Expert Systems With Applications 70 (2017) 20–36

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Clustering stock price time series data to generate stock trading recommendations: An empirical study

Binoy B. Nair[a,*], P.K. Saravana Kumar[a], N.R. Sakthivel[b], U. Vipin[a]

[a] Department of Electronics and Communication Engineering, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, Amrita University, India

[b] Department of Mechanical Engineering, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, Amrita University, India

ARTICLE INFO

Article history:
Received 30 May 2015
Revised 17 August 2015
Accepted 1 November 2016
Available online 2 November 2016

Keywords:
Stock
Trading
Recommender
Clustering
Time-series

ABSTRACT

Predicting the stock market is considered to be a very difficult task due to its non-linear and dynamic nature. Our proposed system is designed in such a way that even a layman can use it. It reduces the burden on the user. The user's job is to give only the recent closing prices of a stock as input and the proposed Recommender system will instruct him when to buy and when to sell if it is profitable or not to buy share in case if it is not profitable to do trading. Using soft computing based techniques is considered to be more suitable for predicting trends in stock market where the data is chaotic and large in number. The soft computing based systems are capable of extracting relevant information from large sets of data by discovering hidden patterns in the data. Here regression trees are used for dimensionality reduction and clustering is done with the help of Self Organizing Maps (SOM). The proposed system is designed to assist stock market investors identify possible profit-making opportunities and also help in developing a better understanding on how to extract the relevant information from stock price data.

1. Introduction

The price of a stock keeps changing depending on the supply and demand of the stocks. Due to the extremely nonlinear nature of stock price movements, forecasting the stock prices and timing the buy/sell decisions becomes an extremely challenging task. It is this risk that tends to keep a vast majority of people from trading in stocks. In this study, a stock trading recommender system that learns patterns from the historical stock price data and recommends when to buy/sell stocks and thus, can help the laypeople invest in equity markets profitably is proposed.

Initial studies carried out in the past century questioned the very predictability of stock markets (Cowles, 1933), (Cowles, 1944). Several other studies claimed the stock price movements to be random (Cootner, 1964), (Fama, 1965), with the efficient market hypothesis (Fama, 1970) ruling out any possibility of making excess returns from the market. However, recent studies, e.g. (Atsalakis & Valavanis, 2009), (Nair & Mohandas, 2015a), (Brabazon, O' Neill, & Dempsey, 2008), (Brabazon & O' Neill, 2006) and (Nair & Mohandas, 2015), have indicated that it is in fact possible to forecast the stock price movements and use the forecasts to generate excess returns. It is also observed that application of data mining techniques tend to generate good forecasting accuracy, as in (Nair et al., 2011), (Atsalakis & Valavanis, 2009), (Nair, Mohandas, & Sakthivel, 2011) and (Nair, Minuvarthini, Sujithra, & Mohandas, 2010). Different aspects of data mining have been explored in the earlier attempts at design of stock trading recommender, e.g. in (Nair & Mohandas, 2015b), a classifier based recommender system is proposed while (Nair et al., 2015) proposes a recommender based on mining of temporal association rules. However, it was observed that clustering based stock trading recommender systems have not been explored. In the present study, an attempt has been made to design a clustering based stock trading recommender system that can employ the historical stock price data to generate buy/sell recommendations.

Technical analysis has traditionally been very popular with stock traders and is still widely used to forecast stock price movements. Technical analysis involves identification of future trends in stock price movements based on the historical stock price values. There are a large variety of technical indicators in use today (Eng, 1988), however, selection of the optimal set of technical indicators for the given market conditions and identification of the optimal technical indicator parameters is quite challenging, thus limiting their utility. There have been a few studies that attempt to identify the optimal technical indicator parameters, such as in (De Brito &

* Corresponding author.
E-mail addresses: b_binoy@cb.amrita.edu (B.B. Nair), pksaravanakumar@gmail.com (P.K.S. Kumar), nr_sakthivel@cb.amrita.edu (N.R. Sakthivel), vipunne@gmail.com (U. Vipin).

http://dx.doi.org/10.1016/j.eswa.2016.11.002
0957-4174/

Oliveira, 2012), (Bodas Sagi, Soltero, Hidalgo, Fernández, & Fernández, 2012), (Soltero, Bodas, Hidalgo, Fernández, & De Vega, 2012) and (Briza & Naval, 2011), however, selection of the relevant indicators themselves is an issue. There are very few studies, e.g. (Nair & Mohandas, 2015a), which take this aspect into account. The second most common approach followed, is to employ statistical techniques such as Autoregressive (AR) and Generalized Autoregressive Conditional Heteroskedasticity models or simple linear regression (Clare & Miffre, 1995), (McKibben, 1972) models for forecasting the stock prices or stock price returns. These techniques have been found to offer poor forecast accuracy when compared to more recent soft computing based forecasting techniques. Soft computing based techniques are now being widely used for forecasting of nonlinear time series, as seen from (Atsalakis & Valavanis, 2009) and (Nair & Mohandas, 2015a). Hence, soft computing based techniques appear to be highly suitable for the proposed stock trading recommender system.

In the present study, separation of trend and cyclic components of the stock price time series is carried out as a preprocessing step. As suggested in (Hodrick & Prescott, 1997), the stock price time series can be considered to be made up of a (linear) trend component and a (nonlinear) cyclic component. Since the cyclic component governs the short-term price fluctuations, an accurate forecast of the cyclic component is essential for a trader wishing to profit by executing short-term trades. It has also been proven based on the theory of dynamical systems that a system exhibiting chaotic behaviour in time domain tends to behave deterministically in its phase-space representation (Huffake, 2010). This property has been successfully utilized in (Wan & Chai, 2014), (Nair et al., 2011), (Zhang & Li, 2010) etc. and it was reported to generate better results. In the present study as well, rather than using the stock price time series as-it-is, the phase space representation of the cyclic component (of the stock price time series) is used with a view to improving the overall accuracy of the system.

Two parameters to be identified for generating the phase space representation of a time series are the delay (d) and the embedding dimension (M). The most common approach to determining d is the technique proposed by Fraser and Swinney in (Fraser & Swinney, 1986), while False Nearest Neighbours (FNN) technique, proposed in (Kennel, Brown, & Abarbanel, 1992) is widely used for identifying M. In this study, however, a novel technique for identification of d and M using regression trees (described in detail in Section 2.2.2), has been presented and empirically validated. The proposed technique has also been compared to the traditional method proposed in (Fraser & Swinney, 1986).

The second novelty in the proposed stock trading recommender systems is to employ clustering algorithms to identify similarity short term price movements. Traditionally, time series clustering algorithms like Symbolic Aggrigate Approximation (SAX) (Lin, Keogh, Wei, & Lonardi, 2007) have been used for the purpose, e.g. in (Nair, Xavier, Mohandas, Anusree, & Kumar, 2014) and (Canelas, Neves, & Horta, 2013), however, there appears to be very little available literature that explores the possibility of employing clustering algorithms to time series data. In this study, a novel technique of transforming the time series data and then clustering it in a higher dimensional space to find similarities in the stock price movements over time, is employed for generating trading recommendations.

Clustering algorithms are very widely used in data mining with k-means clustering algorithm and its variants being the most popular of the clustering algorithms. Though better than the simple hierarchical clustering algorithms such as the single and complete linkage algorithms, k-means clustering technique is unable to yield good quality clusters with high dimensional data (Xu & Wunsch II, 2005). Neural network based clustering algorithms such as Self Organising Maps (SOM) appear to be better suited in this regard. In (Bação, 2005), performance of k-means clustering and SOM was compared. It was observed that the mean quadratic error, the standard deviation of the quadratic error and the structural error were lesser for SOM when compared to k-means clustering. SOMs have been used for predicting stock prices (Afolabi & Olude, 2007), (Sugunsil & Somhom, 2009), stock picking (Khan, Bandopadhyaya, & Sharma, 2008), identification of liquid stocks (Widiputra & Christianto, 2012), trading preference discovery (Tsai, Lin, & Wang, 2009) and study of stock price bubbles (Gao & Xu, 2009). SOMs have also been combined with SVMs (Ismail, Shabri, & Samsudin, 2011), (Hsu, Hsieh, Chih, & Hsu, 2009), Genetic programming (Hsu, 2011), wavelets (Li & Kuo, 2008) for better performance. It was observed from the literature that SOM based systems can be successfully used for mining information from financial data. Hence, in the present study, SOMs have been employed for clustering the stock price data. The technique is described in Sections 2.4 and 2.5 below.

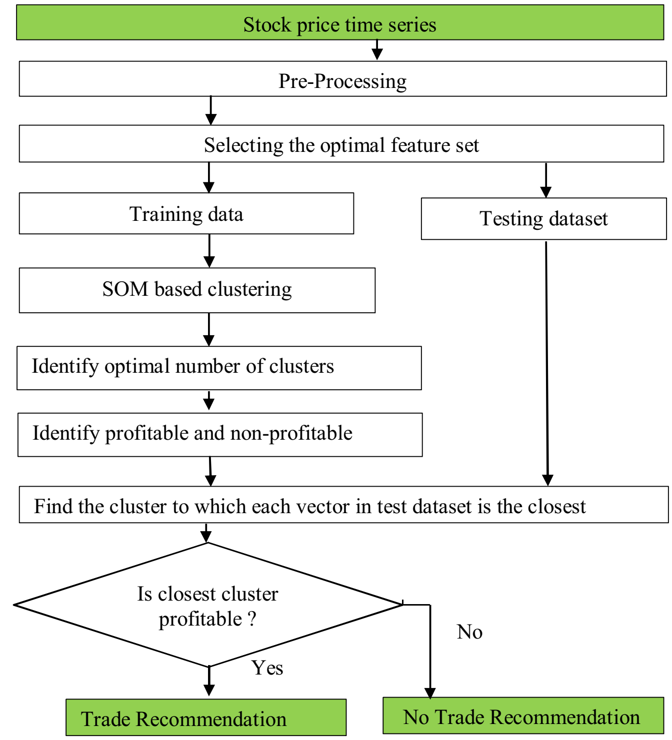

Fig. 1. Top level block diagram of the proposed clustering based stock trading recommender system.

Once the time series data points are clustered, the clusters are identified as profitable or non-profitable, based on the technique as detailed in Section 2.6. For every new data point obtained, the closest cluster to the point is identified and the proposed recommender system generates 'trade' (buy/sell) or 'no-trade' (hold) recommendations depending on whether the point belongs to profitable or non-profitable cluster, as explained in Section 2.7.

The top level block diagram of the proposed system is given in Fig. 1.

Based on the techniques employed for pre-processing, optimal feature set selection and the training and testing techniques used, a total of sixteen different variants of the proposed recommender system are generated. In order to establish the efficacy of the proposed system, all sixteen variants are validated on stocks drawn from US, UK, Brazilian and Indian stock markets (stocks are selected based on the additional criterion that they must belong to different sectors of industry) using eight different performance metrics.

Rest of the paper is organized as follows: Section 2 presents the system description including the working of the proposed system and its components. Section 3 presents the results and the conclusions are presented in Section 4.

2. System description

Major components of the proposed recommender systems are described in the sections below.

2.1. Preprocessing

Consider the time series

$$\boldsymbol{Y} = \{y_t,\ y_{t-1},\ y_{t-2},\ y_{t-3},\ \ldots,\ y_{t-N+1}\} \tag{1}$$

Each sample in $\boldsymbol{Y}$ is made up of a trend (growth) and a cyclic component such that $\boldsymbol{Y} = \boldsymbol{C} + \boldsymbol{G}$ where:

$$\boldsymbol{C} = \{c_t,\ c_{t-1},\ c_{t-2},\ c_{t-3},\ \ldots,\ c_{t-N+1}\},\quad (\boldsymbol{C}) \to 0,\ \ 0 \le i \le N-1 \tag{2}$$

$$\boldsymbol{G} = \{g_t,\ g_{t-1},\ g_{t-2},\ g_{t-3},\ \ldots,\ g_{t-N+1}\} \tag{3}$$

The trend and the cyclic components of the time series are separated using Hodrick–Prescott (HP) filter (Hodrick & Prescott, 1997). The HP filter accomplishes this by attempting to solve

$$\min_{\boldsymbol{G}} \left\{ \sum_{i=0}^{N-1} {c_{t-i}}^2 + \lambda \sum_{i=1}^{N-2} \left(g_{t-i+1} - 2g_{t-i} + g_{t-i-1}\right)\right)^2 \right\} \tag{4}$$

In the present study, the λ value is chosen to be 1600 as suggested by Hodrick and Prescott in (Hodrick & Prescott, 1997).

Two variants are evaluated. The first is to perform no preprocessing, i.e., the stock price time series itself is used as input while in the second case, the cyclic component of the stock price time series (obtained using the HP filter, described previously) is used as the input.

2.2. Selecting the optimal feature set

It was observed from the survey that the stock prices tend to be chaotic in nature (Chen, 1996), (LeBaron, 1994). According to the theory of dynamical systems, a system that exhibits chaotic behaviour in time domain tends to demonstrate deterministic behavior in its phase-space representation (Huffake, 2010) .In the present study, an attempt has been made to use the state space representation of stock price time series data to generate the input features for the proposed recommender system. The following sections represent the two techniques employed for the purpose.

2.2.1. Identification of optimal feature set from delay and embedding dimension

Mutual information was found to be more suitable for finding the optimal delay for short time series for reconstruction of a chaotic attractor in (Zduniak & Lusakowski, 1995) than the traditional autocorrelation function. (Fama, 1965). The technique proposed by Fraser and Swinney in (Fraser & Swinney, 1986) is used to identify the optimal delay.

