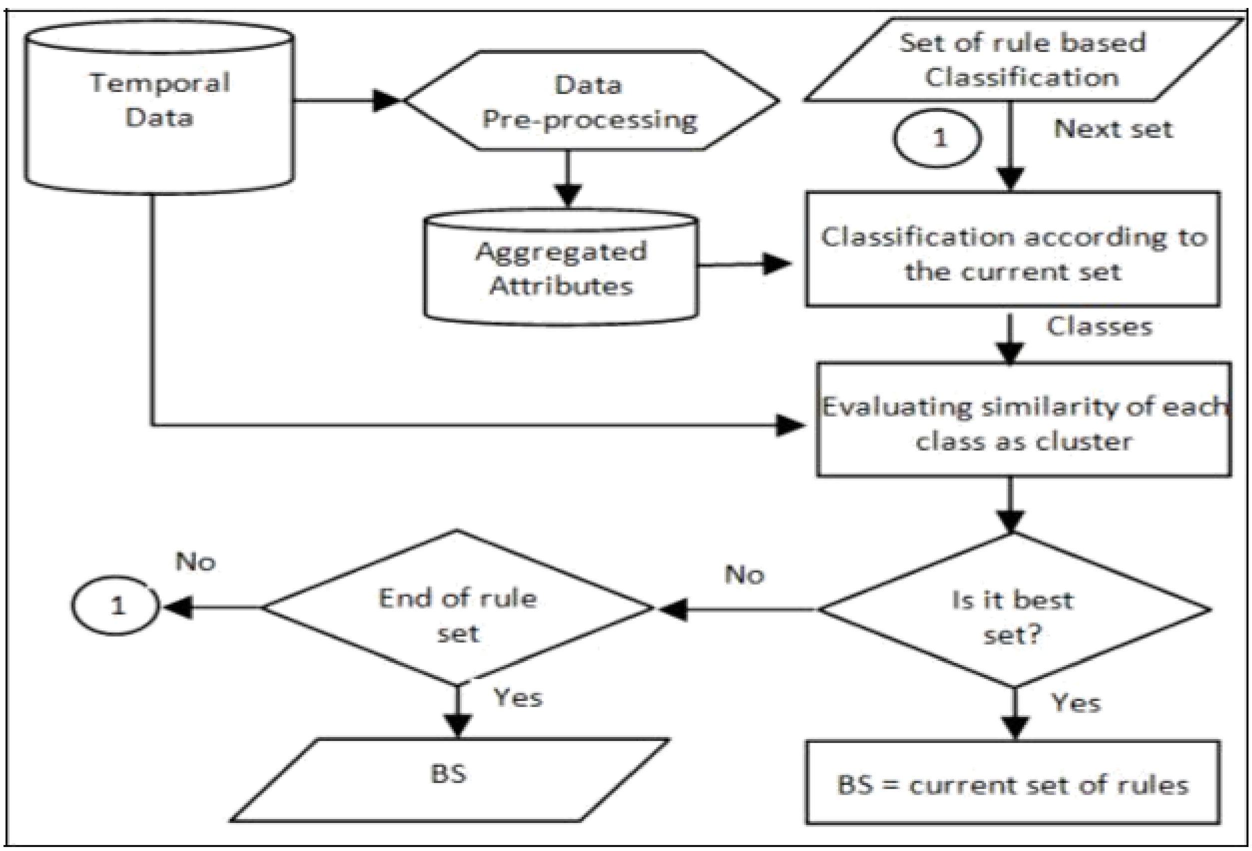

Figure 1

Flowchart for the rule-based classifier.

