International Journal of Computational Intelligence and Informatics, Vol. 5: No. 4, March 2016

Predictive Regression Forecasting for Share Trading Based on Total Part Summation

K. Jayanthi
Department of Computer Science
Thiruvalluvar Govt. Arts College
Rasipuram, Nammakkal (DT)
jaysureshlaya@gmail.com

P. Suresh
Department of Computer Science
Salem Sowdeswari College
Salem - 636 010.
sur_bh0071@rediffmail.com

***Abstract-* In today's changing economic scenario, there is adequate opportunity to influence the varied sources of time series data. These time series data are now easily available and accessible to the intelligent decision maker. Many research works have been contributed towards the interdisciplinary notion of data mining for forecasting of share trading when used to solve time series problems. Though, diversified ensemble of individual predictors was made, but individual time series forecasting was not made in an efficient manner. Also the presence of irrelevant attributes posed significant problems. In this paper to address individual time series forecasting and reduce the forecasting error a framework called Predictive Regression Forecast based on Total Part Summation (PRF-TPS) is designed. PRF-TPS framework is developed on share trading dataset to forecast time series of each component in the presence of irrelevant attributes and therefore reduce the forecasting error. PRF operates on large collections of events which summarizes the events with individual time series forecasting. Total Part Summation (TPS) forecasts discretely, the Dividend Stock Price Ratio, Income Growth Ratio and Stock Price Income Growth of share trading. Total Part Summation effectively learns the irrelevant attributes at different time series properties of the components. The TPS produces statistically and economically significant gains for investors and performs better out of sample in predictive regressions. An intensive and comparative study shows the efficiency of these enhancements and shows better performance in terms of run-time, forecast efficiency, test error rate, and predictive accuracy on shares. Experimental analysis shows that PRF-TPS framework is able to reduce the run-time for share rate forecasting by 36.86% and reduce the test error rate by 50.79% compared to the state-of-the-art works.**

Keywords- Diversified Ensemble, Individual Predictors, Time Series Forecasting, Predictive Regression Forecast and Total Part Summation.

I. INTRODUCTION

Data mining plays an important role in forecasting to solve the issues related to time series data. The target of talk is to explain how to get the most value out of the innumerable of obtainable time series data by making use of data mining techniques. Multilevel Architecture for Time Series Prediction (MA-TSP) [1] constructed diversified ensemble of individual predictors aiming at improving the rate of prediction.

Diversity on Online Ensemble Learning (DOEL) [2] minimized the prediction error using mutually exclusive and non-heterogeneous categories. Full Functional Dependency (FFD) [3] with cross attribute correlation resulted in low information loss. Though prediction was improved in the above method, but run-time remained unsolved. Issues related to run-time are addressed in PRF-TPS framework using Exponent-based Predictive Regression Forecast algorithm.

Various prediction models for interdisciplinary works were presented. Real-time solar forecasting model [5], Direct and Indirect discrimination [6] using legitimate classification rules and genetic programming [7] resulted in the improvement of prediction rate. Time-stamped observational data [8] was applied to sequence of observations resulting in the improvement of running time and accuracy. Clustering methods [9] using correlated probabilistic graphs was designed at improving the pruning techniques. Characterization of events was made in an efficient manner by applying multivariate reconstructed phase space (MRPS) [10]. This in turn resulted in the improvement of prediction accuracy.

All the above methods resulted in the improvement of prediction accuracy, but error rate remained unaddressed. The error rate was reduced in PRF-TPS framework using Total Part Summation model.

Various mining models were applied by several researchers for efficient forecasting. In [11], Intrinsic and Extrinsic Domain Relevance was applied to opinion mining to forecast online reviews, resulting in the improvement of forecasting accuracy. For efficient time series classification, Probabilistic Sequence Translation Alignment (PSTA) [12] model was applied to improve the probability rate.

In [13], temporal patterns were obtained for efficient event characterization in dynamic data stream. This in turn resulted in the improvement of prediction rate. Another model for runtime optimization [14] was introduced for prediction optimization using tree-based models. In [15], stock market forecasting using rough set model was introduced that resulted in the improvement of precision rate.

Efficient prediction of stock market has become an important issue to be handled. In [16], bi-clustering method was applied to improve the prediction of stock rate. Fuzzy time series [17] was applied to Taiwan Stock Index to improve the forecasting rate. In [18], deterministic echo state networks were applied to improve the accuracy and efficiency using randomized model building stages.

Spatio temporal correlations [19] were applied for solar forecasting that resulted in efficient forecasting using forecast models. Classification rules for time series analysis [20] was made to improve the forecasting rate using segmentation and piecewise polynomial modeling.

