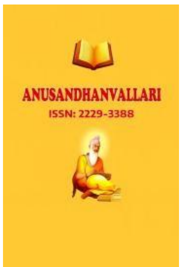

Anusandhanvallari
Vol 2025, No.1
January 2025
ISSN 2229-3388

BSE 30 Financial Forecasting Using Regression Techniques

Jitendra Singh Narwariya[1], Dr. Ragini Shukla[2], Dr. Vineet Kumar Awasthi[3]
Research Scholar, Dr. C V Raman University, Bilaspur (CG), India[1]
Professor, Dr C V Raman University, Bilaspur (CG), India[2]
Assistant Professor, Guru Ghasidas Vishwavidyalaya, Bilaspur (CG), India[3]
Email: jitendras21001@gmail.com[1], raginishukla008@gmail.com[2], vineet99kumar@gmail.com[3]

ABSTRACT

Forecasting financial markets is particularly difficult due to the system's inherent non-linearity and unpredictability. Traditional methods for creating stock market predictive models have several limitations, leading to the widespread adoption of intelligent techniques. This study explores the application of neural network–based regression models utilizing various technical indicators recommended by researchers. Stock index data, including Open, High, Low, and Close values, are used to derive features, which are then selected using numerous existing rank-based feature selection methods. A new index data set with a reduced feature subset is then provided to the neural network–based regression models for predicting future values. Empirical results indicate that feature extraction followed by feature selection can significantly enhance the efficiency of the ANN model. For the experimental study, one years of historical daily Indian stock data were analyzed, From the original data, ten technical-indicator-based features were initially extracted. A rank-based feature selection technique was then applied, resulting in a reduced subset of five optimal features. Using this reduced feature set, the proposed neural network–based regression models achieved a Mean Absolute Percentage Error (MAPE) of 0.594300 on the test dataset.

KEYWORDS
Feature Extraction, Feature Selection, Technical Indicator, Neural Network–based regression models.

1. INTRODUCTION

Stock market is a complex and dynamic system with noisy, non-linear and chaotic data series. Prediction of a financial stock market is more challenging due to non-linearity and uncertainty of the system. Conventionally, technical analysis methods, that predicts stock prices are based on historical prices and volume, Advanced intelligent techniques that are ranging from complete mathematical models and expert systems to fuzzy logic networks [1], have also been used by many financial predictive systems for predicting stock prices based on historical data [6]. In recent years, artificial neural networks have been widely adopted for stock market prediction due to their ability to model complex and non-linear relationships; however, prediction accuracy remains highly sensitive to feature redundancy and irrelevant inputs.

Neural networks have also been found very useful in stock market prediction. Artificial Intelligent techniques are used by various authors [6]. Due to many drawbacks of traditional techniques for developing predictive model, the approaches of computational intelligent methods have gain importance to predict the stock prices. Many researchers utilized ANN [7] and Fuzzy logic for this purpose. This paper proposes an neural network–based regression models stock market prediction framework that integrates technical-indicator-based feature extraction with rank-based feature selection to reduce redundancy and improve prediction accuracy. neural network–based regression models is utilized with special reference to feature extraction and selection. The original feature available in the Indian stock index is used to extract features based on five technical indicators suggested by many authors. The feature selection is a technique that choose the best features among many and removes the unimportant features to be finally utilized by neural network–based regression models. five out of ten features are finally selected for stock market prediction.

2. STOCK DATA AND TECHNICAL INDICATORS

This study utilizes daily BSE 30 stock market data, which was obtained from (www.yahoofinance.com). The dataset includes Date, Open, High, Low and Close, but only the first four features are employed to derive new features using various technical indicators, as illustrated in Table 1. These technical indicators are commonly applied in the development of numerous models, as recommended by financial expert.