The mutual information between $\boldsymbol{Y}$ and its lagged version with lag L denoted by $\boldsymbol{Y_L}$ is given by

$$I(\boldsymbol{Y}, \boldsymbol{Y_L}) = \sum_{i=0}^{N-L} \sum_{j=L}^{N} p(y_{t-i}\ , y_{t-j}) \log \frac{p\big(y_{t-i},\ y_{t-j}\big)}{p(y_{t-i})\ p\big(y_{t-j}\big)} \tag{5}$$

Here *p(.)* denotes the probability. The mutual information for different lags is calculated and the lag or delay at which the first minimum of this mutual information is obtained, is taken as the optimal delay value d.

The Takens theorem (Takens, 1981) presents the conditions under which a smooth attractor can be generated. Taken proposed in (Takens, 1981) that for an optimal value of embedding dimension, M, and optimal delay d, the phase space representation of a system with chaotic dynamics, will tend to behave deterministically. Once the optimal delay d has been obtained using the technique proposed in (Fraser & Swinney, 1986), the optimal embedding dimension is then identified using the False Nearest Neighbours (FNN) technique, proposed by Kennel et al. in (Kennel et al., 1992). In an M-dimensional space, a point $k(t)$ is taken, a neighbor $k(i)$ is determined which will satisfy the condition $||k(i)\ -\ k(t)||\ <\ \epsilon$. Here, ϵ is a small constant, which is no greater than the standard deviation of the data (Perc, 2006) and $k(t) = \{y_t,\ y_{t-d},\ y_{t-2d},\ y_{t-3d},\ \cdots,\ y_{t-(M-1)d}\}$ and $k(i) = \{y_{t-i},\ y_{t-i-d},\ y_{t-i-2d},\ y_{t-i-3d},\ \cdots,\ y_{t-i-(M-1)d}\}$, $0 < i \le N-(M-1)d$. The distance R_i between the $(M+1)$st embedding coordinate of points $k(t)$ and $k(i)$ are calculated using the equation:

$$R_i = \frac{|y_{i+Md} - y_{t+Md}|}{\|k(i)\ -\ k(t)\|} \tag{6}$$

2.2.2. Identification of optimal feature set using regression trees

Regression trees are typically constructed by recursively splitting (partitioning) the training sample into smaller subsets in such a way that each split improves the 'purity' of the subsets generated (Torgo, 1999). Regression tree considered in the present study attempts to minimize the Mean Squared Error (*MSE*) between the actual and predicted value.

The process of partitioning stops when any of these conditions are satisfied:

1. The node (l) is 'pure' i.e. $MSE_l < Err_{tol}.MSE_{total}$

   Where
   MSE_l is the MSE for the observed response in l,
   MSE_{total} is the total MSE for the observed response in the entire data
   Err_{tol} is the tolerance on quadratic error per node.

2. l has fewer than Min_Parents data points.
3. Any split on l generates children with fewer than Min_Leaves number of data points.

Min_Leaves – Each leaf has at least Min_Leaves observations. The present study takes it as 1.

Min_Parents – Each branch node in the tree has at least Min_Parents observations. The present study takes it as 10.

In the present study, the predictor variables are the stock prices for the past 30 days and the target *duration* is the 1–5 day ahead to 1–20-day ahead stock prices. Representing the predictor as

$$\boldsymbol{X} = \begin{bmatrix} y_t & y_{t-1} & y_{t-2} & \cdots & y_{t-29} \\ y_{t-1} & y_{t-2} & y_{t-3} & \cdots & y_{t-30} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{t-N+30} & y_{t-N+29} & y_{t-N+28} & \cdots & y_{t-(N-1)} \end{bmatrix}$$

and the target vector as

$$\boldsymbol{t}_{duration} = \begin{bmatrix} y_{t+duration} \\ y_{t+duration-1} \\ \vdots \\ y_{t-N+30+duration} \end{bmatrix}$$

Here, $1 \le duration \le max\_duration$ where $max\_duration = 5$ in the first variant and $max\_duration = 20$ in the second variant.

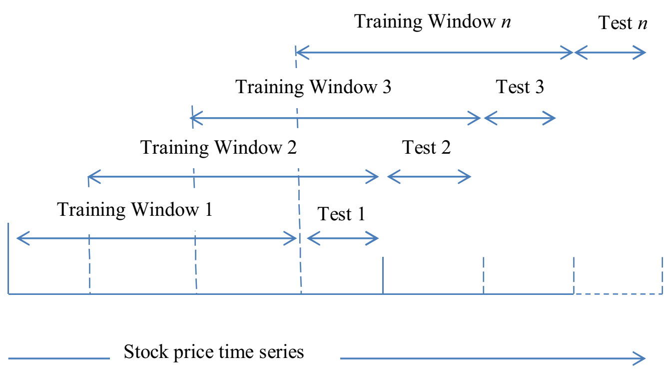

Fig. 2. Sliding window based training.

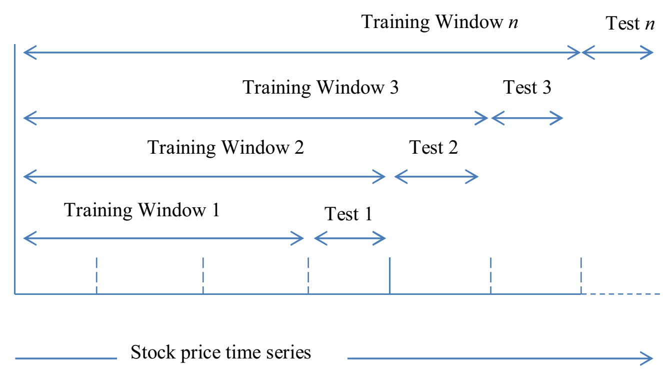

Fig. 3. Expanding window training technique.

Each column of the predictor $\boldsymbol{X}$ is considered to be a feature. Hence, there are 30 features in the predictor $\boldsymbol{X}$, i.e. $\boldsymbol{X} = \{\boldsymbol{Y}_0, \boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_{29}\}$ where, $\boldsymbol{Y}_0 = [y_t, y_{t-1}, \ldots, y_{t-N+30}]^{\mathrm{T}}$, $\boldsymbol{Y}_1 = [y_{t-1}, y_{t-2}, \ldots, y_{t-N+29}]^{\mathrm{T}}$ and so on. The suffixes 0, 1, …, 29 denoting the features $\boldsymbol{Y_0}$, $\boldsymbol{Y_1}$, …,$\boldsymbol{Y_{29}}$ indicate the relative time delay of each feature with respect to the first feature $\boldsymbol{Y_0}$.

The regression tree is used to identify the relationship between $\boldsymbol{X}$ and $\boldsymbol{t}$. The resultant tree will consist of only those features that are required for mapping **f:** $\boldsymbol{X} \rightarrow \boldsymbol{t}_{duration}$.

In the present study, two variants of all recommender systems are considered, one getting re-trained on a weekly basis (a maximum of 5 days) and the other on a monthly basis (a maximum of 20 days). Hence, in the first case, five regression trees are generated for duration = 1, 2, 3, 4 and 5 (i.e. one-day ahead forecast, two-day ahead forecast, …, five-day ahead forecast) and in the second case, twenty regression trees are generated for duration = 1, 2, 3, …, 20 (i.e., one-day ahead forecast, two-day ahead forecast, …, twenty-day ahead forecast).

The features that do not form the part of the tree are considered to be irrelevant.

The set of features selected using the regression tree are denoted by $\boldsymbol{X}_{selected}$, where $\boldsymbol{X}_{selected} \subseteq \boldsymbol{X}$. Assuming that M features are found to be relevant,

$$\boldsymbol{X}_{selected} = \{\boldsymbol{Y}_{\boldsymbol{F}_1}, \boldsymbol{Y}_{\boldsymbol{F}_2}, \cdots, \boldsymbol{Y}_{\boldsymbol{F}_M}\} \quad \text{Where } 0 \le \boldsymbol{F_i} \le 29, \quad \text{and } \boldsymbol{F_i} \ne \boldsymbol{F_j} \text{ with } 1 \le \boldsymbol{i}, \boldsymbol{j} \le M \text{ and } \boldsymbol{i} \ne \boldsymbol{j} \tag{7}$$

The forecast vector can be represented as

$$\boldsymbol{t} = \{\boldsymbol{t}_1, \boldsymbol{t}_2, \ldots, \boldsymbol{t}_{max\_duration}\} \tag{8}$$

2.3. Selection of training and testing data

Two input dataset selection mechanisms are evaluated. The first technique employs the sliding window technique wherein the training window moves to include the new data points (which presumably contain new information) while at the same time discarding equal number of the oldest data points. The second technique employs the expanding window technique, wherein the new data points are simply added to the training set, making the training set ‘expand’. Fig. 2 represents the sliding window training technique.

Fig. 3 presents the expanding window training technique.

As far as the testing time-frames are considered, here too, two techniques are evaluated, the first one employed a weekly trading system, in which, at the maximum, the testing window size would be one week (5 trading days) while the second variant will have a maximum testing window size of one month (20 trading days), after which, the test window gets incorporated into the training data and the system gets re-trained. However, it must be noted that once a round trade is carried out in either of the variants, the stock price data from the beginning of the test dataset to the sell date is incorporated into the new training data and the recommender gets trained.

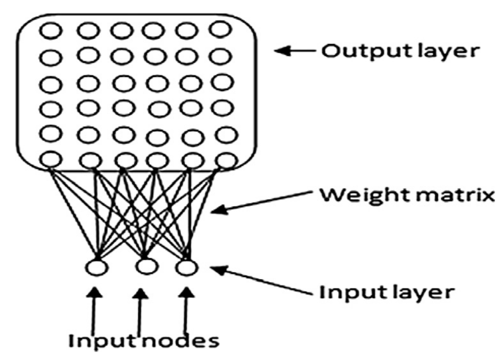

Fig. 4. Typical SOM Network Architecture (Ismail et al., 2011).

2.4. SOM based clustering

The SOM utilizes unsupervised learning technique and is efficient in clustering large multi-dimensional objects to low-dimensional objects. It aims at maximizing similarity within clusters and minimizes similarity between objects belonging to different cluster. It has two layers of nodes called as input and output layer. First the weights for each neuron are assigned randomly. Then when an input is fed into the SOM, a neuron which is closest to the input vector is selected as the winning neuron. Now this winning neuron will update its weight by learning. The neurons which are closer to the winning neuron and within its neighbourhood will also get their respective weights updated. This makes it possible for this region to become active the next time when similar input is entered. This localization is implemented by adjusting the learning rate $\alpha(t)$ and the neighborhood size R(t). Each neuron's amount of learning is $\alpha(t) * e^{-d/R(t)}$. Here distance $d = \|y - w\|$ where y is input and w is its weight. Here as $\alpha(t)$ and $R(t)$ decreases with time, the amount of learning will also decrease and after some iterations the updating will reach saturation. It is also observed that the amount of learning will be highest for winning neuron and decreases as the distance between neuron and. the winning neuron increases. The architecture of a typical SOM network is presented in Fig. 4.

The input to the SOM is $\boldsymbol{I_{SOM}} = \{\boldsymbol{X_{selected}}, \boldsymbol{t}\}$

It must also be noted that for training purpose, the latest time series sample present is y_t, hence $\boldsymbol{t}_{max\_duration} = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-N+30} \end{bmatrix}$, $\boldsymbol{t}_{max\_duration-1} = \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-N+29} \end{bmatrix}$ and so on while for testing, these vectors will be the 1-day-ahead to *max_duration* day-ahead forecasts.

In the present study, since the optimal number of clusters and the cluster configuration of the form $a \times b$ is unknown at the outset, we use the training data obtained using the techniques discussed above, to generate a total of 16 possible cluster combinations: 1×2, 1×3,1×4, 1×5, 2×2, 2×3, 2×4, 2×5, 3×2, 3×3, 3×4, 3×5, 4×2, 4×3, 4×4, 4×5. The total number of clusters generated is, hence, 140 (sum of all the clusters of all the combinations taken together). The optimal cluster combination is identified from these sixteen possible configurations using the technique described below.

2.5. Identification of optimal number of clusters

In the present study, the optimal number of SOM clusters is identified with the help of Silhouettes (Kaufman & Rousseeuw, 1990). Considering the SOM configuration, say, $a \times b$, the number of clusters in this configuration will be $a + b = k$ and according to (Kaufman & Rousseeuw, 1990), the silhouette $S_j(i)$ for the i th point in the j th cluster C_j, denoted by $p_j(i)$, is defined as:

$$S_{C_j}(i) = \frac{b_j(i) - a_j(i)}{\max\{a_j(i), b_j(i)\}} \tag{9}$$

Where,

$$a_j(i) = \frac{1}{|C_j| - 1} \sum\nolimits_{h \neq i,\ h \in j} \| p_j(i) - p_j(h) \| \tag{10}$$

and

$$b_j(i) = \min_{C_i} \left\{ \sum\nolimits_{h=1}^{|C_i-1|} \| p_j(i) - p_C(h) \| \right\},\ \ i = 1, 2, .., k \tag{11}$$

It can be seen that $-1 \le S_j(i) \le 1$. A high value of $S_{C_j}(i)$, indicates that the point $p_j(i)$ is well clustered and that the point indeed belongs to cluster C_j.

