OPEN ACCESS
International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-2, July 2019

Prediction of Stock Index of Tata Steel using Hybrid Machine Learning Based Optimization Techniques

Mohammed Siddique, Debdulal Panda

***Abstract:** The trend of stock price prediction has always been in the focal point of analytical activity in financial domain for both the researchers and investors. Prediction with accuracy is very essential for improved investment decisions that imbibe minimum risk factors. Due to this, majority of investors depend upon that intelligent trading system which generates better forecasting results. As forecasting stock market price with high accuracy is quite a challenging task for the analysts, machine learning has been adopted as one of the popular techniques to predict future trends. Even if there are many recognized analytical time series analysis that are categorized either under soft computing or under conventional statistical techniques like fuzzy logic, artificial neural networks and genetic algorithms, researchers have been looking for more appropriate techniques which can exhibit improved results. In this paper, we developed different hybrid machine learning based prediction models and compared their efficiency. Dimension reduction techniques such as orthogonal forward selection (OFS) and kernel principal component analysis (KPCA) are used separately with support vector regression (SVR) and teaching learning based optimization (TLBO) to predict the stock price of Tata Steel. The performance of both the proposed approach is evaluated with 4143days daily transactional data of Tata steels stocks price, which was collected from Bombay Stock Exchange (BSE). We compared the results of both OFS-SVR-TLBO and KPCA-SVR-TLBO hybrid models and concludes that by incorporating KPCA is more practicable and performs better results than OFS.*

***Index Terms:** Forecasting of stock market; Orthogonal forward selection; Kernel principal component analysis, Support vector regression; Teaching–learning-based optimization.*

I. INTRODUCTION

Forecasting of financial market has an essential part for the growth of any country. Most of investors are presently depending upon intelligent trading systems for forecasting of share market price. Precision of these forecast systems is necessary for better investment decisions with minimum risk factors. Prediction of stock price has been beneficial for both the individual and institutional investors. Technological investigation is an accepted progress towards the stock market analysis. Many authors were used different statistical and machine learning models to forecast future trends or price. ANN, SVM, GA and LR have been used by many authors for this kind of forecasting tasks. Among all these SVM is considered to one of the best performing technique provided appropriate initialization of its regularization parameters is made. In the current scenario, support vector machine (SVM) (Vapnik, 1995; Gilbert, 2007; Haykin, 2009), commonly known as SVR, is one of the demanding and commanding techniques in the field of forecasting to carry high precision results. The use of dimension reduction techniques SVR improved the overview forecasting ability. Using novel machine learning based hybrid models using SVR, many researchers have been benefitted in the field of prediction of stock market. Since all features in a dataset do not influence the output equally, features which are less relevant may be discarded. In such cases, it is advantageous to choose a subset of features which extremely influence the results. Use of orthogonal forward selection (Mao, 2004) has executed better performance with data having high correlations among attributes. The advantage of OFS is that it reduces the uncorrelated features due to orthogonal transformation.

The role of OFS in the hybrid model is to identify the constructive features that improve the performance of the model where as the role of KPCA in this hybrid model is to extract the productive features which enhance the accuracy of the proposed model. KPCA reduce the dimension of the input data. As feature extraction is a technique to extract a feature subset from all the input features to prepare the constructed model better, it is desired that OFS and KPCA to be applied at the initial stages to the whole dataset to simplify and rearrange the original data structure.

We used the Support Vector Regression for core prediction mechanism to predict stock price of Tata Steel. SVR requires its hyper parameters (C and γ) to be optimized to perform better prediction results. Hence TLBO is used to optimize its free parameter C and γ. Historical trading records of Tata Steel are used as the technological indicators. Lagged data in the time series domain have always been influencing the forecasting accuracy. The rest of this paper is organized as follows. Literature review is highlighted in Section-2 and a brief description of OFS, KPCA, SVR and TLBO are given in Section 3. Then in Section-4, the methodology and the process involved in the hybrid model under study is explained.

Revised Manuscript Received on 30 July 2019.
* Correspondence Author
Mohammed Siddique*, Department of Mathematics, Centurion University of Technology and Management, Odisha, India
Debdulal Panda, Department of Mathematics, KIIT Deemed to be University, Bhubaneswar, India

In Section-5, Implementation of models are explained and in chapter 6. experimental analysis is presented and finally, the paper is concluded in Section-7.

II. LITERATURE REVIEW

The main intention of studying stock market is to reduce the risk, which is essential to a financial institutions dealing with either a company or individual one. The uncertainty in trade industry and their forecasting have motivated lots of researchers to work in the most volatile area like stock market. The stock market prediction is a very complex dynamic task. Cao, L [1] investigated the kernel functions by combining SVMs with self organizing feature mapping through two stage architecture. Pai and Hong [2] proposed a model with SVM and SA algorithms for software reliability forecasting. The study used SVM to address the non-linear regression and time series problems and SA to choose the three positive parameters of SVM model. Lin, S.W. et al. [3] used the RBF kernel function in SVM-PSO model in public data set. Huang and Dun [4] proposed a hybrid SOFM-SVR model which condensed the training time and improves the accuracy of prediction in the financial stock index. Lu [5] proposed hybrid model named as ICA-SVR model using ICA and SVR in financial time-series forecasting. In the model, ICA was used for detecting and removing the noise from the financial time-series data whereas SVR was applied for forecasting task. The proposed model was appraised using Nikkei 225 starting index and TAIEX ending index and the experimental results showed that, the proposed model (ICASVR) outperformed the traditional SVR and random walk benchmark model according to predictive error and predictive accuracy.

