Scientia Iranica D (2023) 30(5), 1625–1644

SCIENTIA IRANICA

Sharif University of Technology
Scientia Iranica
Transactions D: Computer Science & Engineering and Electrical Engineering
http://scientiairanica.sharif.edu

Stock market prediction using hybrid multi-layer decomposition and optimized multi-kernel extreme learning machine

P. Kumar Mallick[a], A. Ranjan Panda[a], A. Kumar Parida[a,*], M. Ranjan Panda[a], and S. Rani Samanta[b]

a. *School of Computer Engineering, Kalinga Institute of Industrial Technology Deemed to be University, Bhubaneswar, Odisha-751024, India.*
b. *Vice-Chancellor and Kalinga Institute of Industrial Technology Deemed to be University, Bhubaneswar, Odisha, India.*

Received 27 October 2021; received in revised form 20 May 2022; accepted 15 April 2023

KEYWORDS
Stock market closing price;
Variational mode decomposition;
Empirical mode decomposition;
Kernel extreme learning machine;
Firefly algorithm.

Abstract. The financial time series data is a highly nonlinear signal and hence difficult to predict precisely. The prediction accuracy can be improved by linearizing the signal. In this paper, the nonlinear data sample is linearized by decomposing it into several Intrinsic Mode Functions (IMFs). A hybrid multi-layer decomposition technique is developed. The decomposition method proposed in this paper is composed of both Empirical Mode Decomposition (EMD) and Variational Mode Decomposition (VMD) methods individually. As a new contribution to the previous literature in this study, the VMD is used to decompose further the higher frequency signals obtained from the EMD-based decomposed signal. The result analysis shows that the double decomposition technique improves prediction accuracy. This is a new introduction to the field of stock market prediction. The prediction accuracy of the proposed model is verified by applying it to three different stock market data. Historical data (closing price) is implemented to obtain one day ahead predicted closing price. Comparative analysis of other previously implemented methods like Back Propagation Neural Network (BPNN), Support Vector Machine (SVM), Artificial Neural Network (ANN), and Extreme Learning Machine (ELM), along with the proposed method, is performed. Firefly algorithm is implemented for optimizing the kernel factors. It is observed that the proposed hybrid model outperformed other methods discussed in this study.

1. Introduction

The main difficulty in predicting the accurate price of the stock market is due to the nonlinearity present in the stock market data. It is an important aspect from the investor's point of view as it will reduce the risk caused due to future trends.

The highly fluctuating nature of the stock market makes it difficult to predict future trends in investment. A precise computational technique is needed to predict the stock price movement.

An accurate prediction model is attracting the interest of researchers and financial investors. An accurate prediction of the future trend in the stock

*. *Corresponding author.*
E-mail addresses: pradeep.mallickfcs@kiit.ac.in (P. Kumar Mallick); amiya.pandafcs@kiit.ac.in (A. Ranjan Panda); erakparida@gmail.com (A. Kumar Parida); mohit.pandafcs@kiit.ac.in (M. Ranjan Panda); drsasmita@kiit.ac.in (S. Rani Samanta)

doi: 10.24200/sci.2023.59307.6168

market reduces the risk of the investors for probable loss [1,2].

1.1. Literature review

To increase the prediction accuracy, different prediction model has been developed in recent years for predicting the financial market. Various artificial intelligence models have shown promising contributions to stock market prediction. Prediction models can be categorized as statistical, artificial intelligence, and hybrid models. In the previous study, the mostly implemented model was the statistical model. Traditional approaches extensively used previously are Autoregressive Moving Average Model (ARMA), Autoregressive Integrated Moving Average (ARIMA), Autoregressive Moving Average with eXogenous inputs (ARMAX), Generalized Autoregressive Conditional Heteroskedasticity (GARCH), etc [3–5]. The main drawback of these models is that they are incapable of handling nonlinear data, the computational complexity is more, and the problem of local minima, therefore, slow convergence rate. These limitations led to the introduction of other new models for more accurate predictions. These new models include Back Propagation Neural Network (BPNN), Artificial Neural Network (ANN), Feed-Forward Neural Network (FFNN), Support Vector Machine (SVM), Recurrent Neural Network (RNN), and Extreme Learning Machine (ELM) [6–11]. The Single layered feed-forward neural network gained more importance because of less computational complexity and hence more accurate prediction. ELM depends on the propitious selection of the hidden neurons and the activation function for stability and generalization. The limitation of ELM is the selection of the number of neurons in the hidden layer and weight matrix. To solve this problem, the kernel function came into existence. This improved the stability and generalization of the method. The kernel function enhances the system stability, and when applied in combination with the ELM technique, it is termed a Kernel-based Extreme Learning Machine (KELM) [12]. The basic concept of KELM is such that the kernel matrix is formulated to describe the feature mapping where no random weights are used for its processing. The only concern of this process is the choice of kernel function and its parameters. The limitations of the parameter choice are performed by an optimization technique. Previous literature shows that different optimization techniques like the Particle Swarm Optimization (PSO), Genetic Algorithm (GA), fish swarm optimization, fruit fly, firefly, etc., were used for both weight and parameter optimization [13–18]. This study compares different optimization techniques and shows that the Firefly Algorithm (FA) gives better results in less time than other optimization techniques.

1.2. Objective

As the main objective of the study is to have better prediction accuracy, the traditional prediction model lacks the ability of precise predictions because of the nonlinearity present in the financial data. Previously researchers have tried solving the data's complexity with the help of decomposition techniques. The nonlinearity present in the original dataset is handled using different decomposition techniques before proceeding to the prediction model. The previous literature study shows that a single-layer decomposition can deal with the nonlinear signal but is sometimes limited to a certain extent. Hence not able to deal with the irregular nature of the time-series data. Therefore, this paper proposes a new hybrid model combining two decomposition techniques and a parameter-optimized prediction technique. Previously different single prediction methods were implied by the previous researchers, but the main challenge is handling the data's nonlinearity. This further led to the introduction of decomposition techniques, which increased the predictive accuracy. In [19], the authors described a hybrid model for wind speed prediction where wavelet packet transform was implemented along with the Least Squares Support Vector Machines (LSSVM) technique. Various research was performed in different fields of the application implementing different decomposition techniques like Wavelet Decomposition (WD), Empirical Wavelet Transform (EWT), Empirical Mode Decomposition (EMD), and Variational Mode Decomposition (VMD) combined with various prediction techniques for reducing the noise or nonlinearity present in the signal [12,20–24]. These studies showed the ability of the prediction technique up to a certain extent, as the single decomposition layer cannot reduce the signal nonlinearity, and there lies a probability of further improving the model's prediction ability. Considering the decomposition property in this paper, a multi-layer decomposition technique is introduced.

1.3. Contribution

The main contribution of the paper can be summarised in three parts: in the first part, multi-layer decomposition is performed; initially, a single decomposition technique was implemented to decompose the entire data set, but as newness to the already existing technique in this paper multi-layer decomposition concept is used which further decomposes the higher frequency Intrinsic Mode Function (IMF) for providing more accurate predicted result. Secondly, the kernel extreme learning machine is implemented, and it is observed that a multi-kernel function or combination of more than one kernel function provided more accurate results without much effect on the computational time. Lastly, the kernel parameters were chosen with the help of firefly optimization. At first, the EMD technique is employed to decompose the original nonlinear stock market data into several IMFs and a residual where each IMF is of a different frequency band. The higher frequency IMFs may cause a problem in prediction accuracy; hence, another decomposition model termed VMD is applied to higher frequency IMF for further decomposition, providing more accurate prediction. The IMFs obtained after multi-layer decomposition are fed as the input to the kernel-based ELM; as discussed earlier, the main problem of the kernel function is the choice of parameter, which is solved using the FA-based optimization technique. Based on the aforementioned literature survey, the main contribution of the paper can be segregated into the following three aspects: firstly, the hybrid model composed of a two-layered decomposition technique and FA-MKELM for financial market prediction. Although the single decomposition process EMD has an extensive application in stock market prediction, as in [25] and [26], the researchers have used the EMD-based decomposition to solve the financial market prediction issue. However, another issue has been emphasized by the researchers where the first IMF with higher frequency gets influenced by various stochastic factors, thereby making the prediction more difficult. Hence, this study gives a unique model with a two-layered decomposition technique tackling the aforementioned issues. The VMD technique specifically decomposes the higher frequency IMF into several IMFs, thus reducing the nonlinearity of the time-series model; as shown in [12], the VMD shows better performance than the EMD; hence it is chosen to decompose the higher frequency IMF. Secondly: the FA is applied to optimize the kernel parameters of the multi-kernel technique giving a more accurate prediction result. And lastly, four different stock market data (BSE, NSE, S&P500, and SSE) of different regions were considered for evaluating the proposed model's performance capability. This type of prediction is not performed in the field of the stock market and can be proved beneficial for investors.

The rest of the paper is organized as follows: Section 2 describes the different decomposition techniques implemented in this paper for error minimization purposes. Section 3 briefly describes different prediction techniques, whereas Section 4 demonstrates the proposed hybrid (EMD-VMD-FA-MKELM) model. Section 5 demonstrates the experimental result and analysis result; lastly, Section 6 concludes the paper along with the future scope.

2. Data pre-processing

2.1. Empirical Mode Decomposition (EMD)

EMD is a comparatively easily implemented time series decomposition technique whose main motive is to decompose the nonlinear data unless converted to a stationary signal. The main disadvantage, i.e., the uncertainty limitation factor of the previously implemented EWT technique, is solved by this technique. With the help of this technique, the primary signal is decomposed into several small and finite number of IMFs and a residue using the Hilbert-Huang Transform (HHT) [21].

IMF is a hidden oscillation mode that satisfies certain conditions as mentioned below:

(1) The total number of zero crossing and extrema should be equal, or else the difference between them should not exceed one;

(2) At any instant, the mean value denoted by the local maxima and minima is always zero.

The IMF-based decomposition can be provided as follows:

(a) Each local extrema should be pinpointed;

(b) The technique for producing the lower envelope is repeated after connecting the local maxima with a cubic spline line for the higher envelope since the upper and lower envelope should cover all the data between them [13].

The mean is designated as $mean_n(t)$, and the difference between the signal and mean, i.e., $x_n(t)$ and $mean_n(t)$ behaves as the first component $H_n(t)$, where $H_n(t)$ can be denoted as:

$$H_n(t) = x_n(t) - mean_n(t). \tag{1}$$

In ideal conditions, the definition of IMF is satisfied by $H_n(t)$. This applies only to some cases, as a new extremum gets introduced every time while changing the local zero from a rectangular to a curvilinear system. Corresponding adjustments are made, and hence the repetition of the process continues [12].