Table I: List of Technical Indicators

S. No.	Name of Technical Indicator	Description
1	Exponential Moving Average (EMA)	EMA reduces the lag by applying more weight to recent price.
2	Moving average Convergence-Divergence (MACD)	MACD measures the difference between security's of 26-days and 12 days EMA's).
3	Relative Strength Index (RSI)	It shows price strength by comparing upward and downward close - to - close movements.
4	Stochastic Oscillator	It shows the location of the close relative to the high-low range over a set number of periods.
5	Rate Of Change (ROC)	It measures the percentage change in price from one period to the next.
6	Money Flow Index (MFI)	MFI is an oscillator that uses both price and volume to measure buying and selling pressure.
7	Willium %R	It is a momentum indicator that is the inverse of the Fast Stochastic Oscillator.
8	Accumulation Distribution Line(A/D)	A/D is a Volume based indicator designed to measure the cumulative flow of money into and out of security.
9	On balance Volume (OBV)	It measures Buying and selling pressure as a cumulative indicator that ads volume on up days and subtract volume on down days.
10	Chaikin Oscillator (CHO)	It is the difference between the 3-day EMA and the 10-day EMA of the A/D.

The technical indicators were computed using standard parameter settings commonly adopted in financial literature. Specifically, EMA was calculated using a 12-day period, MACD was computed as the difference between 12-day and 26-day EMAs, RSI was calculated over a 14-day window, the stochastic oscillator used a 14day look-back period, and ROC was calculated as the percentage price change over successive periods.

3. STOCK MARKET PREDICTION TECHNIQUE

A. Radial Basis Function (RBF) Neural Network

Radial basis functions are effective tools for interpolation in spaces with multiple dimensions [9]. An RBF is a function that is constructed based on a distance criterion relative to a center. One advantage of RBF neural networks is their faster convergence compared to traditional multilayer perceptrons, although their performance remains dependent on center selection and spread parameters. RBF Networks are essentially two-layer feedforward networks [11]. The hidden nodes execute a series of radial basis functions, such as Gaussian functions, while the output nodes perform linear summation functions similar to those in an MLP. The training of the network is split into two phases: initially, weights are determined from the input layer to the hidden layer, followed by weights from the hidden layer to the output layer.

B. Error Back Propagation Network (EBPN)

EBPN is a widely recognized prediction method recommended by [5], which can be trained using the well-known Error Back Propagation Algorithm and serves as a generalization of the delta rule. For a specific input vector, the delta value compares the output vector to the correct response. This technique is particularly beneficial for feedforward networks. Back propagation [4] necessitates that the activation function employed by the artificial neurons be differentiable. The backward propagation [8] involves transmitting the output activations through the neural network using the training pattern target to produce the deltas for all output and hidden neurons. In this study, the EBPN model was implemented as a multilayer feed-forward network consisting of one input layer, one hidden layer, and one output layer. The hidden layer employed a sigmoid activation function, while the output layer used a linear activation function. The network was trained using gradient descent back-propagation for a fixed number of epochs until convergence.

C. Feature Extraction and Feature Selection

Feature extraction involves creating a new set of features from the existing ones, which can be crucial for analysis and classification [2][10]. This method essentially transforms the attributes, resulting in new features that are combinations of the original ones. The aim is to enhance the effectiveness and efficiency of analysis and classification [12]. Conversely, feature selection focuses on identifying the optimal subset of features. It is important to eliminate unnecessary features, as they can compromise the quality of the patterns discovered and hinder accurate and efficient predictions [3]. This is often due to some columns being noisy or redundant. Such noisy data complicates the process of uncovering meaningful patterns, and to find quality patterns, most data mining algorithms require significantly larger training datasets when dealing with high-dimensional data. Feature selection techniques automatically identify the most valuable features, addressing the issue of having excessive data with minimal value.

4. FRAMEWORK FOR STOCK MARKET PREDICTION

The proposed framework for predicting stock market trends is illustrated in Figure 1 and can be divided into two main components: feature extraction and selection, followed by stock value prediction. The initial component focuses on data preprocessing, which involves the following steps:

Step 1.1: BSE 30 index data is gathered from a financial website.

Step 1.2: Ten features were derived from five technical indicators based on different parameter combinations, as detailed in Table 1.

Step 1.3: The index data, along with the extracted features, is normalized using a straightforward normalization formula to scale the values between [0 1].

Normalization was performed using min–max scaling:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Step 1.4: A ranking-based feature selection method was employed, where features were ranked based on their individual contribution to prediction accuracy. Features with lower ranks were iteratively eliminated, resulting in a reduced subset of the most informative technical indicators.