Zhang, X. et al. [6] optimized the SVR parameter using ant colony optimization algorithm. Huang [7] developed a hybrid GA–SVR model where GA was optimized the SVR parameters. Kara, Y. et al. [8] analyzed the direction of movement of stock price of the Istanbul Stock Exchange using ANN and SVM. Rao, R.V. et al. [9] analyzed that the optimization method TLBO perform better than other nature inspired optimization models like PSO, ACO. Rao, R. V. et al. [10] introduced adaptive teaching factor by modifying the teaching factor of TLBO by assigning more than one teacher for the learners. He, Y. et al. [11] analyzed the performance of feature selection methodology of PCA and Sequential Forward Selection (SFS) with SVR and concluded that PCA perform better accuracy than SFS. Hsu, C.M. [12] proposed back propagation neural and genetic programming technique to predict the futures stock price. Many authors are used KPCA with the machine learning models. Chandrashekar, G., and Sahin, F. [13] reduced features by using the feature selection techniques. Chang, P. C. and Wu, J. L. [14] improved the prediction performance of stock trading using kernel principal component analysis technique in SVR model. Ballings, M. et al. [15] compared the performance of RF, AKF with NN, SVM, K-nearest neighbors and logistic regression for prediction of stock price. Patel, J. [16] planned a two step fusion approach with SVR model perform better results than a single stage ANN with random forest model. Das et al. [17] proposed an ensemble model named DR-SVM-TLBO for predicting the energy commodity futures index in financial time series. This model used PCA, KPCA, and ICA for performing the dimensionality reduction and SVM-TLBO hybrid component for prediction on the reduced features. In SVM-TLBO, the SVM generalizes the results obtained from learning's where as TLBO evaluates the optimal output of the free parameters of SVM. Jaffel, I. et al. [18] analyzed that moving window reduced kernel principal component analysis MW-RKPCA are perform better results than and MW-KPCA applied to monitoring the nonlinear dynamic system. Das, S. P. [19] designed a hybrid model of DR-SVM-TLBO and applied different dimension reduction method i.e PCA, KPCA and ICA to reduce the features of supplied data. Patel, V. K. and Savsani, V. J. [20] analyzed the results of TLBO algorithm and matched it with existing methods presented in the literature. Mohanty, S. and Padhy, S. [21] proposed hybrid OFS–TLBO–SVR model to predicting gross value added at factor cost. Zhong, X. and Enke, D. [22] analyzed the performance of three dimension reduction techniques PCA, FRPCA and KPCA on financial data. Das, S. P and Padhy, S. [23] analyzed the feasibility of Teaching Learning Based Optimization to optimize the free parameters of SVR for time-series financial forecasting. Henrique, B.M. et al. [24] developed a random walk-based SVR model to predict financial status or condition of a company.

III. METHODOLOGY USED

A. *Orthogonal Forward Selection*

Orthogonal forward selection (OFS) is an advanced approach for feature subset selection using Gram–Schmidt orthogonal transformations in forward selection [25]. Feature selection methods have been improved the performance of machine learning models by neglecting least relevant and uncorrelated features of the datasets. Orthogonal forward selection moves towards a new approach in feature subset selection which was implemented by using Gram-Schmidt orthogonalization. The OFS procedure for classification problems is described below.

Let X represents the feature matrix which consists of 'n' features and 'm' samples. So matrix X has n feature vectors and each has 'm' dimensional i.e. $X = [x_1, x_2, \ldots ..., x_n]$. With Gram-Schmidt and givens orthogonal transformation, the feature matrix specified by X = QR, where Q is the orthogonal transformation of X represented as $Q = [q_1, q_2, \ldots, q_n]$ and R represent the upper triangular matrix. The orthogonalization matrix Q has been by applying Gram-Schmidt orthogonalization process stated in equations (1) and (2)

$$q_i = x_i - \sum_{j=1}^{i-1} \alpha_{ji} q_j \quad (1)$$

$$\text{Where } \alpha_{ji} = \begin{cases} \frac{q_j^T x_i}{q_j^T q_i}, for\ j = 1, 2, 3, \ldots \ldots i-1 \\ 1, \quad for\ j = i \end{cases} \quad (2)$$

Class separation process used to estimate the data features in classification problems. The greater class separation indicates the higher chance to get selected. The OFS procedure is explained in given algorithm.

1) Let all features $x_1, x_2 \dots x_n$ are entering to the feature matrix X

Put $q_1^{(i)} = x_i$

and calculate $q_1^{(1)}, q_1^{(2)} \dots q_1^{(n)}$

Assume x_j gives the highest class reparability.

Put $q_1 = x_j$

2) For remaining (n – 1) features, compute

$q_2^{(i)} = x_i - \alpha_{12}^{(i)} q_1$

where $1 \le i \le n$, $i \ne j$ and $\alpha_{12}^{(i)} = \frac{q_1^T x_i}{q_1^T q_1}$

Submit your manuscript electronically for review.

B. *Support Vector Machine for Regression*

SVM is well known for binary classifiers. SVM produce a decision boundary where the majority of the data points of same kind falls on one side of the boundary while other kind fall on the another side. Consider an p-dimensional feature vector

$x = (x_1, x_2, \dots, x_p)$. We can describe a hyperplane

$$\alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_p x_p = \alpha_0 + \sum_{k=1}^{p} \alpha_k x_k = 0$$

We construct a label, $\alpha_0 + \sum_{k=1}^{p} \alpha_k x_k =$, where y ϵ {-1, 1} is the label classifier. We can redefine the hyperplane equation using inner products

$$y = \alpha_0 + \sum_{k=1}^{p} \beta_k y_k x(k) * x,$$

Here $*$ denotes the inner product operator and each inner product are weighted by its label. We maximize the distance from data point to the plane to obtain the margin of the optimal hyperplane. The maximum marginal hyperplane (MMH) splits the input data very well. So that the data point neighboring to the boundary of the hyperplane are participated in selection and discarded the remaining points. These relevant points are called the support vectors, and the partition hyperplane is known as support vector classifier (SVC), which maximized the distance through the support to the hyperplane.