Changing local zero from a rectangular to a curved system introduces new extrema, necessitating changes, so that doesn't happen in most circumstances, necessitating repeat and shifting of the method [12]. $H_n(t)$ is now considered to be a proto-IMF, and the next iteration of data is expressed as follows:

$$H_{n1}(t) = H_n(t) - mean_{n1}(t). \tag{2}$$

The process of shifting continues for k number of times unless H_{nk} becomes the true IMF, that is:

$$H_{nk}(t) = H_{n(k-1)}(t) - mean_{nk}(t), \tag{3}$$

and finally, it is designated as:

$$c_n(t) = H_{nk}(t). \tag{4}$$

The stopping criterion for this shifting process was suggested by Guo et al. [21], a normalized squared difference between the two successive shifting operations is required.

$$S_1 D_k = \sum_{t=0}^{T} \frac{\left|H_{n(k-1)}(t) - H_{nk}(t)\right|^2}{H_{n(k-1)}^2(t)}. \quad (5)$$

$c_n(t)$ along with the rest of the data, can be separated as:

$$r_n(t) = x_n(t) - c_n(t). \quad (6)$$

Therefore the residue obtained is termed as $r_n(t)$, that consists of a component with a more extended period than the previous component. This is considered as the new signal, repetition of the process determines the second and further IMFs and the residue ($c_{n2}(t)$ and $r_{n2}(t)$). The repetition of the procedure results in:

$$r_{ni}(t) = r_{n(i-1)}(t) - c_{ni}(t), \quad (7)$$

where $i = 1, 2, 3, ..., m$.
A predetermined value ends the shifting process, thus reconstructing the final signal after the cumulative summation of all the IMFs. The final output signal is determined as follows:

$$x_n(t) = \sum_{i=1}^{m} c_{ni}(t) + r_{nm}(t). \quad (8)$$

2.2. Variational Mode Decomposition (VMD)

The main aim of the VMD technique is a non-recursive decomposition of a valued signal into various modes. The original signal (X) is decomposed into several modes within a limited bandwidth. These individual modes are termed X_K. It is needful for each mode K to be compact about a center pulsation (W_K) in addition to the decomposition. Each mode maintains a specific sparsity. Sparsity is defined as the bandwidth within its spectral domain. The following algorithm determines the bandwidth of each mode. The performance capability of the VMD is shown in [12,13]:

1. The unilateral frequency spectrum is attainted by applying HHT to the mode X_K;
2. The frequency spectrum of a specific mode is reprogrammed to the baseband region; this is performed by mixing an exponential tuned to its corresponding approximated center frequencies;
3. The bandwidth of each mode is estimated using the demodulated signal's Gaussian smoothness, which is also known as the norm of the gradient.

The following solution helps in obtaining the VMD [27,28], where each mode denoted as $X_K(t)$ is associated with an analytical signal with the help of HHT:

$$\left(\delta(t) + \frac{j}{\pi t}\right) * X_K(t). \quad (9)$$

Each mode's frequency spectrum is shifted to its estimated central frequency, as given below:

$$\left[\left(\delta(t) + \frac{j}{\pi t}\right) * X_K(t)\right]_e^{-j\omega_K t}. \quad (10)$$

After the bandwidth estimation, the resulting constrained variational problem is expressed as:

$$\min_{\{X_K\}\cdot\{W_K\}} \left\{ \sum_{K=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi\, t}\right) * X_K(t) \right] e^{-j w_K t} \right\|_2^2 \right\}. \quad (11)$$

$$\text{Such that:} \quad \sum_{K=1}^{K} X_K = X. \quad (12)$$

The original signal (X) is decomposed $\{X_K\}$ into a set of modes termed as $\{X_1, X_2, ..., X_K\}$. W_K is the set of center pulsation, i.e., $\{W_1, W_2, ..., W_K\}$. K is the number of modes, $*$ is the convolution. δ is known as the Dirac distribution. The above-mentioned constrained problem can be modified as:

$$L(\{X_K\}, \{W_K\}, \lambda) := \alpha \sum_{K=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t}\right) * X_K(t) \right] e^{-j w_K t} \right\|_2^2 + \left\| X(t) - \sum_{K=1}^{K} X_K(t) \right\|_2^2 + \left\langle \lambda(t),\, X(t) - \sum_{K=1}^{K} X_K(t) \right\rangle. \quad (13)$$

Eq. (11) is obtained as a progression to the iterative measures commonly termed ADMM (Alternate Direction Method of Multipliers). ADMM is applied to obtain the saddle point of the previously derived augmented Lagrange expression:

$$X_K^{n+1} = \underset{X_K \in X}{\arg} \min \left\{ \alpha \left\| \partial(t) \left[\left(\delta(t) + \frac{j}{\pi t}\right) * X_K(t) \right] e^{-j\omega_K t} \right\|_2^2 + \left\| X(t) - \sum_i X_i(t) + \frac{\lambda(t)}{2} \right\|_2^2 \right\}. \quad (14)$$

By applying Perceval Fourier isometry to Eq. (12) and W_K and $\sum_i X_i(t)$ correspond to W_k^{n+1} and $\sum_{iK} X_i(t)^{n+1}$, respectively.

$$X_K^{n+1} = \underset{\hat{X}_K, X_K \in X}{\arg} \left\{ \alpha \left\| j\omega \left[(1 + \operatorname{sgn}(\omega + \omega_K)) \hat{X}_K(\omega + \omega_K) \right] \right\|_2^2 + \left\| \hat{X}(\omega) - \sum_i \hat{X}_i(\omega) + \frac{\lambda(\omega)}{2} \right\|_2^2 \right\}. \quad (15)$$

The new solution obtained for X_K and W_K are as follows:

$$\omega_K^{n+1} = \underset{\omega_K}{\arg\min} \left\{ \int_0^\infty (\omega - \omega_K) \left| \hat{X}_K(\omega) \right|^2 d\omega \right\}. \quad (16)$$

Central frequency ω_K is further converted to different frequency domains:

$$\hat{X}_K^{n+1}(W) = \frac{\hat{X}(W) - \sum_{i \neq K} \hat{X}_i(W) + \frac{\hat{\lambda}(W)}{2}}{1 + 2\alpha(W - W_K)^2}, \quad (17)$$

$$W_K^{n+1} = \frac{\int_0^\infty W \left| \hat{X}_K(W) \right|^2 dw}{\int_0^\infty \left| \hat{X}_K(W) \right|^2 dw}, \quad (18)$$

here $\hat{X}(W)$, $\hat{X}_i(W)$, $\hat{\lambda}(W)$, and $\hat{X}_K^{n+1}(W)$ are the Fourier transforms of $X(t)$, $X_i(t)$, $\lambda(t)$, and $X_K^{n+1}(t)$, respectively, and n is the total number of iterations.

3. Prediction techniques

This section describes different prediction techniques included in this paper and the proposed method for comparison.

3.1. Multi-Kernel Extreme Learning Machine (M-KELM)

ELM basically performs as an Single-Layer Feedforward Neural Network (SLFN), the output function of the basic ELM with L hidden nodes is shown in Figure 1. Mathematically, it can be presented as:

$$f_k(x) = \sum_{i=1}^{L} \beta_i h_i(x) = h(x)\beta. \quad (19)$$

The vector between the hidden neurons (L) and the output neuron is termed the output vector (β), $\beta = [\beta_1, \beta_2, ..., \beta_L]$, the ELM feature mapping function is expressed as $h(x) = [h_1(x), h_2(x), ..., h_L(x)]$, where the number of the input sample is represented as X and $x = [x_1, x_2, ..., x_N]$; N is the number of patterns. m inputs for each pattern, and the hidden layer consist of L neurons that do not need to be tuned, and thus the activation function of the hidden neurons could be comprised of almost any nonlinear piecewise continuous functions. In this paper, the cos, sine, and tanh functions were tested, and it was observed that the tanh function gives a better prediction when compared with the other two. Thus, the following equation can be observed using the tanh function as the activation function:

$$h_i(x) = \tanh(w_{i0} + w_{i1}x_{i1} + w_{i2}x_{i2} + ... + w_{im}x_{im}). \quad (20)$$

Thus, the hidden layer randomized matrix is written as:

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} h(x_1) & \cdots & h_L(x_1) \\ \vdots & \vdots & \vdots \\ h(x_N) & \cdots & h_L(x_N) \end{bmatrix}, \quad (21)$$

and the target vector is given by:

$$T = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_N \end{bmatrix}. \quad (22)$$

Eq. (19) is expressed in a matrix form as:

$$H\beta = T. \quad (23)$$

To obtain the solution for the output vector (β), the overfitting problem needs to be defined. Thus, a

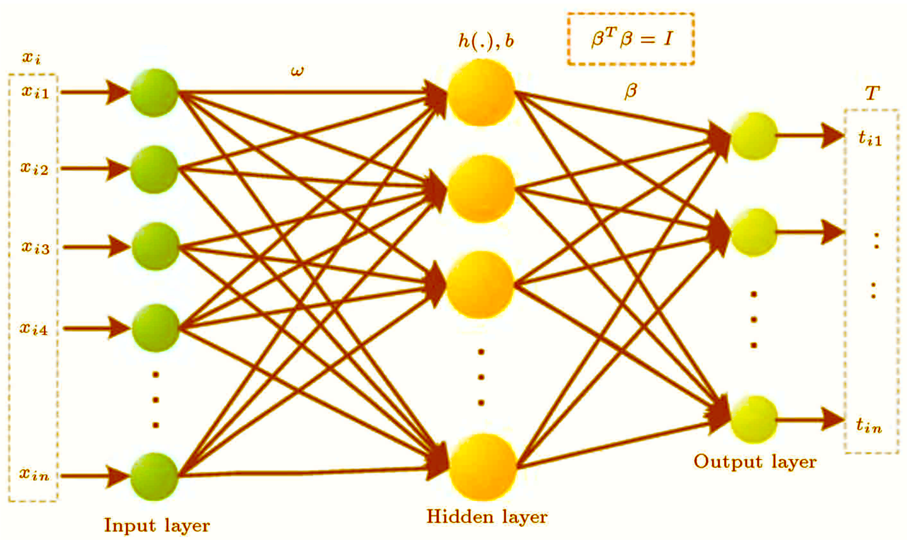

Figure 1. Architecture of basic ELM model.

constrained optimization problem is solved that resolves the overfitting problem, contributing to a better generalization ability than the original basic ELM. This resembles the structural risk minimization of the statistical learning theory and is defined as:

$$L_{C_{ELM}} = \frac{1}{2}||\beta||^2 + \frac{1}{2}C||\tau||^2. \tag{24}$$

Such that $H\beta = T - \tau$.
The error vector τ is obtained as: $\tau = [\tau_1, \tau_2, ..., \tau_N]$, $C \rightarrow$ regularization parameter.