In the present study, the next step is to calculate the average silhouette value for each SOM cluster C_i in all configurations (i.e. 140 clusters, as calculated in Section 2.4 above):

$$\mu_{C_i} = \frac{1}{|C_i|} \sum\nolimits_{j=1}^{||C_i||} S_{C_i}(j), \quad i = 1, 2 \ldots 140 \tag{12}$$

Once the individual silhouette values for all the points in all the SOM clusters are obtained, the median silhouette value considering all the points in all clusters is calculated.

$$median_{\mu_C} = median\{\mu_{C_1},\ \mu_{C_2},\ \mu_{C_3},\ ..,\mu_{C_{140}}\} \tag{13}$$

The SOM cluster combination with all the constituent clusters having average silhouette greater than $median_{\mu_C}$ is considered to be the optimal cluster combination. In case more than one SOM cluster configuration satisfies this criterion, the configuration that has higher total number of clusters, is chosen as the optimal, i.e. say, if the two configurations $a \times b$, represented by $SOM_{a \times b}$ and the configuration $c \times d$, represented by $SOM_{c \times d}$ have $\mu_{C_i} > median_{\mu_C}$, $\forall C_i \in SOM_{a \times b}$, $i = 1, 2, ..,\ (|a| + |b|)$ and $\mu_{C_j} > median_{\mu_C}$, $\forall C_j \in SOM_{c \times d}$, $j = 1, 2, ..,\ (|c| + |d|)$, then choose the configuration $SOM_{a \times b}$ if $(|a| + |b|) > (|c| + |d|)$, otherwise choose $SOM_{c \times d}$.

2.6. Identification of profitable clusters

Once the optimal SOM cluster configuration is selected, the profitable clusters are identified. Each cluster is represented using its centroid. The clusters are first divided into profitable and non-profitable clusters based on the centroid values. The process is as follows:

The centroid of each cluster has $W = M + \max\_duration$ coordinates. Assuming that the SOM configuration $SOM_{a \times b}$ is found to be optimal as per the technique described in Section 2.5, the total number of clusters will then be $V = |a| \times |b|$. The centroids of these clusters are represented as

$$\boldsymbol{SOM_{centroids}} = \begin{bmatrix} \boldsymbol{D}_1 \\ \boldsymbol{D}_2 \\ \vdots \\ \boldsymbol{D_V} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_{\boldsymbol{selected}_{C_1}} & \vdots & \boldsymbol{t}_{\boldsymbol{selected}_{C_1}} \\ \boldsymbol{X}_{\boldsymbol{selected}_{C_2}} & \vdots & \boldsymbol{t}_{\boldsymbol{selected}_{C_2}} \\ \vdots & \vdots & \vdots \\ \boldsymbol{X}_{\boldsymbol{selected}_{C_V}} & \vdots & \boldsymbol{t}_{\boldsymbol{selected}_{C_V}} \end{bmatrix}$$
$$= \begin{bmatrix} \bar{y}_{F_{1C_1}} & \bar{y}_{F_{2C_1}} & \bar{y}_{F_{3C_1}} & \ldots & \bar{y}_{F_{MC_1}} & \vdots & \bar{y}_{t+1_{C_1}} & \ldots & \bar{y}_{t+\max\_duration_{C_1}} \\ \bar{y}_{F_{1C_2}} & \bar{y}_{F_{2C_2}} & \bar{y}_{F_{3C_2}} & \ldots & \bar{y}_{F_{MC_2}} & \vdots & \bar{y}_{t+1_{C_2}} & \ldots & \bar{y}_{t+\max\_duration_{C_2}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots \\ \bar{y}_{F_{1C_V}} & \bar{y}_{F_{2C_V}} & \bar{y}_{F_{3C_V}} & \ldots & \bar{y}_{F_{MC_V}} & \vdots & \bar{y}_{t+1_{C_V}} & \ldots & \bar{y}_{t+\max\_duration_{C_V}} \end{bmatrix} \tag{14}$$

Where $\boldsymbol{D_1}, \boldsymbol{D_2}, \ldots, \boldsymbol{D_V}$ are the centroids.

The profitable and non-profitable are identified based on the following algorithm:

Algorithm Identify_Cluster_Type

Inputs: $\boldsymbol{SOM_{centroids}}$
Outputs: $\boldsymbol{SOM_{Profitable\_centroid}}$, $\boldsymbol{SOM_{Non\_Profitable\_centroid}}$, *pro_clusts,loss_clusts*
Begin
count_profitable←1
count_non_profitable ←1
i ←1
While $i \le V$ **do**
  If $max\ (\boldsymbol{D}_i) \ge min\ (\boldsymbol{D}_i)$ **Then**
    $\boldsymbol{D}_{\boldsymbol{P}_{count\_profitable}} \leftarrow \boldsymbol{D_i}$
    $\boldsymbol{SOM_{Profitable\_centroid}}\ (count\_profitable) \leftarrow \boldsymbol{D}_{\boldsymbol{P}_{count\_profitable}}$
    $count\_profitable \leftarrow count\_profitable + 1$
  Else
    $\boldsymbol{D}_{\boldsymbol{L}_{count\_non\_profitable}} \leftarrow \boldsymbol{D_i}$
    $\boldsymbol{SOM_{Non\_Profitable\_centroid}}\ (count\_non\_profitable) \leftarrow \boldsymbol{D}_{\boldsymbol{L}_{count\_non\_pofitable}}$
    $count\_non\_profitable \leftarrow count\_non\_profitable + 1$
  End
i ← *i*+ *1*
End
pro_clusts ← *count_profitable*
loss_clusts ← *count_non_profitable*
End Algorithm

2.7. Generation of trading recommendations

Using the algorithm above, the set of profitable clusters can be represented by their centroids as follows:

$$\boldsymbol{SOM_{Profitable\_centroid}} = \begin{bmatrix} \boldsymbol{D_{P_1}} \\ \boldsymbol{D_{P_2}} \\ \vdots \\ \boldsymbol{D_{P_{pro\_clusts}}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X_{selected_{P_1}}} & \vdots & \boldsymbol{t_{selected_{P_1}}} \\ \boldsymbol{X_{selected_{P_2}}} & \vdots & \boldsymbol{t_{selected_{P_2}}} \\ \vdots & \vdots & \vdots \\ \boldsymbol{X_{selected_{P_{pro\_clusts}}}} & \vdots & \boldsymbol{t_{selected_{P_V}}} \end{bmatrix}$$

$$= \begin{bmatrix} \bar{y}_{F_{1P_1}} & \bar{y}_{F_{2P_1}} & \bar{y}_{F_{3P_1}} & \cdots & \bar{y}_{F_{MP_1}} & \vdots & \bar{y}_{t+1_{P_1}} & \cdots & \bar{y}_{t+\max\_duration_{P_1}} \\ \bar{y}_{F_{1P_2}} & \bar{y}_{F_{2P_2}} & \bar{y}_{F_{3P_2}} & \cdots & \bar{y}_{F_{MP_2}} & \vdots & \bar{y}_{t+1_{P_2}} & \cdots & \bar{y}_{t+\max\_duration_{P_2}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \bar{y}_{F_{1P_{pro\_clusts}}} & \bar{y}_{F_{2P_{pro\_clusts}}} & \bar{y}_{F_{3P_{pro\_clusts}}} & \cdots & \bar{y}_{F_{MP_{pro\_clusts}}} & \vdots & \bar{y}_{t+1_{P_{pro\_clusts}}} & \cdots & \bar{y}_{t+\max\_duration_{P_{pro\_clusts}}} \end{bmatrix} \quad (15)$$

where $\boldsymbol{SOM_{Profitable\_centroid}} \subseteq \boldsymbol{SOM_{centroids}}$ and $1 \le pro\_clusts \le V$. and the loss making clusters are represented as follows:

$$\boldsymbol{SOM_{Non\_profitable\_centroid}} = \begin{bmatrix} \boldsymbol{D_{L_1}} \\ \boldsymbol{D_{L_2}} \\ \vdots \\ \boldsymbol{D_{L_{loss\_clucts}}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X_{selected_{L_1}}} & \vdots & \boldsymbol{t_{selected_{L_1}}} \\ \boldsymbol{X_{selected_{L_2}}} & \vdots & \boldsymbol{t_{selected_{L_2}}} \\ \vdots & \vdots & \vdots \\ \boldsymbol{X_{selected_{L_{loss\_clucts}}}} & \vdots & \boldsymbol{t_{selected_{L_V}}} \end{bmatrix}$$

$$= \begin{bmatrix} \bar{y}_{F_{1L_1}} & \bar{y}_{F_{2L_1}} & \bar{y}_{F_{3L_1}} & \cdots & \bar{y}_{F_{ML_1}} & \vdots & \bar{y}_{t+1_{L_1}} & \cdots & \bar{y}_{t+\max\_duration_{L_1}} \\ \bar{y}_{F_{1L_2}} & \bar{y}_{F_{2L_2}} & \bar{y}_{F_{3L_2}} & \cdots & \bar{y}_{F_{ML_2}} & \vdots & \bar{y}_{t+1_{L_2}} & \cdots & \bar{y}_{t+\max\_duration_{L_2}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \bar{y}_{F_{1L_{loss\_clucts}}} & \bar{y}_{F_{2L_{loss\_clucts}}} & \bar{y}_{F_{3L_{loss\_clucts}}} & \cdots & \bar{y}_{F_{ML_{loss\_clucts}}} & \vdots & \bar{y}_{t+1_{L_{loss\_clucts}}} & \cdots & \bar{y}_{t+\max\_duration_{L_{loss\_clucts}}} \end{bmatrix} \quad (16)$$

Considering the latest stock price value to be y_t, the delayed values that are considered to be relevant, as identified using the technique described above are: $y_{t-F_1}, y_{t-F_2}, y_{t-F_3}, \ldots, y_{t-F_M}$. Here, F_1 represents the smallest delay value, F_2, the next higher delay and so on with F_M being the highest delay value. The future values (unknown at time y_t) are denoted by $y_{t+1}, y_{t+2}, \ldots, y_{t+\max\_duration}$.

As can be observed, at time *t*, only the historical values are available. Thus, the vector $\boldsymbol{X_{input}} = \{y_{t-F_1}, y_{t-F_2}, y_{t-F_3}, \ldots, y_{t-F_M}\}$ will be used to forecast the approximate future prices and recommend trading decision based on the forecast prices. This is accomplished using the algorithm as follows:

Algorithm Identify_Trading_Recommendation

Inputs: $\boldsymbol{SOM_{centroids}}$
Outputs: *trade_flag, buy_day, sell_day*
Begin
$i \leftarrow 1$
$j \leftarrow 1$
$min\_dist \leftarrow 0$
$\boldsymbol{X_{min}} \leftarrow \{\}$
$\boldsymbol{cluster_{closest}} \leftarrow \{\}$
$dist_{profitable} \leftarrow 0$

(*continued on next page*)

While $i \le V$ **do**
$dist \leftarrow \|\boldsymbol{X_{input}} - \boldsymbol{X_{selected_{C_i}}}\|$
  If $dist \le min\_dist$ **Then**
    $min\_dist \leftarrow dist$
    $\boldsymbol{X_{min}} \leftarrow \boldsymbol{X_{selected_{C_i}}}$
  End
End
While $i \le pro\_clusts$ **do**
  $dist_{profitable} \leftarrow \|\boldsymbol{X_{min}} - \boldsymbol{X_{selected_{P_i}}}\|$
   If $dist_{profitable} = 0$ **Then**
     $\boldsymbol{cluster_{closest}} \leftarrow [\ y_t \ \vdots \ \boldsymbol{t_{selected_{P_i}}}\ ]$
     $temp\_\max \leftarrow \max(\boldsymbol{cluster_{closest}})$
     $temp\_\min \leftarrow \min(\boldsymbol{cluster_{closest}})$
   End
End
While $j \le (\max\_duration + 1)$ **do**
  If $\boldsymbol{cluster_{closest}}(j) = temp\_min$ **Then**
    $buy\_day \leftarrow j+1$
  Else If $\boldsymbol{cluster_{closest}}(j) = temp\_max$ **Then**
    $sell\_day \leftarrow j$
  End
End
If $buy\_day < sell\_day$ **Then**
  $trade\_flag \leftarrow 1$
Else
  $trade\_flag \leftarrow 0$
  $buy\_day \leftarrow 0$
  $sell\_day \leftarrow 0$
End
End Algorithm