The principle of normal SVM model was developed by Vapnik et al. is described below.

Considering k numbers of set of data S = (x_i, y_j), i = 1, 2, … …, k and j = 1, 2,… …, k with each $x_i \epsilon$ x is a subset of R^m, denotes the m dimensional input data and corresponding target values $y_j \epsilon$ y is a subset of R. For the construction of a regression model, we initiated a function f: $R^m \rightarrow R$, to estimate the target value y for hidden data x, which are belongs to the testing data and not trained by the model. Using a nonlinear mapping ϕ, the input data transformed from R^m to R^p, for p > m. We define an estimation function f as

$$f(x) = w^T \phi(x) + b \quad (3)$$

Here $w \epsilon R^p$ is the coefficient of regression vector, b ϵ R, is the bias. Main objective of the support vector regression is to construct a mapping f which has most ϵ-deviation from the target y_j. For which we compute 'w' and 'b' so that the risk of f(x) can be minimized.

$$R_{reg}(w) = \frac{1}{2}||w||^2 + C \sum_{j=1}^{k} L\epsilon\{y_j, f(x_j)\} \quad (4)$$

Where, "C" has maximized the margin between data sets and the hyperplane. $L\epsilon\{y_j, f(x_j)\}$ is the ϵ-intensive loss function, defined as

$$L\epsilon\left(y_j, f(x_j)\right) = \begin{cases} |y_j - f(x_j)| - \epsilon, & |y_j - f(x_j)| \ge \epsilon \\ 0, & |y_j - f(x_j)| < \epsilon \end{cases} \quad (5)$$

Equation (4) can be rewrite by introducing slack variables γ_i and ξ_i as $R_{reg}(w, \gamma_i, \xi_i) = Minimize\ \frac{1}{2}||w||^2 + C \sum_{j=1}^{k}(\gamma_j + \xi_j)$ (6)

Subject to constraints

$$\begin{cases} y_j - w^T x_j - b \le \epsilon + \gamma_j \\ w^T x_j + b - y_j \le \epsilon + \xi_j \\ \gamma_j, \xi_j \ge 0 \end{cases} \quad (7)$$

Where $\frac{1}{2}||w||^2$ is the regularization term which prevents over learning. $(\gamma_j + \xi_j)$ is the pragmatic risk.

By introducing Lagrange multipliers α_j, β_j, μ_j and η_j the quadratic optimization problem (6) and (7) can be reformulated by

$$L = \frac{1}{2}||w||^2 + C\sum_{j=1}^{k}(\gamma_j + \xi_j) - \sum_{j=1}^{k}\alpha_j(\epsilon + \gamma_j - y_j + w^T x_j + b) - \sum_{j=1}^{k}\beta_j(\epsilon + \xi_j + y_j - w^T x_j - b) - \sum_{j=1}^{k}(\mu_j \gamma_j + \eta_j \xi_j) \quad (8)$$

The dual of the corresponding optimization problem (6) and (7) is represented as

$$\text{Maximize } -\frac{1}{2}||w||^2 + C\sum_{j,n=1}^{k}(\alpha_j - \beta_j)(\alpha_n - \beta_n)(x_j)^T x_n - \epsilon \sum_{j=1}^{k}(\alpha_j + \beta_j) + \sum_{j=1}^{k}(\alpha_j - \beta_j)$$

Subject to constraints $\begin{cases} \sum_{j=1}^{k}(\alpha_j - \beta_j) = 0 \\ \alpha_j, \beta_j \epsilon [0, C] \end{cases}$

By changing the equation $w = \sum_{j=1}^{k}(\alpha_j - \beta_j) x_j$, the function f(x) can be rewrite as

$$f(x) = \sum_{j=1}^{k}[(\alpha_j - \beta_j) x_j]^T \phi(x) + b \quad (9)$$

Consequently by applying Lagrange and Kuhn-Tucker condition, the general SVR function can be modified as

$$f(x) = \sum_{j=1}^{k}\sum_{l}^{k}(\alpha_j - \beta_j) K(x_j, x_l) + b \quad (10)$$

Here $K(x_j, x_l)$ known as kernel function.

The kernel can be obtained by taking the inner product of $\phi(x_j)$ and $\phi(x_l)$ in the feature space

$$\phi(x_j, x_l) = \phi(x_j)^T \phi(x_l) \quad (11)$$

C. *Kernel Principal Component Analysis*

Kernel principal component analysis extracts principal components (PC), which are nonlinear in nature, by using a nonlinear kernel method. The key concept behind this is to transform the data into a higher dimensional space, followed by extraction using PCA. KPCA is a very familiar and well known data analysis technique. New variables are able to be produced by transformation of original variable, which is known as principal components. Extraction of a subset of variables from a larger data set depends upon the highest correlations of the principal component with the original variables. KPCA has several application areas in science and engineering.

Consider a non linear transformation ϕ(x) from the original d-dimensional feature space to a D-dimensional feature space, where D >> d, then each data point x_i is projected to a point $\phi(x_i)$. Vapnic-chervononkis stated that kernel mapping provides greater classification power by transferring the dimension of input space into a higher dimensional space. KPCA extracts finite number of non linear principal components. It is useful when input data lie on a low dimensional nonlinear hyperplane.