Applying Karush-Kuhn-Tucker (KKT) in Eq. (24) it can be rewritten as:

$$L_{D_{ELM}} = \frac{1}{2}||\beta||^2 + \frac{1}{2}C||\tau||^2 - \alpha(H\beta - T + \tau), \tag{25}$$

vector of Lagrange multipliers is obtained as $\alpha = [\alpha_1, \alpha_2, ..., \alpha_N]$. By solving Eq. (25) for the optimality condition, the value of the vector β is obtained as:

$$\beta = \left(\frac{I}{C} + H^T H\right)^{-1} H^T T \qquad \text{when} \quad N > L, \tag{26}$$

$$\text{or } \beta = H^T\left(\frac{I}{C} + HH^T\right)^{-1} T \quad \text{when} \quad N < L. \tag{27}$$

In this study, the condition $N > L$ is considered, and I denotes the unity matrix.

The output for the x number of samples comprising m inputs can be denoted as:

$$f(x) = h(x)\left(\frac{I}{C} + H^T H\right)^{-1} H^T T \qquad \text{for} \quad N>. \tag{28}$$

The condition taken into consideration in this paper is $N > L$.
The output vector with N number of patterns is given as:

$$O = H\left(\frac{I}{C} + H^T H\right)^{-1} H^T T \qquad \text{for} \quad N > L.$$

The number of neurons present in the hidden layer and an appropriate activation function determines the stability and generalization performance of the ELM, making it an unsolved problem. Bringing in the concept of kernel function. When the feature mapping function is unknown, Kernel functions can be used for the ELM to provide better stability and generalization. This kernel function-based ELM is called KELM . Thus, the Kernel matrix for the KELM using the Mercer theorem can be written as:

$$KELM = HH^T, \quad \text{and} \quad KELM(x_i, x_j) = h(x_i)h(x_j). \tag{29}$$

Thus, the output function is obtained as:

$$f(x) = h(x) H^T \left(KELM + \frac{I}{C}\right)^{-1} T. \tag{30}$$

Eq. (13) is written in an expanded form as:

$$f(x) = [(x, x_1), (x, x_2), ..., (x, x_N)] \left(KELM + \frac{I}{C}\right)^{-1} T. \tag{31}$$

There are different Kernel functions satisfying the Mercer condition and are best suited for use. Various types of kernel functions are namely: Polynomial kernel, Gaussian Kernel, tan hyperbolic kernel (Sigmoid kernel), and Wavelet kernel. As a new addition to this study, a new variety of kernels known as the multi-kernel is also proposed, whereas multi-kernel functions are a weighted combination of two or more kernels. Different kernel functions are described below:

(1) Gaussian kernel:

$$K_g(x_i, x_j) = \exp\left(-||(x_i - x_j)||^2\right)/(2\delta^2). \tag{32}$$

(2) Sigmoid kernel:

$$K_s(x_i, x_j) = \tanh(bx_i^T x_j + c). \tag{33}$$

(3) Polynomial kernel:

$$K_p(x_i, x_j) = (1 + x_i^T x_j/\sigma^2)^d, d = 2. \tag{34}$$

(4) Wavelet kernel:

$$K_w(x_i, x_j) = \cos\left(\frac{e\ ||(x_i - x_j)\,||}{f}\right) \exp\left(-\frac{||(x_i - x_j)\,||^2}{g}\right). \tag{35}$$

(5) Multi-kernel:

$$K_m(x_i, x_j) = \mu_1 K_p(x_i, x_j) + (1 - \mu_1) K_w(x_i, x_j). \tag{36}$$

Parameters σ, b, c, e, f, and g are to be chosen appropriately to enhance the performance of the kernel-based ELM forecasting model. In the case of the multi-kernel ELM, a global and a local kernel are combined, the polynomial kernel and sigmoid kernel being the global kernel, whereas the Gaussian kernel and wavelet kernel falls under the local kernel category. This paper combines two best-performing kernel functions to form the multi-kernel (polynomial and wavelet kernel). For better performance, the choice of parameters should be made more accurately, necessitating using an optimization technique. The different parameters of the kernel functions are optimized using an optimization technique, namely the firefly optimization. Later in this section, a comprehensive discussion of the FA is performed.

3.2. Firefly optimization algorithm

It is a meta-heuristic algorithm initially developed by Yang in the year 2008. The basic characteristic of the firefly is the flashlight's occurrence due to the biochemical process termed bioluminescence, these flash signals are generated for a twofold function, and it is used for potential mating purposes and also for warding off predators. It is based on the principle of swarm intelligence. The main advantage of this technique is demonstrated as follows:

- Capable of automatic subdivision, thus provides satisfactory results for highly nonlinear problems;
- The multi-modal characteristics present in the algorithm can solve the multi-modal problems at a much higher convergence rate.

The aforementioned advantage leads to its implementation in various fields of application, like clustering [29], image processing [30], and improvement of seeker optimization [31].

Some of the sophisticated performance criteria of the fireflies for better performance are stated as follows:

All the flies are considered unisex, and thus the attractiveness between the flies does not depend upon their sex but rather depends upon the brightness of the flies. The flies with a lesser intensity of light move towards the brighter ones, brightness is indirectly proportional to the distance between the flies, i.e., with the decreases in distance, brightness increases. Brightness is determined by the value of the objective function; for maximization problems, the brightness increases with the increase in the value of the objective function. In this algorithm, the fireflies get attracted due to the flash intensity, which is associated with the objective function. In other words, a firefly is always attracted to another firefly with greater intensity. Different parameters of the FA, like the absorption coefficient, attraction coefficient, population size, and the number of iterations, must be accurately chosen. Detailed steps of the FA are presented below:

Step 1. Initialization of the parameters of FA, absorption coefficient γ, attraction coefficient β.

Step 2. Initialization of the population P_i for $i = 1, 2, 3, ..., n$

Step 3. Obtain the objective function O for each population of P.

Step 4. Obtain the best population and best cost function P_{best} and O_{best} successive to the minimum O.

Step 5. Increase the counter of iteration $i = i + 1$.

Step 6. For P_i do for P_j where $P_i \neq P_j$ do.

Step 7. If $O_j < O_i$ then do:

$$r_{ij} = \sqrt{\sum\nolimits_{k=1}^{q} \left(P_{ik} - P_{jk}\right)^2}.$$

Step 8. If $O_j < O_i$ then do:

$$r_{ij} = \sqrt{\sum\nolimits_{k=1}^{q} \left(P_{ik} - P_{jk}\right)^2}.$$

Step 9. $\beta = \beta_0 e^{-\gamma . r_{ij}}$
$P_{i\_new} = P + \beta . rand\,() . \left(P_j^k - P_i^k\right)$ for $k = 1, 2, ..., q$.

Step 10. $O_{i\_new} = f\left(P_{i\_new}\right)$.

Step 11. If $O_{i\_new} < O_i$ then
instead of P_i, put P_{i_new}.

Step 12. If $O_{i\_new} < O_{Best}$ then
replace P_{best} with P_{i_new}.
End if
End if
End for
End for
End while
Return best P

The best P obtained is the light intensity value of the firefly obtained at the end; it is also known as the best optimization solution. In this paper, different kernel parameters are optimized using the FA optimization technique.

4. Proposed technique

This section demonstrates the proposed hybrid (EMD-VMD-FA-MKELM) model based on multi-layer decomposition and optimized M-KELM for the prediction of the stock market. The main structure of this model is established on the mechanism of the ensemble two-layered decomposition. The first layer is the EMD layer that decomposes the signal into different IMFs and a residue, whereas the second layer is implemented to handle the nonlinearity of the first IMF (IMF1) by further decomposing it into the number of modes, which in turn decreases the prediction complication. The FA-based optimized M-KELM is used to predict different modes of the nonlinear signal. These modes/ IMFs further ensemble to predict the stock market data, IMF1,..., IMF $n-1$, Res. The original decomposed IMF using the EMD technique whereas IMF'1,..., IMF' $n-1$, Res' represents their consecutive predicted values. The details of this model can be well understood in Figure 2. The detailed performance of the two-layered decomposition-based ensemble prediction model can be written in the following steps:

Step 1: Decomposition of the original signal into the number of IMFs and residue using the EMD technique;

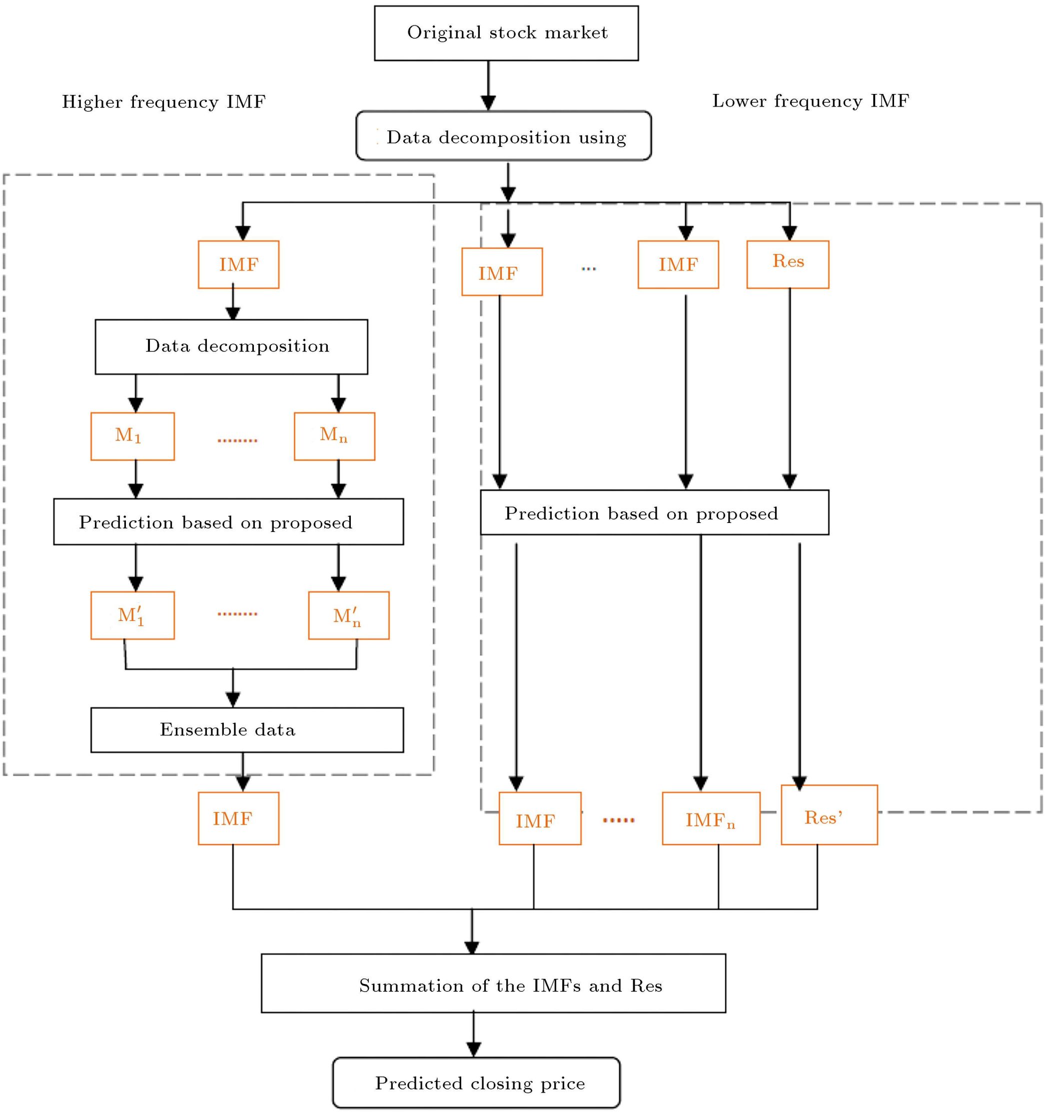

Figure 2. Schematic representation of the proposed model.

