Journal of Statistics Applications & Probability
*An International Journal*

http://dx.doi.org/10.18576/jsap/150314

# Leveraging Deep Learning with Alpine Skiing Optimizer in Financial Performance Forecasting and Financial Statement Analysis: A Case Study from Saudi Arabia

*Manal M. Khayyat [1,*], Babikir Mubarak O. Elsheikh [2], Ahmed Edris Abdu [2], Buthayna G. E. Tamim[3], Alhan H. A. Awadallah [2], Abdelgalal O. I. Abaker [4], and Nadia Bushra Mohammed Ali [5]*

[1] Department of Computer Science and Artificial Intelligence, College of Computing, Umm Al-Qura University, Makkah, Saudi Arabia
[2] Department of Business Administration, College of Business Administration, Najran University, Najran 66462, Saudi Arabia
[3] Department of Marketing and E-Commerce, College of Business Administration, Najran University, Najran 66462, Saudi Arabia
[4] Department of Administrative Sciences, Applied College in Khamis Mushait, King Khalid University, Abha, Saudi Arabia
[5] Department of Finance, Faculty of Business, Imam Mohammad Ibn Saud Islamic university, Saudi Arabia

Received: 20 Dec. 2025, Revised: 25 Feb. 2026, Accepted: 9 April. 2026
Published online: 1 May 2026

**Abstract:** Financial predicting depends upon the usage of present and past financial data to create the best forecast of the future economic condition, evade higher risk states, and upsurge benefits. Such forecasts are main to anyone who needs to get the condition of possible finances in the future, with decision-makers and investors. However, the intricate nature of financial data makes it challenging to acquire precise forecasts. Artificial intelligence (AI), which was exposed to be appropriate for analyzing very intricate issues, can be employed for financial prediction. Currently, Machine Learning (ML) scholars have come up with numerous methods and a massive amount of studies have been published consequently. As such, a major quantity of surveys exist that cover ML for financial predicting studies. In recent times, Deep Learning (DL) systems started seeming in the field, with outcomes that significantly outperform conventional ML counterparts. This study develops a Deep Learning in Financial Performance Forecasting and Financial Statement Analysis (DLFPF-FSA) model. The DLFPF-FSA model relies on improving the prediction of financial performance using state-of-the-art optimization algorithms. To accomplish that, the data normalization stage is initially applied by z-score normalization to standardize data by scaling it to a common range. Next, the feature selection process has been executed by the bald eagle search optimization (BESO) algorithm to identify the most relevant features from input data. For the prediction of financial performance, the DLFPF-FSA system designs a hybrid of long short-term memory and gated recurrent unit (LSTM+GRU) method. Eventually, the alpine skiing optimizer (ASO) adjusts the hyperparameter values of the LSTM+GRU algorithm optimally and outcomes in greater prediction performance. The experimental evaluation of the DLFPF-FSA algorithm can be tested on a benchmark dataset. The extensive outcomes highlight the significant solution of the DLFPF-FSA approach to the financial prediction process

**Keywords:** Deep Learning; Financial Performance Forecasting; Financial Statement Analysis; Feature Selection; Hyperparameter Tuning

## 1 Introduction

Accounting, market information, and basic aggregated financial data are the essential inputs required for planning and financial analysis; regression analysis, statistical methods, computer programming knowledge, and operation research programming models are significant devices for attaining prediction and financial planning [1] While implementing planning and financial analysis, it is significant to recognize how to utilize the proper devices to examine the significant data [2]. The predicting nature revolves around upcoming expectations. Economic predictions are inputs into the process of financial planning, subsequently existing sources might be employed effectively to attain corporate entity [3]. Financial decisions are becoming more complicated, with a broad array of alternate resources and usages of funds, and an unstable prominence away from security to higher growth and profit performance. Precise financial prediction, aiming

* Corresponding author e-mail: mmkhayat@uqu.edu.sa

at both internal financial variables and economic environment has become a significant input into the process of decision-making [4]

The financial forecasting method is a multi-faceted function that examines possible portfolio decisions through a few planning horizons [5]. This function generally necessitates the prediction of the upcoming external financial environment facing internal financial variables. An existing financial position is the consequence of previous decisions to gain funds from other resources and advance these funds in alternate investment possibilities like bonds and loans. An existing decision on investing and acquiring funds will distress financial position [6]. These decisions are the consequence of economic planning based on the current financial position and the predicted external setting, with the possibility of meeting economic performance principles in the structure of management objectives [7]. Primarily, the predictive function is involved in the financial variables related to the upcoming external setting that is either the regional or national levels.

In recent times, Artificial Intelligence (AI) had a significant effect on the global business setting and has established applications in multiple domains [8]. Economic data is stimulating to examine because it takes a lot of uncertainty. AI might be utilized to classify economic data for study, permitting the team to analyze and screen data more rapidly and to assist them create more accurate decisions, substantially decreasing human error, bringing better returns to customers, and enabling risk control. Numerous surveys are published depending on Machine Learning (ML) methods with moderately enhanced performances compared with conventional time series prediction models [9]. Therefore, new implementations and publications keep pouring into computational and financial intelligence works. In recent years, Deep Learning (DL) has developed strongly as the finest-performing predictor class in the area of ML in several implementation domains [10]. Economic time series prediction is no exception like multiple prediction models depend on many DL approaches.

This study develops a Deep Learning in Financial Performance Forecasting and Financial Statement Analysis (DLFPF-FSA) model. To accomplish that, the data normalization stage is initially applied by z-score normalization to standardize data by scaling it to a common range. Next, the feature selection process has been executed by the bald eagle search optimization (BESO) algorithm to identify the most relevant features. For the prediction of financial performance, the DLFPF-FSA system designs hybrid of long short-term memory and gated recurrent unit (LSTM+GRU) method. Eventually, the alpine skiing optimizer (ASO) adjusts the hyperparameter values of the LSTM+GRU algorithm optimally and outcomes in greater prediction performance. The experimental evaluation of the DLFPF-FSA algorithm can be tested on a benchmark dataset.

## 2 Related Works

Ben Mrad et al. [11]projected a solution to the challenge of predicting many possible resultant variables for stock market development utilizing Bayesian Networks. This model progresses 3 approaches depending on Bayesian networks and examines their performance employing 7 principles: Model Complexity, Ease of Deployment, Extensibility, Generality or Flexibility, Model Building, ease of Development, and Interpretability. Shi et al. [12] emphasize the Hybrid Financial Risk Predictor (HFRP) method, by applying the LSTM and CNN systems to enhance economic risk prediction. Integration of qualitative and quantitative ratings originated from the study of economic textual outcomes in higher stability and accuracy compared to the HFRP model. In [13], the one-dimensional CNN (1D-CNN) and LSTM techniques are formed by examining the possibility of predicting corporate economic performance with DL techniques, utilizing the corporate economic aspects and surroundings. 5 evaluation metrics were utilized to evaluate predicting effects, and 4 competing ML approaches were constructed to verify the development to forecast precision made by the DL methods. Additionally, the Accumulated Local Effects model is also presented to discover the predicting processes of the DL methodologies.