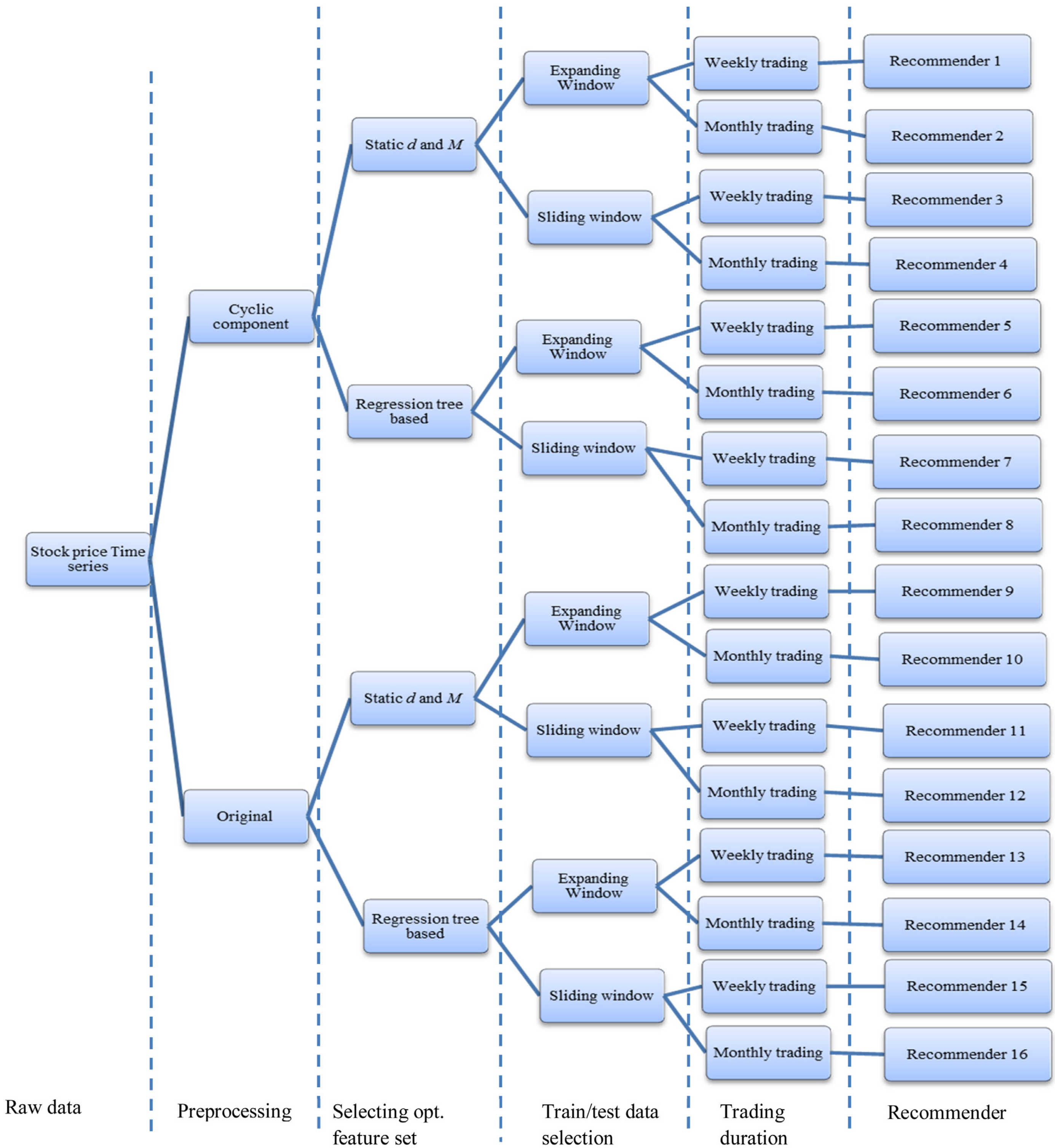

Fig. 5. Recommender system variants evaluated.

Based on the results obtained from the algorithm Identify_Trading_Recommendation, the following decisions can be made:

If the $trade\_flag = 0$, then no trade must be executed till *max_duration* number of days in future from the current day.

If the $trade\_flag = 1$, then buy the stock at the opening price, *buy_day* number of days from the current day and sell the stock *sell_day* number of days from the current day, at that day's closing price (i.e. just before the trading stops).

2.8. Recommender system variants evaluated

It was observed from the literature survey that the performance of such recommender systems is greatly influenced by (a) the input dataset selection mechanism for training, (b) the preprocessing of the inputs and (c) the testing timeframe considered.

Table 1
Recommendersystem performance measures.

S. No.	Performance Measure	Formula
1	Total Profit (TP)	$TP = \sum_{i=1}^{TT} profit(i)$
2	Average Profit (AP)	$AP = \frac{TP}{TT}$
3	Profit per Success Trade ($\frac{P}{ST}$)	$(\frac{P}{ST}) = \frac{Total\ Profit\ only}{PT}$

	Profitable Trades (PT)	$PT \leftarrow 0$
		i ← 1
		While i < TT **Do**
		  If $profit(i) > 0$ **Then**
		    $PT \leftarrow PT + 1$
		  End If
		i ← i+1
		End While

	Total Profit only	$Total\ Profit\ only = \sum_{i=1,\ profit(i)>0}^{TT} profit(i)$
4	Loss per Loss Trade (L/LT)	$(\frac{L}{LT}) = \frac{Total\ Loss\ only}{LT}$

	Loss Trades (LT)	LT ← 0
		i ← 1
		While i < TT **Do**
		  If $profit(i) < 0$ **Then**
		    $LT \leftarrow LT + 1$
		  End If
		i ← i+1
		End While

	Total Loss only	$Total\ Loss\ only = \sum_{i=1,\ profit(i)<0}^{TT} profit(i)$
5	Maximum Drawdown (MD)	$MD = \min(Total\ Loss\ only)$
6	Profit Factor (PF)	$PF = \frac{Total\ Profit\ only}{Total\ Loss\ only}$
7	Win Ratio (WR)	$WR = \frac{PT}{LT}$

Hence, combining all the possibilities as shown in Figure, sixteen recommender system variants are evaluated, as follows:

It must be noted that in the subsequent sections, the sixteen recommender systems evaluated in the present study are referred to as Recommender 1, Recommender 2, …, Recommender 16, where, the configuration of each of these recommenders can be identified from Fig. 5.

3. Results

A total of twenty stocks are considered for evaluating the proposed recommender systems. Due to the computationally expensive nature of the proposed recommender systems, the dataset size had to be confined to twenty. However, the stocks were chosen so as to represent markets belonging to four different geographies (continents) such as BSE from Asia (India), NYSE from North America (U.S.A), LSE and FTSE from Europe (UK) and Sao Palo from South America (Brazil) are chosen. Stocks are considered from various industries such as Aircraft Manufacturing, Banking, Communication, Food, Petroleum, Consumer Retail, Textile, Travels, etc. Stocks that are considered from India are ACC, GAIL, ICICI, NTPC and TATA ELEXSI (TE). Stocks that are considered from Europe are BT-Group (BT), Carnival (CCL), GSK, RDS and Sainsbury (SBRY). From US, the stocks chosen are Bank of America (BoA), Boeing (BA), GE, M3, Pepsico (PEP). From South America the stocks are Duratex (DTEX), Embraer (EMBR), ITSA4, Petrobras (PETR) and Telefonica (TFA). The historical daily stock price data for all the above stocks was obtained from (Yahoo! Finance, 2014). For each stock, sixteen variants of recommender systems are considered, as illustrated in Fig. 5.

The results are analysed using eight performance measures (Brabazon & O' Neill, 2006). The eight performance measures on which the performance of the recommender systems is evaluated are based on the assumption that the trader will follow the recommendations made, as it is, over the testing time-frame. It must also be noted that the profit or loss for each transaction is calculated after taking the transaction cost of 0.5% into consideration. The performance measures, as can be seen from Table 1, are: the total profit made over the entire testing timeframe (TP); the average profit (AP) which is the total profit or loss averaged over the total number of round trades executed; profit per successful trades (P/ST) which is the average profit generated by considering only profit-making trades (PT) and the corresponding profits from profit-making trades alone (*Total Profit only*); loss per loss-making trades (L/LT) which is the average loss generated by considering only loss-making trades (LT) and the corresponding losses from loss-making trades alone (*Total Loss only*); maximum drawdown (MD) which can be thought of as the worst result generated by the recommender over the testing time-frame is simply the loss generated in the largest loss-making trade over the testing time-frame or in case there were no loss-making trades, the lowest profit generated over the testing time-frame; profit factor (PF) which is the ratio of gross profit to the gross losses and finally, the win-ratio (WR) which is the ratio of the number of profit-making trades to loss-making trades. In addition to these seven, the total number of round trades executed in the testing time frame (TT) is also considered as a performance measure. A ‘good’ stock trading recommender system should be able to generate buy/sell recommendations that result in a high TP, AP, P/ST, PF and WR while at the same time minimizing the L/LT, MD and TT.

The performance measures are in the units of the currencies used in the respective countries, i.e. for India- Indian Rupees(INR), for US- US Dollars (US$), for UK- Great Britain pence (GBp) and for Brazil- Brazilian Real (BRL).

The results are compared with the traditional benchmark Buy and Hold (B&H) strategy. The B&H returns are calculated as follows:

$$B\&H = y_{Last} - y_{First} - 0.005(y_{Last} + y_{First}) \tag{17}$$

Where:

y_{First} = the closing price on the first day of testing set;
y_{Last} = closing price on the last day of testing set.

Table 2
Stocks and time-frames considered.

Market	Stock	Training	Testing
India	**ACC**	01/01/2009–01/09/2014	01/10/2014–11/19/2014
	GAIL	01/01/2009–01/09/2014	01/10/2014–12/22/2014
	ICICI	01/01/2009–01/15/2014	01/16/2014–11/19/2014
	NTPC	01/01/2009–01/29/2014	01/30/2014–11/19/2014
	TE	01/01/2009–01/09/2014	01/10/2014–12/22/2014
US	**BoA**	01/02/2009–03/04/2014	03/05/2014–12/17/2014
	BA	01/02/2009–03/04/2014	03/05/2014–11/20/2014
	GE	01/02/2009–03/04/2014	03/05/2014–11/20/2014
	M3	01/02/2009–03/04/2014	03/05/2014–11/20/2014
	PEP	01/02/2009–03/04/2014	03/05/2014–12/22/2014
UK	**BT**	01/01/2009–01/09/2014	01/10/2014–12/17/2014
	CCL	01/01/2009–01/09/2014	01/10/2014–12/17/2014
	GSK	01/01/2009–01/07/2014	01/08/2014–11/20/2014
	RDS	01/01/2009–01/07/2014	01/08/2014–11/20/2014
	SBRY	01/01/2009–01/08/2014	01/09/2014–11/20/2014
Brazil	**DTEX**	01/01/2009–02/26/2014	02/27/2014–12/15/2014
	EMBR	01/01/2009–01/29/2014	01/30/2014–12/15/2014
	ITSA4	01/02/2009–03/21/2014	03/24/2014–11/19/2014
	PETR	01/01/2009–01/27/2014	01/28/2014–11/20/2014
	TFA	01/02/2009–08/05/2014	09/05/2014–12/15/2014

It must be noted that a transaction cost of 0.5% was considered in this case as well, as can be seen in Eq. (17).

Apart from TT, which is just the total number of round trades executed (here, a 'buy' followed by a 'sell' constitutes a round trade, also referred to as TT in Table 1), the remaining performance measures used are presented in Table 1.

The training and testing time frames considered for each of the stock is presented in Table 2.

The performance of the recommender systems are presented in the following tables.

Table 3 presents the trading performance for the recommender with HP filter based preprocessing of the stock price data, feature selection based on assumption of static lag and embedding dimension with expanding window based learning and trading a maximum of weekly once (Recommender 1).

Table 4 presents the trading performance for the recommender with HP filter based preprocessing of the stock price data, feature selection based on assumption of static lag and embedding dimension with expanding window based learning and trading a maximum of monthly once (Recommender 2).

Table 5 presents the trading performance for the recommender with HP filter based preprocessing of the stock price data, feature selection based on assumption of static lag and embedding dimension with sliding window based learning and trading a maximum of weekly once (Recommender 3) .

Table 6 presents the trading performance for the recommender with HP filter based preprocessing of the stock price data, feature selection based on assumption of static lag and embedding dimension with sliding window based learning and trading a maximum of monthly once (Recommender 4).

Table 7 presents the trading performance for the recommender with HP filter based preprocessing of the stock price data, feature selection based on regression tree with expanding window based learning and trading a maximum of weekly once (Recommender 5).

Table 8 presents the trading performance for the recommender with HP filter based preprocessing of the stock price data, feature selection based on regression tree with expanding window based learning and trading a maximum of monthly once (Recommender 6).

Table 9 presents the trading performance for the recommender with HP filter based preprocessing of the stock price data, feature selection based on regression tree with sliding window based learning and trading a maximum of weekly once (Recommender 7).

Table 10 presents the trading performance for the recommender with HP filter based preprocessing of the stock price data, feature selection based on regression tree with sliding window based learning and trading a maximum of monthly once (Recommender 8).

Table 11 presents the trading performance for the recommender without preprocessing of the stock price data, feature selection based on assumption of static delay and embedding dimension with expanding window based learning and trading a maximum of weekly once (Recommender 9).

Table 12 presents the trading performance for the recommender without preprocessing of the stock price data, feature selection based on assumption of static delay and embedding dimension with expanding window based learning and trading a maximum of monthly once (Recommender 10).

Table 13 presents the trading performance for the recommender without preprocessing of the stock price data, feature selection based on assumption of static delay and embedding dimension with sliding window based learning and trading a maximum of weekly once (Recommender 11).

Table 14 presents the trading performance for the recommender without preprocessing of the stock price data, feature selection based on assumption of static delay and embedding dimension with sliding window based learning and trading a maximum of monthly once (Recommender 12).

Table 15 presents the trading performance for the recommender without preprocessing of the stock price data, feature selection based on regression tree with expanding window based learning and trading a maximum of weekly once (Recommender 13).

Table 16 presents the trading performance for the recommender without preprocessing of the stock price data, feature selection based on regression tree with expanding window based learning and trading a maximum of monthly once (Recommender 14).

Table 17 presents the trading performance for the recommender without preprocessing of the stock price data, feature selection based on regression tree with sliding window based learning and trading a maximum of weekly once (Recommender 15).

Table 18 presents the trading performance for the recommender without preprocessing of the stock price data, feature selection based on regression tree with sliding window based learning and trading a maximum of monthly once (Recommender 16).

One of the key parameters that illustrates the effectiveness of the stock trading recommender system is the amount of excess profit generated by the system when compared to the benchmark B&H strategy. The excess profit can be calculated as from the tables above using Eq. (18):

$$Excess\,Profit = \mathrm{TP} - \mathrm{B\&H} \tag{18}$$

The excess profit generated by each recommender is graphically presented in Figs. 6–9.