Step 2: Categorization of the IMFs into high-frequency (IMF1) and low-frequency IMFs (IMF 2-IMF$n-1$ and Res);

Step 3: Further decomposition of the higher frequency IMF1 into several modes using VMD;

Step 4: Prediction of the modes based on FA-M-KELM;

Step 5: Obtaining the predicted value of IMF1 is termed IMF1' by adding up the predicted values obtained in Step 4;

Step 6: Prediction of the low-frequency IMFs are performed based on FA-M-KELM; these predicted IMFs are termed IMF'2,..., IMF' $n-1$, and Res;

Step 7: Final prediction of the stock data is obtained by summing up each predicted IMFs and Res.

5. Experimental result analysis

In this section, a different aspect of the data is shown; it gives a wide knowledge of the data source and different input variables implemented to demonstrate the effectiveness of the proposed prediction model.

5.1. Data acquisition

This paper considers four different stock markets for experimental purposes, which can be categorized under developing and developed markets. BSE & NSE can be categorized under the developing market, whereas S&P 500 and Shanghai stock markets are categorized as underdeveloped. BSE is one of the leading Indian markets that comprise 30 blue-chip companies. NSE comprises 50 companies with core products. And the other two developed markets considered in this paper are SSE, which comprises 1615 companies indexed in it, and the S&P 500, which is based on 500 companies. The experimental verification of four different markets shown in the result analysis section proves the robustness in performance. It is observed from the previous studies that variation in the sample marks an impact on the prediction technique; hence in this paper, different markets are considered to observe the performance of the proposed prediction model. In this study, different market data is considered for prediction purposes ranging from 3rd January 2014 to 4th January 2020; the data is collected based on a 1-day interval. Out of the total data, 70% was used for training, while the rest was used for testing. The result analysis section gives detail on the number of data. The data is collected from online sources or to be more precious from yahoo finance [32]. Different markets are considered different cases in this study for a better understanding of the proposed hybrid model for determining individual market data [33–35]. The closing price is regarded as the input variable in this paper.

5.2. Performance index

The efficiency of the proposed model is testified by using different performance indices; these performance indices measure the prediction accuracy by comparing the difference between the predicted output and the desired output. The result analysis section will show both tabular and graphical representations of the different evaluation matrices' values. Among various performance matrices in this paper, we have considered the following for measuring the performance of the proposed model:

Mean Absolute Percentage Error (MAPE)

$$MAPE = \left(\frac{1}{N} \sum_{i=1}^{N} \frac{|T_i - P_i|}{T_i} \right) * 100. \tag{37}$$

Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} |T_i - P_i|^2}. \tag{38}$$

Mean Absolute Error (MAE):

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |T_i - P_i|. \tag{39}$$

Correlation Coefficient (R):

Eq. (40) is shown in Box I.

5.3. Result analysis

In this section, details regarding the evaluation of the model's performance are illustrated, and the four different data sets are used to explore the predicting capability of the prediction model. The models implemented in this paper are based on coding performed in MATLAB 15, using a system with the specification i5, 64 bit, 4 GB ram, and CPU 7200U @2.50 Ghz. Each data set is divided into training and testing data, where 70% of the data is fed as the training data while the rest 30% is considered the testing set. Four different financial market data are treated as four different case studies, and a comparative study is performed using different other prediction models like ELM, EMD-KELM, VMD-KELM, and the proposed hybrid model (EMD-VMD-FA-MKELM). A clear illustration of the prediction performed using the proposed model is performed by initially decomposing the original stock market signal; the FA-MKELM is utilized for predicting the IMFs and the residue. Initially, the EMD model decomposes the nonlinear original financial data series into several IMFs and a residue; further from the EMD-based decomposed signal, as shown in (Figure 4), it can be well observed that the IMF1 is of higher frequency as compared to the other IMFs. Hence it may lead to inferior execution of the proposed hybrid model; thus, a better decomposition technique is implemented to decompose the IMF1, thus performing the second decomposition of the multi-layer decomposition, thereby providing a better prediction result.

Different parameters of the FA are chosen, which imparts better prediction output; parameters like the number of iterations, the population of fireflies, randomness factor (α), attraction coefficient (β), and absorption coefficient (Υ) need to be appropriately selected for obtaining the better result, the parameter values are demonstrated in Table 1, these parameters are constantly maintained throughout the paper for

$$R = \frac{\left(N \sum_{i=1}^{N} T(x_i).P(x_i) - \sum_{i=1}^{N} T(x_i) \sum_{i=1}^{N} P(x_i) \right)^2}{\left(N \sum_{i=1}^{N} T(x_i)^2 - \left(\sum_{i=1}^{N} P(x_i) \right)^2 \right) \left(N \sum_{i=1}^{N} T(x_i)^2 - \left(\sum_{i=1}^{N} T(x_i) \right)^2 \right)}. \tag{40}$$

Box I

Table 1. Parameters of firefly optimization algorithm.

Specifications	Values
Number of iteration	100
Population of firefly	50
Attraction coefficient (β_0)	0.9
Absorption coefficient (γ)	1

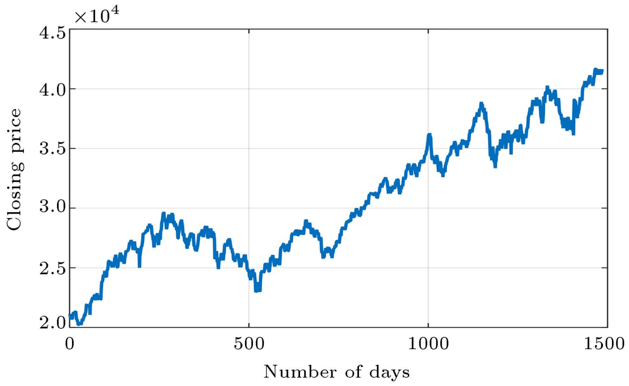

Figure 3. Original closing price of BSE market.

keeping the uniformity of the model. The optimized and non-optimized values of different kernel parameters are described for different case studies. The whole data is normalized between a common scale (0 to 1 is considered in this study) to reduce data redundancy, thereby increasing the data consistency; Eq. (41) is implemented to normalize the data further. It is helpful to perform the comparative study when different markets are taken into consideration; for example, in this paper, the closing price of the BSE and NSE market is obtained in Indian Rupee (INR) while the S&P 500 and Shanghai Stock Exchange (SSE) market is in US dollars; hence the normalized values gives uniformity to the market data. Different comparative studies are performed to mark the outperforming nature of the hybrid model in comparison to other models embedded with single decomposition techniques, further discussed in the following sub-section:

$$N' = \frac{N - N_{\min}}{N_{\max} - N_{\min}}, \tag{41}$$

here N' is considered to be the normalized value of the whole dataset, whereas N is the total data that is normalized while $N_{\min}$ and $N_{\max}$ represents the minimum and maximum stock price, respectively.

The first stock market considered for implementing the proposed prediction technique is a developing market. The original BSE data for a one-day interval is shown in Figure 3, which clearly shows the nonlinearity of the data; hence the necessity of the decomposition technique is well understood; the x-axis determines the number of days, whereas the y-axis gives the original closing price of the BSE market in INR.

Table 2a. Optimized kernel parameters of the multi-kernel function for the BSE stock market.

Time interval	Optimized kernel parameters				
	σ	g	μ_1	e	f
One-day interval	0.25	0.58	1.77	2.56	0.97

Table 2b. Non-optimized kernel parameters of single kernel function for BSE stock market.

Time interval	Non-optimized kernel parameters					
	σ	g	c	e	f	δ
One-day interval	0.8	1.82	0.84	2.3	0.85	1.63

Figure 4 shows the EMD-based decomposition where the IMFs ranged from higher to lower frequency. The original signal is decomposed into 10 IMFs (9 IMFs and a residue), IMF1 being the highest frequency signal with maximum details of the stock market and residue being the lowest frequency signal. IMF1 with maximum frequency is further decomposed using the VMD technique described in Figure 5, which shows 12 different modes of the IMF1 obtained from EMD-based decomposition, a total number of 1486 data is chosen (Jan-2014–Jan 2020), of which 1040 observations are considered for training and the rest 446 observations are considered testing data.

The FA is used to optimize different kernel parameters of the M-KELM; the optimized values are shown in Table 2a, whereas the un-optimized parameters of other kernel functions are mentioned in Table 2b; this is obtained by trial-and-error method with the help of authors previous experience, a total number of 10 trials are performed, and the values with minimum error is considered.

Convergence graphs for the BSE market are shown in Figure 6, and the minimum cost function corresponding to a particular number of iterations is presented on the y-axis. The total number of iterations considered is 100 in this study and the minimum value of the RMSE (cost function) is observed; thus, the parameter values corresponding to the minimum cost function are the desired optimized value.

Table 3 shows a comparative study of different kernel functions to show the best-performing kernel function; it can be observed that among the single kernel models (without any decomposition and optimization), the best-performing models are polynomial and wavelet. Hence these two better-performing kernel functions are combined to obtain the multi-kernel function.