D. *Teaching* Learning Based Optimization

The Teaching Learning based Optimization TLBO is metaheuristic population based optimization model, which imitate the transmission of knowledge in a classroom atmosphere. It perform through two stage one is learner phase another is teacher phase. In the teacher phase a teacher transform the knowledge between the students, and during learner phase the students are acquire the knowledge by interaction among themselves. A set of student's represents the initial population and known as learners, and then the best student among all is recognized and leveled as the new teacher. This two phased process of acquire knowledge repeated until we get the desired accuracy.

Initially, recognized the decision variables and set as the features of individual's population. Let us consider "n" number of learners with "m" number of decision variables. Every decision variables were assigned in a range with upper limit U and lower limit L and the features for each learner represented as Y[j], which initializing randomly in the specified ranges.

Y [j] [k] = L[k] + RAND ( ) * {U[k] - L[k]} (12)

where, j = 1, 2, 3, … …, n; k = 1, 2, 3, … …, m and RAND ( ) produces number randomly in (0, 1).

Set the termination criteria of TLBO as the number of iteration or mean fitness value of learners

Calculate the population mean of teacher phase using

$Y_{mean}[k] = Y_{mean}[k] = \frac{1}{n}\{\sum_{i=1}^{n} Y[j][k]\}$ (13)

for j = 1, 2, 3, … …, n; k = 1, 2, 3, … …, k. Then, estimate the fitness function of all learners Y[j]. The learners having best fitness value has been recognized as the new teacher $Y_{teacher}$. Then $Y_{teacher}$ improves the knowledge of the entire class through quality of teaching and learners. Selection of new candidates depending on the difference between teacher and average of the learners for all decision variable. The difference mean can be calculated by

DM [k] = RAND ( ) * {$Y_{teacher}$ [k] – TF * Y_{mean} [k]} (14)

Where, k = 1, 2, 3… m and teaching factor TF = round {1 + RAND (0, 1)} which gives random values. Ultimately, the new teachers increase the knowledge of other learners through Y_{new} [j] [k] = Y [j] [k] + DM [k] (15)

For j = 1, 2, 3, … …, n; k = 1, 2, 3, … …, m; and if the current result X_{new} [j] gives better solution in comparison to previous X [j] then, the X_{new} [j] is accepted, otherwise it is discarded.

In the learner phase, each learner improves their level by exchanges of knowledge with each other randomly. Two different learners Y[p] and Y[q] are chosen arbitrarily and depending upon their fitness values their performance are updated. If fitness value of Y[p] > Y[q] then the new candidate is identified as Y_{new} [p] [k] = Y [p] [k] + RAND ( ) * (Y [p] [k] - Y[q] [k]) (16)

or else Y_{new} [p] [k] = Y[p] [k] + RAND ( ) * (Y [q] [k] – Y [p] [k]) (17)

For j = 1, 2, 3, ……, n; k = 1, 2, 3, ……, m

By applying same practice to all learners to obtained upgrade fitness value.

This process of teacher phase and learner phase are continued till we reach to the termination criteria. Then the fitness of the learners is evaluated and the optimum fitness value among all the learners identified as the final solution.

IV. PROPOSED MODEL

The proposed models is built using three components, i.e., OFS, SVR, TLBO and KPCA, SVR, TLBO . In this model OFS and KPCA are used to extract the most relevant features from the dataset, SVR is applied to address the forecasting mechanism and TLBO optimizes the hyper parameters of support vector regression. Proper selection of kernel type, regularization parameter, and the ϵ-insensitive loss of SVR greatly influences the efficiency of the prediction model. Therefore, radial basis function (RBF) kernel has been used due to the nonlinearity nature of time series dataset under study. Mathematically, RBF kernel is defined as $K(u,v) = e^{-\gamma||u-v||^2}$, where $\gamma = \frac{1}{2\sigma^2}$ .The hyperparameters C and γ of SVR that requires to be optimized by TLBO. The flowchart shown in Figure-1 describes the important steps involved in the prediction mechanism of the hybrid model. The OFS and KPCA are the first component that receives the dataset. Once the task of dimension reduction is over, the selected features filtered is forwarded to the SVR, and then optimized by TLBO. The flow chart describes the steps and mechanism of the given hybrid models.

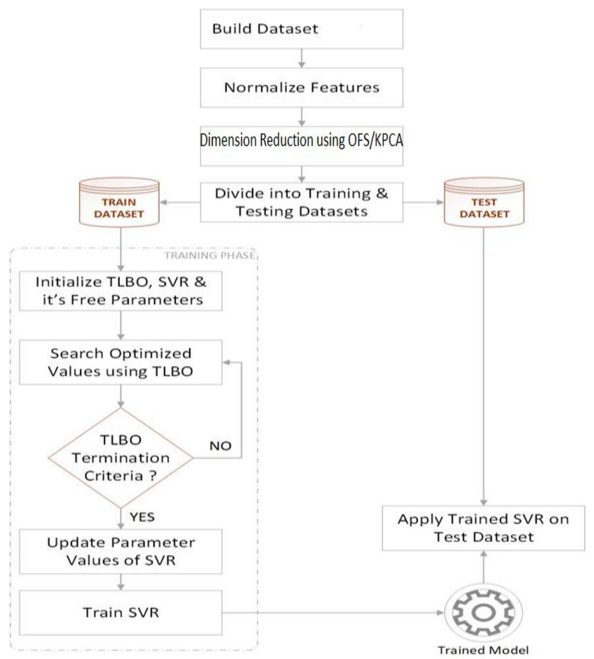

Figure-1: Flowchart of OFS/KPCA-SVR-TLBO Model

The dataset under study comprises of 7 features (shown in Table-1) that are daily recordings of time-series data, along with lagged (past period) values of the last 5 days.