Step 1.5: The data is split into two segments: training and testing. The training data is used for model development, while the testing data is utilized for validating the predictive model.

The second component of this framework involves model construction and validation using ANN techniques, which include the following steps:

Step 2.1: 80% of the data is used to build the model, while 20% is allocated for model validation using both the RBFN and EBPN ANN techniques.

Step 2.2: Obtained forecasting next-day-close price from RBFN and EBPN are analyzed with the help of following error measures:

$$\text{Mean Absolute Error (MAE)} = \frac{1}{n}\sum_{i=1}^{n}\left|Y_{a_i i} - Y_{p_i i}\right| \quad \text{...............(1)}$$

$$\text{Mean Absolute Percentage Error (MAPE)} = \left(\frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_{a_i i} - Y_{p_i i}}{Y_{a_i i}}\right|\right) \times 100\% \quad \text{...............(2)}$$

$$\text{Root Mean Square Error (RMSE)} = \sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n}\left(Y_{a_i i} - Y_{p_i i}\right)^2} \quad \text{...............(3)}$$

$$\text{R – Square } (R^2) = 1 - \frac{\sum_{i=1}^{n}\left(Y_{a_i i} - Y_{p_i i}\right)^2}{\sum_{i=1}^{n}\left(Y_{a_i i} - \bar{Y}\right)^2} \quad \text{...............(4)}$$

Where, $Y_{a_i i}$ and $Y_{p_i i}$ are respectively actual and predicted values of stock price on a particular day i,

$\bar{Y}$ represents the mean (average) of the actual values and

n is the total number of days.