The prime focus of this paper is to address individual time series forecasting and reduce the forecasting error using Predictive Regression Forecast and Total Part Summation models. The problem studied here is about the forecasting of share trading for investors' usage. We have chosen Predictive Regression Forecast model for improving the predictive accuracy and Total Part Summation model for improving forecasting efficiency. The proposed framework extracts knowledge in the form of rules from the Istanbul Stock Exchange daily dataset that would guide investors whether to buy, sell, or hold a share.

The paper is structured as follows. Section II explains about the framework Predictive Regression Forecast based on Total Part Summation with the aid of diagrams and algorithm. Experimentation is enclosed in Section III. Section IV include data preparation, analysis, results, and discussion of results with the aid of table and graph form. Finally, concluding remarks are provided in Section V.

II. RELATED WORK

In this section, first the block diagram for Predictive Regression Forecast based on Total Part Summation is described, then the novel framework for forecasting of share is proposed and finally the PRF and TPS models are presented. Figure 1 shows the block diagram of Predictive Regression Forecast based on Total Part Summation.

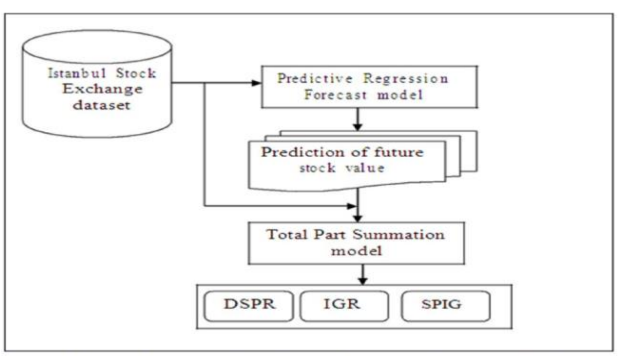

Figure 1 Block diagram of Predictive Regression Forecast based on Total Part Summation

As shown in the figure, the block diagram of Predictive Regression Forecast based on Total Part Summation includes two steps. The first step is the construction of Predictive Regression Forecast model that efficiently predicts the future stock based on the values of Moving Average Estimation and

Exponent Moving Average Estimation. Predicting the future stock based on these two values improves the predictive accuracy on shares and also reduces the runtime for prediction.

The second step involves the design of Total Part Summation model. This step evaluates three parameters, Dividend Stock Price Ratio (DSPR), Income Growth Ratio (IGR) and Stock Price Income Growth (SPIC) of share trading. This in turn reduces the test error rate and therefore improves the forecast efficiency. The elaborate description of the framework is provided in the forthcoming sections.

III. DESIGN OF PREDICTIVE REGRESSION FORECAST MODEL

Trend analysis and forecasting of share trading play a vital role in complex events. On the basis of the past observations, stock traders efficiently predict the future stock's price based. In this section, a Predictive Regression Forecast (PRF) model is designed aiming at improving the prediction accuracy on shares at minimum time interval. Figure 2 given below shows the block diagram of Predictive Regression Forecast model.

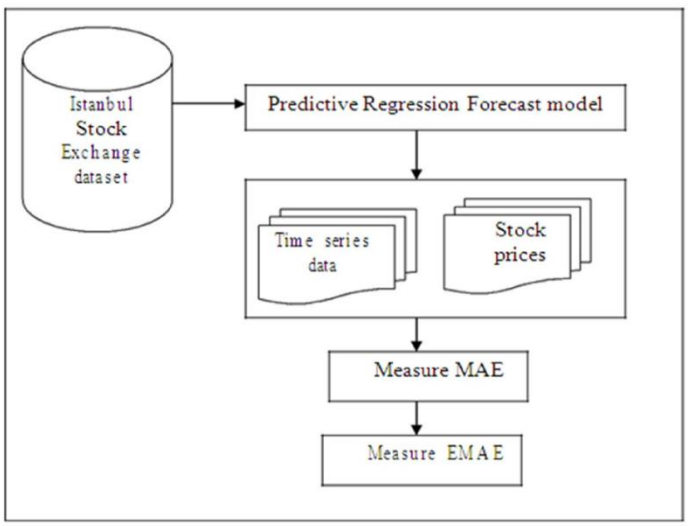

Figure 2 Block diagram of Predictive Regression Forecast model

As shown in the figure, the block diagram includes measurement of two values MAE and EMAE that considers time series data from the stock value listed on the Istanbul Stock Exchange Dataset. The above said two values are used aiming at improving the prediction accuracy of stock values and minimize the prediction runtime using Predictive Regression Forecast model.