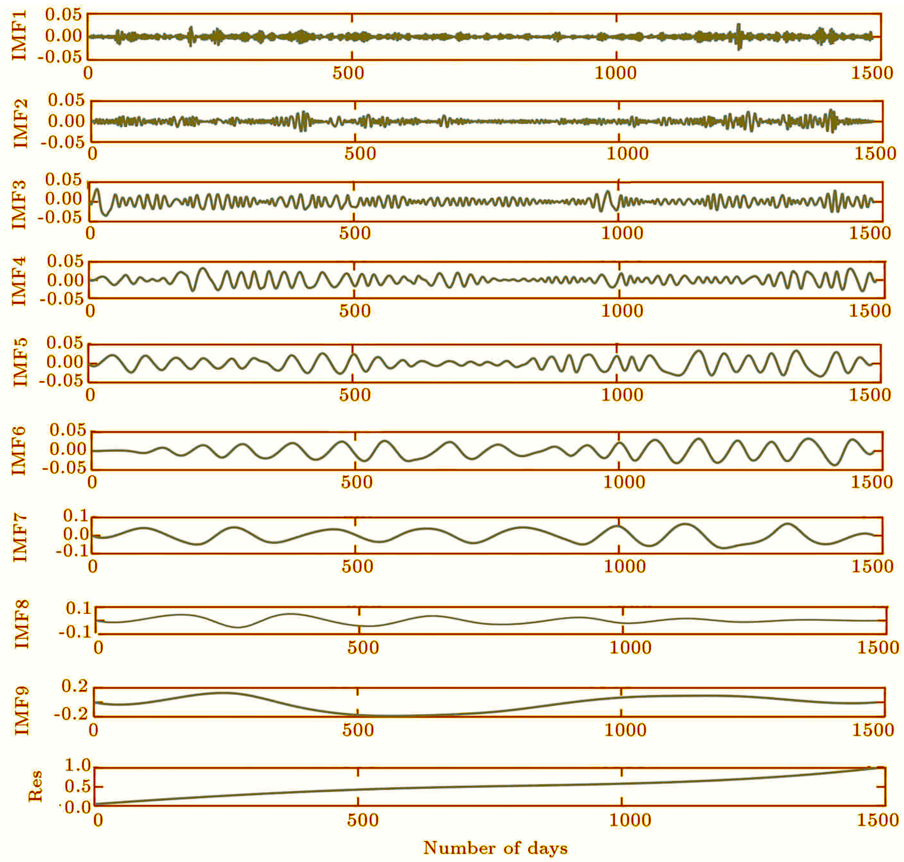

Figure 4. EMD-based decomposition of the BSE stock market data.

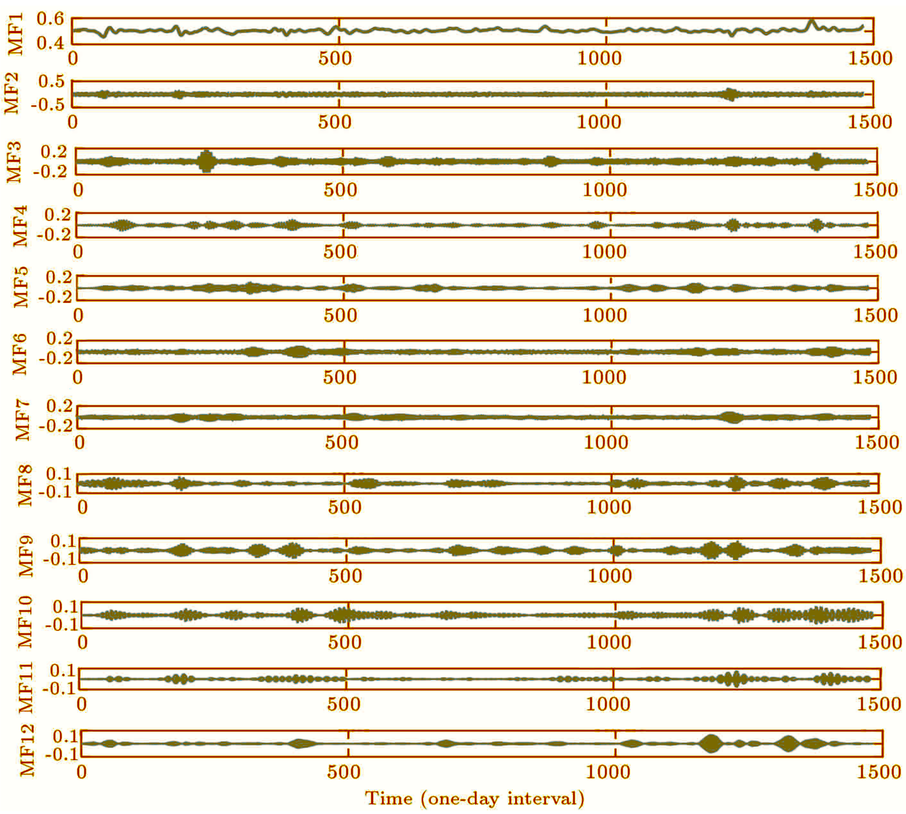

Figure 5. VMD-based decomposition of the BSE stock market data.

Table 3. Comparative study of different kernel functions for the BSE stock market.

Performance indices	G-KELM	P-KELM	W-KELM	S-KELM
MAPE	3.478	**2.926**	**2.653**	3.229
MAE	0.339	**0.274**	**0.268**	0.301
RMSE	0.36	**0.47**	**0.273**	0.39

Table 4. Performance evaluation of EMD-VMD-FA-MKELM compared to other prediction models based on different values examined on the BSE stock market data.

Methods	MAPE (%)	MAE (P.U)	RMSE (P.U)	R	Ex (sec)
ELM	3.756	0.125	0.135	0.966	98.04
M-KELM	3.134	0.104	0.105	0.971	102.45
EMD-MKELM	1.656	0.052	0.055	0.995	110.66
VMD-MKELM	1.177	0.039	0.041	0.996	119.58
FA-MKELM	2.471	0.082	0.083	0.979	105.75
EMD-VMD-FA-MKELM	0.274	0.009	0.008	0.998	121.35

Note. Ex: Execution Time.

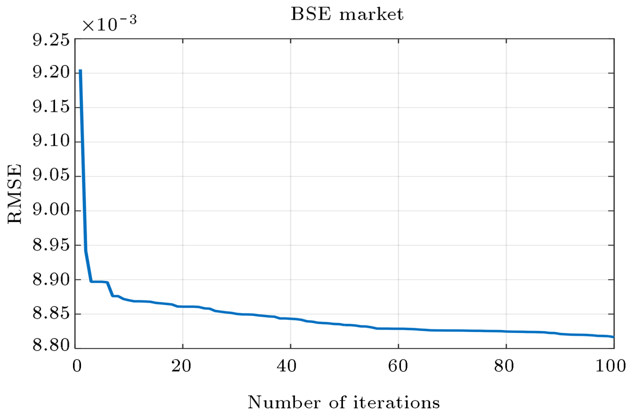

Figure 6. Convergence graph of BSE stock market data using FA.

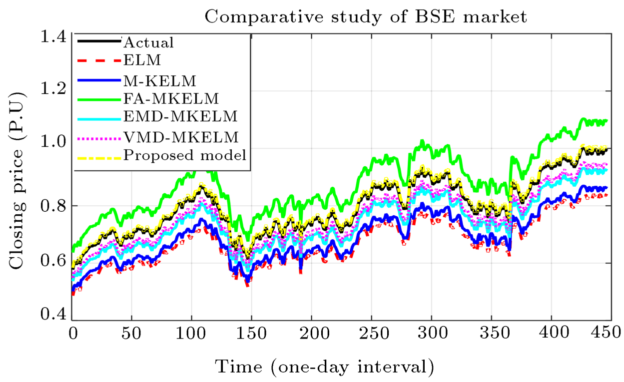

Figure 7. Comparative study of different one-day ahead prediction techniques.

- A comparative study of the different prediction models and the proposed hybrid model in accordance with varying indices of performance are shown in Table 4; the minimum error is observed in the case of the hybrid model with a MAPE value of 0.274%, followed by M-KELM, FA-MKELM, EMD-MKELM, and VDM-MKELM with a MAPE value of 3.134%, 2.47%, 1.65%, and 1.77% respectively, the MAPE value describes the statistical measurement of the prediction models accuracy, and the computational time needed to perform the test is 121.35 sec which is not much more than the previously implemented methods. The RMSE value describes the distance between the actual and the predicted value, and the correlation coefficient (R) shows the relationship between the actual and the predicted nearer to 1 means holding a suitable relation; we can observe from the table that the *R*-value of the proposed model is 0.99 which means a more significant relation between the actual and the predicted. The corresponding graphical representation is given in Figure 7. From the figure, it can be well observed that the proposed hybrid model performs better than various single decomposition-based models;
- A similar trend is followed for the NSE, SSE, and S&P 500 market data. The original data is shown in Figure 8(a), (b), and (c), respectively.

Figure 9 describes the two-layered decomposition of the NSE market where Figure 9(a) represents the EMD-based decomposition of the NSE data set, and Figure 9(b) represents the further decomposition of the higher frequency IMF (IMF1) using VMD obtained from the previous EMD-based technique.

Figure 10 shows the FA-based convergence graph of the NSE market, followed by a comparative study of

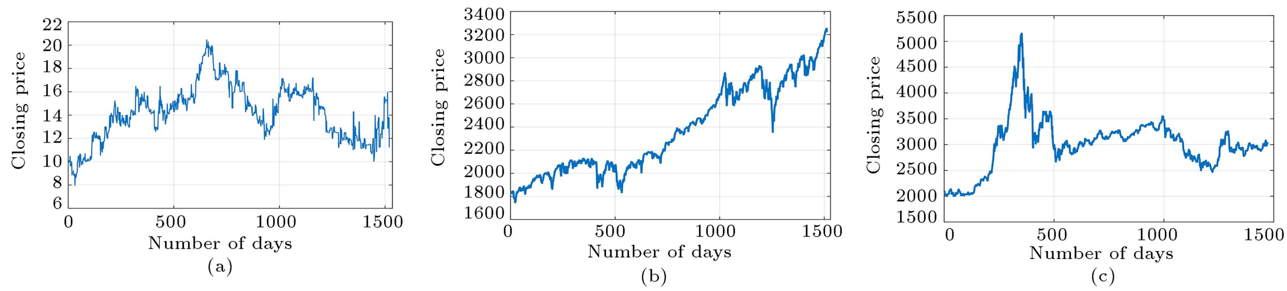

Figure 8. Original closing price data of different stock markets for a range of January 2014-January 2020; (a) NSE market closing price, (b) S&P 500 market closing price, and (c) SSE market closing price.