Table-1: Features

Sl.	Variable	Description
1	Open Price	The opening share price on a day.
2	Highest Price	Maximum share value throughout the day.
3	Lowest Price	Lowest value of stock during a day.
4	Close price	The closing price of share on a day.
5	No. of Shares	Total transaction of shares on a day.
6	No. of Trades	Total number of trades happened on a day.
7	Turnover	Total value of stock traded on a day.

As the quantity of data values is very huge, so we first normalized the data. Normalization of data has been adopted to avoid numerical difficulties during computation and overcome the dominance of features with greater numerical ranges over smaller numerical ranges. Normalization of data removes the gross influences of the data. It minimizes the redundancy and maximizes the integrity as a result we achieved the improved accuracy of the algorithm. In this process of normalization the range of original data was transferred to another scale, which reduces its range and proceeds to carry the data more rapidly. In this proposed hybrid models the data set is normalized using the equation

$$NV_k = \frac{A_k - A_{min}}{A_{max} - A_{min}}, for\ k = 1,\ 2,\ 3,\ \ldots\ldots, l$$

Here, A_k denotes original cost of the kth features, l denotes represents the count of entire data processed, A_{max} is the highest value and A_{min} is the lowest value among whole data set. NV_k refers to the corresponding normalized value. After normalization the data lies in the closed interval [0.0, 1.0]. Here, the dataset is divided into training and testing datasets with a predefined proportion. After initializing the parameters of TLBO and SVR, the model building process starts with the help of the training dataset. The optimized values of the hyper parameters of SVR are obtained using TLBO and the termination criteria determine when to terminate this process of searching. Once the values of the free parameters are obtained, the same is used to build the SVR. Now to determine the efficiency of the trained model, the same is applied to the testing dataset.

V. IMPLEMENTATION OF MODELS

In this application of machine learning, the features under consideration may consists of irrelevant or dependent features and hence feature extraction can provide a newly created set of features by removing unrelated or unnecessary attributes from the existing data. We have applied dimension reduction techniques of OFS and KPCA to extract the important features from the input data build up the respective OFS-SVR-TLBO and KPCA-SVR-TLBO hybrid models. It is desired that feature extraction techniques to be applied at the initial stages to the whole dataset to simplify and rearrange the original data structure. The data generated after the application of OFS and KPCA is divided into 2 sets i.e., training and testing, with a predefined proportion. Then, the parameters of TLBO and SVR are initialized; following which, the model training procedure begins by the application of the training dataset. The optimized values of "C" and "γ" of SVR are obtained using TLBO and the termination criteria determine when to terminate the process of searching. Once the values of the hyper-parameters are obtained, the same is applied to build the SVR. Now to compute the prediction accuracy, the testing datasets were applied to the model. For both the proposed models, root mean squared error has been treated as fitness function during both the phases. In the testing phase of OFS-SVR-TLBO, the range of C varies from 0.25 to 20000 and the range of 'γ' has considered between 0.00006 and 20.

The value of ϵ-insensitive loss has fixed to 0.001. The population size of learners in TLBO has taken 20 and the maximum number of iterations is 60, where as in the testing phase of KPCA-SVR-TLBO, the range of C has configured between 0.008 and 10000 and the range of 'γ' has lies between 0.00001 and 50. ϵ-insensitive loss has fixed to 0.001. The learner's population in TLBO has taken 25 and the maximum number of iterations has been taken 75.

The parameters of both the models are shown in Table 3. With the mentioned parameter values, the given model have been optimized as per the process describe in Figure 1.

Radial basis function kernel has been used in SVR and defined as $K(u, v) = e^{-\gamma\|u-v\|^2}$, where $\gamma = \frac{1}{2\sigma^2}$. After completion of feature extraction, the selected filtered features are forwarded to SVR, which is optimized by TLBO. The dataset under study comprises of 7 features (shown in Table-1) that are daily recordings of time-series data, along with lagged (past period) values of the last 5 days. To power the machine learning technology a set of open source libraries have been plugged into the system. Weka (version 3.9.0), which is supported with a wide range of ML libraries, has enriched the analytical capability of this application. In addition to this some specialized machine learning libraries are also available for specific kind of tasks. For an additional implementation of SVR component, LibSVM (Version 1.0.8) is used. The KPCA component was organized from the open source library 'GitHub tools' and the OFS and TLBO algorithm have been developed by us. The system used in the implementation with Intel®Corei3-4005U_ 1.7GHz_4GB RAMS.

Table-2: Parameters values used in experiment.

SVR Parameters	Parameter value in OFS-SVR-TLBO	Parameter value in KPCA-SVR-TLBO
SVM Type	Epsilon SVR	Epsilon SVR
Kernel Type	Radial Basis Function	Radial Basis Function
C	0.25 to 20000	0.008 to 10000
γ	0.00006 to 20	0.00001 to 50
ϵ	0.001	0.001 (fixed)

TLBO Parameters	Value of TLBO supported by SVR in OFS-SVR-TLBO	Value of TLBO supported by SVR in KPCA-SVR-TLBO
Population size	20	25
Max. iterations	60	75
Optimization type	Minimization	Minimization
Lower limit	10, 0.00001	10, 0.00001
Upper limit	100000000, 100	100000000, 100
Fitness function	RMSE	RMSE

VI. EXPERIMENTAL RESULTS AND DISCUSSIONS

A. *Evaluation Criteria*

To evaluate the performance of the regression model, we have considered three standard statistical metrics. They are mean absolute error (MAE), root mean squared error (RMSE), and mean absolute percentage error (MAPE) and their details have been described in Table 3. As MAE, RMSE, and MAPE indicate variants of the differences between the actual and predicted values, it is important to note that smaller the error value, better the performance.