Let us consider the time series data '$TS_i = TS_1, TS_2, \ldots, TS_n$' with a given window length 'wl' and stock prices 'pi' formulated as given below.

$$TS_1 \rightarrow \{p_1, p_2, \ldots, p_{wl}\} \quad (1)$$

$$TS_2 \rightarrow \{p_2, p_3, \ldots, p_{wl+1}\} \quad (2)$$

$$TS_n \rightarrow \{p_n, p_{n+1}, \ldots, p_{wl+n-1}\} \quad (3)$$

From (1), (2) and (3) the prices '$p_1, p_2, .., p_{wl}$', '$p_2, p_3, .., p_{w+1}$' for time series data 'TS_1' and 'TS_2' are obtained at time 'i'. Then, the discrete features generated using Predictive Regression Forecast model is as given below.

$$User\left(p_i, t\right) \rightarrow \left\{f_1, f_2, \ldots, f_n\right\} \qquad (4)$$

From (4), 'pi' represents the price history for some company 'i' where 'f_1, $f_2, \ldots, f_n$' denotes the discrete features generated by a user at time 'i'. Therefore, a trend in any time series data '$TS = TS_1, TS_2, .., TS_n$' represents a period of time where

'$TS_i \in TS$'

The Predictive Regression Forecast model with the view of improving the prediction accuracy rate, identifies the trending direction of a time series 'TS'. This is identified in order to create a smoothed representation, 'S' that describes 'TS'. Next, the initial difference 'ID' of the smoothed representation 'S' is formulated as given below.

$$ID = S(i) - S(i-1) \qquad (5)$$

The Predictive Regression Forecast model uses Moving Average Estimation (MAE) to efficiently derive the final smoothed representation. MAE is running averages with finite window size over a dataset used for time series data.

The MAE with size 'n' for a data point, (i.e. stock values) '$TS_i \in TS$' is the un-weighted mean for 'n' past data points (i.e. past stock values). The value of MAE is then formulated as given below.

$$MAE\left(TS_i, n\right) = \left(\frac{TS_i + TS_{i-1} + TS_{i-2} \ldots + TS_{i-n-1}}{n}\right) \qquad (6)$$

From (6) the value of 'MAE' is obtained based on the time series data 'TS_i' over a period of time 't' with 'n' observations being made in a significant manner. With MAE, whenever a new observation (i.e. a new stock value) is available, the oldest observation (i.e. the old stock value) in the window is dropped and the new observation (i.e. new stock value) is included.

With the objective of reducing the runtime, Predictive Regression Forecast model in addition to Moving Average Estimation uses the Exponent Moving Average Estimation (EMAE). The Exponent Moving Average Estimation in Predictive Regression Forecast model is weighted moving average where each time series data 'TS_i' is scaled by an exponential predictive factor 'a'.

$$EMAE\ (TS_i, n) = (\alpha * TS_{i-1}) + (1 - \alpha * TS_{i-2}) \ldots\ldots \qquad (7)$$

Therefore by applying EMAE several trends in stock market are analyzed that not only minimizes the runtime for prediction, but also improves the prediction accuracy in an extensive manner. Figure 3 shows the algorithmic description of Exponent-based Predictive Regression Forecast.

Input: Time Series data '$TS_i = TS_1, TS_2, \ldots, TS_n$', stock prices '$p_i = p_1, p_2, \ldots, p_n$', User '$User_i = User_1, User_2, \ldots, USer_n$', time '$i$'.
Output: Predictive forecast efficient
Step 1: Begin Step 2: For each user $User_i$ Step 3: Measure discrete features using (4) Step 4: Measure initial different smoothed representation using (5) Step 5: Evaluate moving average estimation using (6) Step 6: Evaluate exponent moving average estimation using (7) Step 7: End for Step 8: End

Figure 3 Exponent-based Predictive Regression Forecast algorithm

As shown in the above figure, the structure of the algorithm includes four steps. The first step extracts the discrete features from Oslo Benchmark Index listed on the Oslo stock exchange. Followed by this, the second step evaluates the initial different smoothed representation based on the time series data. The third data then measures the value of moving average estimation to record the new and old observation (i.e. new and old stock values). Finally, to arrive at smoother trend line, exponent moving average estimation is obtained. This in turn improves the prediction accuracy of stock values and at the same time reduces the run time in predicting the stock values.

IV. CONSTRUCTION OF TOTAL PART SUMMATION

In this section, the Total Part Summation with the objective of improving the forecast efficiency and at the same time to reduce the test error rate is designed. The share trading returns are divided into three elements using Total Part Summation (TPS). Using TPS, the proposed method forecasts discretely, the Dividend Stock Price Ratio (DSPR), Income Growth Ratio (IGR) and Stock Price Income Growth (SPIG).