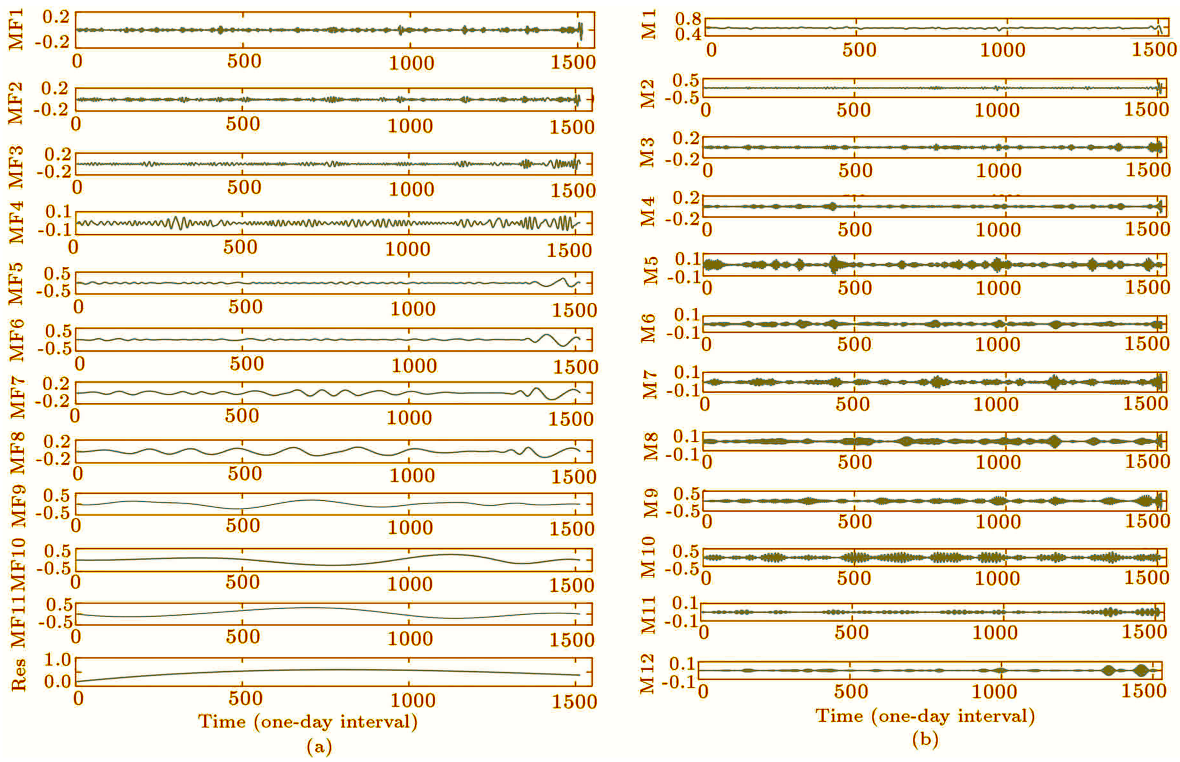

Figure 9. The two-layer decomposition based on EMD and VMD, (a) EMD-based decomposition of the NSE market and (b) VMD-based decomposition of IMF1 obtained using EMD for NES market.

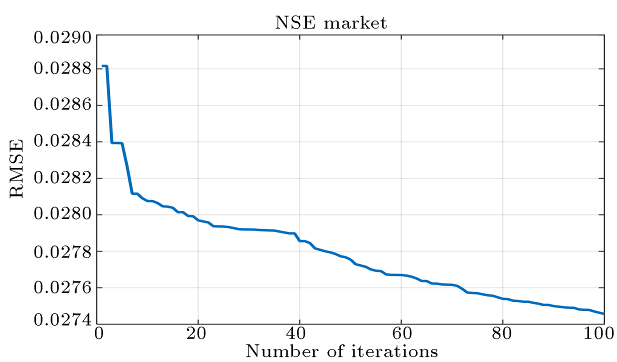

Figure 10. Convergence graph of NSE stock market data using FA.

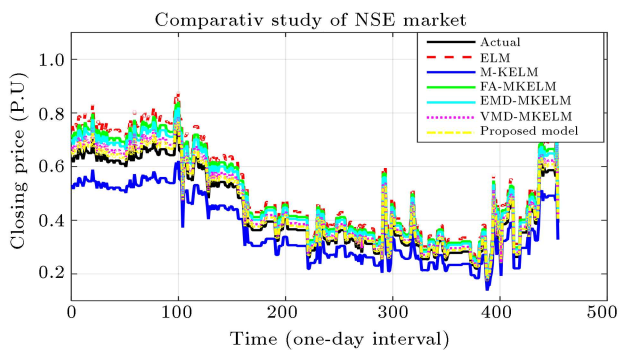

Figure 11. Comparative study of different one-day ahead prediction techniques.

different prediction models for one-day ahead prediction of the closing price of the NSE market in Figure 11. The corresponding error values are explained in Table 5, where it can be observed that the proposed model outperforms all other models. In this case, a total number of 1513 data is considered for one-day ahead

Table 5. Performance evaluation of EMD-VMD-FA-MKELM compared to other prediction models based on different values examined on the NSE stock market data.

Methods	MAPE (%)	MAE (P.U)	RMSE (P.U)	R	Ex (sec)
ELM	3.280	0.081	0.086	0.890	97.64
M-KELM	2.879	0.071	0.075	0.972	95.58
EMD-MKELM	1.897	0.047	0.049	0.986	109.44
VMD-MKELM	1.194	0.027	0.029	0.996	110.76
FA-MKELM	2.398	0.059	0.063	0.994	94.85
EMD-VMD-FA-MKELM	0.662	0.016	0.027	0.998	122.60

Note. Ex: Execution Time.

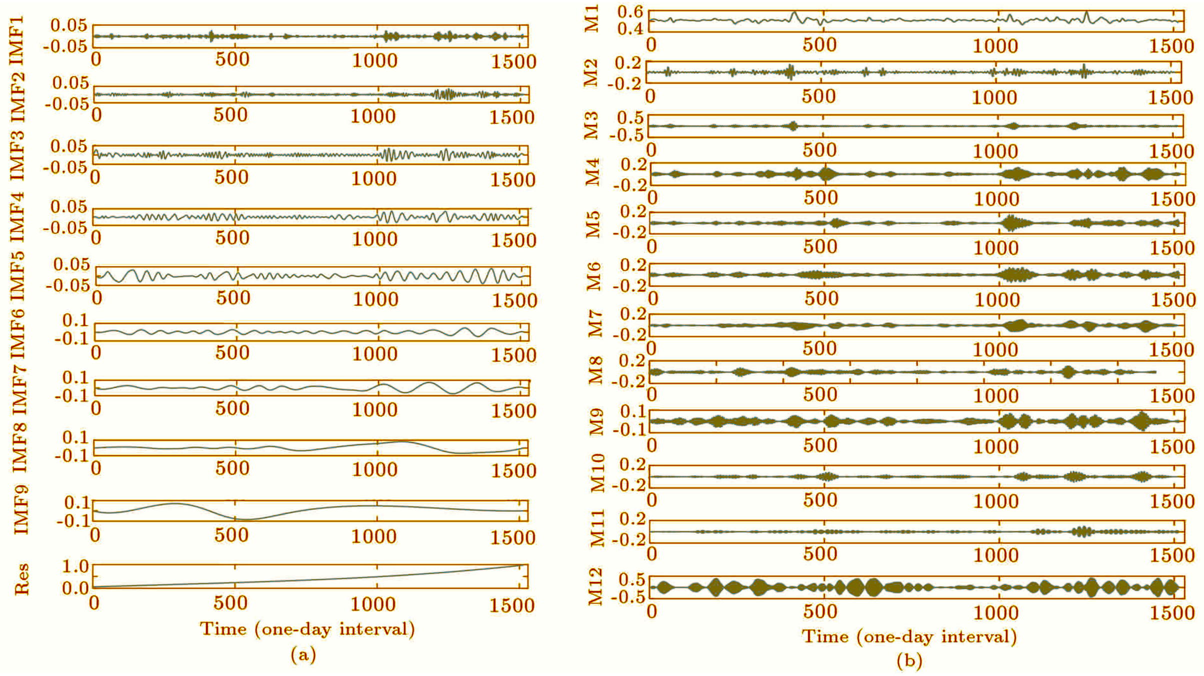

Figure 12. The two-layer decomposition based on EMD and VMD, (a) EMD-based decomposition of the S&P 500 market and (b) VMD-based decomposition of IMF1 obtained using EMD for the S&P 500 market.

price prediction, out of which 1059 data were considered for training and the rest 454 data were executed for testing, and the best MAPE value is observed to be 0.66% followed by the rest of the models.

The optimized values of the kernel parameters corresponding to the minimum cost function are as follows: $\sigma = 0.32$, $g = 0.66$, $\mu_1 = 1.53$, $e = 2.42$, and $f = 1.02$. Figure 12 describes the two-layered decomposition method, where Figure 12(a) shows the EMD-based decomposition of the whole S&P 500 market data set, and Figure 12(b) shows the VMD-based decomposition of the EMDs first IMF (IMF1).

Figure 13 shows the convergence graph of the S&P 500 data set where the optimized kernel parameters obtained are $\sigma = 0.55$, $g = 0.72$, $\mu_1 = 1.05$, $e = 2.11$, and $f = 1.52$. It consists of a total of 1513 data, from which 1059 data are treated for training, whereas the rest 454 data are treated for testing, and the MAPE value obtained using the proposed hybrid model is 0.461%.

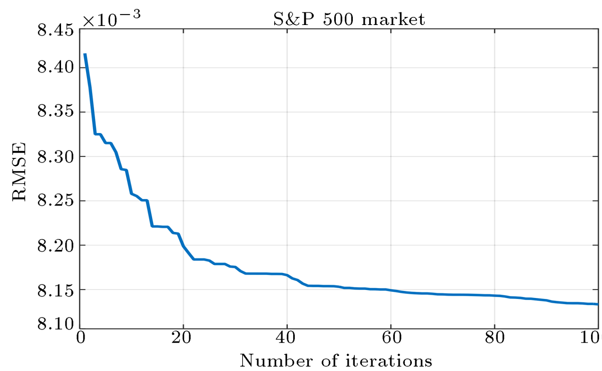

Figure 13. Convergence graph of S&P 500 stock market data using FA.

Figure 14 shows the comparison between different prediction models, and it is observed that the proposed model performs better, irrespective of the stock market. A detail of different error values is given in Table 6.

A similar trend is observed for the SSE market Figure 15 shows the EMD and VMD-based decompo-

Table 6. Performance evaluation of EMD-VMD-FA-MKELM compared to other prediction models based on different values examined on the S&P 500 stock market data.

Methods	MAPE (%)	MAE (P.U)	RMSE (P.U)	R	Ex (sec)
ELM	3.524	0.117	0.119	0.965	96.58
M-KELM	2.704	0.090	0.091	0.978	129.43
EMD-MKELM	1.797	0.059	0.061	0.991	135.62
VMD-MKELM	1.339	0.0447	0.045	0.994	137.26
FA-MKELM	2.322	0.117	0.119	0.983	98.42
EMD-VMD-FA-MKELM	0.461	0.017	0.008	0.998	140.33

Note. Ex: Execution Time.

Table 7. Performance evaluation of EMD-VMD-FA-MKELM compared to other prediction models based on different values examined on the SSE stock market data.

Methods	MAPE (%)	MAE (P.U)	RMSE (P.U)	R	Ex (sec)
ELM	3.574	0.049	0.050	0.978	94.35
M-KELM	3.340	0.045	0.047	0.976	95.79
EMD-MKELM	2.201	0.030	0.031	0.985	98.75
VMD-MKELM	1.608	0.022	0.024	0.996	102.63
FA-MKELM	2.439	0.033	0.034	0.986	96.39
EMD-VMD-FA-MKELM	0.437	0.007	0.006	0.998	110.20

Note. Ex: Execution Time.