Anusandhanvallari
Vol 2025, No.1
January 2025
ISSN 2229-3388

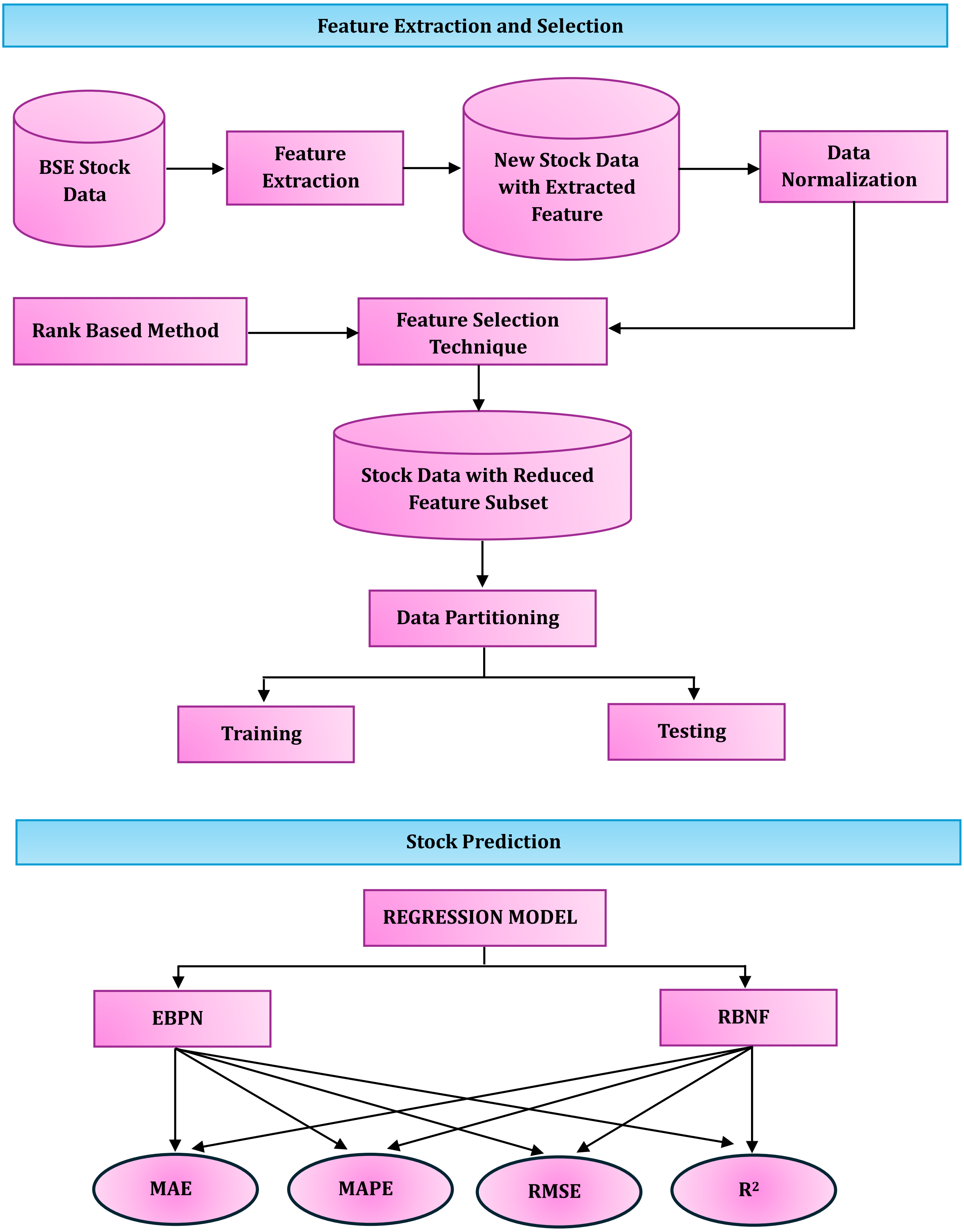

Fig 1: Process of implementing stock market prediction with technical indicators

5. RESULT AND DISCUSSION

Initially the normalized BSE Sensex dataset was evaluated using two neural network–based regression models, namely EBPN and RBNF. The result is evaluated in the form of MAE, MAPE, RMSE and R Square that is presented in Table II. As per the table it is clearly seen that the accuracy of prediction is not as expected. The value of RMSE and R square is highly deviated from robust prediction. The non-linear behaviour of Actual Vs Predicted is also identified by figure 2, that motivates us to improve experimental stock data.

In second phase of experimental work is initiated by generating new dimensions in stock data using technical indicators. After applying feature selection techniques, the normalized BSE Sensex dataset was again evaluated using two neural network-based regression model, namely EBPN and RBNF, and the results are presented in Table III. The performance is measured using MAE, MAPE, RMSE, and R Square for two different feature sets containing 10 and 5 selected features. From the table, it can be clearly observed that feature selection improves the overall prediction performance of both models. In the EBPN model, reducing the number of features from 10 to 5 results in lower error values and a significant improvement in the R Square value, indicating better prediction capability. Similarly, the RBNF model shows better performance compared to EBPN, achieving lower MAE and RMSE values along with higher R Square values, especially when 5 features are used. This demonstrates that selecting an optimal number of relevant features helps in reducing unnecessary information in the dataset and allows the models to learn the underlying stock market patterns more effectively, leading to improved prediction accuracy.

Table II: Comparative Evaluation of neural network-based Regression Model from Normalized BSE Sensex Dataset

Regression Models	MAE	MAPE	RMSE	R Square
EBPN	0.128392	0.144774	14.689191	-71.180132
RBNF	0.758681	0.006473	0.005724	0.807451