Table 3
Recommender 1 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	**ACC**	348.05	372.73	8.10	35.55	−24.57	−55.77	46.00	1.72	1.19
	GAIL	101.12	39.95	0.93	9.01	−9.28	−54.33	43.00	1.23	1.26
	ICICI	429.64	352.53	7.34	36.84	−19.79	−65.27	48.00	1.71	0.92
	NTPC	24.46	35.62	0.79	2.80	−2.86	−14.97	45.00	1.78	1.81
	TE	218.83	126.02	3.07	14.99	−12.16	−38.71	41.00	1.58	1.28
US	**BoA**	−0.30	−0.86	−0.02	0.20	−0.23	−0.71	43.00	0.83	0.95
	BA	−2.64	−3.41	−0.07	1.55	−1.56	−4.13	48.00	0.91	0.92
	GE	0.36	−0.09	0.00	0.24	−0.30	−1.14	20.00	0.97	1.22
	M3	8.75	15.02	0.33	1.08	−0.85	−2.84	46.00	1.98	1.56
	PEP	15.71	16.07	0.35	0.80	−0.57	−1.29	46.00	2.87	2.07
UK	**BT**	−3.65	−33.43	−0.78	5.10	−6.39	−15.02	43.00	0.76	0.95
	CCL	−247.88	−108.46	−2.26	41.75	−46.27	−144.28	48.00	0.90	1.00
	GSK	−34.38	−95.52	−2.03	14.75	−14.46	−39.30	47.00	0.76	0.74
	RDS	282.03	133.83	3.04	14.93	−15.83	−41.29	44.00	1.50	1.59
	SBRY	−32.22	−56.02	−1.19	4.56	−6.25	−28.16	47.00	0.64	0.88
Brazil	**DTEX**	−1.23	−0.36	−0.01	0.21	−0.26	−1.03	46.00	0.94	1.19
	EMBR	4.10	4.19	0.09	0.40	−0.28	−0.97	46.00	1.71	1.19
	ITSA4	1.15	1.53	0.05	0.27	−0.23	−0.51	30.00	1.51	1.31
	PETR	3.73	3.58	0.08	0.64	−0.68	−2.56	47.00	1.26	1.35
	TFA	0.73	4.03	0.09	1.00	−0.72	−2.51	47.00	1.22	0.88

Table 4
Recommender 2 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	ACC	394.72	414.17	34.51	73.05	−42.56	−55.27	12.00	3.43	2.00
	GAIL	109.48	53.58	4.47	13.99	−14.58	−37.26	12.00	1.92	2.00
	ICICI	462.32	−47.51	−3.96	31.13	−29.03	−69.10	12.00	0.77	0.71
	NTPC	16.35	−10.45	−0.95	9.34	−6.83	−18.06	11.00	0.78	0.57
	TE	181.61	105.47	8.79	47.06	−18.55	−31.99	12.00	1.81	0.71
US	BoA	−2.13	−0.84	−0.07	0.28	−0.32	−0.55	12.00	0.63	0.71
	BA	3.93	−1.37	−0.11	2.45	−2.68	−4.98	12.00	0.91	1.00
	GE	0.36	2.74	0.25	0.59	−0.34	−0.66	11.00	2.99	1.75
	M3	9.41	−3.00	−0.25	2.59	−2.28	−5.52	12.00	0.81	0.71
	PEP	11.35	13.44	1.12	1.25	−0.33	−0.33	12.00	41.94	11.00
UK	BT	−2.45	34.18	2.85	11.55	−5.85	−20.20	12.00	1.97	1.00
	CCL	−145.39	−154.23	−12.85	78.80	−78.32	−179.10	12.00	0.72	0.71
	GSK	−176.67	57.71	4.81	44.78	−23.74	−45.77	12.00	1.35	0.71
	RDS	207.90	161.19	14.65	40.08	−29.85	−50.75	11.00	2.35	1.75
	SBRY	−38.78	15.92	1.33	3.49	−3.01	−5.57	12.00	2.32	2.00
Brazil	DTEX	−2.30	−2.60	−0.22	0.30	−0.39	−0.98	12.00	0.26	0.33
	EMBR	0.71	−0.04	0.00	0.42	−0.43	−0.90	12.00	0.98	1.00
	ITSA4	1.15	0.38	0.04	0.35	−0.43	−0.79	10.00	1.22	1.50
	PETR	3.73	1.86	0.16	0.53	−0.22	−0.59	12.00	2.39	1.00
	TFA	0.68	−3.76	−0.31	0.84	−1.46	−3.11	12.00	0.57	1.00

Table 5
Recommender 3 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	**ACC**	322.28	144.13	3.13	30.39	−21.85	−59.80	46.00	1.27	0.92
	GAIL	108.78	56.81	1.26	10.54	−6.86	−17.61	45.00	1.35	0.88
	ICICI	462.32	34.18	0.78	24.50	−17.25	−69.10	44.00	1.08	0.76
	NTPC	13.11	37.91	0.81	3.65	−1.92	−5.92	47.00	1.82	0.96
	TE	45.69	113.38	2.58	16.63	−14.29	−95.42	44.00	1.40	1.20
US	**BoA**	−0.49	−2.93	−0.06	0.19	−0.24	−1.00	48.00	0.57	0.71
	BA	−8.08	−12.71	−0.26	1.19	−1.61	−3.65	48.00	0.68	0.92
	GE	−1.70	−2.74	−0.06	0.27	−0.21	−0.54	48.00	0.60	0.45
	M3	23.31	9.40	0.20	1.47	−1.52	−4.79	47.00	1.31	1.35
	PEP	10.45	7.17	0.16	0.76	−0.60	−1.00	45.00	1.60	1.25
UK	**BT**	−0.66	1.59	0.04	4.21	−4.34	−19.30	43.00	1.02	1.05
	CCL	−377.23	−186.07	−4.13	29.29	−39.08	−115.42	45.00	0.78	1.05
	GSK	−193.58	−158.21	−3.37	16.20	−15.51	−54.73	47.00	0.65	0.62
	RDS	267.10	322.88	6.87	19.64	−13.71	−44.78	47.00	2.31	1.61
	SBRY	−33.81	13.13	0.29	5.81	−3.96	−11.04	46.00	1.13	0.77
Brazil	**DTEX**	−2.61	−1.27	−0.03	0.21	−0.23	−1.00	42.00	0.76	0.83
	EMBR	3.19	3.59	0.08	0.31	−0.28	−0.67	46.00	1.71	1.56
	ITSA4	1.15	2.08	0.05	0.21	−0.17	−0.49	43.00	1.67	1.39
	PETR	6.20	3.05	0.07	0.57	−0.31	−1.25	44.00	1.39	0.76
	TFA	1.41	2.63	0.06	0.78	−0.77	−4.21	47.00	1.16	1.14

Table 6
Recommender 4 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	**ACC**	491.33	540.78	45.07	78.04	−20.88	−29.00	12.00	7.47	2.00
	GAIL	80.33	133.83	11.15	23.14	−12.82	−25.32	12.00	3.61	2.00
	ICICI	496.40	494.42	41.20	63.37	−25.31	−69.10	12.00	7.51	3.00
	NTPC	9.08	−11.24	−0.94	5.53	−5.56	−18.26	12.00	0.71	0.71
	TE	180.62	0.15	0.01	32.47	−18.54	−45.02	11.00	1.00	0.57
US	**BoA**	−1.12	−1.76	−0.15	0.51	−0.48	−1.48	12.00	0.54	0.50
	BA	−2.82	−14.91	−1.24	3.56	−2.20	−4.25	12.00	0.32	0.20
	GE	−0.06	2.04	0.17	0.54	−0.57	−1.00	12.00	1.90	2.00
	M3	16.59	1.17	0.10	3.03	−2.00	−4.23	12.00	1.08	0.71
	PEP	4.36	6.98	0.58	1.30	−0.86	−1.76	12.00	3.02	2.00
UK	**BT**	−8.62	−8.66	−0.72	11.06	−9.14	−23.18	12.00	0.86	0.71
	CCL	−347.38	−239.80	−19.98	33.46	−126.86	−179.10	12.00	0.53	2.00
	GSK	496.40	−250.24	−20.85	22.89	−35.43	−62.69	12.00	0.22	0.33
	RDS	267.10	157.71	13.14	32.46	−25.50	−50.75	12.00	2.55	2.00
	SBRY	260.97	−22.98	−1.92	7.08	−10.91	−16.52	12.00	0.65	1.00
Brazil	**DTEX**	−1.24	−2.07	−0.17	0.24	−0.59	−2.33	12.00	0.41	1.00
	EMBR	5.09	2.95	0.25	0.62	−0.50	−0.78	12.00	2.46	2.00
	ITSA4	1.15	−0.52	−0.05	0.46	−0.27	−0.59	10.00	0.73	0.43
	PETR	4.77	1.21	0.10	0.63	−0.43	−0.82	12.00	1.47	1.00
	TFA	1.23	7.20	0.65	1.36	−0.57	−1.56	11.00	4.15	1.75

Table 7
Recommender 5 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	**ACC**	475.41	485.81	10.12	40.19	−25.42	−74.28	48.00	1.87	1.18
	GAIL	99.68	122.58	2.61	11.75	−10.86	−53.63	47.00	1.59	1.47
	ICICI	404.27	448.30	9.34	33.42	−19.12	−67.56	48.00	2.07	1.18
	NTPC	10.28	14.23	0.30	4.11	−3.85	−19.45	48.00	1.16	1.09
	TE	203.70	117.91	2.46	21.13	−16.22	−38.31	48.00	1.30	1.00
US	**BoA**	−0.48	−0.13	0.00	0.26	−0.36	−0.86	47.00	0.98	1.35
	BA	2.02	5.64	0.12	1.18	−1.31	−3.76	47.00	1.21	1.35
	GE	−0.47	0.22	0.00	0.25	−0.26	−0.93	48.00	1.04	1.09
	M3	6.34	6.79	0.14	1.17	−0.98	−2.84	48.00	1.30	1.09
	PEP	11.55	10.04	0.21	0.85	−0.82	−3.13	47.00	1.68	1.61
UK	**BT**	−2.55	1.59	0.03	4.91	−6.55	−13.03	47.00	1.01	1.35
	CCL	−149.37	−108.46	−2.26	41.43	−49.75	−165.17	48.00	0.91	1.09
	GSK	−56.77	−113.43	−2.41	18.32	−15.29	−50.75	47.00	0.74	0.62
	RDS	231.28	130.84	2.84	13.26	−18.67	−52.24	46.00	1.47	2.07
	SBRY	−41.47	−55.52	−1.21	4.83	−6.75	−28.76	46.00	0.66	0.92
Brazil	**DTEX**	−1.57	−0.75	−0.02	0.13	−0.17	−0.39	40.00	0.78	1.00
	EMBR	3.41	2.58	0.05	0.42	−0.44	−1.03	47.00	1.29	1.35
	ITSA4	1.67	1.35	0.03	0.21	−0.15	−0.52	40.00	1.46	1.00
	PETR	−3.70	−8.34	−0.69	0.22	−0.88	−2.12	12.00	0.05	0.20
	TFA	2.16	3.07	0.09	1.00	−0.94	−2.11	34.00	1.20	1.13

Table 8
Recommender 6 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	**ACC**	329.39	238.75	19.90	51.90	−24.91	−51.89	12.00	2.92	1.40
	GAIL	84.61	90.30	7.52	15.29	−8.00	−20.55	12.00	3.82	2.00
	ICICI	443.77	174.32	14.53	46.66	−17.60	−30.60	12.00	2.65	1.00
	NTPC	16.35	28.76	2.61	7.58	−6.07	−16.57	11.00	2.18	1.75
	TE	200.12	210.34	19.12	28.50	−23.08	−43.13	11.00	5.56	4.50
US	**BoA**	−0.75	−0.31	−0.03	0.45	−0.50	−1.04	12.00	0.90	1.00
	BA	1.54	69.65	6.97	55.89	−66.42	−108.46	10.00	1.26	1.50
	GE	−0.19	1.25	0.14	0.43	−0.23	−0.50	9.00	2.37	1.25
	M3	4.89	2.19	0.18	2.04	−1.68	−4.76	12.00	1.22	1.00
	PEP	12.53	1.03	0.09	1.65	−1.48	−3.00	12.00	1.12	1.00
UK	**BT**	3.12	−38.36	−3.20	6.60	−8.10	−19.60	12.00	0.41	0.50
	CCL	−199.12	86.57	7.21	49.50	−77.36	−160.20	12.00	1.28	2.00
	GSK	−183.14	−227.86	−18.99	15.12	−43.35	−125.87	12.00	0.25	0.71
	RDS	28.30	329.35	27.45	50.47	−41.62	−60.20	12.00	3.64	3.00
	SBRY	−42.36	−17.11	−1.43	8.16	−6.22	−10.35	12.00	0.66	0.50
Brazil	**DTEX**	−1.57	−0.82	−0.07	0.41	−0.41	−1.00	12.00	0.72	0.71
	EMBR	0.04	0.22	0.02	0.48	−0.63	−1.13	12.00	1.07	1.40
	ITSA4	0.27	0.67	0.07	0.22	−0.11	−0.22	9.00	2.46	1.25
	PETR	−0.37	−0.75	−0.07	0.28	−0.35	−0.74	11.00	0.65	0.83
	TFA	4.31	−5.06	−0.42	0.46	−0.86	−1.85	12.00	0.27	0.50