Let us split the total return of the share market consider into 'Ret_{i+1}', representing the return yielded from time 'i *to* $i+1$', 'Div_{i+1}' representing the dividend from time 'i *to* $i+1$', and 'p_{i+1}', from time 'i *to* $i+1$'. The Dividend Stock Price Ratio (DSPR) is mathematically evaluated using the dividend value in varied time period and is as given below.

$$DPR = \left(\frac{Div_{i+1}}{p_i} \right) \qquad (8)$$

From (8), Dividend Stock Price Ratio '*DPR*' is measured based on the dividend per share 'Div_{i+1}' in stock market over price '*pi*' respectively. The Income Growth Ratio is the ratio of returns obtained to the price of the share and is formulated as given below

$$IGR = \left(\frac{Ret_{i+1}}{p_i} \right) \qquad (9)$$

From (9), the Income Growth Ratio '*IGR*', is measured using the returns 'Ret_{i+1}' over price '*pi*'. Finally, the third part Stock Price Income Growth is measured as given below.

$$SPIG = DPR * IGR \qquad (10)$$

From (10), the Stock Price Income Growth '*SPIG*' is the product of Dividend Stock Price Ratio '*DPR*' and Income Growth Ratio '*IGR*'. On the basis of the above three elements, the proposed framework Predictive Regression Forecast based on Total Part Summation effectively learns the irrelevant attributes at different time series properties of the components.

The idea behind Total Part Summation at different time series properties of the component is that each Exponent Moving Average Estimation forecasts the trend at different time scales. Hence, when a shorter exponent moving average increases than a longer exponent moving average, the TPA states that the long-term trend as measured by the increasing moving average may reverse to the decreasing trend as measured by the short exponent moving average.

Consequently, a buy trend (i.e. purchase share) is generated until the short exponent moving average goes below the long exponent moving average. The TPA produces statistically and economically significant gains for investors and performs better out of sample at crossover point. This is due to the fact that buying trend is higher at the crossover point because the closing price of the share over the past '*a*' days has exceeded the average closing price over the past '*b*' days. Therefore, the crossover point '*CP*' is formulated as given below.

$$CP = MAX\,(a, b) \qquad (11)$$

Let us consider two company shares with different values, 'a' and '*b*', where '$a < b$' are predicted simultaneously and buying and selling shares are generated at points where the two moving averages of two shares intersect. Buying shares are generated when '*a*' rises above '*b*', and selling shares are generated when '*a*' falls below the '*b*' value. Therefore, the crossover point with short exponent and long exponent moving average is formulated based on the following rules.

$$CP_{ab}(i) = Buy\ Share\ ;\ EMAE_a < EMAE_b \tag{12}$$

$$CP_{ab}(i) = Sell\ Share\ ;\ EMAE_a > EMAE_b \tag{13}$$

From (12) and (13), depending on the resultant values of exponent moving average estimation '*EMAEa and EMAEb*', either share is purchased or sale is sold. Figure 4 shows the algorithmic description of Summation Crossover.

Input: Time Series data '$TS_i = TS_1, TS_2, \ldots, TS_n$', stock prices '$p_i = p_1, p_2, \ldots, p_n$', User '$User_i = User_1, User_2, \ldots, USer_n$', time '$i$',
Output:
Step 1: Begin Step 2: For each user $User_i$ Step 3: Measure Dividend Stock Price Ratio using (8) Step 4: Measure Income Growth Ratio using (9) Step 5: Measure Stock Price Income Growth using (10) Step 6: Measure crossover point using (11) Step 7: If $EMAE_a < EMAE_b$ then Step 8: $CP_{ab}(i) = Buy\ Share$ Step 9: End if Step 10: If $EMAE_a > EMAE_b$ then Step 11: $CP_{ab}(i) = Sell\ Share$ Step 12: End if Step 13: End for Step 14: End

Figure 4 Summation Crossover algorithm

The summation crossover algorithm given above includes four steps. The first step evaluates the Dividend Stock Price Ratio. Followed by which the Income Growth Ratio is measured based on the dividend per share. The third step evaluates the Stock Price Income Growth to obtain the price of the stock on the basis of income growth. Finally, buying of share or selling of share is efficiently carried out. This in turn improves the forecasting efficiency in a significant manner.

V. EXPERIMENTAL SETTINGS

The framework Predictive Regression Forecast based on Total Part Summation (PRF-TPS) is developed for effective forecasting of share trading using JAVA platform. The PRF-TPS framework uses the Istanbul Stock Exchange data set [4] to obtain daily price data for the indices from http://finance.yahoo.com and http://imkb.gov.tr and converted the prices the returns.

All stocks listed on the indices from the two websites are easily transferable which makes our framework simpler to validate. The data used is available from January 5, 2009 to February 22, 2011. The days for which the Turkish stock exchange was closed was excluded. In the case of missing data for the other indices, the previous day's value was used. Stock exchange returns the Istanbul stock exchange national 100 index, Standard & poorâ€™s 500 return index, Stock market return index of Germany, Stock market return index of UK, Stock market return index of Japan, Stock market return index of Brazil, MSCI European index, MSCI emerging markets index. The performance of the PRF-TPS framework is evaluated for parameters such as time (year), run-time, forecast efficiency, test error rate, and predictive accuracy on shares.