In [14], the DL approach in economic market prediction that comprises BRICS economies as the test case is advanced. Financial markets are rife with instability which is influenced by a bed of complexity, coddled by distal and local features. To utilize these massive databases DL methods like LSTM, CNN together with hybrid structures are utilized in this paper. Particularly, DL models are implemented in case studies in the BRICS to emphasize the application of DL to different country-specified concerns like market shocks and liquidity crises. Srivastava et al. [15] developed a hybrid method by combining association rule mining with DL approaches. The recognized associated companies describe the data to the neural methods. The existing model applies a multi-variate LSTM neural structure for the stock price prediction challenge. In [16], a stock market index prediction method named SenT-In incorporated with a sentiment awareness method is presented. A sentiment awareness method utilizing GRU and CNN is projected to evaluate the sentiment index of a huge amount of news articles gathered from reputable financial websites. Moreover, a sentiment attention model is projected to integrate stock data and news sentiment index as the input for predicting and training utilizing the SenT-In system that is either simple or efficient.

Che et al. [17] developed a regularized and attentive DL model for forecasting economic distress utilizing multi-modal data comprising economic indicators, interfirm networks, and existing reports. Particularly, deliberating inside heterogeneity and modalities, 3 modality-explicit attentions, that is report-, neighbour-, and ratio-aware attentions are intended for adaptably removing essential information from current reports, interfirm networks, and financial indicators correspondingly. Knifo and Alzubi [18] projected a dual-recurrent neural network with a tri-channel attention mechanism (DR-Z2AN) for precise prediction. The projected method incorporates the multi-head attention and tri-channel attention mechanism with dual-RNN that improves the interpretability and robustness of methods. The multi-head attention learns the complicated relations among data that progresses the generalization ability of the model in prediction challenges.

## 3 Materials and Methods

This study develops a DLFPF-FSA approach. The DLFPF-FSA model relies on improving the prediction of financial performance using state-of-the-art optimization algorithms. To achieve this, the proposed DLFPF-FSA model involves various levels such as data normalization, feature selection, prediction, and hyperparameter tuning model. Fig. 1 represents the overall working process of DLFPF-FSA model.

### *3.1 Data Normalization*

Initially, the data normalization stage is applied by $z-corse$ normalization to standardize data by scaling it to a common range. One significant model for assuring relevance and consistency is $z-corse$ normalization [19]. With its help, every feature is regularized and contains a standard deviation $SD$ of 1 and mean($\mu$) of 0. It functions particularly well with dissimilar kinds of data patterns and changing cloud source customs. For every feature($\sigma'$), this normalization has been formulated utilizing Eq. (1):

$$Z_{normalize} = \frac{(X_f - \mu)}{\sigma_r} \tag{1}$$

Where a value of original data is signified by X_f, the value of standardized is signified by $Z_{normalize}$, $\mu$ is denoted as mean, $\sigma^{'}$ refers to standard deviation. By certifying that the data recollects its relative position in the distribution, this alteration generates it appropriate for contrast.

### *3.2 Feature Selection Process*

Next, the FS process is executed by the BESO algorithm to recognize the most relevant features from the input data. BESO is a new add-on to the area of metaheuristic techniques [20]. It appeals to stimulation from the bald eagle's behavior of hunting, exactly their skill in finding and capturing prey. In the optimizer procedure, BESO employs numerous hunt strategies demonstrated in these behaviors of predatory. These tactics permit the searching agents to discover the searching space efficiently, eventually foremost to the classification of optimum solutions for intricate real-time issues.

Step1: Pretending Exploration: Picking the Search Region

The BESO technique mimics this stage of a bald eagle's search over its initial phase. The technique defines an appropriate exploration region for a potential solution, similar to how a bald eagle may spy on a massive area for prey. Its mathematical formulation is given below:

$$P_{new}^{\,t} = P_{Best}^{\,t} + ar\left(P_{mean}^{\,t} - P_i^{\,t}\right). \tag{2}$$

A control parameter ($a$) is fixed among 1.5 and 2, which directs the size of the hunt area, whereas a randomly generated number ($r$) amongst 0 and 1, averting stagnation in local goals. This exploration is additionally induced by dual important factors. The technique reflects the best solution discovered ($P_{Besi}^{\,t}$) for directing its search near promising areas. Moreover, the present eagle population in an average location ($P_{mean}^{\,t}$) certifies exploration beyond the instant district of the finest solution and explores novel regions.

Step2: Increasing the Hunt

This phase strengthens the hunt for optimum solutions in the exploration region. In this, the searching agents exactly scan the distinct area. This attentive hunt pattern permits the technique to effectively discover the chosen region and recognize probable solutions with larger accuracy. The mathematical formulation is given below:

$$P_{new}^{\,t} = P_i^{\,t} + y_i X(P_i^{\,t} - P_{i+1}^{\,t}) + x_i X(P_i^{\,t} - P_{mean}^{\,t}), \tag{3}$$