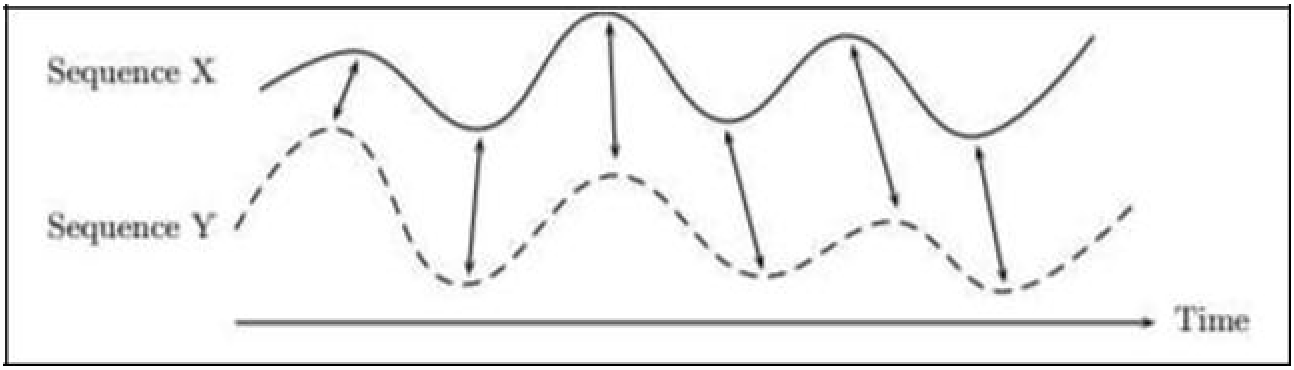

Figure 2

Time alignment of two-time series using DTW. Aligned points are indicated by arrows [9].

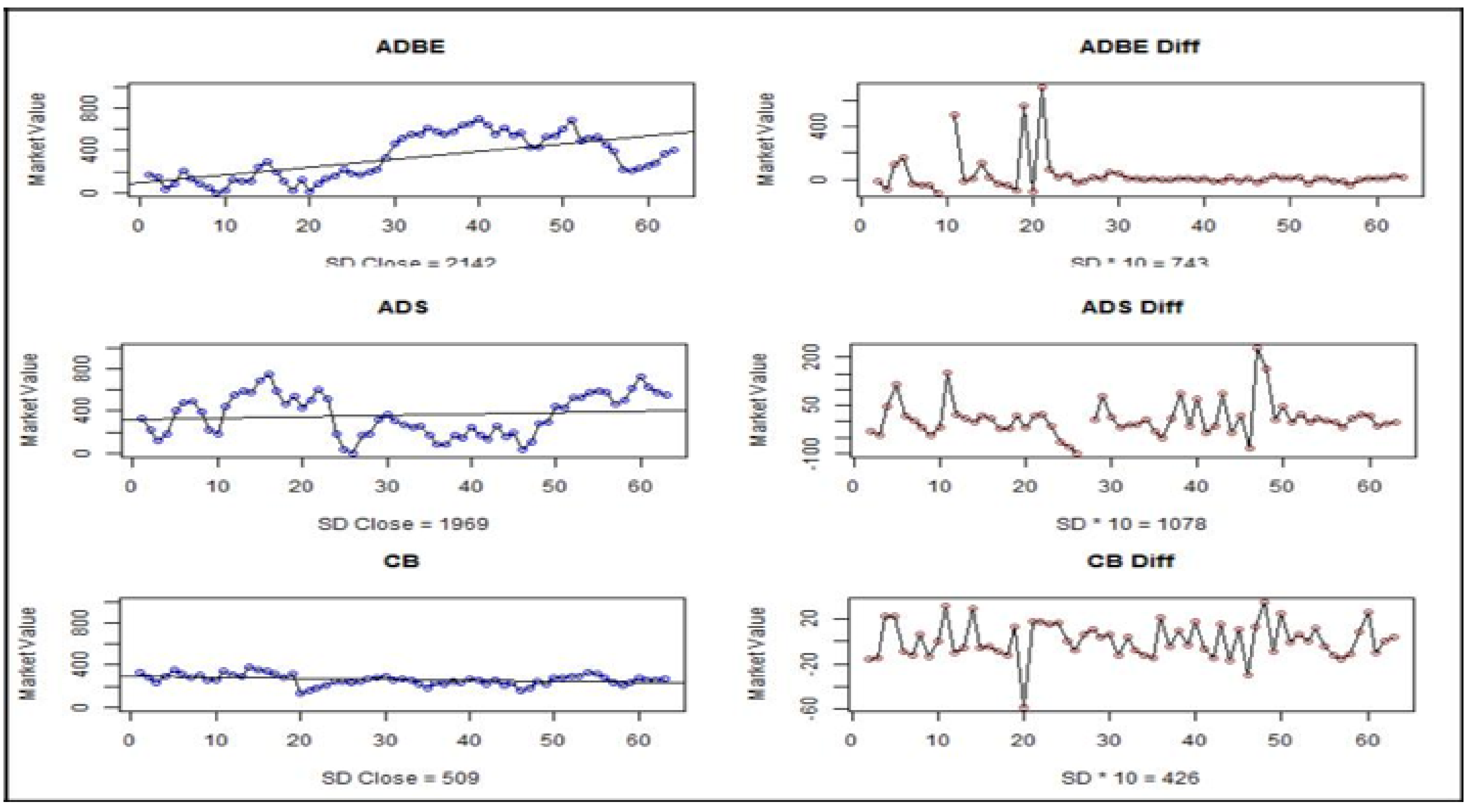

Figure 3

Sample of visualized profiles for stock market symbols