VI. DISCUSSION

Predictive Regression Forecast based on Total Part Summation (PRF-TPS) framework is compared with the existing Multilevel Architecture for Time Series Prediction (MA-TSP) [1] and Diversity on Online Ensemble Learning (DOEL) [2] using Istanbul Stock Exchange dataset extracted from UCI repository. Certain measurements made are run-time, forecast efficiency, test error rate and predictive accuracy on shares with respect to returns and different years.

Impact of Run-time: Run-time refers to the time taken to measure the share price with respect to number of returns. The mathematical formulation for run-time is as given below.

$$RT = n * Time\ (p_i) \tag{14}$$

From (14), the run-time '*RT*' is measured using the returns '*n*' and time to obtain the share price '*pi*' respectively. The run-time is measured in terms of milliseconds (ms). Lower the run-time more efficient the method is said to be.

Table 1 Tabulation for run-time

Returns	Run-time (ms)		
	PRF-TPS	MA-TSP	DOEL
10	2.4	3.9	4.4
20	3.8	4.9	6.6
30	5.5	6.6	8.3
40	7.3	8.4	10.1
50	5.8	6.9	8.6
60	9.4	10.5	12.2
70	11.3	12.4	14.1

In table 1 we evaluate the performance of run-time using the PRF-TPS framework and compared with two other methods MA-TSP and DOEL. Number of returns used in this experiment ranges from 10 to 70.

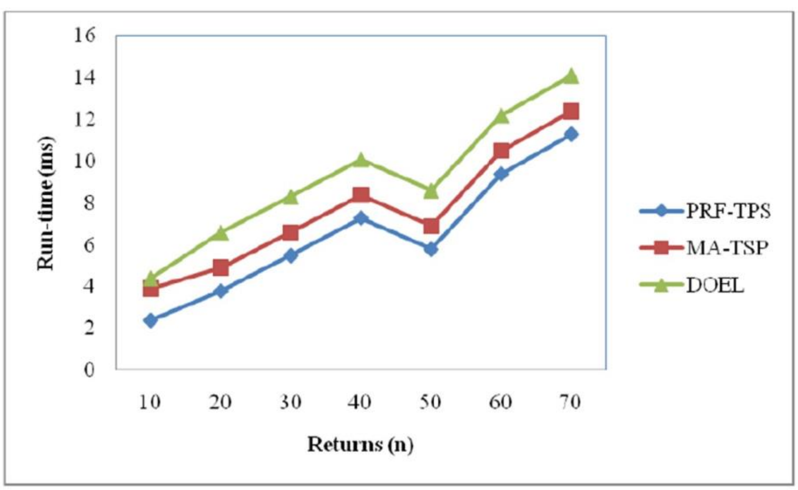

Figure 5 Measure of run-time

Figure 5 shows the run-time for different forecasting method as a function of different returns. Compared to the existing Multilevel Architecture for Time Series Prediction (MA-TSP) and Diversity on Online Ensemble Learning (DOEL) [2], the proposed PRF-TPS framework consumes less run-time through moving average estimation operations. This is because by applying Predictive Regression Forecast (PRF) model for determining the varied stock prices in order to create smoothed representation, the run-time decreases using PRF-TPS framework by 23.84% compared to MA-TSP. In addition, by measuring the moving average estimation and exponent moving average estimation, for '*n*' past data points significantly reduces the run-time by 49.87% compared to DOEL respectively.

Impact of forecast efficiency: Forecast efficiency for predicting share is measured on the basis of the relevant attributes retrieve from the total attributes. Forecast efficiency is mathematically evaluated as given below.

$$FE = \left(\frac{R_a}{TA}\right) * 100 \qquad (15)$$

From (15), the forecast efficiency '*FE*' is obtained using relevant attributes '*Ra*' and total attributes 'TA' respectively. Higher the forecast efficiency, more efficient the method is said to be.

Table 2 Tabulation for forecast efficiency

Year	Forecast efficiency (%)		
	PRF-TPS	**MA-TSP**	**DOEL**
Jan 2009	75.35	63.45	51.43
Feb 2009	81.38	70.33	61.25
Mar 2009	74.19	63.14	54.06
Apr 2009	78.32	67.27	58.19
May 2009	82.49	71.44	62.38
June 2009	80.27	69.22	60.14
July 2009	84.33	73.28	64.20

In order to improve the forecast efficiency for share trading, total part summation is used where irrelevant attributes are removed in a significant manner that helps in increasing the forecast efficiency. In the experimental setup, the forecasting efficiency is measured for the period from Jan 2009 to July 2009. The results of 7 different forecast efficiencies are listed in table 2. As listed in table 2, the PRF-TPS framework measures the forecast efficiency in terms of percentage (%).