Table 9
Recommender 7 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	**ACC**	439.29	456.01	9.50	42.11	−23.11	−84.62	48.00	1.82	1.00
	GAIL	100.53	114.43	2.43	12.63	−10.19	−38.21	47.00	1.53	1.24
	ICICI	381.13	378.90	8.06	37.62	−31.85	−98.75	47.00	1.59	1.35
	NTPC	13.11	34.58	0.82	3.47	−3.47	−15.62	42.00	1.62	1.63
	TE	218.38	235.91	4.91	21.14	−15.94	−53.48	48.00	1.70	1.29
US	**BoA**	−0.34	0.23	0.00	0.25	−0.33	−1.12	47.00	1.03	1.35
	BA	−2.52	−0.77	−0.02	1.54	−1.45	−3.84	48.00	0.98	0.92
	GE	−0.58	0.03	0.00	0.22	−0.36	−1.03	48.00	1.00	1.67
	M3	23.81	26.52	0.63	1.66	−0.62	−2.39	42.00	3.25	1.21
	PEP	14.91	17.74	0.37	0.87	−0.85	−2.99	48.00	2.48	2.43
UK	**BT**	−3.65	10.85	0.27	5.71	−5.16	−12.64	40.00	1.11	1.00
	CCL	−327.48	−311.44	−6.63	39.34	−63.54	−171.14	47.00	0.77	1.24
	GSK	−36.87	−7.46	−0.16	20.83	−18.63	−37.81	47.00	0.98	0.88
	RDS	257.65	243.78	5.19	19.67	−20.37	−58.71	47.00	1.70	1.76
	SBRY	−117.79	−66.76	−1.42	4.91	−7.49	−28.76	47.00	0.63	0.96
Brazil	**DTEX**	−2.28	−0.26	−0.01	0.19	−0.21	−1.03	35.00	0.93	1.06
	EMBR	5.55	2.64	0.07	0.45	−0.41	−0.85	40.00	1.36	1.22
	ITSA4	2.04	1.58	0.03	0.18	−0.16	−0.57	46.00	1.50	1.30
	PETR	−1.62	−3.93	−0.28	0.42	−0.67	−2.12	14.00	0.35	0.56
	TFA	−1.24	2.84	0.06	0.96	−1.00	−4.21	46.00	1.13	1.19

Table 10
Recommender 8 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	**ACC**	423.82	168.50	14.04	54.00	−25.91	−62.93	12.00	2.08	1.00
	GAIL	149.53	1.54	0.13	18.42	−9.02	−32.64	12.00	1.02	0.50
	ICICI	430.48	89.90	7.49	48.32	−21.67	−57.11	12.00	1.59	0.71
	NTPC	15.45	−4.48	−0.41	5.39	−5.24	−16.57	11.00	0.86	0.83
	TE	180.82	140.49	11.71	41.07	−29.39	−48.06	12.00	1.96	1.40
US	**BoA**	−2.26	−0.31	−0.03	0.20	−0.34	−0.85	12.00	0.82	1.40
	BA	1.54	6.58	0.73	3.05	−2.17	−4.39	9.00	1.76	1.25
	GE	−1.70	0.21	0.02	0.27	−0.34	−0.59	12.00	1.12	1.40
	M3	2.24	3.86	0.32	1.68	−2.40	−4.34	12.00	1.40	2.00
	PEP	10.39	−0.41	−0.03	1.37	−1.03	−4.52	12.00	0.94	0.71
UK	**BT**	−11.91	0.55	0.05	13.74	−6.80	−14.73	12.00	1.01	0.50
	CCL	−389.17	−240.79	−20.07	49.75	−89.88	−171.14	12.00	0.55	1.00
	GSK	−183.14	−60.70	−5.06	50.50	−32.84	−61.19	12.00	0.77	0.50
	RDS	202.43	214.92	17.91	39.92	−26.12	−37.81	12.00	3.06	2.00
	SBRY	−40.08	−18.11	−1.51	6.02	−5.27	−11.04	12.00	0.57	0.50
Brazil	**DTEX**	−2.28	−1.82	−0.15	0.19	−0.63	−1.15	12.00	0.42	1.40
	EMBR	3.69	2.68	0.22	0.72	−0.48	−0.78	12.00	2.12	1.40
	ITSA4	1.15	3.02	0.27	0.48	−0.27	−0.67	11.00	4.75	2.67
	PETR	6.47	−0.96	−0.08	0.42	−0.58	−1.09	12.00	0.73	1.00
	TFA	0.10	−0.50	−0.04	1.06	−1.15	−1.96	12.00	0.93	1.00

Table 11
Recommender 9 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	**ACC**	367.85	249.75	10.41	25.39	−25.99	−84.62	24.00	2.37	2.43
	GAIL	97.29	71.24	6.48	9.28	−1.00	−1.09	11.00	24.87	2.67
	ICICI	457.90	29.95	0.88	23.41	−19.15	−55.72	34.00	1.09	0.89
	NTPC	13.01	41.54	0.94	3.32	−3.64	−16.57	44.00	1.76	1.93
	TE	237.63	276.66	6.29	24.91	−14.11	−36.62	44.00	1.93	1.10
US	**BoA**	−2.10	−2.57	−0.06	0.18	−0.23	−1.00	45.00	0.56	0.73
	BA	−2.12	2.98	0.06	0.87	−1.12	−3.22	47.00	1.14	1.47
	GE	0.09	1.24	0.03	0.18	−0.14	−0.39	48.00	1.38	1.09
	M3	6.21	7.78	0.16	0.96	−0.96	−1.86	48.00	1.41	1.40
	PEP	6.43	4.31	0.10	0.61	−0.49	−1.09	43.00	1.44	1.15
UK	**BT**	−4.54	12.64	0.27	3.15	−3.29	−8.86	47.00	1.18	1.24
	CCL	−467.77	−182.09	−4.05	47.76	−49.38	−166.17	45.00	0.85	0.88
	GSK	23.33	1.99	0.05	16.17	−13.39	−33.33	44.00	1.01	0.83
	RDS	282.03	133.83	3.04	14.93	−15.83	−41.29	44.00	1.50	1.59
	SBRY	−101.67	−9.35	−0.35	4.11	−3.41	−13.13	27.00	0.83	0.69
Brazil	**DTEX**	−2.61	1.76	0.05	0.25	−0.17	−0.49	39.00	1.53	1.05
	EMBR	3.16	1.52	0.04	0.37	−0.37	−1.00	42.00	1.22	1.21
	ITSA4	1.15	1.36	0.03	0.28	−0.23	−0.60	41.00	1.30	1.05
	PETR	−0.34	−0.90	−0.02	0.72	−1.00	−2.41	37.00	0.94	1.31
	TFA	−1.24	6.32	0.27	1.24	−1.53	−3.18	23.00	1.52	1.88

Table 12
Recommender 10 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	**ACC**	439.29	435.66	36.31	93.84	−44.25	−85.02	12.00	2.97	1.40
	GAIL	93.76	1.74	0.15	10.94	−21.44	−48.80	12.00	1.02	2.00
	ICICI	523.22	164.37	13.70	28.05	−15.00	−31.04	12.00	3.74	2.00
	NTPC	21.82	4.78	0.43	6.69	−4.78	−12.24	11.00	1.17	0.83
	TE	187.53	103.18	8.60	15.70	−5.60	−9.40	12.00	5.61	2.00
US	**BoA**	−2.26	−1.90	−0.16	0.28	−0.38	−1.42	12.00	0.37	0.50
	BA	2.37	−0.58	−0.05	1.27	−1.90	−4.73	12.00	0.94	1.40
	GE	−0.87	1.08	0.09	0.43	−0.38	−0.95	12.00	1.57	1.40
	M3	9.52	5.32	0.44	1.51	−1.05	−1.86	12.00	2.02	1.40
	PEP	3.93	2.70	0.22	0.55	−0.44	−0.77	12.00	2.55	2.00
UK	**BT**	−1.46	30.45	2.77	4.78	−2.59	−5.07	11.00	4.92	2.67
	CCL	−160.32	3.98	0.36	48.26	−27.01	−46.77	11.00	1.02	0.57
	GSK	15.87	17.41	1.58	14.84	−14.33	−24.38	11.00	1.24	1.20
	RDS	273.57	91.54	9.15	38.31	−34.58	−91.04	10.00	1.66	1.50
	SBRY	−21.77	−13.93	−1.16	8.39	−14.53	−27.46	12.00	0.81	1.40
Brazil	**DTEX**	−1.17	−0.05	0.00	0.23	−0.35	−0.66	10.00	0.96	1.50
	EMBR	1.81	−0.02	0.00	0.64	−0.37	−1.14	11.00	0.99	0.57
	ITSA4	1.15	0.48	0.05	0.28	−0.40	−0.82	9.00	1.39	2.00
	PETR	5.26	6.07	0.55	1.36	−0.42	−0.95	11.00	3.89	1.20
	TFA	−1.24	2.21	0.25	2.41	−1.48	−2.96	9.00	1.30	0.80

Table 13
Recommender 11 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	**ACC**	437.40	379.89	8.63	25.28	−13.27	−40.20	44.00	2.51	1.32
	GAIL	71.27	23.33	0.53	6.49	−6.00	−15.42	44.00	1.19	1.10
	ICICI	404.07	268.15	5.83	36.68	−22.45	−82.78	46.00	1.50	0.92
	NTPC	10.92	−0.40	−0.01	3.36	−3.10	−16.52	48.00	0.99	0.92
	TE	138.23	184.12	4.28	23.19	−13.77	−52.98	43.00	1.61	0.95
US	**BoA**	−1.41	0.62	0.01	0.22	−0.23	−0.77	46.00	1.13	1.19
	BA	7.16	3.24	0.07	1.06	−1.16	−3.16	47.00	1.13	1.24
	GE	−0.28	0.53	0.01	0.21	−0.22	−0.70	48.00	1.11	1.18
	M3	6.21	8.54	0.18	0.96	−0.92	−1.86	48.00	1.46	1.40
	PEP	6.95	3.46	0.08	0.53	−0.51	−1.48	46.00	1.34	1.30
UK	**BT**	−5.85	12.64	0.27	3.15	−3.29	−8.86	47.00	1.18	1.24
	CCL	−389.17	−31.84	−0.68	40.30	−36.74	−160.20	47.00	0.97	0.88
	GSK	−49.81	−38.81	−0.83	11.33	−12.48	−26.37	47.00	0.87	0.96
	RDS	214.86	94.53	2.05	15.48	−15.40	−43.78	46.00	1.31	1.30
	SBRY	−29.93	−38.90	−0.86	2.78	−4.35	−28.16	45.00	0.61	0.96
Brazil	**DTEX**	−2.30	−0.50	−0.01	0.13	−0.16	−1.03	43.00	0.85	1.05
	EMBR	0.70	0.73	0.02	0.26	−0.21	−0.67	44.00	1.15	0.91
	ITSA4	2.24	1.86	0.04	0.17	−0.14	−0.51	48.00	1.66	1.40
	PETR	−4.07	0.26	0.01	0.30	−0.29	−0.94	44.00	1.04	1.00
	TFA	0.10	6.14	0.13	0.91	−1.08	−3.14	46.00	1.32	1.56

Table 14
Recommender 12 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	**ACC**	348.05	375.07	31.26	78.42	−34.78	−71.14	12.00	3.16	1.40
	GAIL	13.81	7.96	0.72	6.90	−4.42	−16.96	11.00	1.30	0.83
	ICICI	405.11	141.29	11.77	39.00	−15.46	−24.18	12.00	2.52	1.00
	NTPC	4.10	19.85	1.99	10.42	−6.45	−11.64	10.00	1.62	1.00
	TE	136.14	168.16	15.29	26.64	−4.59	−9.40	11.00	10.16	1.75
US	**BoA**	−1.87	−3.30	−0.28	0.13	−0.41	−1.43	12.00	0.11	0.33
	BA	5.15	−6.65	−0.55	1.14	−1.40	−4.73	12.00	0.41	0.50
	GE	−0.12	−0.31	−0.03	0.33	−0.28	−0.57	12.00	0.84	0.71
	M3	4.10	11.40	0.95	2.48	−1.19	−1.86	12.00	2.92	1.40
	PEP	5.69	1.07	0.09	0.75	−0.38	−1.05	12.00	1.40	0.71
UK	**BT**	−1.46	30.45	2.77	4.78	−2.59	−5.07	11.00	4.92	2.67
	CCL	−136.44	−3.98	−0.36	36.57	−21.46	−46.77	11.00	0.97	0.57
	GSK	−21.95	79.10	6.59	18.05	−9.45	−20.40	12.00	2.67	1.40
	RDS	214.86	193.03	16.09	44.78	−24.08	−75.62	12.00	2.60	1.40
	SBRY	−10.43	−28.76	−2.40	3.81	−8.61	−27.06	12.00	0.44	1.00
Brazil	**DTEX**	−1.17	−0.08	−0.01	0.23	−0.29	−0.66	11.00	0.94	1.20
	EMBR	3.19	3.15	0.39	1.26	−0.47	−1.46	8.00	2.67	1.00
	ITSA4	−0.02	0.50	0.04	0.28	−0.20	−0.44	12.00	1.42	1.00
	PETR	−3.22	2.99	0.30	0.65	−0.22	−0.62	10.00	4.34	1.50
	TFA	1.18	−6.90	−0.57	0.75	−1.24	−3.31	12.00	0.30	0.50