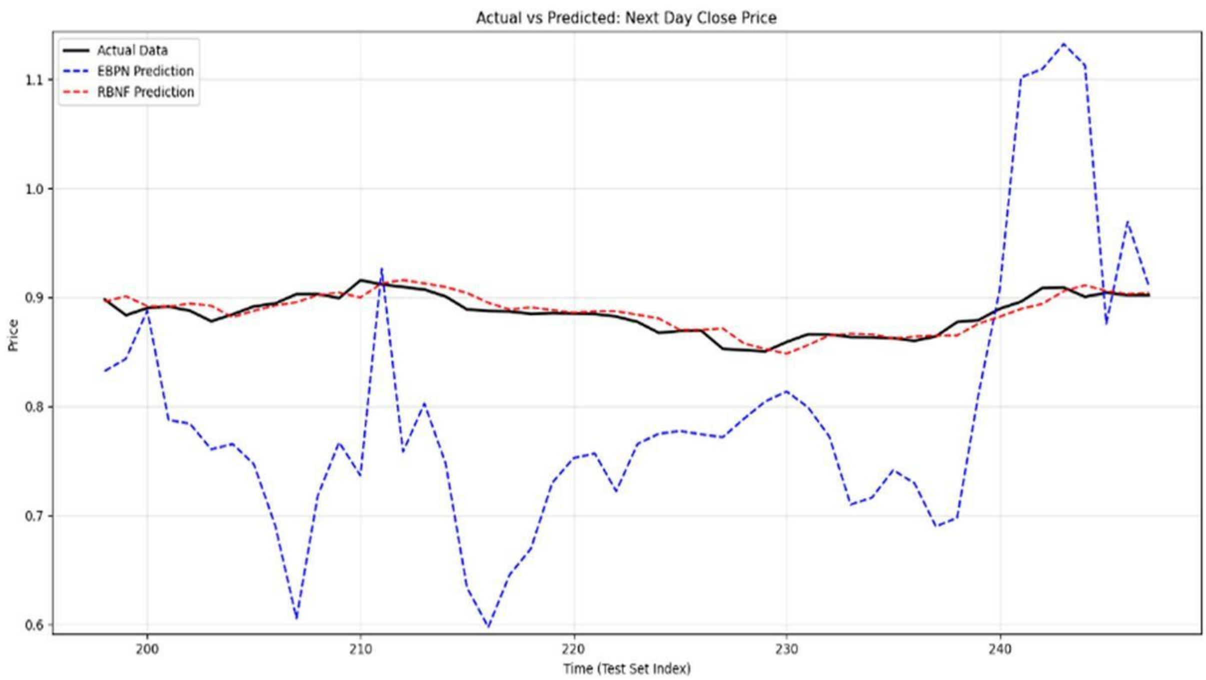

Fig 2: Performance Comparison of EBPN and RBFN for Next-Day Stock Price Prediction

Table III: Results after applying feature selection techniques with EBPN and RBNF

Regression Models	No of Features selected	MAE	MAPE	RMSE	R Square
EBPN	10	0.020641	0.023344	2.503159	-1.096037
	5	0.008184	0.924800	0.010262	0.647735
RBNF	10	0.005724	0.758681	0.006473	0.849697
	5	0.005255	0.594300	0.006391	0.863373