where

$x_i = \frac{xr_i}{\max|xr|}$, $y_i = \frac{yr_i}{\max|yr|}$,

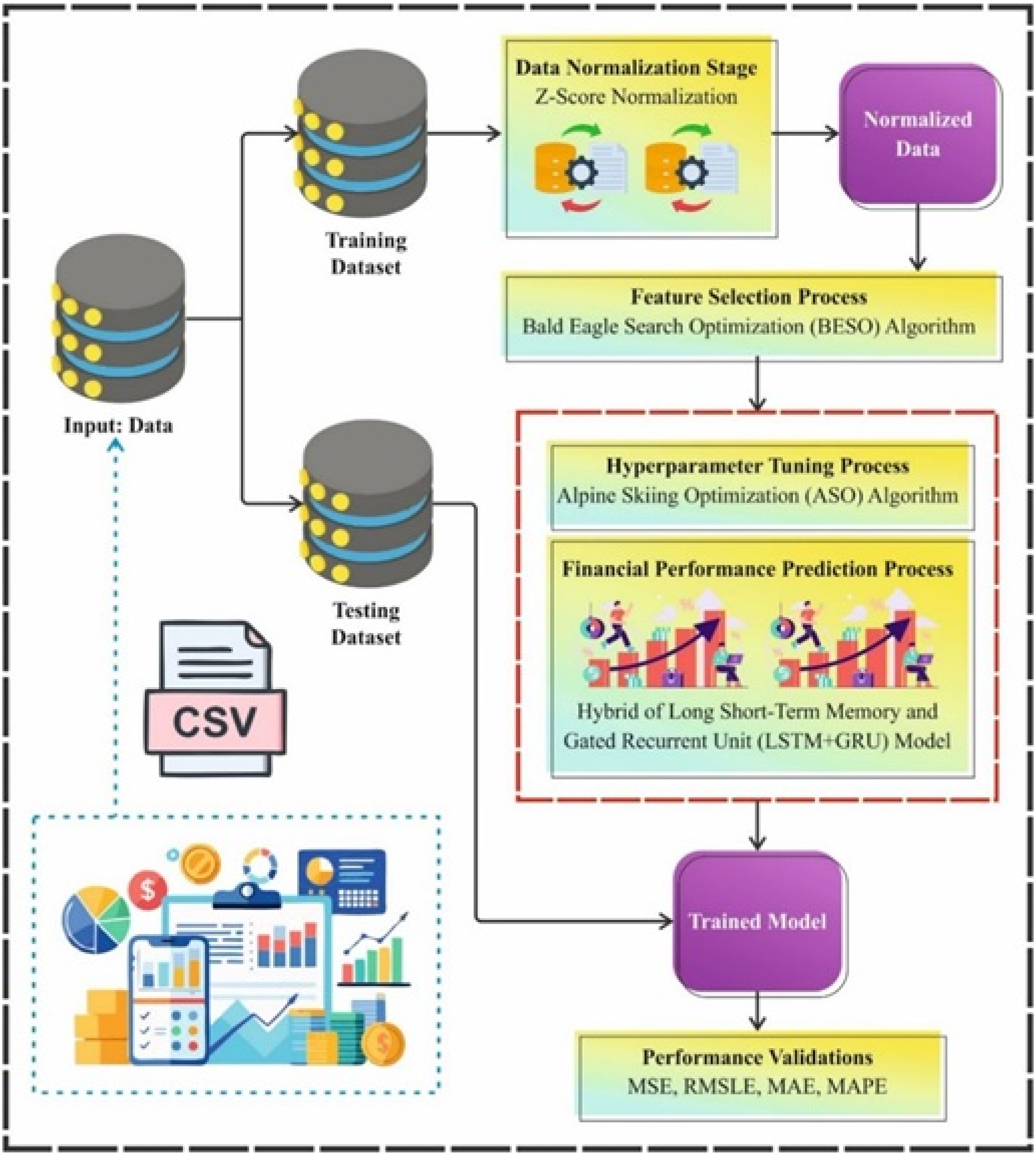


**Fig. 1:** Working process of DLFPF-FSA methodology

$$xr_i = r_i X \sin(\theta_i), \quad yr_i = r_i X \cos(\theta_i),$$
$$\theta_i = a\pi \times \text{rand}, \quad a \in (5,10),$$
$$r_i = \theta_i + R + rand \ , \ R \in [0.5, 2].$$

The dual randomly generated parameters are signified as $R'$ and $a'$, which control the pattern in the selected search region. While they influence the amount of spirals taken by an agent and the variant in their spiraling form. This component of randomness aids in preventing the model from stuck into a sub-optimal solution. By presenting variants in the pattern of search, BESO inspires the search of dissimilar locations in the search area, finally leading to the classification of precise solutions.

Step3: Convergence

This phase signifies the convergence near an optimum solution. Here, the search agents incline near the most encouraging location recognized. The computation formulation of this procedure is inscribed below:

$$P_{new}^{t} = \text{rand} X P_{Best}^{t} + x1_i X(P_i^{\,t} - C_1 X P_{mean}^{\,t}) + y1_j X(P_i^{\,t} - C_2 X P_{Best}^{t}),$$

$x1_i \; = \frac{rx_i}{\max|xr|}$ , $y1_j \; = \frac{yr_j}{\max|yr|}$,

$rx_i \; = r_i X \sinh(\theta_i), \;\; yr_i = r_i X \cosh(\theta_i),$

$r_i = \theta_i, \;\; \theta_i \; = X a\pi X \text{ rand}, \;\; a \in (5, 10).$

In this stage, the dual randomly generated parameters ($C_1$ and $C_2$) play a vital part. While, they generally range among 1 and 2, which controls the intensity of the searching agents' drive near the finest solutions.

In the BESO approach, the fitness function (FF) is intended to have a balance among the amount of chosen features in every solution (least) and the classifier accuracy (largest) attained by employing these preferred features, Eq. (5) embodies the FF to assess a solution.

$$Fitness = \alpha\gamma_R(D) + \beta\frac{|R|}{|C|}, \tag{4}$$

while $\gamma_R(D)$ signifies the classification rate of error of a set classifier. $|R|$ refers to the cardinality of the chosen sub-set and $|C|$ denotes the total amount of features in the data, $\alpha$ and $\beta$ are dual parameters that reflect the significance of classifier quality and sub-set length $\in [1,0]$ and $\beta \; = \; 1-\alpha$.

### *3.3 Hybrid Prediction Process*

For the prediction of financial performance, the proposed DLFPF-FSA system designs a hybrid of the LSTM+GRU method. For predicting the LSTM+GRU method is used, in this stage. Let $X_{top} \; \in \; \mathbb{R}^{(n\times k)}$ characterize the decreased feature set gained from Phase 2, while $n1Stn_{\varepsilon}$ sample counts and k denotes the number of where $\Theta^{(0)}$ refers to the primary location of the ith earthworm, preferred attributes [21].

LSTM Layer: The LSTM layer handles the input sequence $X_{top}$ by calculating the hidden layer (HL) $h_t^{RSTM}$ and cell state $c_t^{LSTM}$ at every time step $t$:

$$i_t = \sigma(W_i x_t + U_i h_{(t-1)}^{LSTM} + b_i), \tag{5}$$

$$f_t = \sigma(W_f x_t + U_f h_{(t-1)}^{LSTM} + b_f), \tag{6}$$

$$o_t = \sigma(U_o x_t + U_o h_{(t-1)}^{LSTM} + bo), \tag{7}$$

$$c_{(t-1)}^{LSTM} = f_t \odot c_{(t-1)}^{LSTM} + \; i_t \; \odot \tanh(\; W_c x_t + U_c h_{(t-1)}^{LSTM} + b_c), \tag{8}$$

$$h_{(t-1)}^{LSTM} = o_t \; \odot \tanh(\; c_t^{LSTM}), \tag{9}$$

whereas, $x_t$ denotes input at time-step $t$. $b_j$,$b_f$,$b_o$, and $b_c$ represent biased vectors. $W_j$,$W_f$,$W_o$,$W_c$,$U_i$,$U_f$,$U_o$,$U_c$ denote weighted matrices. $f_t$,$i_t$,$o_t$ represents forget, input, and output gates, correspondingly. $\sigma$ symbolizes the function of sigmoid activation, and Error::$0x$0000 signifies element-wise multiplication.

GRU Layer: The resultant from the LSTM layer $h_t^{LSTM}$ is then given to the GRU layer that calculates the HL $h_t^{GRU}$ at every time step $t$:

$$z_t = \sigma(W_z h_t^{LSTM} + U_Z h_{(t-1)}^{GRU} + b_z), \tag{10}$$

$$r_t = \sigma(W_r h_t^{LSTM} + U_r h_{(t-1)}^{GRU} + b_r), \tag{11}$$

$$\tilde{h}_t^{GRU} = \tanh(W_h h_t^{LSTM} + U_h(r_t \odot h_{(t-1)}^{GRU} + b_h), \tag{12}$$

$$h_t^{GRU} = 1(-z_t) \odot h_{(t-1)}^{GRU} + z_t \odot \tilde{h}_t^{GRU}), \tag{13}$$

whereas $z_t$ denotes the update gate, and $r_t$ represents the reset gate. • $\tilde{h}_t^{GRU}$ refers to candidate HL. $W_Z$,$W_r$,$W_h$,$U_z$,$U_r$, and $U_h$, represent weighted matrices. $b_z$,$b_r$, and $b_h$ symbolize biased vectors.

Prediction: The last HL from the layer of GRU $h_t^{GRU}$, while processing the complete series, is penetrated an FC layer using softmax activation for classification:

$$\hat{y} = softmax\ (W_{outh}\ h_t^{GRU} + b_{out}), \tag{14}$$

whereas, $b_{out}$ refers to the output layer's bias vector. $b_{out}$ stands for the output layer's weight matrix. $\hat{y}$ represents predicted probability distribution under the classes.