Table 15
Recommender 13 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	**ACC**	439.29	542.72	12.62	40.68	−19.65	−59.80	43.00	2.38	1.15
	GAIL	130.87	35.57	1.32	9.29	−8.65	−30.00	27.00	1.34	1.25
	ICICI	523.37	479.49	10.42	37.48	−24.75	−69.10	46.00	1.97	1.30
	NTPC	14.64	44.33	1.23	4.96	−2.49	−12.19	36.00	1.99	1.00
	TE	216.64	191.14	4.25	24.89	−17.34	−48.06	45.00	1.50	1.05
US	**BoA**	−0.51	−2.21	−0.06	0.21	−0.27	−1.00	40.00	0.63	0.82
	BA	−3.83	−3.75	−0.08	1.41	−1.56	−3.76	48.00	0.90	1.00
	GE	0.13	0.36	0.01	0.18	−0.19	−0.48	47.00	1.09	1.14
	M3	15.50	11.14	0.23	1.77	−1.44	−4.79	48.00	1.34	1.09
	PEP	7.67	7.93	0.17	0.62	−0.47	−1.70	48.00	1.84	1.40
UK	**BT**	−0.17	12.14	0.26	5.25	−7.09	−24.68	47.00	1.09	1.47
	CCL	−201.11	−81.59	−1.70	42.34	−58.33	−175.12	48.00	0.93	1.29
	GSK	−174.68	−147.26	−3.07	19.70	−19.33	−67.66	48.00	0.73	0.71
	RDS	250.68	258.20	5.87	17.61	−19.30	−54.23	44.00	1.96	2.14
	SBRY	−88.43	−72.73	−1.58	6.14	−6.11	−19.80	46.00	0.59	0.59
Brazil	**DTEX**	−3.37	−2.95	−0.08	0.22	−0.33	−1.10	37.00	0.56	0.85
	EMBR	3.16	2.34	0.05	0.43	−0.31	−0.66	45.00	1.33	0.96
	ITSA4	1.15	0.91	0.02	0.33	−0.37	−0.85	37.00	1.15	1.31
	PETR	−0.12	−6.29	−0.17	0.54	−0.80	−2.77	38.00	0.61	0.90
	TFA	−1.24	0.24	0.01	1.18	−1.02	−3.14	30.00	1.01	0.88

Table 16
Recommender 14 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	**ACC**	394.72	490.63	49.06	99.63	−26.79	−71.29	10.00	5.58	1.50
	GAIL	25.40	−5.02	−0.46	3.75	−3.96	−9.55	11.00	0.79	0.83
	ICICI	555.71	85.47	7.77	92.92	−40.89	−150.05	11.00	1.30	0.57
	NTPC	−10.62	2.34	0.19	2.40	−1.38	−2.99	12.00	1.24	0.71
	TE	197.68	161.44	14.68	58.85	−22.14	−63.33	11.00	2.22	0.83
US	**BoA**	−0.26	1.90	0.17	0.57	−0.51	−1.09	11.00	1.92	1.75
	BA	−3.35	−3.48	−0.29	2.66	−4.42	−11.11	12.00	0.84	1.40
	GE	−1.13	−0.03	0.00	0.49	−0.28	−0.66	11.00	0.98	0.57
	M3	23.81	15.63	1.74	3.73	−5.23	−5.84	9.00	2.49	3.50
	PEP	8.60	6.90	0.63	1.35	−1.31	−2.57	11.00	2.75	2.67
UK	**BT**	−19.17	−13.93	−1.16	12.12	−10.65	−19.90	12.00	0.81	0.71
	CCL	−328.47	86.57	7.87	31.27	−33.08	−103.48	11.00	1.65	1.75
	GSK	−174.68	−126.37	−10.53	39.20	−46.05	−178.11	12.00	0.61	0.71
	RDS	16.36	−143.28	−13.03	19.33	−69.65	−162.19	11.00	0.49	1.75
	SBRY	−20.77	−31.14	−2.60	11.29	−9.54	−28.76	12.00	0.59	0.50
Brazil	**DTEX**	−3.37	−2.05	−0.23	1.12	−0.61	−1.04	9.00	0.52	0.29
	EMBR	3.68	0.60	0.10	0.74	−0.54	−0.95	6.00	1.37	1.00
	ITSA4	−0.25	1.77	0.15	0.36	−0.16	−0.32	12.00	3.28	1.40
	PETR	3.84	0.07	0.01	0.60	−0.83	−2.06	12.00	1.02	1.40
	TFA	−1.24	2.86	0.36	2.60	−1.89	−2.98	8.00	1.38	1.00

Table 17
Recommender 15 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	**ACC**	399.05	217.56	5.31	31.40	−24.90	−60.60	41.00	1.46	1.16
	GAIL	149.53	142.24	4.06	9.98	−8.84	−34.63	35.00	2.46	2.18
	ICICI	517.90	512.33	11.39	32.17	−26.28	−69.10	45.00	2.22	1.81
	NTPC	15.45	37.31	0.93	4.18	−2.66	−8.26	40.00	1.74	1.11
	TE	203.85	171.89	3.74	21.58	−19.46	−48.06	46.00	1.44	1.30
US	**BoA**	−1.76	−0.42	−0.01	0.24	−0.32	−1.02	45.00	0.94	1.25
	BA	−7.36	−7.34	−0.15	1.69	−1.99	−4.40	48.00	0.85	1.00
	GE	−1.21	−0.48	−0.01	0.19	−0.20	−0.54	47.00	0.90	0.96
	M3	21.02	23.05	0.48	1.80	−1.36	−4.79	48.00	1.85	1.40
	PEP	6.99	8.50	0.18	0.62	−0.45	−0.93	48.00	1.94	1.40
UK	**BT**	0.53	3.78	0.09	4.70	−7.39	−23.18	42.00	1.03	1.63
	CCL	−410.06	−282.58	−6.01	42.25	−56.37	−173.13	47.00	0.78	1.04
	GSK	−23.44	−49.75	−1.08	15.25	−13.64	−26.37	46.00	0.86	0.77
	RDS	248.20	298.50	6.63	16.59	−15.42	−58.71	45.00	2.38	2.21
	SBRY	−123.46	−78.51	−2.07	6.97	−7.33	−31.34	38.00	0.55	0.58
Brazil	**DTEX**	−3.41	−3.22	−0.07	0.30	−0.31	−1.19	44.00	0.61	0.63
	EMBR	1.61	0.67	0.02	0.34	−0.29	−0.92	44.00	1.10	0.91
	ITSA4	1.15	0.51	0.01	0.30	−0.32	−0.99	37.00	1.09	1.18
	PETR	−1.62	−4.10	−0.16	0.63	−0.61	−1.83	25.00	0.58	0.56
	TFA	−1.24	3.32	0.10	1.02	−0.82	−2.25	34.00	1.24	1.00

Table 18
Recommender 16 trading performance.

Market	Stock	B&H	TP	AP	P/ST	L/LT	MD	TT	PF	WR
India	**ACC**	448.35	670.28	60.93	94.28	−27.99	−60.30	11.00	8.98	2.67
	GAIL	86.10	−34.23	−3.11	14.07	−12.93	−48.76	11.00	0.62	0.57
	ICICI	381.13	−107.31	−10.73	59.89	−57.81	−119.30	10.00	0.69	0.67
	NTPC	−12.01	−4.98	−0.45	1.16	−1.37	−2.89	11.00	0.48	0.57
	TE	224.45	154.37	12.86	57.27	−18.86	−38.41	12.00	2.17	0.71
US	**BoA**	−0.73	0.17	0.01	0.31	−0.58	−1.33	12.00	1.07	2.00
	BA	−1.00	3.73	0.31	2.92	−1.56	−5.40	12.00	1.34	0.71
	GE	−0.47	−0.77	−0.07	0.45	−0.36	−0.70	11.00	0.70	0.57
	M3	23.81	15.55	1.94	3.29	−2.09	−4.01	8.00	4.71	3.00
	PEP	11.35	10.03	0.84	0.91	−0.03	−0.03	12.00	337.00	11.00
UK	**BT**	5.31	−20.40	−1.70	5.95	−7.16	−19.90	12.00	0.59	0.71
	CCL	−363.30	−54.73	−4.56	25.87	−19.78	−42.79	12.00	0.65	0.50
	GSK	−223.43	−27.86	−2.32	56.22	−31.59	−128.36	12.00	0.89	0.50
	RDS	199.94	112.44	10.22	19.53	−14.59	−22.89	11.00	3.57	2.67
	SBRY	−50.32	−12.14	−1.10	3.52	−4.96	−8.86	11.00	0.59	0.83
Brazil	**DTEX**	−3.37	−1.85	−0.19	0.31	−0.40	−1.00	10.00	0.33	0.43
	EMBR	5.20	1.88	0.16	0.64	−0.80	−0.93	12.00	1.59	2.00
	ITSA4	1.06	0.94	0.08	0.31	−0.24	−0.47	12.00	1.78	1.40
	PETR	−1.62	−5.82	−0.53	0.45	−1.34	−5.45	11.00	0.28	0.83
	TFA	−1.24	4.52	0.45	1.18	−1.25	−1.83	10.00	2.20	2.33

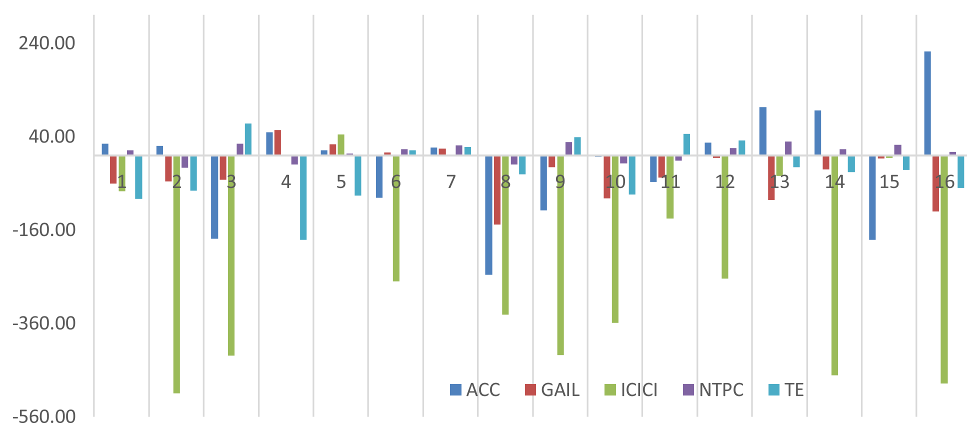

Fig. 6. Excess profit (in INR) generated by the sixteen recommender systems compared to B&H for Indian stocks.

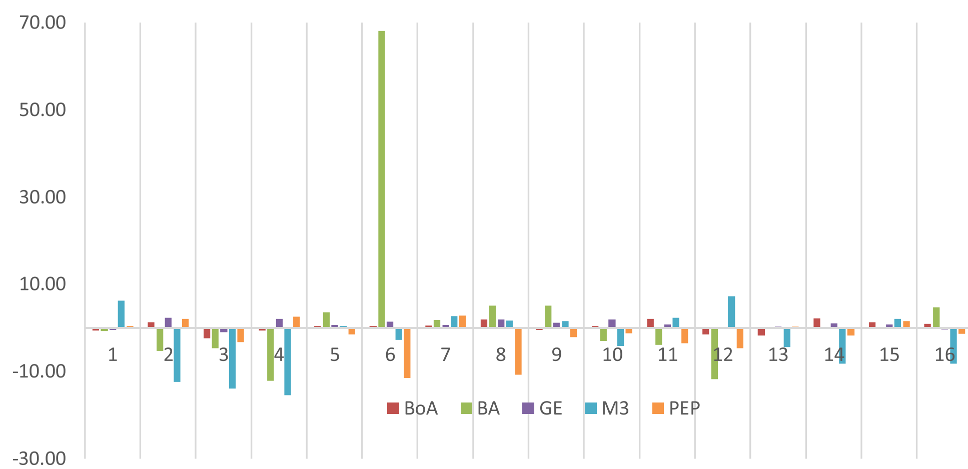

Fig. 7. Excess profit (in US$) generated by the sixteen recommender systems compared to B&H for US stocks.

From the Tables 3–18, it is observed that the recommender systems trading on a weekly basis tend to generate total profits that are higher or at least, as good as similar recommenders trading on a monthly basis. It is also observed that weekly trading recommenders are able to do this at the cost of much higher TT and poorer P/ST, L/LT and PF. It is also observed that Recommender 7 is able to generate better results than the B&H strategy (as seen by the positive excess profits in Figs. 6–9) in fifteen out of the total twenty stocks considered. This ability to generate positive excess returns is the most important aspect of the trading recommender system, since this measure indicates if following the trading recommendations will be helpful in beating the passive B&H strat-

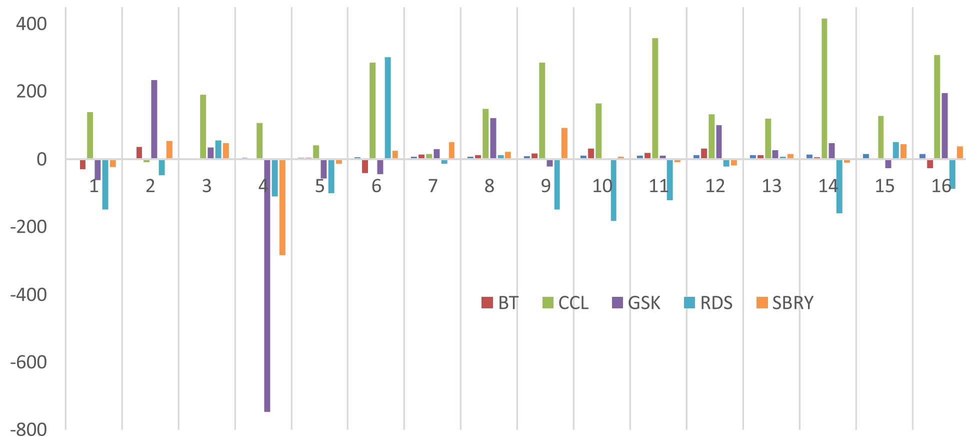

Fig. 8. Excess profit (in GBp) generated by the sixteen recommender systems compared to B&H for UK stocks.