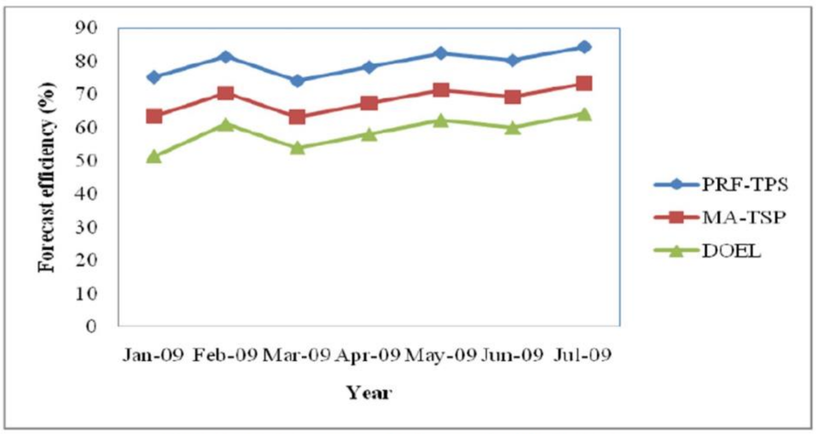

Figure 6 Measure of forecast efficiency

Figure 6 shows the forecast efficiency with respect to the stock prices in the year Jan 2009 to July 2009. Figure measures the forecast efficiency trend at different time scale. The results are consistent which shows that forecast efficiency with respect to stock prices is comparatively better using the proposed PRF-TPS framework than compared to the existing MA-TSP [1] and DOEL [2] respectively. This is because of the application of total part summation using three elements Dividend Stock Price Ratio (DSPR), Income Growth Ratio (IGR) and Stock Price Income Growth (SPIG). This there element helps in removing the irrelevant attributes and therefore results in the forecast efficiency. In addition, the PRF-TPS framework though uses the total part summation, to emphasize forecast efficiency, only the values obtained from moving average estimation and exponent moving average estimation is used which increases the forecast efficiency by 14.09% compared MA-TSP [1] and 26.09% compared to DOEL [2] respectively.

Impact of test error rate: The test error rate is the ratio of difference between the actual share value and the predicted share value to the actual share value. The test error rate using relative absolute error is formulated as given below.

$$RAE = \left(\frac{A-P}{A}\right)*100 \qquad (16)$$

From (16) the test error rate (relative absolute error) '*RAE*' is measured using actual share value '*A*' and the predicted share value '*P*' respectively. Lower the test error rate more efficient the method is said to be.

Table 3 Tabulation for test error rate

Year	Test error rate (%)		
	PRF-TPS	MA-TSP	DOEL
Jan 2009	15.55	23.48	28.20
Feb 2009	18.36	26.33	31.36
Mar 2009	21.35	29.32	34.35
Apr 2009	17.22	25.19	30.22
May 2009	24.33	32.30	37.33
June 2009	27.89	35.86	40.89
July 2009	25.21	33.18	38.21

In table 3 we further compare the test error rate for prediction of share trading using three elements. The experiments were conducted at different time period which is measured in terms of percentage (%).

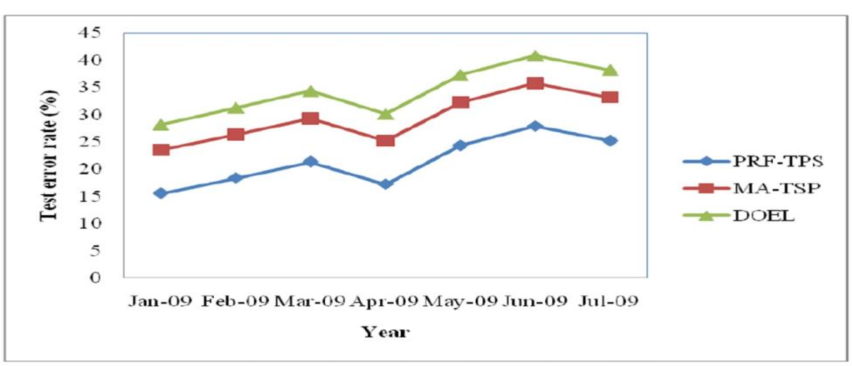

Figure 7 Measure of test error rate

To better understand the test error rate of the proposed framework, PRF-TPS, GM-SRFIS, share trading data were extracted at different time periods between Jan 2009 and July 2009. Compared to the existing Multilevel Architecture for Time Series Prediction (MA-TSP) and Diversity on Online Ensemble Learning (DOEL) [2], the test error rate in the proposed PRF-TPS framework is lower. This is because of the application of summation crossover algorithm that significantly decides the buying or selling of share through Exponent Moving Average Estimation at different time scales (i.e. Year). This in turn decreases the test error rate using PRF-TPA framework. Furthermore, to have a significant impact of test error rate, the proposed PRF-TPA framework, uses the crossover point for effective forecasting of share trading. This extensively reduces the test error rate using PRF-TPA framework by 38.70% compared to MA-TSP and 62.87 % compared to DOEL respectively.