Table-3: Error evaluation

Sl.	Metric	Definition
1	Mean Absolute Error (MAE)	$\frac{1}{l}\sum_{i=1}^{l}\lvert y_i - d_i\rvert$
2	Root Mean Squared Error(RMSE)	$\sqrt{\frac{1}{l}\sum_{i=1}^{l}(y_i - d_i)^2}$
3	Mean Absolute Percentage Error (MAPE)	$\frac{1}{l}\left(\sum_{i=1}^{l}\left\lvert\frac{y_i - d_i}{d_i}\right\rvert\right)100$

where,
l is the total number of instances or records under evaluation,
d_i is the desired output value, i.e., actual or true value of interest, and
y_i is the estimated value obtained using a prediction algorithm.

B. *Comparison of Results*

In this paper, the performance of our hybrid model i.e., OFS-SVR-TLBO and KPCA-SVR-TLBO compared. Here, the total dataset are divided into two parts, training and testing datasets after the application of the dimension reduction mechanism using OFS and KPCA. The training and testing datasets are applied to the above hybrid model for training and testing phases for prediction of the next day opening price. Out of 4143 numbers of data of Tata Steel (from 24-July-2001 to 19-March-2018) three-fourth of the data are used for building the training dataset and rest one-fourth for the testing dataset. In OFS-SVR-TLBO model, errors evaluated with MAE, RMSE, and MAPE in training phase are 0.856451993, 1.895755314, and 0.4575 % (approx) respectively and the errors in testing phase are 1.202996162, 2.043010993, 0.4930 % (approx.) respectively. In KPCA-SVR-TLBO model, errors evaluated with MAE, RMSE, and MAPE in training phase are 0.194310815, 0.30439309, and 0.143517405 % (approx) respectively and the errors in testing phase are 0.233326719, 0.351664082, 0.271837721 % (approx.) respectively. The Table-4 shows the error measures found for both the models. This empirical study shows that KPCA-SVR-TLBO performed better than OFS-SVR-TLBO in all the three evaluation criteria.

Table-4: Comparison of Performance of OFS-SVR-TLBO and KPCA-SVR-TLBO Models on Training and Testing Datasets

		Models	
		OFS-SVR-TLBO	KPCA-SVR-TLBO
Training	MAE	0.856451993	0.194310815
	RMSE	1.895755314	0.30439309
	MAPE	0.457514028 %	0.143517405 %
Testing	MAE	1.202996162	0.233326719
	RMSE	2.043010993	0.351664082
	MAPE	0.493052126 %	0.271837721 %

The Figures- 2 to 9 shows the comparison of the actual stock value and prediction of stock values using OFS-SVR-TLBO and KPCA-SVR-TLBO. It also includes the absolute error.

Figure-2: Actual Verses Prediction of OFS-SVR-TLBO on complete dataset.