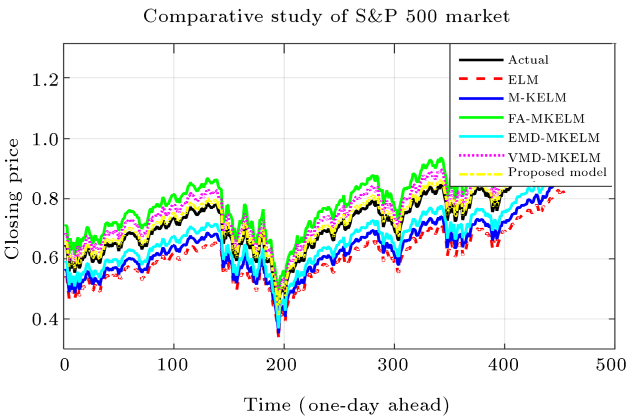

Figure 14. Comparative study of different one-day ahead prediction techniques.

sition of the SSE market for higher frequency IMF. Figure 16 shows the convergence graph, and the optimal values observed are listed as follows: $\sigma = 0.12$, $g = 0.74$, $\mu_1 = 1.22$, $e = 2.81$, and $f = 1.04$.

Figure 17 shows a comparative study of the different prediction models for predicting the closing price of the SSE market, and the corresponding values are shown in Table 7.

Figure 18 compares the performance of various prediction models at the different stock markets per MAPE value. All the result shows the superiority of the proposed technique. Table 8 shows a comparative analysis of different optimization techniques in combination with the multi-kernel ELM, thus stating the reason behind choosing the FA algorithm. It can be observed that the FA fits better with the MKELM compared to other techniques for the error minimization problem.

6. Conclusion and future scope

Stock market analysis and precise prediction are of utmost importance. In this paper, the prediction of four different stock markets is performed using the two-layer decomposition technique, as shown in the result analysis. BSE, NSE, Shanghai Stock Exchange (SSE), and S&P 500. The comparison study between Extreme Learning Machine (ELM), MKELM, FA-MKELMEMD-MKELM, VMD-MKELM, and the proposed hybrid model shows that the double decomposition performs better than the single decomposition and basic model. The hybrid model offers a Mean Absolute Percentage Error (MAPE) value of 0.274%, which is relatively low compared to 0.662%, 0.461% and 0.437% for BSE, NSE, S&P 500, and SSE stock

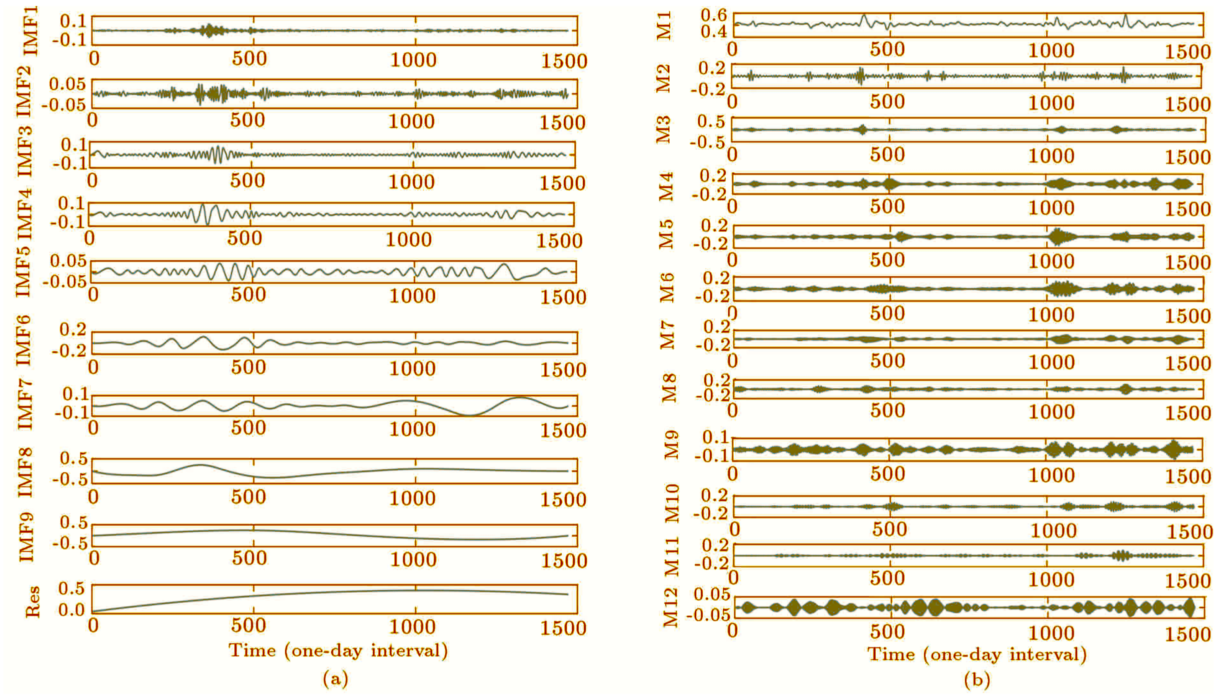

Figure 15. The two-layer decomposition based on EMD and VMD, (a) EMD-based decomposition of the SSE market, (b) VMD-based decomposition of IMF1 obtained using EMD for SSE market.

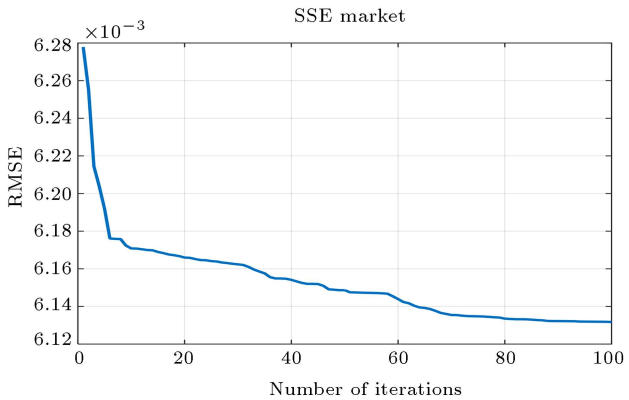

Figure 16. Convergence graph of SSE stock market data using FA.

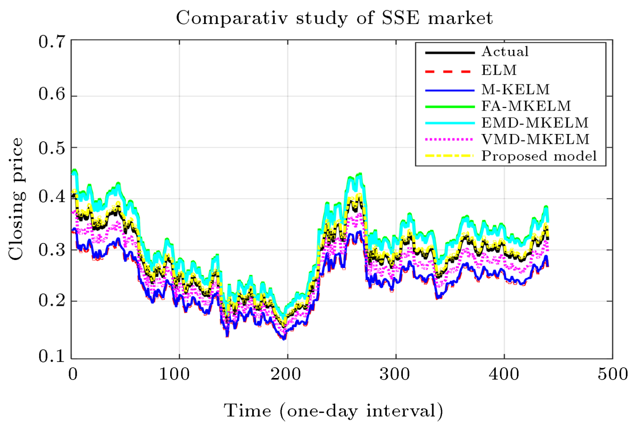

Figure 17. Comparative study of different one-day ahead prediction techniques.

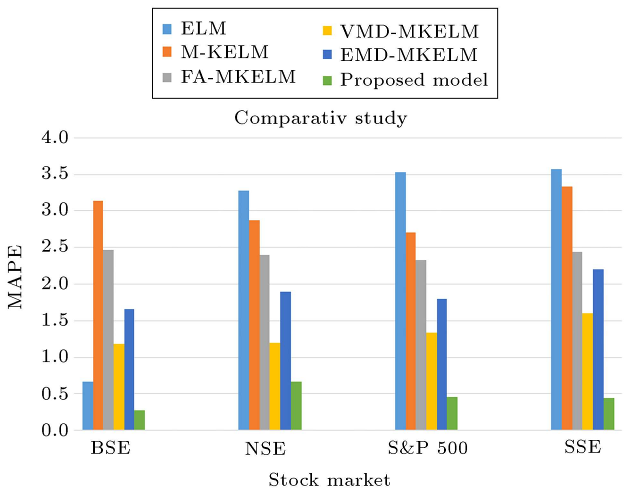

Figure 18. Comparative study of MAPE value for different prediction techniques.

markets, respectively, which is quite less than other models.

The primary drawback of the kernel-based ELM is the selection of different kernel parameters, which is further managed with the help of FA, which provides the optimized values of the kernel parameters, thereby providing a more accurate stock market prediction. Although there is an increase in execution time, the proposed model outperforms the basic single models. The statistical representation of the MAPE values in the different markets shows the uniformity and robustness of the proposed technique. In the future, this study can be extended further by providing input

Table 8. Performance evaluation of different optimization techniques in combination with M-KELM implemented at the different stock market.

Stock market	Optimization technique	MAPE (%)	No. of iteration	Computational time (sec)
BSE	PSO	2.735	100	112.02
	GA	2.721	100	111.34
	FFA	2.687	100	99.54
	FA	2.471	100	105.75
NSE	PSO	2.724	100	111.47
	GA	2.711	100	100.01
	FFA	2.547	100	89.33
	FA	2.398	100	94.85
S&P 500	PSO	2.873	100	111.42
	GA	2.809	100	104.26
	FFA	2.465	100	93.76
	FA	2.322	100	98.42
SSE	PSO	2.723	100	108.22
	GA	2.653	100	102.54
	FFA	2.392	100	91.85
	FA	2.439	100	96.38

variables at different market influencing conditions and performing in an online performance space.

Nomenclature

EMD	Empirical Mode Decomposition
VMD	Variational Mode Decomposition
FA	Firefly Algorithm
IMF	Intrinsic Mode Function
KELM	Kernel Extreme Learning Machine
SVM	Support Vector Machine
BPNN	Back Propagation Neural Network
GA	Genetic Algorithm
ANN	Artificial Neural Network
ELM	Extreme Learning Machine
PSO	Particle Swarm Optimization
GKELM	Gaussian Kernel Extreme Learning Machine
PKELM	Polynomial Kernel Extreme Learning Machine
SKELM	Sigmoidal Kernel Extreme Learning Machine
WKELM	Wavelet Kernel Extreme Learning Machine
M-KELM	Multi Kernel Extreme Learning Machine

References

1. Poon, S.H. and Granger, C.W. "Forecasting volatility in financial markets: A review", *Journal of Economic Literature*, **41**(2), pp. 478–539 (2003).
2. Wang, Y.F. "Predicting stock price using fuzzy grey prediction system", *Expert Systems with Applications*, **22**(1), pp. 33–38 (2002).
3. Awartani, B.M. and Corradi, V. "Predicting the volatility of the S&P-500 stock index via GARCH models: the role of asymmetries", *International Journal of Forecasting*, **21**(1), pp. 167–183 (2005).
4. Bollerslev, T. "Generalized autoregressive conditional heteroskedasticity", *Journal of Econometrics*, **31**(3), pp. 307–327 (1986).
5. Rojas, I., Valenzuela, O., Rojas, F., et al. "Soft-computing techniques and ARMA model for time series prediction", *Neurocomputing*, **71**(4–6), pp. 519–537 (2008).
6. Trippi, R.R. and Turban, E., *Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real World Performance*, McGraw-Hill, Inc (1992).
7. Weng, B., Ahmed, M.A., and Megahed, F.M. "Stock market one-day ahead movement prediction using disparate data sources", *Expert Systems with Applications*, **79**, pp. 153–163 (2017).
8. Guresen, E., Kayakutlu, G., and Daim, T.U. "Using artificial neural network models in stock market index

prediction", *Expert Systems with Applications*, **38**(8), pp. 10389–10397 (2011).