Impact of predictive accuracy on shares: Finally, table 4 given below provides the predictive accuracy on shares of PRF-TPS framework for seventy different returns rate that is measured in terms of percentage (%).

Table 4 Tabulation for predictive accuracy on shares

Methods	Predictive accuracy on shares (%)
PRF-TPS	78.32
MA-TSP	65.16
DOEL	59.35

To measure the effectiveness of the proposed PRF-TPS framework, the impact of predictive accuracy on shares efficiency with respect to three different methods, PRF-TPS, MA-TSP and DOEK is shown in table 4.

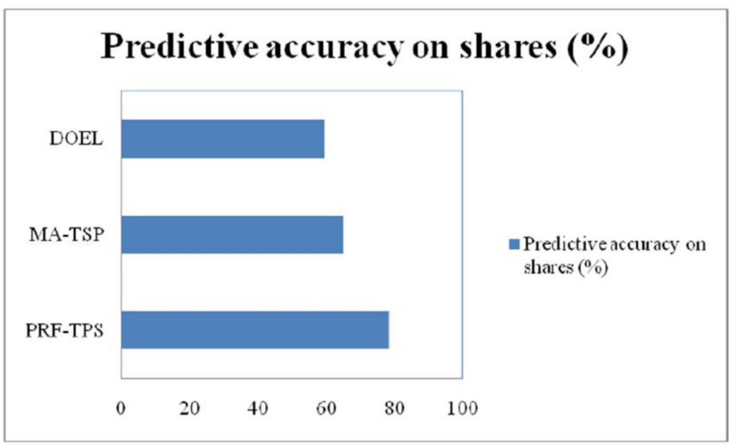

Figure 8 Measure of predictive accuracy on shares

Lastly the predictive accuracy on shares efficiency is measured via different returns obtained through different shares at different time scale for implementation purpose. From the figure 8 it is illustrative that the proposed PRF-TPS framework potentially yields better results than the existing MA-TSP [1] and DOEL [2]. The significant results achieved using the PRF-TPS framework is because of the application of Exponent-based Predictive Regression Forecast algorithm to obtain higher predictive accuracy on shares using PRF-TPS framework by 16.80% when compared to MA-TSP. As a result, the predictive accuracy on shares is improved to a coarser construction, because the Exponent-based Predictive Regression Forecast algorithm estimates the exponent moving average estimation and thereby increasing the predictive accuracy on shares by 8.91% when compared to DOEL [2].

VII. CONCLUSION

Forecasting of share trading is becoming more challenging, due to the unpredictable nature of the fluctuations in the inflation rate. In this work, Predictive Regression Forecast based on Total Part Summation (PRF-TPS) framework is designed with the objective of attaining an effective predictive accuracy on shares. The resulting share trading forecasting problem has been formulated as a predictive regression forecast based average estimation and solved through a novel Exponent-based Predictive Regression Forecast algorithm. The initially selected share values from Istanbul Stock Exchange dataset obtain trend at different time scales using moving average and exponent moving average estimation. Next, the proposed summation crossover algorithm for effective analysis of trend analysis by estimating three elements, Dividend Stock Price Ratio (DSPR), Income Growth Ratio (IGR) and Stock Price Income Growth (SPIG). This in turn enhances the forecast efficiency aspect while making effective analysis while buying or selling shares. Finally, the forecast error rate at different time scales is reduced by applying Summation Crossover algorithm based on the crossover point. The performance of PRF-TPS framework was compared to two other time series prediction model, MA-TSP and DOEL respectively. We compared the performance with many different system parameters, and evaluated the performance in terms of different metrics, such as run-time, forecast efficiency, test error rate and predictive accuracy on shares. The results show that PRF-TPS framework offers better performance with an improvement of forecast efficiency by 20.09% and predictive accuracy on shares by 25.71% compared to MA-TSP and DOEL respectively.

REFERENCES

[1] Dymitr Ruta, Bogdan Gabrys, and Christiane Lemke, "A Generic Multilevel Architecture for Time Series Prediction", IEEE Transactions on Knowledge and Data Engineering, Volume 23, Issue 3, Pages 350-359, March 2011.

[2] Leandro L. Minku, Allan P. White, and Xin Yao, "The Impact of Diversity on Online Ensemble Learning in the Presence of Concept Drift", IEEE Transactions on Knowledge and Data Engineering, Volume 22, Issue 5, Pages 730-742, May 2010.