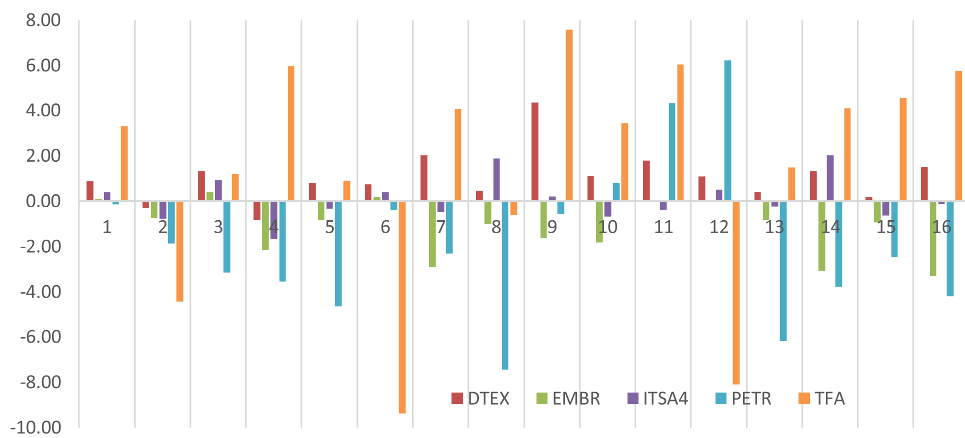

Fig. 9. Excess profit (in BRL) generated by the sixteen recommender systems compared to B&H for Brazil stocks.

egy at all. From the performance measures, it is seen that Recommender 7, while generating the highest TP in only two of the stocks considered, is still able to generate positive excess profits. This is attributable to the low L/LT and MD values that the system is able to maintain for the stocks considered. However, on the downside, the recommender is not able to generate very high P/ST either, leading to the inference that the recommender 7 is not affected by sharp movements in stock prices.

4. Conclusions

Totally sixteen variants of the proposed recommender system were considered. Economic performance of these recommender systems was evaluated on eight different performance measures, as suggested in (Brabazon & O' Neill, 2006) . The best performance (based on the excess profit generated, was demonstrated by the recommender system trained using the cyclic component extracted from the stock price time series with regression tree based feature selection and sliding window based adaptive learning and trading weekly (Recommender 7). It is believed that taking cyclic components helped the recommender system capture the short-term fluctuations in the stock prices better, thus generating better results when compared to considering the entire dataset. The sliding window technique for adaptive learning helps the recommender system learn patterns from the latest samples while at the same time discarding the oldest samples, unlike the expanding window technique wherein the new data just gets added to the training data. The use of an expanding window can lead to undesirable levels of generalization during the training process which could have resulted in the poorer performance of the variants adapting through this technique. Recommender systems trading weekly are re-trained on a weekly basis, while those trading monthly are re-trained on a monthly basis. It must be noted that the re-training happens as soon as one round trade is executed, i.e. all the new stock price data up to the day when the stock is sold is immediately taken for re-training and the next week (or month) begins from the very next day. Hence, the worst case training scenario in the weekly trading case is to buy on the first day of the week and sell on the last day (typically 5-day week). The recommender thus, has to wait for five days before it can take these five data points into its training set and get retrained. For the monthly trading case, the worst case training scenario is to buy on the first day of the month and sell on the last day (typically 20-day month). The recommender thus, has to wait for twenty days before it can take these twenty data points into its training set and get retrained. Due to the more frequent re-training, the recommenders trading weekly as seen from the results above, tend to give better results.

References

Afolabi, M. O., & Olude, O. (2007). Predicting stock prices using a hybrid Kohonen self organizing map (SOM). In R. H. Sprague, Jr. (Ed.), *Fortieth annual hawaii international conference on system sciences* p. 48. Waikoloa,Hawaii: IEEE.

Atsalakis, G., & Valavanis, K. (2009). Forecasting stock market short-term trends using a neuro-fuzzy based methodology. *Expert Systems with Application, 36*(7), 10696–10707.

Atsalakis, G., & Valavanis, K. (2009). Surveying stock market forecasting techniques – part II: Soft computing methods. *Expert Systems with Applications, 36*(3), 5932–5941.

Bação, F. V. (2005). Self-organizing maps as substitutes for k-means clustering. *Computational science* (pp. 476–483). Berlin Heidelberg: Springer.

Bodas Sagi, D. J., Soltero, F. J., Hidalgo, J. I., Fernández, P. C., & Fernández, F. C. (2012). A technique for the optimization of the parameters of technical indicators with multi-objective evolutionary algorithms. In *Proceedings of 2012 IEEE congress on evolutionary computation* (pp. 1–8). Brisbane: IEEE.

Brabazon, A., & O'Neill, M. (2006). *Biologically inspired algorithms for financial modelling* (1st ed.). New York: Springer-Verlag.

Brabazon, A., O'Neill, M., & Dempsey, I. (2008). An introduction to evolutionary computation in finance. *IEEE Computational Intelligence Magazine, 3*(4), 42–55.

Briza, A. C., & Naval, P. C. (2011). Stock trading system based on the multi-objective particle swarm optimization of technical indicators on end-of-day market data. *Applied Soft Computing, 11*(1), 1191–1201.

Canelas, A., Neves, R., & Horta, N. (2013). A SAX-GA approach to evolve investment strategies on financial markets based on pattern discovery techniques. *Expert Systems with Applications, 40*(5), 1579–1590.

Chen, P. (1996). A random-walk or color-chaos on the stock market? - time-frequency analysis of S&P indexes. *Studies in Nonlinear Dynamics & Econometrics, 1*(2), 87–103.

Clare, A., & Miffre, J. (1995). A note on forecasting the CAC40 and DAX stock index futures. *Applied Economics Letters, 2*(10), 327–330.

Cootner, P. (1964). *The random character of stock market prices*. Cambridge, MA: MIT Press.

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

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

De Brito, R. F., & Oliveira, A. L. (2012). Comparative study of FOREX trading systems built with SVR+GHSOM and genetic algorithms optimization of technical indicators. In *Proceedings of international conference on tools with artificial intelligence: 1* (pp. 351–358).

Eng, W. (1988). *The technical analysis of Stocks, options & Futures- Advanced Trading systems and techniques*. New Delhi: Vision Books.

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business, 38*(1), 34–105.

Fama, E. F. (1965). Random walks in stock market prices. *Financial Analysts Journal, 21*(5), 55–59.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance, 25*(2), 383–417.

Fraser, A. M., & Swinney, H. L. (1986). Independent Coordinates for Strange Attractors from Mutual Information. *Physical review A, 33*(2), 1134–1140.

Gao, Z., & Xu, X. (2009). Stock bubbles' nature:cluster analysis of chinese shanghai a share based on SOM neural network. In S. Wang, L. Yu, F. Wen, S. He, Y. Fang, & K. K. Lai (Eds.), *Proceedings of international conference on business intelligence and financial engineering* (pp. 12–16). Beijing: IEEE.

Hodrick, R., & Prescott, E. (1997). Postwar U.S. Business cycles: An empirical investigation. *Journal of Money, Credit, and Banking, 29*(1), 1–16.

Hsu, C.-M. (2011). A hybrid procedure for stock price prediction by integrating self--organizing map and genetic programming. *Expert Systems with Applications, 38*(11), 14026–14036.

Hsu, S.-H., Hsieh, J. P.-A., Chih, T.-C., & Hsu, K.-C. (2009). A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression. *Expert Systems with Applications, 36*(4), 7947–7951.

Huffake, R. (2010). Phase space reconstruction from time series data: Where history meets theory. In M. Fritz, U. Rickert, & G. Schiefer (Eds.), *Proceedings of the 4th international european forum on system dynamics and innovation in food networks* (pp. 1–9). Innsbruck-Igls: Austria: International Center for Food Chain and Network Research.

Ismail, S., Shabri, A., & Samsudin, R. (2011). A hybrid model of self-organizing maps (SOM) and least square support vector machine (LSSVM) for time-series forecasting. *Expert Systems with Applications, 28*(8), 10574–10578.

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in Data: an introduction to cluster analysis* (1st ed.). Hoboken, NJ, USA: John Wiley & Sons.

Kennel, M., Brown, R., & Abarbanel, H. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A, 45*, 3403–3411.

Khan, A. U., Bandopadhyaya, T. K., & Sharma, S. (2008). Classification and identification of stocks using SOM and genetic algorithm based backpropagation neural network. In *Proceedings of international conference on innovations in information technology* (pp. 292–296).

LeBaron, B. (1994). Chaos and nonlinear forecastability in economics and finance. *Philosophical Transactions: Physical Sciences and Engineering, 348*(1688), 397–404.

Li, S.-T., & Kuo, S.-C. (2008). Knowledge discovery in financial investment for forecasting and trading strategy through wavelet-based SOM networks. *Expert Systems with Applications, 34*(2), 935–951.

Lin, J., Keogh, E., Wei, L., & Lonardi, S. (2007). Experiencing SAX: A novel symbolic representation of time series. *Data Mining and Knowledge Discovery, 15*(2), 107–144.

McKibben, W. (1972). Econometric Forecasting of Common Stock Investment Returns: A New Methodology Using Fundamental Operating Data. *The Journal of Finance, 27*(2), 371–380.

Nair, B. B., & Mohandas, V. P. (2015a). An Intelligent recommender system for stock trading. *Intelligent Decision Technologies, 9*(3), 243–269.

Nair, B. B., & Mohandas, V. P. (2015b). Artificial Intelligence Applications in Financial Forecasting- A Survey and Some Empirical Results. *Intelligent Decision Technologies, 9*, 99–140.

Nair, B. B., Xavier, N., Mohandas, V. P., Anusree, E., & Kumar, P. (2014). A GA-optimized SAX- ANN based stock level prediction system. *International Journal of Computer Applications, 106*(15), 7–12.

Nair, B. B., Minuvarthini, M., Sujithra, B., & Mohandas, V. P. (2010). Stock market prediction using a hybrid neuro-fuzzy system. In *Proceedings of international conference on advances in recent technologies in communication and computing* (pp. 243–247).

Nair, B. B., Mohandas, V. P., & Sakthivel, N. R. (2011). Predicting stock market trends using hybrid ant-colony-based data mining algorithms: An empirical validation on the Bombay stock exchange. *International Journal of Business Intelligence and Data Mining, 6*(4), 362–381.

Nair, B. B., Mohandas, V. P., Nayanar, N., Teja, E. S., Vigneshwari, S., & Teja, K. V. (2015). A Stock Trading Recommender System Based on Temporal Association Rule Mining. *SAGE Open, 5*(2), 1–10.

Nair, B. B., Sai, S. G., Naveen, A. N., Lakshmi, A., Venkatesh, G. S., & Mohandas, V. P. (2011). A GA-artificial neural network hybrid system for financial time series forecasting. In V. Das, G. Thomas, & F. Gaol (Eds.), *Lecture notes in computer science-communications in computer and information science: 147* (pp. 499–506). Nagpur: Springer-Verlag.

Perc, M. (2006). Introducing Nonlinear Time Series Analysis in Undergraduate Courses. *FIZIKA A (Zagreb), 15*(2), 91–112.

Soltero, F. J., Bodas, D. J., Hidalgo, J. I., Fernández, P., & Fernández-de-Vega, F. (2012). Optimization of technical indicators in real time with multiobjective evolutionary algorithms. In *Proceedings of 14th international conference on genetic and evolutionary computation* (pp. 1535–1536). Philadelphia: ACM.

Sugunsil, P., & Somhom, S. (2009). Short Term Stock Prediction Using SOM. In J. Yang, A. Ginige, H. C. Mayr, & R.-D. Kutsche (Eds.), *Proceedings of 3rd international united information systems conference: 20* (pp. 262–267). Sydney: Springer Berlin Heidelberg.

Takens, F. (1981). Detecting strange attractors in turbulence. In D. Rand, & L.-S. Young (Eds.), *Symposium held at the university of warwick: 898* (pp. 366–381). Coventry: Springer.

Torgo, L. F. (1999). Inductive Learning of Tree-based Regression Models. *AI Communications, 13*(2), 137–138.

Tsai, C.-F., Lin, Y.-C., & Wang, Y.-T. (2009). Discovering stock trading preferences by self-organizing maps and decision trees. *International Journal on Artificial Intelligence Tools, 18*(4), 603–611.

Wan, C., & Chai, S. (2014). Research on the prediction of Stock market based on chaos and SVM. *WSEAS Transactions on Information Science and Applications, 11*, 186–195.

Widiputra, H., & Christianto, L. (2012). Indonesia Stock Exchange Liquid Stocks Identification using Self-Organizing Map. In *Proceedings of international conference on uncertainty reasoning and knowledge engineering* (pp. 131–136). Jakarta: IEEE.

Xu, R., & Wunsch, D., II (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks, 16*(3), 645–678.

Yahoo! Finance (2015,. January). Historical Stock Price Data. Retreived from: http://finance.yahoo.com

Zduniak, A., & Lusakowski, J. (1995). Autocorrelation Function and Mutual Information from Short Experimental Time Series. *Acta Physica Polonica-Series A General Physics, 87*(1), 257–260.

Zhang, H., & Li, S. (2010). Forecasting volatility in financial markets. In Y. Wu (Ed.), *Proceedings of international conference on advanced measurement and test: 439-440* (pp. 679–682).