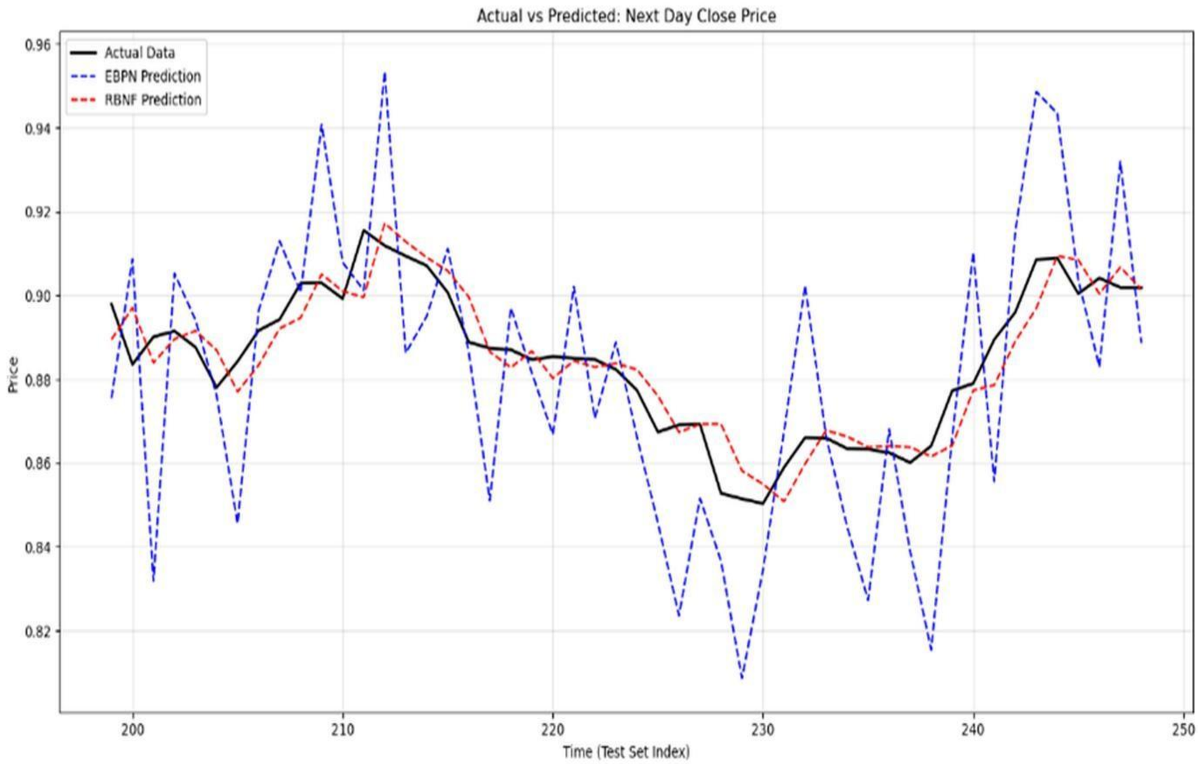

Fig 3: Performance Comparison of EBPN and RBFN for Next-Day Stock Price Prediction (using technical indicators)

6. CONCLUSION

This study proposed a stock market prediction framework using neural network–based regression models combined with technical-indicator-based feature extraction and rank-based feature selection. The experimental results show that prediction accuracy using the original dataset was limited due to redundant and less informative features. After applying feature extraction and selecting the most relevant features, the performance of both models improved significantly.

Among the evaluated models, the RBNF regression model performed better than EBPN by achieving lower prediction errors and higher $R2$ values. The best results were obtained using a reduced subset of five selected features, achieving a MAPE value of 0.594300. The findings demonstrate that selecting optimal features helps the model learn stock market patterns more effectively and improves forecasting accuracy. Overall, the proposed approach provides an efficient and reliable method for financial market prediction.

REFERENCES

[1] Bilbao-Terol, M. Arenas-Parra, and V. Cañal-Fernández, 'Selection of Socially Responsible Portfolios using Goal Programming and fuzzy technology', *Information Sciences*, vol. 189, pp. 110–125, 2012.

[2] N. VenkateswaraRao, A. Srikrishna, B. RaveendraBabu, and R. M. Babu, 'An efficient feature extraction and classification of handwritten digits using neural networks', *International Journal of Computer Science, Engineering and Applications*, vol. 1, no. 5, p. 47, 2011.

[3] C.-F. Tsai and Y.-C. Hsiao, 'Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches', *Decision support systems*, vol. 50, no. 1, pp. 258–269, 2010.

[4] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, 'Learning representations by back-propagating errors', *nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[5] G. Shukla and R. Rudrakshi, 'Survey on prediction using back propagation neural network', *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 4, no. 3, pp. 917–920, Mar. 2014.

[6] E. Guresen, G. Kayakutlu, and T. U. Daim, 'Using artificial neural network models in stock market index prediction', *Expert systems with Applications*, vol. 38, no. 8, pp. 10389–10397, 2011.

[7] K. S. Ravichandran, P. Thirunavukarasu, R. Nallaswamy, and R. Babu, 'Estimation of return on investment in share market through ANN', *Journal of Theoretical and Applied Information Technology*, vol. 3, pp. 44–54, 2005.

[8] C. H. Lee and K. C. Park, 'Prediction of monthly transition of the composition stock price index using recurrent back-propagation', in *Artificial neural networks*, Elsevier, 1992, pp. 1629–1632.

[9] J. A. Leonard and M. A. Kramer, 'Radial basis function networks for classifying process faults', *IEEE Control Systems Magazine*, vol. 11, no. 3, pp. 31–38, 2002.

[10] S. V. Rajashekararadhya and P. V. Ranjan, 'Efficient zone-based feature extraction algorithm for handwritten numeral recognition of four popular south Indian scripts', *Journal of Theoretical & Applied Information Technology*, vol. 4, no. 12, 2008.

[11] Y. Arora, A. Singhal, and A. Bansal, 'A study of applications of RBF network', *International Journal of Computer Applications*, vol. 94, no. 2, pp. 17–20, 2014.

[12] Ø. D. Trier, A. K. Jain, and T. Taxt, 'Feature extraction methods for character recognition-a survey', *Pattern recognition*, vol. 29, no. 4, pp. 641–662, 1996.