[3] Hui Wang, Ruilin Liu, "Privacy-preserving publishing micro data with full functional dependencies", Elsevier, Volume 70, Issue 3, Pages 249-268, March 2011.

[4] Oguz Akbilgic , Hamparsum Bozdogan, M. Erdal Balaban, "A novel Hybrid RBF Neural Networks model as a forecaster", Springer, Volume 24, Issue 3, Pages 365-375, May 2014.

[5] Yinghao Chu, Mengying Li, Hugo T.C. Pedro, Carlos F.M. Coimbra, "Real-time prediction intervals forintra-hour DNI forecasts", Elsevier, Volume 83, Pages 234–244, November 2015.

[6] Sara Hajian and Josep Domingo-Ferrer, "A Methodology for Direct and Indirect Discrimination Prevention in Data Mining", IEEE Transactions on Knowledge and Data Engineering, Volume 25, Issue 7, Pages 1445-1459, July 2013.

[7] Wijnand Nuij, Viorel Milea, Frederik Hogenboom, Flavius Frasincar, and Uzay Kaymak, "An Automated Framework for Incorporating News into Stock Trading Strategies", IEEE Transactions on Knowledge and Data Engineering, Volume 26, Issue 4, Pages 823-835, April 2014,.

[8] Massimiliano Albanese, Cristian Molinaro, Fabio Persia, Antonio Picariello, and V.S.Subrahmanian,"Discovering the Top-k Unexplained Sequences in Time-Stamped Observation Data", IEEE Transactions on Knowledge and Data Engineering, Volume 26, Issue 3, Pages 577-594, March 2014.

[9] Yu Gu, Chunpeng Gao, Gao Cong, and Ge Yu, "Effective and Efficient Clustering Methods for Correlated Probabilistic Graphs", IEEE Transactions on Knowledge and Data Engineering, Volume 26, Issue 5, Pages 1117-1130, May 2014.

[10] Wenjing Zhang and Xin Feng, "Event Characterization and Prediction Based on Temporal Patterns in Dynamic Data System", IEEE Transactions on Knowledge and Data Engineering, Volume 26, Issue 1, Pages 144-156, January 2014.

[11] Zhen Hai, Kuiyu Chang, Jung-Jae Kim and Christopher C. Yang, "Identifying Features in Opinion Mining via Intrinsic and Extrinsic Domain Relevance", IEEE Transactions on Knowledge and Data Engineering, Volume 26, Issue 3, Pages 623-634, March 2014.

[12] Minyoung Kim, "Probabilistic Sequence Translation-Alignment Model for Time-Series Classification", IEEE Transactions on Knowledge and Data Engineering, Volume 26, Issue 2, Pages 426-437, February 2014.

[13] Wenjing Zhang and Xin Feng, "Event Characterization and Prediction Based on Temporal Patterns in Dynamic Data System", IEEE Transactions on Knowledge and Data Engineering, Volume 26, Issue 1, Pages 144-156, January 2014.

[14] Nima Asadi, Jimmy Lin1, Arjen P. de Vries, "Runtime Optimizations for Prediction with Tree-Based Models", IEEE Transactions on Knowledge and Data Engineering, Volume 26, Issue 9, Pages 2281-2292, September 2014.

[15] Shipra Banik, A. F. M. Khodadad Khan and MohammadAnwer, "Hybrid Machine Learning Technique for Forecasting Dhaka Stock Market Timing Decisions", Computational Intelligence and Neuroscience, , Pages 1-7, February 2014.

[16] Yun Xue, Zhiwen Liu, Jie Luo, ZhihaoMa, Meizhen Zhang, Xiaohui Hu, and Qiuhua Kuang, "Stock Market Trading Rules Discovery Based on Bi clustering Method", Mathematical Problems in Engineering, Pages 1-14, January 2015.

[17] Hossein Javedani Sadaei and Muhammad Hisyam Lee, "Multilayer Stock Forecasting Model Using Fuzzy Time Series", The ScientificWorld Journal, Pages 1-11, January 2014.

[18] Jingpei Dan,Wenbo Guo, Weiren Shi, Bin Fang and Tingping Zhang, "Deterministic Echo State Networks Based Stock Price Forecasting", Abstract and Applied Analysis, Pages 1-7, June 2014.

[19] Athanassios Zagouras, Hugo T.C. Pedro, Carlos F.M. Coimbra, "On the role of lagged exogenous variables and spatioetemporal correlations in improving the accuracy of solar forecasting methods", Elsevier, Volume 78, Pages 203–218, June 2015.

[20] Dominik Fisch, Thiemo Gruber, and Bernhard Sick, "SwiftRule: Mining Comprehensible Classification Rules for Time Series Analysis", IEEE Transactions on Knowledge and Data Engineering, Volume 23, Issue 5, Pages 774-787, May 2011.