9. Sahin, S., Tolun, M.R., and Hassanpour, R. "Hybrid expert systems: A survey of current approaches and applications", *Expert Systems with Applications*, **39**(4), pp. 4609–4617 (2012).

10. Li, X., Xie, H., Wang, R., et al. "Empirical analysis: stock market prediction via extreme learning machine", *Neural Computing and Applications*, **27**(1), pp. 67–78 (2016).

11. Grigorievskiy, A., Miche, Y., Ventelä, A.M., et al. "Long-term time series prediction using OP-ELM", *Neural Networks*, **51**, pp. 50–56 (2014).

12. Majumder, I., Dash, P.K., and Bisoi, R. "Variational mode decomposition based low rank robust kernel extreme learning machine for solar irradiation forecasting", *Energy Conversion and Management*, **171**, pp. 787–806 (2018).

13. Mohanty, D.K., Parida, A.K., and Khuntia, S.S. "Financial market prediction under deep learning framework using auto encoder and kernel extreme learning machine", *Applied Soft Computing*, **99**, 106898 (2021).

14. Göçken, M., Özçalıcı, M., Boru, A., et al. "Integrating metaheuristics and artificial neural networks for improved stock price prediction", *Expert Systems with Applications*, **44**, pp. 320–331 (2016).

15. Majhi, R., Panda, G., Majhi, B., et al. "Efficient prediction of stock market indices using adaptive bacterial foraging optimization (ABFO) and BFO based techniques", *Expert Systems with Applications*, **36**(6), pp. 10097–10104 (2009).

16. Zhang, Y. and Wu, L. "Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network", *Expert Systems with Applications*, **36**(5), pp. 8849–8854 (2009).

17. Chou, J.S. and Nguyen, T.K. "Forward forecast of stock price using sliding-window metaheuristic-optimized machine-learning regression", *IEEE Transactions on Industrial Informatics*, **14**(7), pp. 3132–3142 (2018).

18. Liu, H., Mi, X., and Li, Y. "Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM", *Energy Conversion and Management*, **159**, pp. 54–64 (2018).

19. Zhou, F., Zhou, H.M., Yang, Z., et al. "EMD2FNN: A strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction", *Expert Systems with Applications*, **115**, pp. 136–151 (2019).

20. Yaslan, Y. and Bican, B. "Empirical mode decomposition based denoising method with support vector regression for time series prediction: A case study for electricity load forecasting", *Measurement*, **103**, pp. 52–61 (2017).

21. Guo, Z., Zhao, W., Lu, H., et al. "Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model", *Renewable Energy*, **37**(1), pp. 241–249 (2012).

22. Kao, L.J., Chiu, C.C., Lu, C.J., et al. "A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting", *Decision Support Systems*, **54**(3), pp. 1228–1244 (2013).

23. Hong, L. "Decomposition and forecast for financial time series with high-frequency based on empirical mode decomposition", *Energy Procedia*, **5**, pp. 1333–1340 (2011).

24. Jammazi, R. and Aloui, C. "Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling", *Energy Economics*, **34**(3), pp. 828–841 (2012).

25. Wei, L.Y. "A hybrid ANFIS model based on empirical mode decomposition for stock time series forecasting", *Applied Soft Computing*, **42**, pp. 368–376 (2016).

26. Chen, C.F., Lai, M.C., and Yeh, C.C. "Forecasting tourism demand based on empirical mode decomposition and neural network", *Knowledge-Based Systems*, **26**, pp. 281–287 (2012).

27. Dragomiretskiy, K. and Zosso, D. "Variational mode decomposition", *IEEE Transactions on Signal Processing*, **62**(3), pp. 531–544 (2013).

28. Zhang, Y., Liu, K., Qin, L., et al. "Deterministic and probabilistic interval prediction for short-term wind power generation based on variational mode decomposition and machine learning methods", *Energy Conversion and Management*, **112**, pp. 208–219 (2016).

29. Senthilnath, J., Omkar, S.N., and Mani, V. "Clustering using firefly algorithm: performance study", *Swarm and Evolutionary Computation,* **1**(3), pp. 164–171 (2011).

30. Tuba, M. and Bacanin, N. "April. JPEG quantization tables selection by the firefly algorithm", In *2014 International Conference on Multimedia Computing and Systems (ICMCS)*, pp. 153–158, IEEE (2014).

31. Tuba, M. and Bacanin, N. "Improved seeker optimization algorithm hybridized with firefly algorithm for constrained optimization problems", *Neuro Computing*, **143**, pp. 197–207 (2014).

32. https://in.finance.yahoo.com/

33. Chang, P.C. and Liu, C.H. "A TSK type fuzzy rule based system for stock price rediction", *Expert Systems with Applications*, **34**(1), pp. 135–144 (2008).

34. Samanta, S.R., Mallick, P.K., Pattnaik, P.K., Mohanty, J.R., and Polkowski, Z. eds., *Cognitive Computing for Risk Management*, Springer (2022).

35. Mukherjee, A., Singh, A.K., Mallick, P.K., et al., *Portfolio Optimization for US-Based Equity Instruments Using Monte-Carlo Simulation*, In: Mallick, P.K., Bhoi, A.K., Barsocchi, P., de Albuquerque, V.H.C.

(Eds.) Cognitive Informatics and Soft Computing. Lecture Notes in Networks and Systems, **375**, Springer, Singapore (2022). https://doi.org/10.1007/978-981-16-8763-1_57

Biographies

Pradeep Kumar Mallick is currently working as a Senior Associate Professor in the School of Computer Engineering, Kalinga Institute of Industrial Technology (KIIT) Deemed to be University, Odisha, India. He has also served as Professor and Head Department of Computer Science and Engineering at Vignana Bharathi Institute of Technology, Hyderabad. He has completed his Post Doctoral Fellow (PDF) from Kongju National University South Korea, PhD from Siksha Ó Anusandhan University, M. Tech. (CSE) from Biju Patnaik University of Technology (BPUT) and MCA from Fakir Mohan University Balasore, India. Besides academics, he is also involved in various administrative activities, and he is a Member of the Board of Studies at C.V. Ramman Global University Bhubaneswar, a Member of the Doctoral Research Evaluation Committee, Admission Committee, etc. His area of research includes Data Mining, Image Processing, Soft Computing, and Machine Learning. Now he is the editorial member of the Korean Convergence Society for SMB. He has published 22 edited books, one textbook, and more than 115 research papers in national and international journals and conference proceedings in his credit. He also serves as guest editor for special issues of journals like Springer, Nature, and Inderscience.

Amiya Ranjan Panda has seven years of research experience in DRDO and one year of teaching experience. He received the B.Tech. degree in information technology from Biju Patnaik University of Technology, Rourkela, India, in 2009, and the M.Tech. degree in computer science and engineering from the Kalinga Institute of Industrial Technology (KIIT), Bhubaneshwar, India, in 2012. He received a PhD degree from Siksha "O" Anusandhan University, Bhubaneshwar, in the year 2017, working on a real-time project "Design, development and implementation of Software Defined Radio based Flight Termination System" of DRDO, ITR, Chandipur. He has worked as JRF, SRF, and RA in DRDO, ITR, and Chandipur for over ten years. Then he worked as Assistant Professor at Siksha 'O' Anusandhan University for four months, and currently, he is working as Assistant Professor at KIIT Deemed to be University, Bhubaneswar. His research interest is Machine learning, IOT, data acquisition system, and Software Defined Radio. He has published in 15 international journals. Two of the journals are published in IEEE Transactions of Industrial Informatics. One edited book has been accepted for publication in 2019. He is an expert in teaching Data Structure and algorithm, Database Management System (DBMS), Artificial Intelligence, Operating System, Compiler Design, Computer Networks, Machine Learning, etc.

Ajaya Kumar Parida is currently working as an Associate Professor in the School of Computer Engineering, Kalinga Institute of Industrial Technology (KIIT), Deemed to be University, Bhubaneswar, India. He received his B.E degree from KIIT University, India, and his MBA in Systems and Marketing from KIIT University. He did his MTECH from CET, Bhubaneswar, India, and completed his PhD in Computer Science and Engineering from Utkal University, Bhubaneswar, India. He has published over 20 research papers in various international journals and conferences. He is also associated with multiple educational and research societies. He is a state-level jury member of the National Children's Science Congress. His research interest includes the field of Data Mining, Soft Computing, and Machine Learning. He also has more than 15 years of teaching and research experience in several engineering colleges and universities.

Mohit Ranjan Panda is currently working as an Assistant Professor in the School of Computer Engineering, Kalinga Institute of Industrial Technology (KIIT), Deemed to be University, Bhubaneswar, India. He completed his PhD in 2019 in the area "Mobile robot path planning under cluttered environment", M. Tech. in 2009, and B. Tech in 2001 in the branch of Computer Engineering. He has 17 years of teaching experience with both graduate and postgraduate students. He has published many research articles in various international journals and conferences.

Sasmita Rani Samanta is an accomplished academician and academic administrator who is currently the Professor of Management and Vice Chancellor of KIIT Deemed to be University. Her important academic credentials include distinctions such as a Post-Doc in Leadership, Taiwan; PhD in Management; a Stanford LEAD Alumna and Stanford Distinguished Scholar, Stanford Graduate School of Business, Stanford University, USA; Fellow of the Royal Society of Arts (FRSA, UK); and Fellow of Computer Society of India (FCSI, India). She has been trained in leadership at the University of Nebraska Omaha (UNO), USA, and CSC leaders program, Common Purpose Charitable Trust, UK. Professor has published over 60 research articles in peer-reviewed Journals of International repute and Conferences. Dr Samanta has filed ten national and international patents, of which two patents have been granted (international). Dr Samanta is the author of four textbooks and two reference books published by reputed international publishers like University Science Press, Springer, Elsevier, and IGI Global. She is now Chairman of KIIT-TBI, Chairman Center of Excellence for Sustainability and Equity (since 2019), Chairperson Center for Excellence for Public Policy Research (since 2021), Director Center of Excellence on Occupational Safety and Health, President of the World Leadership Academy (since 2020), President, Indian Commerce Association, and Chairperson, Sustainable Outreach Universal Leadership (SOUL) Pvt. Ltd (since 2021). Dr Samanta is actively engaged in research in the area of leadership, management, spiritualism, education, and human development. She has already supervised 25 master and 6 PhD theses.