Model Training: The model is trained by reducing the cross-entropy loss function L utilizing the best hyperparameters $\Theta^*$ gained from the EOA:

$$L(y,\hat{y}) = -\sum_{c=1}^{C} y_c \log(\hat{y}_c) \tag{15}$$

Here $C$ denotes total classes, $\hat{y}_c$ symbolizes true label, and $\hat{y}_c$ signifies predicted probability for class $c$.

Last Prediction: After the model is trained, it is applied to making predictions on novel input data $X_{new}$ :

$$\hat{y}_{new} = \text{Prdict}(X_{new}, \Theta^*). \tag{16}$$

The predictive class for all input is then gained by choosing the class with the maximum predictive probability. Fig. 2 depicts the infrastructure of LSTM+GRU.

$$Class = argmax(\hat{y}_{new}).$$

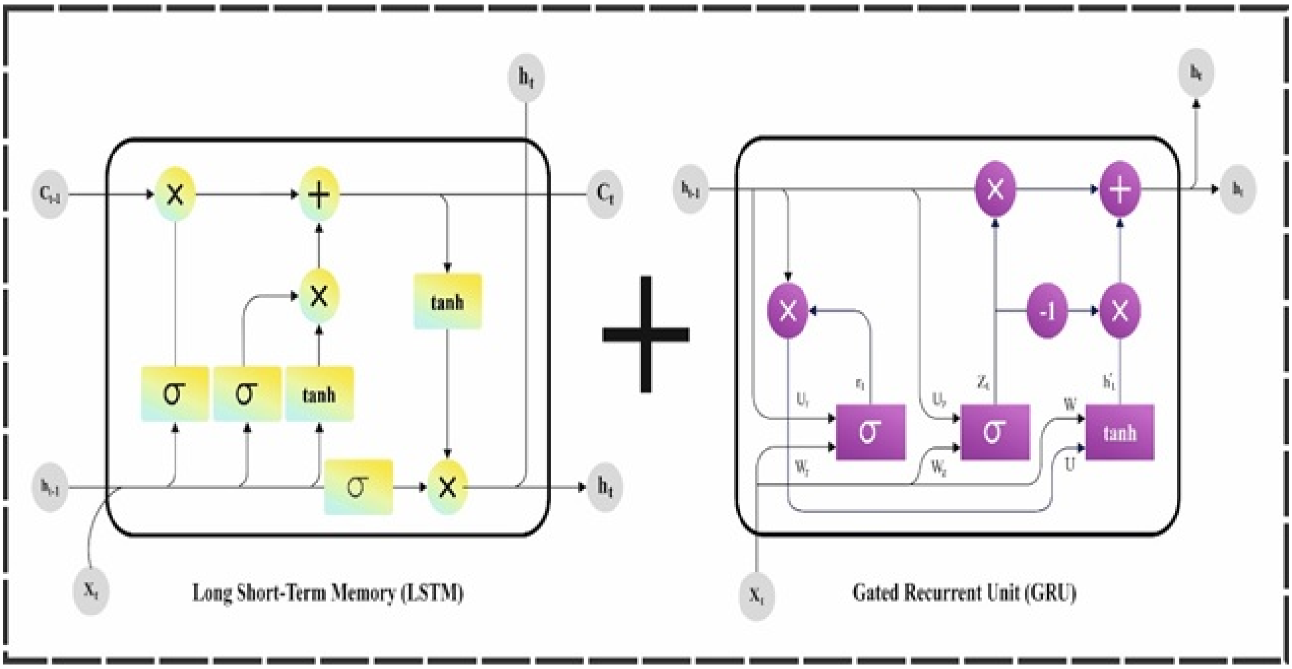


**Fig. 2:** Architecture of LSTM+GRU

### *3.4 Hyperparameter Tuning Process*

Eventually, the ASO algorithm adjusts the hyperparameter values of the LSTM+GRU algorithm optimally and outcomes in greater prediction performance. Skiing is the most challenging sport during winter [22]. Moreover, many skiers take part in this competition to compete with one another, and only each one reaches ultimately the winner. Some skier's last target is to form their best work to attain the tournament inside all matches. In advance of the match, whole skiers should halt on the same latitude and different longitude situations. Skiers yield the benefit of dual models in the race: holding power and sliding on the snow. During static slipping, skiers can slide back and forward in the smaller area, the primary objectives are to avoid declines and keep physical strength, and desire for sports. The leading inspiration of the ASO is skiers' behaviors that struggle for the contest. These skiing behaviors are like the exploration and exploitation performance inside the investigative methods.

Operators of exploration and exploitation

During this ASO, the skier's manner follows numerous basic fundamental beliefs:

- Overall individuals were expected not dependent, and there is no coincidence among dual skiers.
- Assuming that the individuals like appropriate skills to evade slides down derived from convex and concave land.
- Every skier has the ability to move randomly after an important person declines.

In the proposed model, a skiing match has three phases: initialization, iteration, and last sprint. Throughout this stage of initialization, every skier would attain the existing perfect condition, and then complete skiers follow the way to the first place. In the iteration phase, particular skiers could stop by, and others no need to come nearer the person who slides down. During this last stage, entire skiers try to increase the first position. Skiers need to determine their particular state that is randomly allotted at this phase.

Achieved the parameters' quantities, the ASO might generate the initial population to calculate the fitness quantities. Particularly, each skier is held consistent. During this ASO's second phase, for example, iteration, the ASO's iterations are carried out. Each skier goes from their recent states to novel states using particular regularity. It must be noted that the skiers have announced their position and come nearer the initial position. The skier's fitness values are measured and they are in comparison with the prior marks. Formerly, the optimal solution is sustained in the recent iteration. The mental shifts of the skiers are derived from dual important matters: the initial position and eluding dropping.

$$s_i(t) = |(|z_i(t) - z_b(t)|)|, \tag{17}$$

where $z_i(t)$ denotes a skier's condition within the iteration amount t, and the first place's state in the iteration amount t is described by $z_b(t)$. Skiers might have utilized several skiing capabilities to contend for first place in the race. As the sport develops, the skier's physical strength would modify so that is established by the subsequent equation:

$$g_i(t) = \left(\frac{L}{1 + exp\ (-b(z_i(t) - z_0))}\right), \tag{18}$$

whereas, $z_0$ represents the mid-point of the sigmoid $z$ -value, which is configured to 0.5; b turned out to be the logistic growing ratio; $L$ was considered as the high-level. The state of the individuals and the stage vector can be transformed by the subsequent equation.

$$\Delta Z(t) = r \cdot S(t) \cdot G(t), \tag{19}$$

$$Z(t+1) = Z(t) + \Delta Z(t), \tag{20}$$

whereas $G(g)$ and $S(s)$ were, successively, deliberated that matrix of skier's physical ability and distance matrix, $r$ indicates random numbers namely among $(0,1), Z(t+1)$ refers to the individuals' conditions inside the iteration count $t+1$. Some individuals may drop down through skiing owing to chances and imperfect capabilities. Concerning the influence of physical distance and many features, such as rain, wind, and rough glides, individuals are more possibly to drop down after competing for the first place.