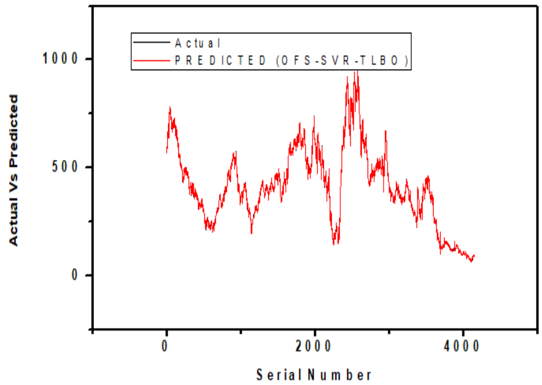

Figure-2: Actual Verses Prediction of OFS-SVR-TLBO on complete dataset

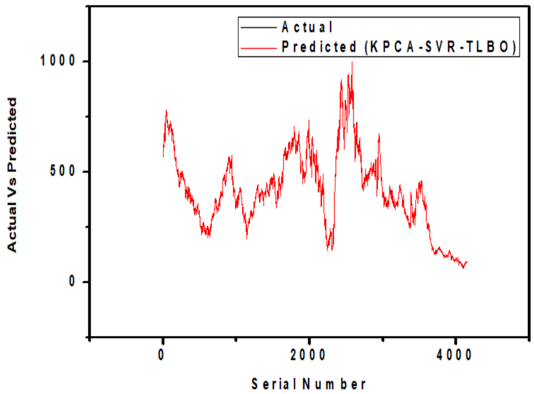

Figure-3: Actual Verses Prediction of KPCA-SVR-TLBO on complete dataset

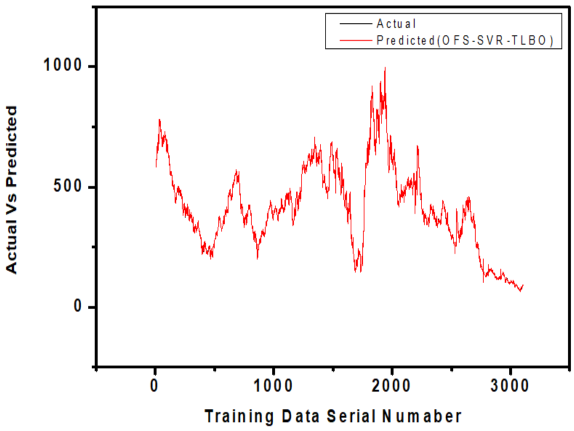

Figure-4: Actual Verses Prediction of OFS-SVR-TLBO on training dataset

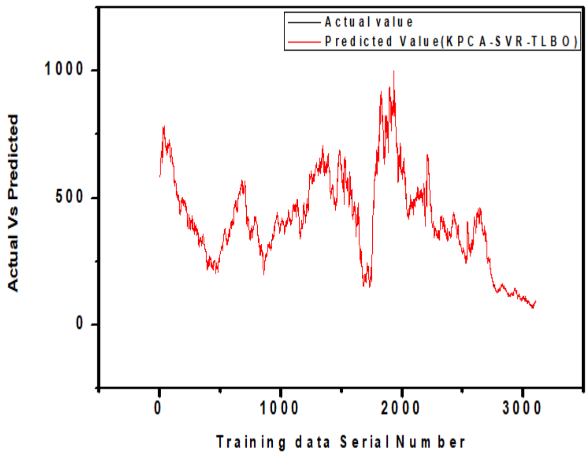

Figure-5: Actual Verses Prediction of KPCA-SVR-TLBO on training dataset

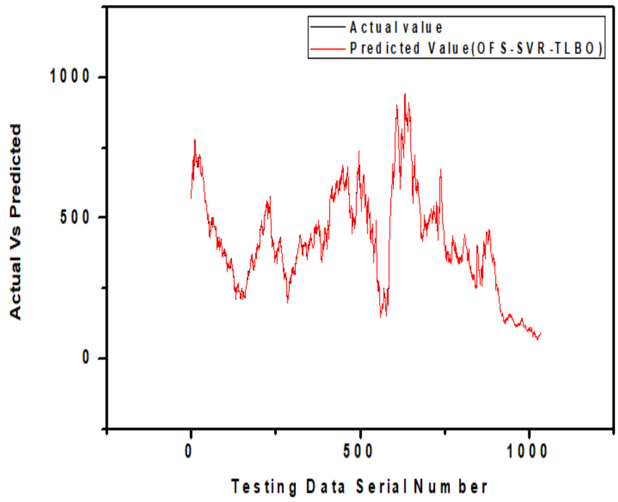

Figure-6: Actual Verses Prediction of OFS-SVR-TLBO on testing dataset

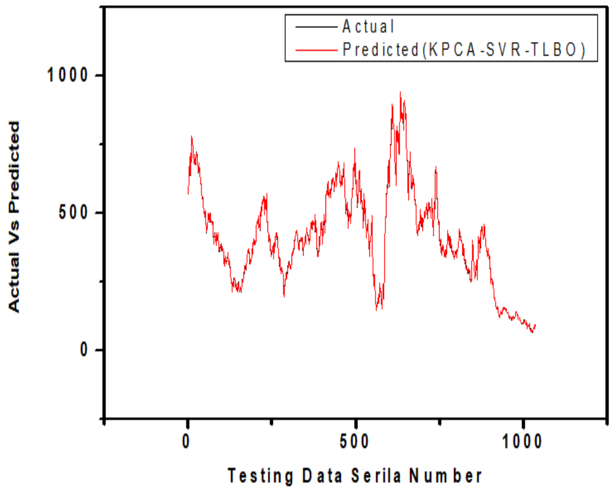

Figure-7: Actual Verses Prediction of KPCA-SVR-TLBO on testing dataset

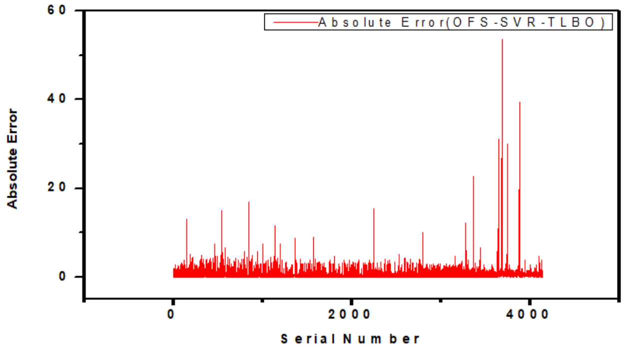

Figure-8: Absolute error of OFS-SVR-TLBO

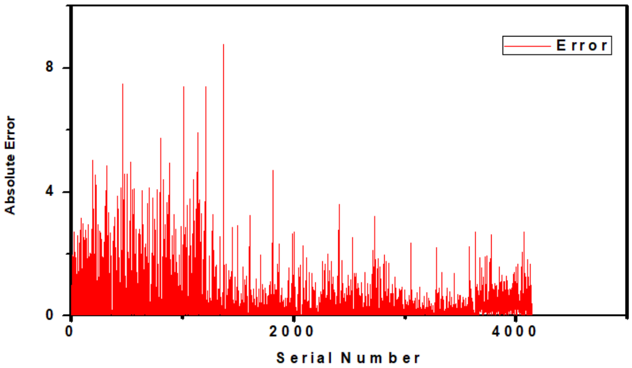

Figure-9: Absolute error of KPCA-SVR-TLBO

VII. CONCLUSION

The stock market is an important and inter-connected economic international business. To address the daily open stock price forecasting of Tata still, we proposed two hybrid models by incorporating OFS and KPCA with SVR-TLBO. Empirical results show that in the testing phase of OFS-SVR-TLBO hybrid model 0.5 % (approx.) mean absolute percentage error (MAPE) where as in KPCA-SVR-TLBO hybrid model shows 0.27 % (approx.) mean absolute percentage error (MAPE), and also it outperformed OFS-SVR-TLBO in MAE, RMSE, and MAPE. We conclude that KPCA-SVR-TLBO improved the performance of the prediction model due to the presence of the feature extraction technique KPCA and

Such remarkable performance is achieved due to the application of KPCA on the lagged time-series dataset and use of teaching learning based optimization, which optimized the hyper parameters of support vector regression (SVR). Based on the outcome of this piece of research work, we proposed KPCA-SVR-TLBO hybrid model for the future applications of regression based forecasting tasks and it can be used for taking better decision and more accurate predictions for financial investors on daily stock market forecasting.

