Assessment of Class Imbalance Data Handling with Attention-Based Deep Learning Approach for Robust Financial Distress Prediction in Enterprises

Alisher Sherov
Department of Economics, Mamun University, Khiva, Uzbekistan | Department of Finance and Tourism, Termez University of Economics and Service, Termez, Uzbekistan
sherov_alisher@mamunedu.uz (corresponding author)

Tukhtabek Rakhimov
Department of Economics, Urgench State University, Uzbekistan
tuxtabek.r@urdu.uz

Hafis Hajiyev
Department of Finance and Audit, Azerbaijan State University of Economics (UNEC), Baku, Azerbaijan
hafiz_hajiyev@unec.edu.az

Maria Zelinskaya
Department of Management, Kuban State Agrarian University named after I.T. Trubilin, Krasnodar, Russia
zelinskaja.m@edu.kubsau.ru

Gavkhar Khidirova
Department of Tourism and Hotel Management, Bukhara State University, Bukhara, Uzbekistan
g.r.xidirova@buxdu.uz

Received: 16 September 2025 | Revised: 3 October 2025 and 16 October 2025 | Accepted: 19 October 2025

 DOI: https://doi.org/10.48084/etasr.14843

ABSTRACT

Bankruptcy prediction and credit risk assessment are two of the most crucial problems in finance, remaining an advanced area of research in the financial sector. Financial distress is when an enterprise faces financial problems. Standard analytical methods often depend on limited financial indicators and expert opinions, which miss complex non-linear patterns. With the expansion of Artificial Intelligence (AI) and Deep Learning (DL), financial stress testing has experienced a fundamental change, utilizing progressive methods of computation to improve the precision of risk prediction. This study presents an Attention-based Deep Learning Approach for Robust Financial Distress Prediction in Enterprises (ADLA-RFDPE) method to improve the detection of financial risks in enterprise management. Initially, min-max normalization is used to maintain data consistency, and SMOTE is utilized to address class imbalance. Then, the Elephant Herding Lion Optimizer (EHLO) is used for the dimensionality reduction process. For financial distress classification, an Attention-based Gated Recurrent Unit (Attention-GRU) technique is employed, optimized with the Pied Kingfisher Optimizer (PKO) technique. The experimental results highlighted a superior accuracy of 98.68% over existing approaches on an Australian credit dataset.

Keywords-financial distress prediction; deep learning; bankruptcy; data analysis; class imbalance; elephant herding lion optimizer

I. INTRODUCTION

Financial Distress Prediction (FDP) is crucial due to changes in the economy and business environment [1]. The ability to predict financial issues is important for real organizations and can improve the financial controllability of businesses [2, 3]. Small and Medium-sized Enterprises (SMEs) are treated differently in credit risk assessment than other business entities [4]. In addition, accurate FDP is crucial for business stability and informed investment. However, incomplete data, financial adaptability, and imprecision complicate this assessment for SMEs, increasing credit risk [5]. Conventional risk valuation models, involving rule- and expert-driven heuristics, and statistical methods, have long been the basis of financial distress tests [6]. With the growth of AI and DL, FDP has experienced a fundamental change, utilizing modern computational approaches to improve the precision of risk forecasting [7]. DL is a subdomain of Machine Learning (ML) that uses several layers of nonlinear processing units [8]. DL techniques, especially Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN), excel in capturing complex patterns and temporal dependences in financial data [9].

In [1], a novel FDP method was proposed, using the Adaptive Whale Optimization Algorithm with a DL (AWOA-DL) method. In [10], a novel multi-class description of FDP was presented from the perspective of financially sound businesses, using MNLogit to integrate the results of the earlier stage. In [11], a modified PSO technique with three development strategies was proposed. In [12], a User-Response-Guided Deep Attention Network (URGDAN) technique was proposed for FDP. In [13], a new Lightspace-SMOTE upsampling model was introduced. In [14], a new Case-Based Reasoning (CBR)-based ensemble learning model included three phases: CBR-based single forecasting, ensemble performance, and missing data imputation. The aim in [15] was to evaluate smart methods in FDP and select the best.

Despite their efficiency, existing studies struggle with high complexity and class imbalance. The major gap is in developing efficient and scalable FDP methods that handle imbalance and improve feature selection. This study presents an Attention-based Deep Learning Approach for Robust FDP in Enterprises (ADLA-RFDPE) method, making the following key contributions:

- Applies the min-max normalization to standardize financial feature scales and ensure consistency across input data. This improves the learning process by preventing the dominance of high-range attributes. This step assists in more stable and effective model training for FDP.
- Uses SMOTE to address class imbalance, improving model sensitivity toward minority class samples. EHLO is introduced for effective dimensionality reduction, detecting the most relevant financial indicators.
- Utilizes an Attention-based GRU classifier for capturing temporal and contextual patterns in financial data for distress prediction. The PKO model is used for fine-tuning its performance, enhancing prediction robustness and generalization.
- The novelty of this study is in integrating EHLO and PKO within the Attention-GRU framework, presenting a unique hybrid DL process to efficiently handle class imbalance and reduce dimensionality. Effective capturing of intrinsic financial patterns is improved, resulting in robust and accurate prediction.

II. PROPOSED METHOD

The proposed method involves Min-Max normalization, EHLO-based feature selection, Attention-GRU-based classification, and PKO-based hyperparameter tuning. Figure 1 shows the flow of the proposed ADLA-RFDPE model.

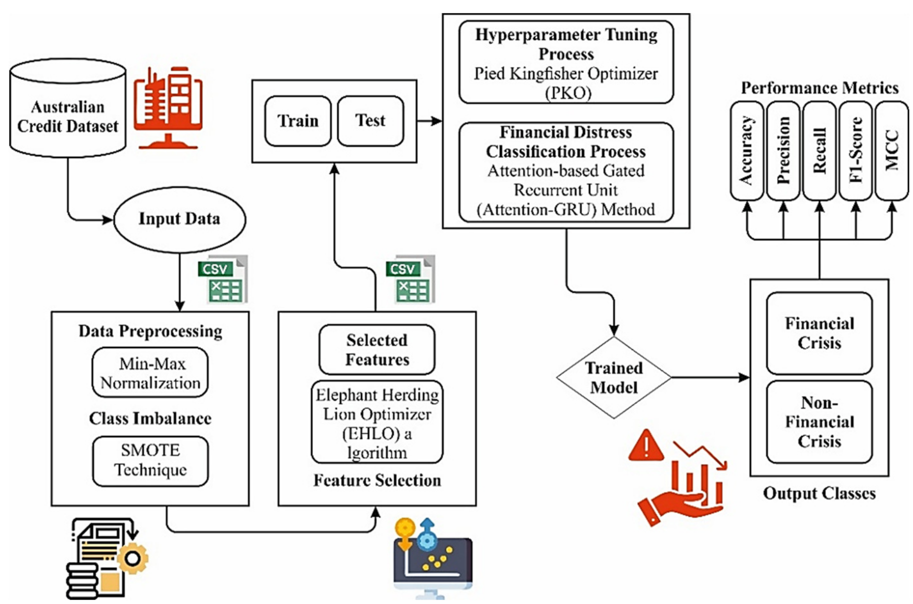

Fig. 1. General structure of the ADLA-RFDPE approach.

A. Feature Scaling Methods

Min-max normalization ensures data consistency and was selected for its simplicity and efficiency in scaling features to a consistent range, which enhances model convergence. SMOTE was selected to address class imbalance, as it generates synthetic minority samples [16]. The class imbalance is effectively addressed without losing valuable data, unlike simple oversampling or undersampling.

The MinMax scaler scales data to a specified range, typically (0,1), by subtracting the feature's minimum value and dividing by its range (max-min).

$$y = \frac{(x - x_{min})}{(x_{max} - x_{min})} \quad (1)$$

where x_{min} implies the minimal feature value, x_{max} refers to the maximal feature value, x denotes the original value, and y refers to the normalized value in the interval of $[0,1]$.

SMOTE is a widely used method for addressing class imbalance. It generates synthetic minority instances by selecting a random instance $x_i \in X$ and finding a nearby instance $x_i' \in X$ within the same class to create new samples in that neighborhood. Then, a line is computed among the x_i and x_i', and the new minority class sample x_{new} is obtained using:

$$x_{new} = x_i + (x_i' - x_i) \cdot \delta,\ \ \delta \in [0,1] \quad (2)$$

where δ refers to a random number in [0,1].

B. Feature Reduction

Then, the EHLO approach is used for the dimensionality reduction process to extract the most significant financial indicators while removing redundant attributes [17]. This method was selected for its ability to balance exploration and exploitation by incorporating elephant herding behavior with lion optimizer strategies, resulting in more effective feature selection. This method outperforms conventional methods by mitigating redundancy while preserving crucial financial indicators, thus enhancing the accuracy of the model and mitigating computational complexity.

This approach is based on elephants' herding behavior, where clusters of female elephants led by a matriarch influence group movement. To enhance this approach, the location is updated by using the Lion optimizer. In addition, this method holds higher EHO exploration efficacy and an active LOA ability, which prolongs lifespan.

$$I_p = e_1, e_2, \ldots \ldots, e_n \quad (3)$$

Next, the fitness is assessed, which reflects the least volume, the lowest amount of delay (D), energy consumption (P_V), and throughput. The fitness function (F_t) is given by:

$$F_t = \mathrm{Min}(t_p), \mathrm{Min}(P_V, D) \quad (4)$$

Every clan contains elephants, and the worst and best likelihoods for each elephant stage are involved in every elephant position P.

$$L_{new,c_{i,j}} = L_{c_{i,j}} + \alpha \left(L_{best,c_{i,j}} - L_{c_{i,j}} \right) \times r \quad (5)$$

where $L_{new,\, c_{i,j}}$ denotes an upgraded location, $L_{c_{i,j}}$ is an old position, $L_{best,c_{i,j}}$ represents the best location, and $\alpha, r \in 0$ to 1. The greatest fitting set member's location upgrade is given by:

$$L_{new,c_{i,j}} = \beta \times L_{center,c_j} \text{ and}$$

$$L_{center,c_j} = \frac{\sum_{Error::0x0000=1}^{n} L_{c_{i,j}}}{n_l} \quad (6)$$

where n_l denotes the overall elephants in all clans and $\beta \in [0,1]$.

The EHO and LOA search behaviors are united, removing potential negative search ability and improving search ability for convergence. The novel part of a female lion is:.

$$FL' =$$

$$FL + 2G \times \eta_d\{S_1\} + w(-1,1) \times \tan(\theta) \times G \times \{S_2\} \quad (7)$$

where FL' denotes the novel female lion location, FL denotes the female lion, G denotes the distance, $\{S_1\}$ is a vector of the female lion, and $\{S_1\}$ is normal to $\{S_2\}$. In addition, nomadic lions are searching for the victim as:

$$N_d(O_{ij}) = \begin{cases} O_{ij} & if (\eta_d)_j > (p_b)_i \\ (\eta_d)_j & \text{otherwise} \end{cases} \quad (8)$$

where $N_d(O_{ij})$ denotes the present location of the i-th nomad lion, j represents a dimension, η_d signifies a uniform randomly generated number in $[0,1]$, and p refers to a likelihood, which is intended for every nomadic lion self-sufficiently.

$$og_j^1 = \tau \times FL_j + \sum \frac{1-\tau}{Z_i} \times ML_j \times Z_i \quad (9)$$

$$og_j^2 = (1-\tau) \times FL_j + \sum \frac{\tau}{Z_i} \times ML_j \times Z_i \quad (10)$$

where j denotes a dimension and Z_i is equivalent to 1. Lastly, the worst or male elephants are removed from their group because they are the lowest-ranking elephants.

$$L_{worst,\, c_{i,j}} = L_{\min} + (L_{\max} - L_{\min} + 1) \times r \quad (11)$$

where $L_{worst,\, c_{i,j}}$ are the clan's worst male elephants, and $L_{\max}$ and $L_{\min}$ denote the maximum and minimum elephant range. The selected features enhance efficacy and responsiveness by optimizing decision-making and resource allocation. The fitness function ($Fitness$) reflects both the classifier's accuracy and the number of chosen features. Hence, the following fitness function was applied to compute an individual solution:

$$Fitness = \alpha \times\ ErrorRate + (1-\alpha) \times \frac{\#SF}{\#All\_F} \quad (12)$$

where $ErrorRate$ is the classifier error rate, determined as the ratio of inappropriate grades to the number of classifications arranged among 0 and 1, $\#SF$ denotes the number of selected attributes, $\#All\_F$ denotes the overall features, and α is employed to determine the impact.

C. Classification Model

An Attention-GRU-based model was used for financial distress classification [18]. The temporal dependencies in financial data are effectually captured by this model, while focusing on the most relevant features through attention mechanisms. Unlike models such as LSTM or transformers, the proposed one has a simpler architecture with lower computational complexity, enabling faster training and inference without losing accuracy. This balance makes it specifically appropriate for financial distress classification tasks where interpretability and efficiency are significant. GRU is an RNN with a gated structure to solve the longer-time sequence memory issue. Compared to LSTM, its architecture is simpler, but it can reach a similar performance. The GRU architecture includes dual gates: reset and update. The update gate (r_T) evaluates which input information is updated or forgotten by:

$$r_T = \sigma\left(W_r \cdot \left[\acute{h}_{T-1}, X_T\right] + b_r\right) \quad (13)$$

whereas b_r and W_r represent the bias and the weighted matrix of the update gate, respectively. The reset gate (Z_T) defines the amount of forgotten data, with a value from 0 to 1, with a value closer to 0 demonstrating a high forgetting degree:

$$z_T = \sigma\left(W_z \cdot \left[\acute{h}_{T-1}, X_T\right] + b_z\right) \quad (14)$$

Here b_z and W_z represent the bias and the weighted matrix of the reset gate, correspondingly. After the input message passes through the reset and the update gates, the candidate hiding state $\tilde{h}_T$ at time step T is upgraded to:

$$\tilde{h}_T = tanh(W_h \cdot [h_{T-1} \odot r_T, X_T] + b_h) \quad (15)$$

where W_h and b_h represent the bias and weight matrices of the hidden state of the cell, respectively, and $\odot$ denotes the Hadamard product. Finally, the output h_T at time step T is gained as:

$$h_T = (1 - z_T) \odot h_{T-1} + z_T \odot \tilde{h}_T \quad (16)$$

The AM principle similarly identifies data of higher significance during processing, focusing attention on the most crucial parts of the input data. The basic concept of AM is to characterize the amount of attention of the model to dissimilar segments of the input data by using scores [19]. This process can be described as:

1. Compute weights by measuring the correlation or similarity between the given query and each key.

$$e_t = \tanh(W_t h_t + b_h) \quad (17)$$

2. Normalize the weights utilizing the softmax function to acquire the weights corresponding to all keys.

$$\alpha_t = \frac{exp(e_t)}{\Sigma_{t=1}^{T} exp(e_t)} \quad (18)$$

3. Weight and sum the consistent values based on the weights to acquire the last attention output.

$$c = \sum_{t=1}^{T} \alpha_t \, h_t \quad (19)$$

D. Parameter Selection Using PKO Approach

PKO is used to fine-tune model parameters and improve classification performance [20]. PKO is a novel metaheuristic optimizer inspired by the Pied Kingfisher's (PK) predatory behavior, divided into four phases: initialization, exploration, development, and symbiosis. This model effectively balances exploration and exploitation in the search space. During initialization, a population of candidate solutions is arbitrarily generated within the search space using:

$$Y_{i,j} = S_l + (S_u - S_l) \times r \quad (20)$$

where $i = 1,2, \dots, N;\ j = 1,2, \dots, M$, $Y_{i,j}$ denotes the location of the i-th individual in the j-th dimension, r refers to a randomly generated number between zero and one, S_l and S_u denote the lower and upper limits of the searching space, respectively, N indicates the population size, and M denotes the proportion of the problem.

In the exploration phase, PKO updates the position of each individual to explore new regions using:

$$\begin{cases} Y_i(t+1) = Y_i(t) + \beta \times A \times [Y_j(t) - Y_i(t)] \\ \beta = 2 \times R(1, M) \end{cases} \quad (21)$$

where $Y_i(t+1)$ and $Y_i(t)$ signify the location of subsequent and present iterations of PK, t denotes the current number of iterations, i and j are numbers from $[1, \dots, N]$ but i is not equivalent to j, and R is a randomly produced numeral. In the symbiosis phase, individuals update their positions cooperatively by interacting with randomly selected peers, enhancing diversity and convergence toward optimal solutions.

The selection of fitness is an important factor for manipulating the performance of the PKO model. The PKO model considers precision as the main standard to design the fitness function, expressed as:

$$Fitness = \max(P) \quad (22)$$

$$P = \frac{TP}{TP + FP} \quad (23)$$

where TP and FP denote true and false positives.

III. RESULTS AND DISCUSSIONS

The experimental performance of ADLA-RFDPE was examined using an Australian credit dataset [21], which contains 690 samples with 14 features, of which 9 key features are usually employed for analysis. The data is separated into dual classes depicting credit risk outcomes: 383 samples depict Financial Crisis cases, whereas 307 samples correspond to Non-Financial Crisis cases. The proposed method was developed using Python 3.6.5 on an i5-8600K CPU, 4GB GPU, 16GB RAM, 250GB SSD, and 1TB HDD, using a learning rate of 0.01, ReLU, 50 epochs, 0.5 dropout, and a batch size of 5.

Figure 2 shows the confusion matrices of the ADLA-RFDPE model under 80:20 and 70:30 train-test splits. The results indicate effective classification of the instances under each class. Table I presents brief financial crisis prediction results for the 80:20 train-test split. On 80% training, the ADLA-RFDPE technique achieved average $accur_y$, $preci_n$, $reca_l$, $F1_{score}$, and MCC of 98.68%, 98.77%, 98.68%, 98.72%, and 97.45%, respectively. On 20% testing, the ADLA-RFDPE model achieved an average $accur_y$, $preci_n$, $reca_l$, $F1_{score}$, and MCC of 97.50%, 97.87%, 97.50%, 97.68%, and 95.36%, respectively.

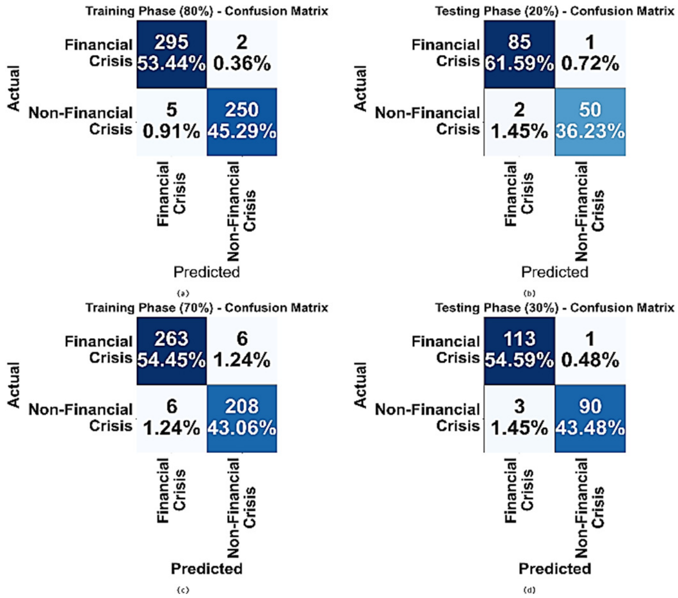

Fig. 2. Confusion matrices for different train-test splits.

TABLE I. FINANCIAL CRISIS PREDICTION FOR 80:20 TRAIN:TEST SPLIT

Class labels	$Accur_y$	$Preci_n$	$Reca_l$	$F1_{Score}$	MCC
Training (80%)					
Financial Crisis	99.33	98.33	99.33	98.83	97.45
Non-Financial Crisis	98.04	99.21	98.04	98.62	97.45
Average	**98.68**	**98.77**	**98.68**	**98.72**	**97.45**
Testing (20%)					
Financial Crisis	98.84	97.70	98.84	98.27	95.36
Non-Financial Crisis	96.15	98.04	96.15	97.09	95.36
Average	**97.50**	**97.87**	**97.50**	**97.68**	**95.36**

Table II presents a detailed comparison of ADLA-RFDPE with other methods [22, 23]. Based on $accur_y$, ADLA-RFDPE achieved a superior $accur_y$ of 98.68%, whereas the DNN-MLP, DT, Light GBM, XGBoost, Extra Trees (ET), CART, and GBDT methods achieved 93.00%, 80.00%, 90.67%, 98.58%, 91.05%, 94.09%, and 91.64%, respectively.

TABLE II. COMPARATIVE ANALYSIS OF ADLA-RFDPE WITH OTHER MODELS

Methods	$Accur_y$	$Preci_n$	$Reca_l$	$F1_{Score}$
DNN-MLP [22]	93.00	92.06	92.75	93.09
DT Classifier [22]	80.00	93.34	91.20	93.42
Light GBM [23]	90.67	96.63	93.59	93.96
XGBoost [24]	98.58	97.12	93.42	97.23
ET [23]	91.05	98.56	98.30	93.64
CART Model [23]	94.09	89.60	92.71	95.60
GBDT [23]	91.64	94.37	92.49	89.06
ADLA-RFDPE (proposed)	98.68	98.77	98.68	98.72

IV. CONCLUSION

This study presented the ADLA-RFDPE model, which uses min-max and SMOTE-based preprocessing, EHLO-based feature reduction, Attention-GRU-based classification, and PKO-based fine-tuning. The experimental results highlighted a superior accuracy of 98.68% over existing approaches on an Australian credit dataset. The limitations include the dependence on a single dataset and potential challenges in generalizing results across diverse financial contexts. Future studies may explore integrating alternative data sources and developing adaptive models that consider evolving market conditions. In addition, real-time data integration and scalable frameworks can be investigated to improve predictive accuracy and practical applicability.

REFERENCES

[1] M. Elhoseny, N. Metawa, G. Sztano, and I. M. El-hasnony, "Deep Learning-Based Model for Financial Distress Prediction," *Annals of Operations Research*, vol. 345, no. 2, pp. 885–907, Feb. 2025, https://doi.org/10.1007/s10479-022-04766-5.

[2] A. Oualid, Y. Qasmaoui, Y. Balouki, B. Itri, and L. Moumoun, "Advancing Credit Risk Management in Open Banking with Enhanced Federated Averaging Algorithm," *Engineering, Technology & Applied Science Research*, vol. 15, no. 3, pp. 22573–22579, Jun. 2025, https://doi.org/10.48084/etasr.10266.

[3] N. Aydin, N. Sahin, M. Deveci, and D. Pamucar, "Prediction of financial distress of companies with artificial neural networks and decision trees models," *Machine Learning with Applications*, vol. 10, Dec. 2022, Art. no. 100432, https://doi.org/10.1016/j.mlwa.2022.100432.

[4] F. T. Kristanti, M. Y. Febrianta, D. F. Salim, H. A. Riyadh, and B. A. H. Beshr, "Predicting Financial Distress in Indonesian Companies using Machine Learning," *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, pp. 17644–17649, Dec. 2024, https://doi.org/10.48084/etasr.8520.

[5] Y. Aker and A. Karavardar, "Using Machine Learning Methods in Financial Distress Prediction: Sample of Small and Medium Sized Enterprises Operating in Turkey," *Ege Academic Review*, vol. 23, no. 2, pp. 145–162, May 2023, https://doi.org/10.21121/eab.1027084.

[6] Y. Tong and Z. Serrasqueiro, "Predictions of failure and financial distress: A study on Portuguese high and medium-high technology small and mid-sized enterprises," *Journal of International Studies*, vol. 14, no. 2, pp. 9–25, Jun. 2021, https://doi.org/10.14254/2071-8330.2021/14-2/1.

[7] D. Wu, X. Ma, and D. L. Olson, "Financial distress prediction using integrated *Z*-score and multilayer perceptron neural networks," *Decision Support Systems*, vol. 159, Aug. 2022, Art. no. 113814, https://doi.org/10.1016/j.dss.2022.113814.

[8] A. Mehmood and F. De Luca, "Financial distress prediction in private firms: developing a model for troubled debt restructuring," *Journal of Applied Accounting Research*, vol. 26, no. 6, pp. 205–222, Nov. 2023, https://doi.org/10.1108/JAAR-12-2022-0325.

[9] A. S. Haya and S. Mishra, "The impact of AI-based cyber security on the banking and financial sectors," *Journal of Cybersecurity and Information Management*, vol. 14, no. 1, pp. 08–19, 2024, https://doi.org/10.54216/JCIM.140101.

[10] X. Chen, C. Wu, Z. Zhang, and J. Liu, "Multi-class financial distress prediction based on stacking ensemble method," *International Journal of Finance & Economics*, vol. 30, no. 3, pp. 2369–2388, 2025, https://doi.org/10.1002/ijfe.3020.

[11] J. Liu and Z. Wang, "Financial distress prediction using an improved particle swarm optimization wrapper feature selection method and tree boosting ensemble," *Journal of the Operational Research Society*, vol. 76, no. 4, pp. 617–640, Apr. 2025, https://doi.org/10.1080/01605682.2024.2385467.

[12] C. Wu, C. Jiang, Z. Wang, and Y. Ding, "Predicting financial distress using current reports: A novel deep learning method based on user-response-guided attention," *Decision Support Systems*, vol. 179, Apr. 2024, Art. no. 114176, https://doi.org/10.1016/j.dss.2024.114176.

[13] Y. Chen, X. Kuang, and J. Guo, "LiFoL: An Efficient Framework for Financial Distress Prediction in High-Dimensional Unbalanced Scenario," *IEEE Transactions on Computational Social Systems*, vol. 11, no. 2, pp. 2784–2795, Apr. 2024, https://doi.org/10.1109/TCSS.2023.3276059.

[14] L. Yu and M. Li, "A case-based reasoning driven ensemble learning paradigm for financial distress prediction with missing data," *Applied Soft Computing*, vol. 137, Apr. 2023, Art. no. 110163, https://doi.org/10.1016/j.asoc.2023.110163.

[15] D. A. Abdullah and N. J. Al-Anber, "Implement data mining and deep learning techniques to detect financial distress," *AIP Conference Proceedings*, vol. 2591, no. 1, Mar. 2023, Art. no. 020009, https://doi.org/10.1063/5.0119272.

[16] S. Knifo and A. Alzubi, "DR-Z2AN: dual-recurrent neural network with a tri-channel attention mechanism for financial management prediction," *Complex & Intelligent Systems*, vol. 11, no. 1, Nov. 2024, Art. no. 53, https://doi.org/10.1007/s40747-024-01613-x.

[17] U. S. Bist and N. Singh, "Enhancing Support Vector Machine Performance: A Hybrid Approach with Davidon-Fletcher-Powell Algorithm and Elephant Herding Optimization (EHO-DFP) for Parameter Optimization," *International Journal of Maritime Engineering*, vol. 1, no. 1, pp. 127–140, Jul. 2024, https://doi.org/10.5750/ijme.v1i1.1345.

[18] I. K. Friday, S. P. Pati, D. Mishra, P. K. Mallick, and S. Kumar, "CAGTRADE: Predicting Stock Market Price Movement with a CNN-Attention-GRU Model," *Asia-Pacific Financial Markets*, vol. 32, no. 2, pp. 583–608, Jun. 2025, https://doi.org/10.1007/s10690-024-09463-w.

[19] K. Mutreja, A. R. Panda, and M. K. Mishra, "Predicting Stock Price using Att-Gru-Ols with Vader Sentimental Analysis," *Indian Journal Of Science And Technology*, vol. 18, no. 25, pp. 2047–2057, Jun. 2025, https://doi.org/10.17485/IJST/v18i25.112.

[20] Q. Yang, "Financial Distress Prediction by using Long Short-Term Memory based Adaptive Whale Optimization Algorithm," in *2025 International Conference on Intelligent Systems and Computational Networks (ICISCN)*, Bidar, India, Jan. 2025, pp. 1–5, https://doi.org/10.1109/ICISCN64258.2025.10934466.

[21] R. Quinlan, "Statlog (Australian Credit Approval)." UCI Machine Learning Repository, 1987, https://doi.org/10.24432/C59012.

[22] M. El-Bannany, M. Sreedharan, and A. M. Khedr, "A robust deep learning model for financial distress prediction," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 2, 2020.

[23] R. Gao, S. Cui, Y. Wang, and W. Xu, "Predicting financial distress in high-dimensional imbalanced datasets: a multi-heterogeneous self-paced ensemble learning framework," *Financial Innovation*, vol. 11, no. 1, Jan. 2025, Art. no. 50, https://doi.org/10.1186/s40854-024-00745-w.