In the proposed method, the drop-down specifies the physical strength of skiers changes considerably regarding the former instant's physical effectiveness. After the change among the dual quantities meets 15%, the individual was stated as an individual, drop-down. After the remoteness value is lower than 0.5, other individual conditions became reorganized to a Levy fight (LF) model. Another individual condition can be calculated by the following equation:

$$Z(t+1) = Z(t) + Levy(z(t)) \times Z(t), \tag{21}$$

Whereas $Levy(z)$ symbolizes the skiers' state's dimension. The LF is calculated as below:

$$Levy(z) = 0.01 \times (\frac{(r_1 \times \delta)}{(|r_2|^{1/\gamma})}), \tag{22}$$

with $r_1$ and $r_2$ representing stochastic quantities within $(0,1)$; $\gamma$ denotes a constant which is equivalent to 1.5, and $\delta$ can be achieved by:

$$\delta = \left( \frac{\Gamma(1+\gamma) \times \sin\left(\frac{\pi\gamma}{2}\right)}{\Gamma(1+\gamma) \times \gamma \times 2^{\left(\frac{\gamma-1}{2}\right)}} \right)^{\frac{1}{\gamma}} \tag{23}$$

where $\Gamma(z) = (z-1)!$.

In the ASO's final sprint phase, the first skier would continue to near the recent way. Another participant would confirm the new skiing motion over again connecting the first position's forward manner associated with their recent orientation and state. The iteration would stop after the standard of convergence remains to be regarded as consistent. In the proposed method, the process of search can be divided into seven stages:

1$^{st}$ Stage: Initialization of the individuals in addition to the phase vectors.

2$^{nd}$ Stage: Calculate the isolation between individuals and first place, the enhanced solution of the iterations' recent quantity is gained.

3$^{rd}$ Stage: Repeat the skiers' first position in the iterations' recent number.

4$^{th}$ Stage: The ASO model outlines whether someone drops. After numerous skiers drop down, other skier conditions are changed with an LF method to aim for first position.

5$^{th}$ Stage. Entire skiers repeat their state according to the first position.

6$^{th}$ Stage. The model identifies whether the stop-searching standard is attained. After the stop standard is attained, continue Stage 7; or else, return to the second Stage.

7$^{th}$ Stage. Output of the best solutions.

Here, the ASO is employed for determining the hyperparameter intricate in the LSTM+GRU technique. The MSE is measured as the objective function and is described below.

$$MSE = \frac{1}{T} \sum_{j=1}^{L} \sum_{j=1}^{M} (y_j^i - d_j^i)^2, \tag{24}$$

while $L$ and $M$ signify the resulting value of data and layer respectively, $y_j^i$ and $d_j^i$ mean the achieved and suitable magnitudes for the $j^{th}$ unit from the resulting layer in time t respectively

## 4 Performance Analysis

This section provides a comprehensive evaluation of the performance of the proposed DLFPF-FSA methodology using a dataset of 502 samples collected from Saudi Arabia between 2020 and 2025. The initial dataset comprised 24 financial attributes; however, through an attribute selection mechanism, 13 key variables were retained: proposed dividends, investments in associates, customer deposits, other liabilities, issued sukuk, investment property (net), capital, legal reserve, financing (net), other reserves, retained earnings, Tier 1 sukuk, and total equity. This reduction in dimensions improves the model's efficiency while preserving the most relevant predictive indicators for forecasting financial performance.

A comparative analysis of the actual and predicted values produced by the DLFPF-FSA model across different training phases demonstrates a clear improvement in prediction accuracy as the learning process progresses. In the early stages (e.g., from training cycle 10 to 30), the model exhibits significant discrepancies between actual and predicted values, indicating an initial phase of poor agreement. However, as the number of training cycles increases (from cycle 40 to 70), predictions increasingly converge with actual observations, reflecting improved model learning and a decrease in prediction error.

Statistically, the variance between actual and predicted values is significantly reduced in the middle training stages, indicating a decrease in both bias and error measures such as mean absolute error (MAE) and root mean square error (RMSE). This stage represents the model's optimal learning zone, where a balance is achieved between bias and variance. In later training cycles (e.g., from cycle 80 to 100), although the model maintains relatively high predictive accuracy, an increase in the variability of predicted values can be observed, indicating the onset of hyper-agreement and increased variance.

Furthermore, a comparison of the training and test results reveals a slight generalization gap, with the model performing slightly better on the training data compared to the unseen test data. However, the overall deviation remains minimal, confirming the robustness and generalizability of the DLFPF-FSA methodology. The results also highlight that the proposed model consistently produces predictions that accurately mimic the actual trend across different operating periods, with minimal variance between the two curves.

In summary, the DLFPF-FSA model demonstrates superior predictive performance, characterized by low variance between actual and predicted values, stable convergence behavior, and strong generalizability. These results confirm the effectiveness of the proposed hybrid framework for forecasting financial performance in the Saudi context.

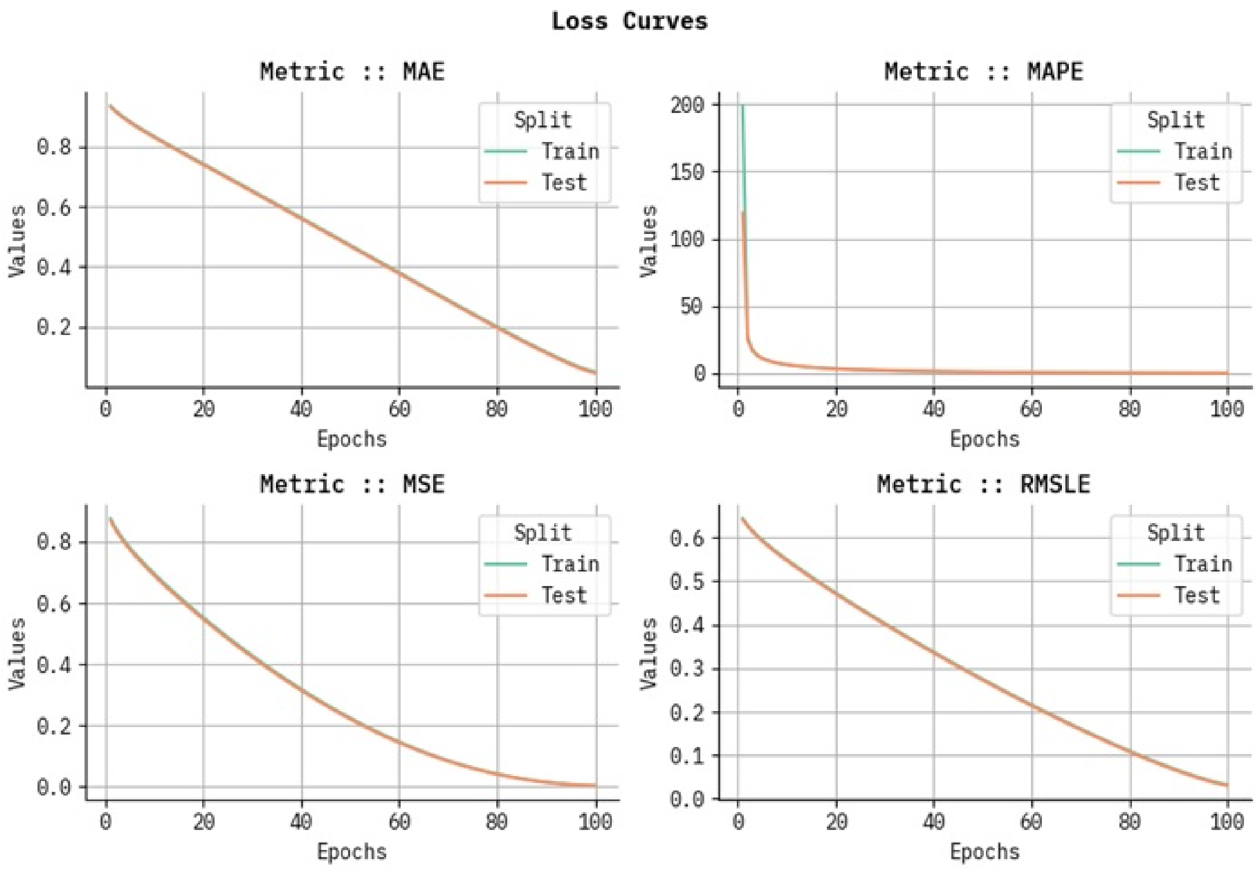


**Fig. 3:** Result analysis for Lose curve for all the metrics Epoch 0-100.

Fig. 3 established result analysis for the loss graph of the DLFPF-FSA method under epoch 0-100. The values of loss are computed across a time period of 0-100 epochs. It is demonstrated that the training outcomes signify a lessening trend, which points out the proficiency of the DLFPF-FSA technique in harmonizing a trade-off among generalization as well as data fitting. The succeeding dilution in values of loss as well as assurances the superior outcome of the DLFPF-FSA approach and tune the prediction results gradually.

Table 1 and Fig. 4 illustrate the Training set (TRAST) and Testing set (TESST) of the DLFPF-FSA system under various metrics. Under TRAST, the DLFPF-FSA method obtains MSE of 0.0037, RMSLE of 0.0312, MAE of 0.0491, and MAPE of 0.0542. Besides, with TESST, the DLFPF-FSA methodology gains MSE of 0.0038, RMSLE of 0.0317, MAE of 0.0526, and MAPE of 0.0579.

Table 2 offers the classifier result of the DLFPF-FSA method with existing algorithms under dissimilar metrics [23, 24, 25].

Fig. 5 provides the MSE outcome of the DLFPF-FSA technique with existing models. The Autoregressive, ARIMA, APSO-MLSSVR, DTR, Random walk, LSTM, and ARMA-CNNLSTM approaches have reached maximum MSE of

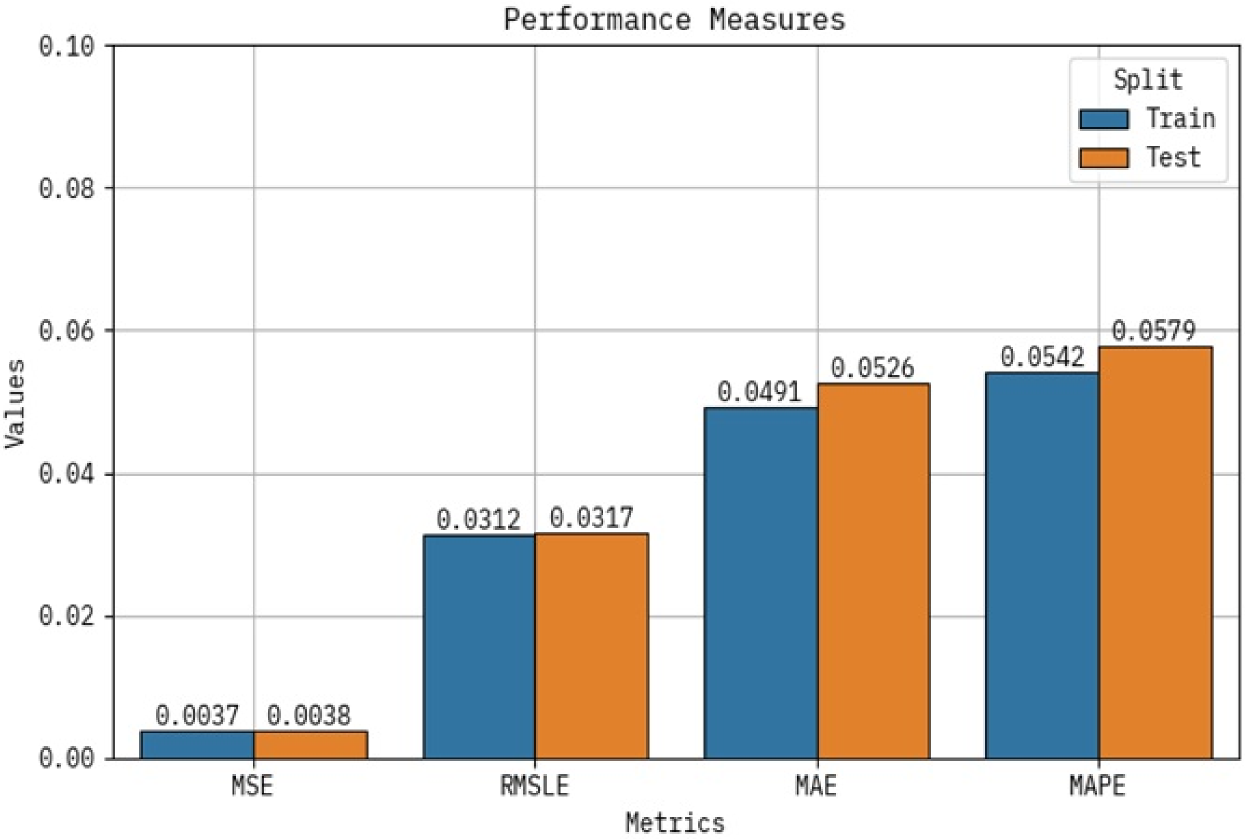


**Fig. 4:** TRAST and TESST outcome of DLFPF-FSA method under various metrics.

**Table 1:** TRAST and TESST outcome of DLFPF-FSA method under various metrics.

| Metrics | Training Set | Testing Set |
|---|---|---|
| MSE | 0.0037 | 0.0038 |
| RMSLE | 0.0312 | 0.0317 |
| MAE | 0.0491 | 0.0526 |
| MAPE | 0.0542 | 0.0579 |

0.4885, 0.3241, 0.5679, 0.5643, 0.2299, 0.5667, and 0.4195, correspondingly. While ANN models have obtained a closer better MSE of 0.1822. Additionally, the proposed DLFPF-FSA method has got lower MSE of 0.0037.

Fig. 6 delivers the RMSLE outcome of the DLFPF-FSA technique with existing methodologies. The Autoregressive, ARIMA, ANN, APSO-MLSSVR, and DTR methods have gained improved RMSLE of 0.3595, 0.2915, 0.2375, 0.1805, and 0.1195, respectively. Whereas, the random walk, LSTM, and ARMA-CNNLSTM techniques have got somewhat better RMSLE of 0.0415, 0.0552, 0.0400, and 0.0312, correspondingly. In addition, the proposed DLFPF-FSA model has got lesser RMSLE of 0.0312.

The MAE outcome of DLFPF-FSA method with exiting algorithms are depicted in Fig. 7. Based on MAE, the proposed DLFPF-FSA method has gained minimal MAE of 0.0491, while the existing methodologies such as Autoregressive, ARIMA, ANN, APSO-MLSSVR, DTR, Random walk, LSTM, and ARMA-CNNLSTM techniques have reached better MAE of 1.2411, 1.1701, 1.1091, 1.0511, 0.9721, 0.9151, 2.3696, and 0.8837, correspondingly.

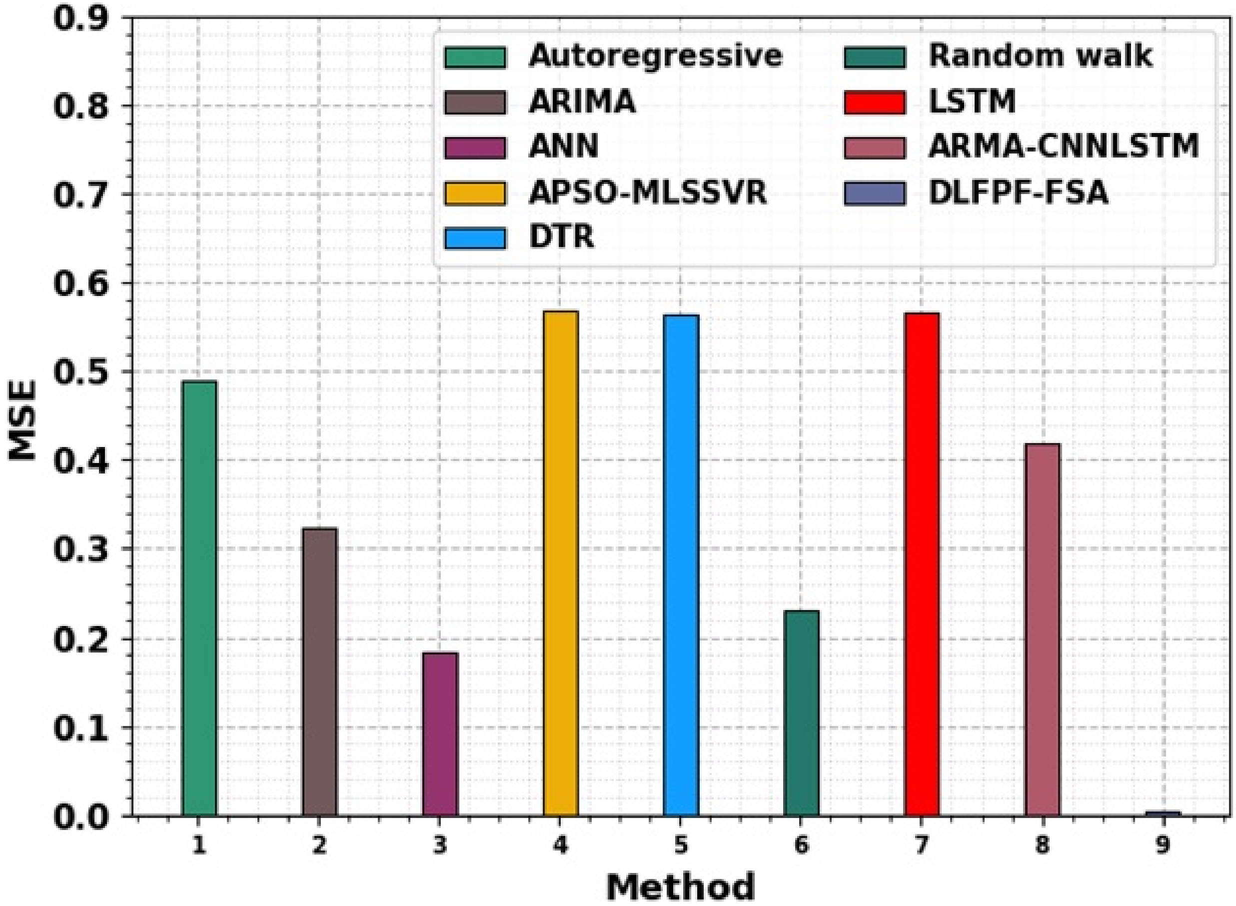


**Fig. 5:** MSE outcome of DLFPF-FSA method with existing models.

**Table 2:** MSE outcome of DLFPF-FSA technique with existing models.

| Method | MSE | RMSLE | MAE |
|---|---|---|---|
| Autoregressive | 0.4885 | 0.3595 | 1.2411 |
| ARIMA | 0.3241 | 0.2915 | 1.1701 |
| ANN | 0.1822 | 0.2375 | 1.1091 |
| APSO-MLSSVR | 0.5679 | 0.1805 | 1.0511 |
| DTR | 0.5643 | 0.1195 | 0.9721 |
| Random walk | 0.2299 | 0.0415 | 0.9151 |
| LSTM | 0.5667 | 0.0552 | 2.3696 |
| ARMA-CNNLSTM | 0.4195 | 0.0400 | 0.8837 |
| DLFPF-FSA | 0.0037 | 0.0312 | 0.0491 |

## 5 Conclusion

This study presents a DLFPF-FSA methodology. The DLFPF-FSA model relies on improving the prediction of financial performance using state-of-the-art optimization algorithms. To accomplish that, the data normalization stage is initially applied by z-score normalization to standardize data by scaling it to a common range. Next, the FS process is executed by the BESO algorithm to identify the most relevant features from the input data. For the prediction of financial performance, the proposed DLFPF-FSA system designs a hybrid of LSTM+GRU method. Eventually, the ASO algorithm adjusts the

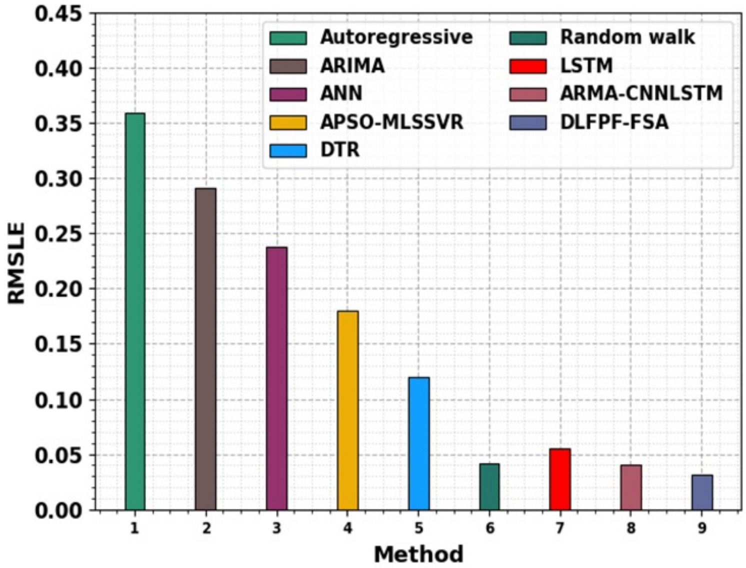


**Fig. 6:** RMSLE outcome of DLFPF-FSA method with existing models.

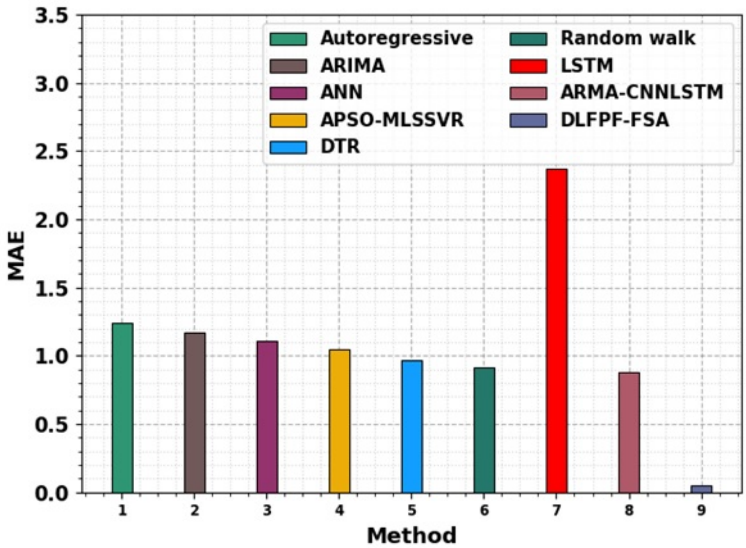


**Fig. 7:** MAE outcome of DLFPF-FSA method with existing models.

hyperparameter values of the LSTM+GRU algorithm optimally and outcomes in greater prediction performance. The experimental evaluation of the DLFPF-FSA algorithm can be tested on a benchmark dataset. The extensive outcomes highlight the significant solution of the DLFPF-FSA approach to the financial prediction process.

**Data Availability Statement**: The authors confirm that the data supporting the findings of this study are available within the article [23-25].

**Acknowledgement**: The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through Large Research Project under grant number RGP2/492/46.

## References


[1] Sezer, O.B., Gudelek, M.U. and Ozbayoglu, A.M., 2020. Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. Applied soft computing, 90, p.106181

[2] Barra, S., Carta, S.M., Corriga, A., Podda, A.S. and Recupero, D.R., 2020. Deep learning and time series-to-image encoding for financial forecasting. IEEE/CAA Journal of Automatica Sinica, 7(3), pp.683-692.

[3] Altan, A. and Karasu, S., 2019. The effect of kernel values in support vector machine to forecasting performance of financial time series. The Journal of Cognitive Systems, 4(1), pp.17-21.

[4] Ozbayoglu, A.M., Gudelek, M.U. and Sezer, O.B., 2020. Deep learning for financial applications: A survey. Applied soft computing, 93, p.106384.

[5] Hu, Z., Zhao, Y. and Khushi, M., 2021. A survey of forex and stock price prediction using deep learning. Applied System Innovation, 4(1), p.9.

[6] Cao, J., Li, Z. and Li, J., 2019. Financial time series forecasting model based on CEEMDAN and LSTM. Physica A: Statistical mechanics and its applications, 519, pp.127-139.

[7] Bradul, O., Varava, L., Turylo, A., Dashko, I. and Varava, A., 2021. Forecasting the effectiveness of the enterprise to intensify innovation and investment development, taking into account the financial component of economic potential.

[8] Dixon, M. and London, J., 2021. Financial forecasting with $\alpha$-RNNs: A time series modeling approach. *Frontiers in Applied Mathematics and Statistics*, 6, p. 551138.

[9] Nirino, N., Santoro, G., Miglietta, N. and Quaglia, R., 2021. Corporate controversies and company's financial performance: Exploring the moderating role of ESG practices. Technological Forecasting and Social Change, 162, p.120341

[10] Aziz, A.S., 2023. Neutrosophic Combinative Distance-based Assessment (CODAS) Method for Evaluating the Financial and Operational Performance of Shipping Companies. Full Length Article, 4(1), pp.28-8.

[11] Ben Mrad, A., Hnich, B., Lahiani, A. and Mefteh-Wali, S., 2025. Comprehensive Stock Market Insight: Bayesian Networks for Multi-output Forecasting. Computational Economics, pp.1-21.

[12] Shi, X., Zhang, Y., Yu, M. and Zhang, L., 2025. Deep learning for enhanced risk management: a novel approach to analyzing financial reports. PeerJ Computer Science, 11, p.e2661.

[13] Liang, L., Liu, B., Su, Z. and Cai, X., 2024. Forecasting corporate financial performance with deep learning and interpretable ALE method: Evidence from China. Journal of Forecasting.

[14] Abir, S.I., Sarwer, M.H., Hasan, M., Sultana, N., Dolon, M.S.A., Arefeen, S.S., Shimanto, A.H., Zakaria, R.M., Al Shiam, S.A., Shoha, S. and Saha, T.R., 2025. Deep Learning for Financial Markets: A Case-Based Analysis of BRICS Nations in the Era of Intelligent Forecasting. Journal of Economics, Finance and Accounting Studies, 7(1), pp.01-15.

[15] Srivastava, T., Mullick, I. and Bedi, J., 2024. Association mining based deep learning approach for financial time-series forecasting. Applied Soft Computing, 155, p.111469.

[16] Chauhan, J.K., Ahmed, T. and Sinha, A., 2025. A novel deep learning model for stock market prediction using a sentiment analysis system from authoritative financial website's data. Connection Science, 37(1), p.2455070.

[17] Che, W., Wang, Z., Jiang, C. and Abedin, M.Z., 2024. Predicting financial distress using multimodal data: An attentive and regularized deep learning method. Information Processing & Management, 61(4), p.103703.

[18] Knifo, S. and Alzubi, A., 2025. DR-Z2AN: dual-recurrent neural network with a tri-channel attention mechanism for financial management prediction. Complex & Intelligent Systems, 11(1), pp.1-16.

[19] Reddy, T.S., Kalaichelvi, T., El-Ebiary, Y.A.B., Rajmohan, V. and Ramesh, J.V.N., 2025. Computing Environments: Employing Recurrent Neural Networks and ELM for Advanced Analysis in Investigation Scenarios. Journal of Advances in Information Technology, 16(1).

[20] Chellappan, D. and Rajaguru, H., 2024. Machine Learning Meets Meta-Heuristics: Bald Eagle Search Optimization and Red Deer Optimization for Feature Selection in Type II Diabetes Diagnosis. Bioengineering, 11(8), p.766.

[21] Gupta, B.B., Gaurav, A., Arya, V., Bansal, S., Attar, R.W., Alhomoud, A. and Psannis, K., 2025. Earthworm optimization algorithm based cascade LSTM-GRU model for android malware detection. Cyber Security and Applications, 3, p.100083.

[22] He, M., Wang, H. and Thwin, M., 2025. A machine learning technique for optimizing load demand prediction within air conditioning systems utilizing GRU/IASO model. Scientific Reports, 15(1), p.3353.

[23] Chou, J.S., Truong, D.N. and Le, T.L., 2020. Interval forecasting of financial time series by accelerated particle swarm-optimized multi-output machine learning system. IEEE Access, 8, pp.14798-14808.

[24] Tutcu, B., Kayakuş, M., Terzioğlu, M., Ünal Uyar, G. F., Talaş, H. and Yetiz, F., 2024. Predicting Financial Performance in the IT Industry with Machine Learning: ROA and ROE Analysis. *Applied Sciences*, 14(17), p. 7459.

[25] He, K., Yang, Q., Ji, L., Pan, J. and Zou, Y., 2023. Financial time series forecasting with the deep learning ensemble model. Mathematics, 11(4), p.1054.