REFERENCES

1. Cao, L. (2003). Support vector machines experts for time series forecasting, Neurocomputing, Elsevier, Vol. 51: 321 – 339.
2. Pai, P. F., and Hong, W. C. (2006). Software reliability forecasting by support vector machines with simulated annealing algorithms. Journal of Systems and Software, Vol. 79(6): 747–755.
3. Lin, S. W., Ying, K. C., Chen, S. C. and Lee, Z. J. (2008). Particle swarm optimization for parameter determination and feature selection of support vector machines," Expert systems with applications, Vol. 35 (4): 1817–1824.
4. Huang, C.L. and Dun, J.F. (2008). A distributed PSO–SVM hybrid system with feature selection and parameter optimization, Applied Soft Computing Elsevier, Vol. 8: 1381–1391
5. Lu, C. J., Lee, T. S and Chiu, C. C. (2009). Financial time series forecasting using independent component analysis and support vector regression, Decision Support Systems, Vol. 47 (2):115–125.
6. Zhang, X., Chen, X. and He, Z. J. (2010). An ACO-based algorithm for parameter optimization of support vector machines, Expert Systems with Applications, Elsevier, Vol. 37: 6618–6628.
7. Huang, C.F. (2011). A hybrid stock selection model using genetic algorithms and support vector Regression, Applied Soft Computing, Elsevier, Vol. 12: 807–818.
8. Kara, Y., Boyacioglu, M. A. and Baykan, O.K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange, Expert Systems with Applications, Elsevier, Vol. 38: 5311–5319.
9. Rao, R. V., Savsani, V. J., & Balic, J. (2012). Teaching–learning-based optimization algorithm for unconstrained and constrained real-parameter optimization problems. Engineering Optimization, Vol. 44(12): 1447–1462.
10. Rao, R.V. and Patel, V. (2012). An improved teaching-learning-based optimization algorithm for solving unconstrained optimization problems, Scientia Iranica, Sharif University of Technology, Vol. 20 (3): 710–720.
11. He, Y., Fataliyev, K. and Wang, L. (2013). Feature Selection for Stock Market Analysis, Springer-Verlag Berlin Heidelberg: 737–744.
12. Hsu, C.M. (2013). A hybrid procedure with feature selection for resolving stock/futures price forecasting problems, Neural Comput & Applic, Vol. 22: 651–671.
13. Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. Computers & Electrical Engineering, Vol. 40(1): 16–28.
14. Chang, P. C. and Wu, J. L. (2015). A critical feature extraction by kernel PCA in stock trading model, Soft Comput, Springer, Vol. 19: 1393–1408.
15. Ballings, M., Poel, D.V., Hespeels, N. and Gryp, R. (2015). Evaluating multiple classifiers for stock price direction prediction, Expert Systems with Applications, Elsevier, Vol. 42: 7046-7056.
16. Patel, J., Shah, S., Thakkar, P. and Kotecha, K. (2015). Predicting stock market index using fusion of machine learning Techniques, Expert Systems with Applications, Elsevier, Vol. 42: 2162–2172.
17. Das, S. P., Achary, N.S. and Padhy, S. (2016). Novel hybrid SVM-TLBO forecasting model incorporating dimensionality reduction techniques, Appl Intell, Springer, Vol. 45:1148–1165.
18. Jaffel, I., Taouali, O., Harkat, M.F. and Messaoud, H. (2016). Moving window KPCA with reduced complexity for nonlinear dynamic process monitoring, ISA transactions, Vol. 64:184–192.
19. Das, S. P., Achary, N.S. and Padhy, S. (2016). Novel hybrid SVM-TLBO forecasting model incorporating dimensionality reduction techniques, Appl Intell, Springer, Vol. 45:1148–1165.
20. Patel, V. K. and Savsani, V. J. (2016). A multi-objective improved teaching–learning based optimization algorithm, Information Sciences, Elsevier, Vol. 357: 182–200.
21. Patel, V. K. and Savsani, V. J. (2016). A multi-objective improved teaching–learning based optimization algorithm, Information Sciences, Elsevier, Vol. 357: 182–200.
22. Mohanty, S. and Padhy, S. (2017). A novel OFS–TLBO–SVR hybrid model for optimal budget allocation of government schemes to maximize GVA at factor cost, Journal of Management Analytics, Taylor & Francis, Vol. 5 (2018): 32-53
23. Das, S. P. and Padhy, S. (2018). A novel hybrid model using teaching–learning-based optimization and a support vector machine for commodity futures index forecasting, International Journal of Machine Learning and Cybernetics, Springer, Vol. 9: 97-111.
24. Henrique, B.M., Sobreiro, V.A. and Kimura, H. (2018). Stock price prediction using support vector regression on daily and up to the minute prices, The Journal of Finance and Data Science, Science Direct, Vol. 4 : 183-201.
25. Mao, K. Z. (2004). Orthogonal forward selection and backward elimination algorithms for feature subset selection. IEEE Transactions on Systems, Man and Cybernetics, Vol. 34(1): 629–634.

Retrieval Number: B3223078219/19©BEIESP
DOI: 10.35940/ijrte.B3223.078219
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication