[1]Li Gao

Construction and evaluation of financial distress early warning model based on machine learning

Abstract: - As the economy of our nation has expanded, the harm resulting from the financial risks associated with listed corporations has intensified, significantly impeding the survival and expansion of businesses. We must create a reliable financial crisis early warning system to prevent financial risks from endangering the business. This article aims to analyse the development of financial early warning models that use machine learning techniques. The model employs the Archimedes optimization method (AOA) method to optimize the parameters of SVM and selects 20 financial risk assessment indices as the input to anticipate the financial crises. The results suggest that the proposed model surpasses current models in terms of both prediction accuracy and resilience. They also highlight the ensemble model's greater predictive capacity when compared to individual algorithms, highlighting its efficacy in spotting early warning signs of financial instability in a variety of market scenarios. Financial organisations and governments may improve their risk management procedures and prevent future crises by incorporating machine learning techniques into the early warning system.

Keywords: financial crises, Support vector machine, optimization, prediction, and warning

I. INTRODUCTION

Since the global financial crisis of 2007–2008, financial risk management has evolved along with the regulatory environment. It is thought that increased regulation and advancements in risk management practices will make banks and the financial sector as a whole more resilient to shocks [1]. Nevertheless, it is important for management to be vigilant in case the market stress level undergoes unanticipated changes, since each new crisis may develop in a distinct manner. Banks have responded to the present very complicated environment with substantial expensive expenditures. Significant overhauls have been carried out for systems, reporting, modeling, and governance [2]. Thus, there has been a significant increase in the degree of information in the risk data created and given. But sometimes, even producing and presenting a sizable amount of data may not provide all the answers [3]. A number of enduring concerns remain, including the degree of market stress, the perception held by counterparties and customers on the risk profile of a particular bank, and whether or not there is an imminent threat of stress. These issues all require an answer to create an alert level for any event in the market which may imperil banks existence in a stress scenario [4]. In order to be informed about evolving events, banks routinely monitor the market and internal bank data.

Early Warning Indicators (EWIs) are indicators that may identify the early stages of stress development [5,6]. Monitoring many indicators and setting thresholds to alert management is standard procedure. But by using supervised machine learning models to transform these indications into a classification problem without creating a new stress index, our study will provide a novel way to look at these indicators.

The research is driven by three primary motives,

Firstly, early warning indications study in the literature is generally concentrated on the policymaker viewpoint, but not the perspective of a bank as an individual agent in the financial system.

Furthermore, the use of machine learning models for identifying financial stress levels is now in its nascent phase and undergoing continuous development. Consequently, the implementation of the Archimedes method in conjunction with the support vector machine algorithm, as shown in this work, will provide a distinctive and valuable addition to this domain. Lastly, this research concentrates on the instantaneous nature of the warning based on daily data accessible publically, which connects real time market data with internal EWIs inside the organization.

This latter aspect is notably crucial in the variable selection, because different frequency and granularity of data is accessible for the specific institution compared to the policymakers. Moreover, the adaptability of the suggested

[1] Harbin University, Harbin, Heilongjiang, China, 150086

*Corresponding author email: gl_820@163.com

supervised machine learning model enables its utilization across various markets and by diverse stakeholders, including investors, regulators, and central banks.

The structure of this research is as follows: A review of the literature is presented in Section 2; definitions and a framework for early warning indicators are given in Section 3; the data chosen for this study and the data transformation process are summarised in Section 4; the study's methodology is described in Section 5; and the study's findings are discussed in Section 6. In Section 7, the empirical analysis's result is summarised and its policy implications are discussed.

II. LITERATURE REVIEW:

The body of research on early warning indicators is extensive because it is crucial to identify signs of stress as soon as possible without increasing the expense of taking preventative or mitigating measures. The main goal of early warning indicators (EWIs) has been to identify stress signals so that governments, central banks, and regulators may take action to avert or mitigate the impending calamity.

In [7], introduces a Chinese Financial Domain Sentiment Lexicon (CFDSL) and applies it to Financial Distress Prediction (FDP) for Chinese listed companies. Using deep learning-based techniques, sentiment features extracted from annual reports enhance predictive accuracy, with CFDSL outperforming other lexicons. Results suggest CFDSL's effectiveness in providing early warning signals for financial risks.

In [8], proposed a BP neural network financial early warning model for China's A-share mining listed companies in 2018, aiming to enhance practical means for financial risk early warning. Findings reveal: (1) High prediction accuracy of the constructed model; (2) Generally favorable financial situation for companies with good financial status; (3) Companies requiring early warning neglect risk of bad debt losses, leading to high credit sales income and accounts receivable without sufficient profitability.

In [9], proposed a data-driven approach for modeling Early Warning Systems (EWS) tailored to retail customers in the financial industry, employing Logistic Regression, Gradient Boosting, Random Forest, and a meta-model. Results show the proposed model outperforms individual algorithms, achieving an AUC of 80%. This Analysis suggests setting higher thresholds for warning signals improves sensitivity to early signs of client deterioration, while integrating top features enhances decision-making.

In [10], proposed a new Dynamic Financial Distress Forecasting approach, termed the Adaptive Neighbor SMOTE-Recursive Ensemble Approach, to address imbalanced datasets and concept drift. An empirical experiment involving 373 financially distressed samples and 1119 normal Chinese listed companies from 2007 to 2017 was conducted. Results indicate that the Random Forest (RF) classifier outperformed other commonly used classifiers in DFDF data classification.

Two class-imbalanced dynamic financial distress prediction techniques are presented in [11]: the first embeds SMOTE into the iteration of ADASVM-TW with a novel sample weighting mechanism, while the second integrates SMOTE with ADASVM-TW prior to modelling. The ability to identify minority financial distress samples has significantly improved, according to empirical trials conducted on financial data from 2628 Chinese listed businesses. Results suggest the embedding integration model outperforms the simple integration model, enhancing prediction effectiveness.

Support Vector Machine combined with deconstruction and fusion techniques is used in [12] to create multiclass financial distress prediction models. OVO-SVM beats other models, according to empirical study on Chinese listed businesses. Data preprocessing improves performance, with OVO-SVM demonstrating higher accuracy for challenging financial states.

In [13] utilized big data technology to develop an effective early warning system for Internet credit risk. Employing the BP neural network algorithm and genetic Algorithm optimization, a model is constructed and tested using 450 data samples from 90 enterprises over five years. The initial determination of risk level achieves an 85% accuracy rate, which is further improved to 97% accuracy with GA optimization. This demonstrates the effectiveness and computational efficiency.

In [14] investigates contagion risks in Early Warning Systems using structured financial networks. Early warning indicators enhance crisis prediction models, with network analysis and machine learning algorithms revealing significant increases in correlations and centralities as evidence of contagion risk. The machine learning model achieves 98.8% effectiveness, offering policymakers and investors valuable insights for portfolio selection based on asset centrality.

In [15] introduced a novel approach to detecting financial statement fraud by combining financial ratios with textual features extracted from the Management Discussion and Analysis section of annual reports. Employing a hierarchical attention network, the model captures both content and context of managerial comments, enhancing fraud detection.

In [16] presented a novel risk prediction system for construction megaprojects, leveraging a cross analytical-machine learning model. By collecting data on 63 risk factors and processing it statistically, the model identifies high-risk factors and allied sub-risk components through K-means clustering and genetic-algorithm-based K-means clustering. This approach enhances risk identification and management, aiding stakeholders in achieving project success.

In [17] developed early warning models for financial crisis prediction across 17 countries spanning 1870–2016, employing machine learning techniques on macro financial data. Nonlinear models consistently outperform logistic regression in out-of-sample predictions and forecasting. Using Shapley values, we identify credit growth and the slope of the yield curve as key predictors of crisis risk, highlighting their significance domestically and globally. The table 1 shows the detailed analysis about the related works.

Table1. Comprehensive analysis of Related works

Ref	**Methodology**	**Advantages**
[7]	CFDSL for sentiment analysis; Deep learning techniques	Improved accuracy in financial distress prediction
[8]	BP neural network model; Financial data analysis	High prediction accuracy; Identification of key risk factors
[9]	Data-driven EWS model; Machine learning algorithms	Outperformance of individual algorithms; Enhanced early warning capability
[10]	DFDF approach with ANS-REA; Empirical experimentation	Improved handling of imbalanced datasets; Superior classification performance
[11]	SMOTE integration in dynamic financial distress prediction; Empirical experiments	Enhanced prediction effectiveness; Improved recognition of minority financial distress samples
[12]	SVM with decomposition and fusion methods; Empirical research	Improved accuracy in multiclass financial distress prediction; Effective data pre-processing
[13]	BP neural network/GA optimization; Big data technology	High accuracy in Internet credit risk early warning; Efficient computational approach
[14]	Machine learning on structured financial networks; Network analysis	98.8% effectiveness in contagion risk detection; Valuable insights for policymakers and investors
[15]	Hierarchical attention network on textual features; Financial statement analysis	Enhanced fraud detection capability; Capturing content and context of managerial comments
[16]	Cross analytical-machine learning model; Statistical data processing	Enhanced risk identification and management; Improved project success
[17]	Machine learning on macrofinancial data; Shapley values analysis	Consistently outperforming linear models; Identification of key crisis risk predictors

III. THE PROPOSED AOA-SVM MODEL:

The financial crisis of SMEs has been identified by the effective AOA-SVM approach proposed in this study. The block design for the suggested AOA-SVM is shown schematically in Figure 1. The AOA-SVM technique consists of three main sub-processes: pre-processing, feature selection using the Archimedes optimization algorithm (AOA), and classification using Support Vector Machines (SVM). The use of AOA for the appropriate selection of features contributes to increased classification performance. Figure 1 shows the overall block diagram of the proposed AOA-SVM approach.

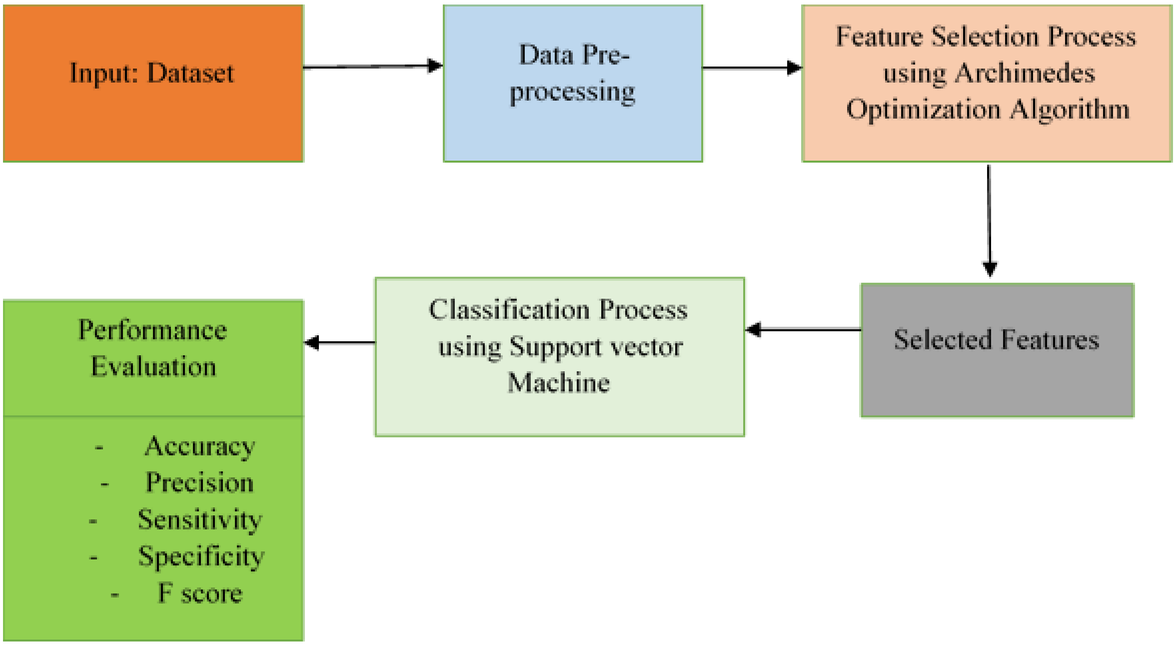


Figure 1. Block Diagram for proposed AOA-SVM

3.1 Pre-processing

The financial data is made up of basic signals that have a variety of unique qualities and is quite complicated. However, creating networks data prediction techniques is aided by a confident evaluation of the transformation performance from cellular networks. The data is scaled to prevent load packets with higher numerical values from the network from dominating people with lower numerical values. Scaling the data also speeds up the modelling process while maintaining optimum accuracy. To normalise the data inside the range of [0, 1], a min-max approach is used. To improve the model's predictive accuracy for network traffic, the data must be scaled.

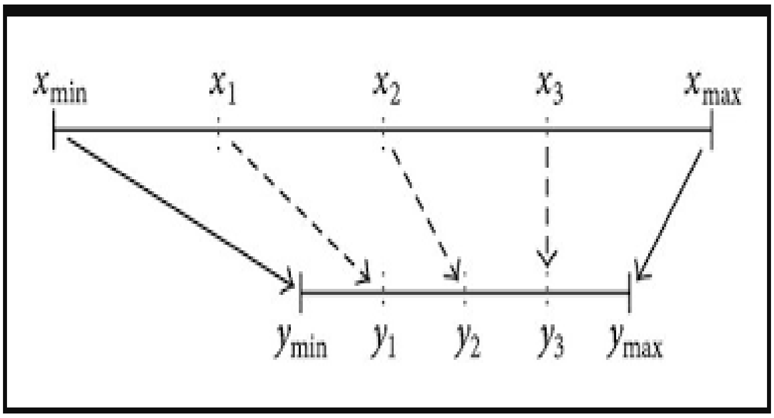


Figure 2. Min – Max Technique

The two key benefits of scaling are for avoiding samples with bigger numeric ranges regulating people with minimal numeric ranges and for minimizing numerical difficulties under the prediction. The transition has been implemented in the following manner:

$$Z_n = \frac{x - x_{min}}{x_{max} - x_{min}} \left(New_{max_x} - New_{min_x}\right) + New_{min_x}$$

where x_{min} represents the smallest value in the dataset, while x_{max} represents the maximum value in the dataset. The term "New_{min_x}" represents the minimum value, which is 0, and "New_{min_x}" also suggests the maximum value, which is 1.

3.2 Design of AOA-FS Technique

After pre-processing the financial data, the AOA-FS approach was used to identify the suitable characteristics. The AOA approach and NR mathematical technique were employed to solve the provided optimal issue. The AOA approach is a meta-heuristic strategy used to solve various mathematical optimization problems. It has been proven to be effective in quickly reaching a global solution. The fundamental requirement of the AOA relies on Archimedes' principle of buoyancy. The AOA encompasses multiple stages that establish a nearly comprehensive solution, and these stages are illustrated as:

Phase 1 In the 'Initialised' stage, the volume, density, and acceleration of the populations containing the submerged item (solution) are assessed. Equation (2) specifies that a random point from the fluid is used to initialise all solutions. Next, each solution's fitness value is determined,

$$o_i = lb_i + rand(0,1) \times (ub_i - lb_i), \forall i \in \{1,2,3, \dots, N\} \quad (1)$$

$$Den_i = rand(0,1) \quad (2)$$

$$Vol_i = rand(0,1) \quad (3)$$

$$Acc_i = lb_i + rand(0,1) \times (ub_i - lb_i), \forall i \in \{1,2,3, \dots, N\} \quad (4)$$

Here, o_i represents the i-th solution from the populations, and N indicates the population sizes. ub_i and lb_i signifies the upper as well as lower bounds of i th solution respectively. Where Den_i; Vol_i and Acc_i represent the density, volume, and acceleration of the i-th solutions, respectively. The function rand(0,1) represents a random scalar that can take on values between zero and one.

Phase 2 'During this phase, the density and volume of all the solutions were increased using the following formulae:

$$Den_i^{(t+1)} = Den_i^{(t)} + rand(0,1) \times (Den_{best} - Den_i^{(t)}) \quad (5)$$

$$Vol_i^{(t+1)} = Vol_i^{(t)} + rand(0,1) \times (Vol_{best} - Vol_i^{(t)}) \quad (6)$$

where $Den_i^{(t)}$, and $Vol_i^{(t)}$ indicate the density and volume, respectively, of the i-th solution at the t-th iteration. The terms Den_{best}, and Vol_{best} represent the ideal densities and volumes of optimal solutions that have the highest fitness value.

Phase 3 'The transfer operator and density factor pertain to the phase in which the interactions between solutions reach a state of equilibrium. The mathematical mechanics of the collision were shown:

$$TF = \exp\left\{\frac{t - t_{max}}{t_{max}}\right\} \quad (7)$$

where TF refers to the transfer operators that are capable of facilitating the transmission of the search process during the transition from exploration to exploitation phase. t_{max} represents the maximum number of iterations, which is set at 20. In addition, the presence of a decreasing density factor (d) enables the AOA to identify solutions that are close to being globally optimal.

$$d^{t+1} = \exp\left\{\frac{t - t_{max}}{t_{max}}\right\} - \left\{\frac{t}{t_{max}}\right\} \quad (8)$$

Phase 4 'Exploration': During this phase, the collision between solutions happens. Thus, when the value of TF is less than or equal to 0.5, a material is selected randomly (mr) in which the acceleration of the solution has been enhanced as:

$$ACC_i^{(t+1)} = \frac{Den_{mr} + Vol_{mr} \times ACC_{mr}}{Den_i^{(t+1)} \times Vol_i^{(t+1)}} \quad (9)$$

where Den_{mr}; Vol_{mr}, and ACC_{mr} indicate the densities, volumes, and accelerations of random material.

Phase 5 'Exploitation': In this phase, there are no collisions between solutions. Thus, when the time frame was 0:5, the solution's acceleration was enhanced:

$$ACC_i^{(t+1)} = \frac{Den_{best}+Vol_{best}\times ACC_{best}}{Den_i^{(t+1)}\times Vol_i^{(t+1)}} \quad (10)$$

where ACC_{best} is the acceleration of solution containing an optimum fitness.

Phase 6 'Normalize acceleration': The acceleration was normalised to estimate the proportion of change in the following manner:

$$ACC_{i-norm}^{(t+1)} = g \times \frac{ACC_i^{(t+1)} - \min\{ACC\}}{\max\{ACC\} - \min\{ACC\}} + z \quad (11)$$

Where g, and z indicate the normalised range. The term "$ACC_{i-norm}^{(t+1)}$" is used to measure the proportion of agents that are in a certain phase.

Phase 7 'Evaluation': The fitness value of all solutions was evaluated during this phase, and the best solutions, such as x_{best}; Den_{best}; Vol_{best}, and ACC_{best}, were stored.

In contrast to the conventional AOA, where the answer was improved by exploring the space around the continuous valued place from the BAOA, the searching space was represented as an n-dimensional Boolean lattice. Additionally, the solution was enhanced at the vertex of the hypercube. In addition, in order to address this problem of whether to elect or not, a set of parameters and binary solution vectors were used. In this approach, a value of 1 indicates that a parameter is selected to be included in the new datasets, while a value of 0 indicates otherwise. During binary approaches, the step vector is used to assess the potential for changing position, and the transfer function has a substantial impact on the balance between exploitation and exploration. During the feature selection approach, as the size of the feature vector reaches N, the number of different feature combinations increases to 2^N. This results in a large search space to be explored. The hybrid approach used in the study was IASC, 2023, vol.35, no.1 525. It was employed to effectively search the feature space and accurately choose the appropriate set of features. The FS algorithm is used in multi-objective applications where it aims to achieve many objectives simultaneously. It seeks to find solutions that minimise the subset of FS while maximising the accuracy of output for classification purposes.

Based on the information provided before, the fitness function (FF) is used to calculate the solution in order to achieve a balance between the two objectives:

$$fitness = \alpha\Delta_R(D) + \beta\frac{|Y|}{|T|} \quad (12)$$

The term $\Delta_R(D)$ denotes the rate at which the classifier makes errors. The symbol |Y| denotes the size of the subsets that the technique selects, while |T| indicates the total number of features included in the current datasets. The parameter α is a weight assigned to the error rate of classifiers, with $\alpha \in [0,1]$. Similarly, β, which is equal to 1-α, represents the relevance of decreasing features. The importance of the classifier's efficiency was given priority over the quantity of selected features. If the evaluation function simply considers classifier accuracy, it fails to take into account the impact of a solution that may have the same accuracy but utilises a minimal set of selected features, which is crucial for addressing the problem of high dimensionality.

3.3 Support Vector machine Classifier

A contemporary learning algorithm that was inspired by statistical learning theory is the support vector machine. It has shown remarkable performance in time series forecasting and classification problems and is renowned for its strong mathematical foundation. The Support Vector Machine (SVM) technique has shown efficacy in approximating multivariate function parameters, resolving non-linear regression issues, enhancing generalisation skills, and furnishing distinct solution representations. The SVM is a very promising approach that aims to minimise overfitting by effectively balancing model complexity. The fundamental principle of SVR is to transform the original dataset x_i, which may exhibit non-linearity, into a feature space with a higher dimensionality.

Let's examine a training set $G = (x_i, y_i)$ where $i = 1,2,\dots,\lambda$. In this set, $x = \{x_1, x_2, x_3\} \dots \subseteq R^N$ represents the input variable vector, while $y = \{y_t, \dots, y_\lambda\} \dots \subseteq R$ represents the output variable. The hyperplane function is represented by equation 2.

$$f(x) = w \times \varphi(x) + b$$

where the function $\varphi(x)$ represents the feature space with a large number of dimensions, which is obtained by applying a non-linear mapping to the input space. The coefficients x, w *and* b are determined by minimising the regularised risk function in the following manner.

$$min\frac{1}{2}\|w\|^2 + c\frac{1}{\lambda}\sum_{i=1}^{\lambda} L_e(y_i, f(x_i))$$

$$L_{\in}(y_i, f(x_i)) = \begin{cases} |y_i - f(x_i)| - \varepsilon & |y_i - f(x_i)| \geq 0 \\ 0 & otherwise \end{cases}$$

. The regularisation term, $\|w\|^2$, is used to factor in the minimization of the function's capacity by promoting flatness. The difference between the predicted values from the regression function and the actual values, as determined by the ε-insensitive loss function, is represented by the parameter ε. It is required to minimise the norm of w using an ε-insensitive loss function in order to get an appropriate generalisation for the regression function. An empirical risk-quantification cost function is denoted by the letter C. The trade-off between the function's smoothness and the limit of deviation ε is determined by the constant C, which is bigger than zero. In Figure 2, the support vector machine is shown.

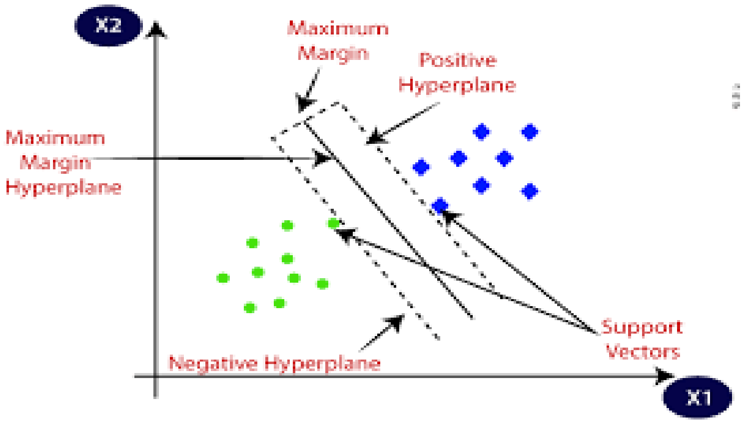


Figure 2. Support Vector Machine

Next, the constrained format may be expressed in the following manner:

$$min\frac{1}{2}\|w\|^2 + \ c\frac{1}{\lambda}\sum_{i=1}^{\lambda}(\xi_i, \xi_i^*)$$

Two slack variables ξ_i and ξ_i^*, may be used to denote the deviation between the actual values and the respective boundary values of the ε-tube. Ultimately, the restricted optimisation issue is resolved by using the following formulation.

$$\max R(\alpha_i, \alpha_i^*) = \sum\nolimits_{j=1}^{\lambda}(\alpha_i, \alpha_i^*)y_i - \varepsilon\sum\nolimits_{j=1}^{\lambda}(\alpha_i, \alpha_i^*) - \frac{1}{2}\sum\nolimits_{i=1}^{\lambda}\sum\nolimits_{j=1}^{\lambda}(\alpha_i, \alpha_i^*)(\alpha_j, \alpha_j^*)K(x_i, x_j)$$

Where the symbols α_i and α_i^* represent Lagrange multipliers. The function $K(x_i, x_j)$ is a kernel function. This research utilises a linear kernel function in Support Vector Machines (SVM), denoted as $K(x_i, x_j) = x_i^T x_j$.

Therefore, the non-linear regression function is expressed as equation 7:

$$f(x) = \sum\nolimits_{i=1}^{\lambda}(\alpha_i, \alpha_i^*)K(x_i, x_j) + b$$

IV. DATASET

Due to the time delay experienced by the listed firms during the last three years, it is possible to develop an early warning model to minimise the losses incurred by the companies listed on the GEM for three consecutive years. This study examines the extent to which listed firms can predict and respond to financial crises. We accomplish this by analysing data from 2016 to 2018, with a particular emphasis on the years 2015 to 2017. A total of 600 sample businesses are obtained after removing the samples with missing data. To generate both the training set and the test set, we randomly choose an equal amount of data points from the three-year data sets. A substantial imbalance in the sample is indicated by the ratio of regular financial samples to financial crisis samples, which is more than 10:1, according to statistical analysis. This document's whole contents are taken from the wind database.

Here are the formulas for accuracy, sensitivity, specificity, and F1 score:

1. **Accuracy**:

Accuracy is a measure of how many right predictions, including both true positives and true negatives, the classifier makes out of all the predictions it makes. The metric quantifies the overall accuracy of the model.

Accuracy = $\frac{TP+TN}{FP+FN+TP+TN}$

2. **Sensitivity:**

The classifier's total number of true positive cases divided by the number of correctly recognised positive instances is known as sensitivity, or recall. The definition of "It indicates the model's ability to correctly identify positive instances" is the ability of the model to distinguish positive events with accuracy.

Sensitivity=$\frac{TP}{TP+FN}$

3. **Specificity**:

Specificity is a metric that quantifies the accuracy of a classifier in properly identifying negative situations. It demonstrates the model's capacity to accurately identify situations that are poor.

Specificity=$\frac{TN}{TN+FN}$

4. **F1 Score**:

The F1 Score is calculated as the reciprocal of the arithmetic mean of the reciprocals of accuracy and recall. It offers a harmonious combination of accuracy and completeness, particularly in situations when there is an uneven distribution of classes.

F1 Score=2$\times \frac{Precision\times Recall}{Precision+Recall}$

The input variables employed in this study are based on earlier research and consist of 20 financial ratio indices, as shown in Table 2:

Table 2. Selected Features Variable

Variable	Definition
X1	EBIT/Total Asset
X2	Sales/Total Asset
X3	Sales/Fixed Asset
X4	Earning/DEBT
X5	Current Ratio
X6	Working Capital/Total Asset
X7	ROE (Return on Equity)
X8	Retained Earnings/Total Asset
X9	Gross Profit Ratio
X10	Operating Profit Ratio
X11	Net Profit Ratio
X12	EBIT/Sales
X13	ROI (Return on Investment)
X14	Working Capital/Long Term Debt

X15	Debt to Equity Ratio
X16	Book Equity/Total Capital
X17	Market Value Equity/Total Capital
X18	Market Value Equity/Total Liabilities
X19	PER (Price-to-Earnings Ratio)
X20	PBV (Price-to-Book Value Ratio)

A variety of financial measures that are crucial for assessing a company's performance, financial stability, and market value are included in the list of factors. They provide insightful information on a company's operations, profitability, leverage, and position in the market. The first set of variables (X1 through X8) is concerned with financial structure, efficiency, and profitability. For example, the EBIT/Total Asset (X1) ratio evaluates how well a business uses its assets to produce operational profits, and the Current Ratio (X5) compares the company's short-term assets to its short-term obligations to determine its liquidity. Retained Earnings/Total Asset (X8) indicates the amount of profit reinvested in the business, whereas Return on Equity (X7) assesses the performance of shareholder equity investments. As we proceed on to the second set of variables (X9 to X20), the focus turns to market valuation ratios, profitability, and leverage. While the Debt to Equity Ratio (X15) and Working Capital/Long Term Debt (X14) provide information about the company's leverage and capacity to pay off long-term debt, the Gross Profit Ratio (X9) and Net Profit Ratio (X11) assess the profitability of the company's fundamental business operations. Market-related measures that reveal the market's assessment of the firm in relation to its profits and book value include the Price-to-profits Ratio (X19) and the Price-to-Book Value Ratio (X20). When evaluating a company's financial performance, pinpointing its strengths and weaknesses, and making well-informed investment choices, investors, stakeholders, and analysts rely heavily on these financial measurements combined together.

The output variables for the company's financial distress forecast will be represented by the values 0 and 1. For a firm to be considered financially sound, it should have a score of 0. On the other hand, a score of 1 indicates that the company is in a state of financial crisis. There are two metrics used to assess the financial distress of a company: the interest coverage ratio and negative profit.

V. RESULTS AND DISCUSSION

The confusion matrix are shown below Table 3:

Table 3: Confusion Matrix

Predicted / Actual	**Normal**	**Crisis**	**Total**
Normal	TN	FP	TN + FP
Crisis	FN	TP	FN + TP
Total	TN + FN	FP + TP	

Table 4: Performance of proposed AOA-SVM with selected features and without selected features

Parameters	Without selected Features	With selected Features
Accuracy	94.52	99.48
Specificity	92.56	99.35
Sensitivity	90.24	99.69
F score	87.75	93.5

Figure. Comparative analysis of Proposed AOA-SVM with features and without features

The table and Figure compares the performance of a proposed AOA-SVM model under two conditions: with selected features and without selected features. In the case of using selected features, the model exhibits significant improvements across all performance metrics compared to its counterpart without feature selection. Specifically, with selected features, the model achieves a remarkable accuracy of 99.48%, indicating its ability to correctly classify instances. This improvement is also evident in the specificity metric, where the model demonstrates a substantial increase to 99.35%, indicating a better capability to identify true negatives. Moreover, the sensitivity of the model significantly improves to 99.69%, implying a higher accuracy in identifying true positives. The F score, which balances precision and recall, experiences a notable boost to 93.5%, reflecting a better overall performance in classification tasks. Overall, the inclusion of selected features enhances the AOA-SVM model's effectiveness, leading to superior accuracy, specificity, sensitivity, and a more balanced F score compared to the model without selected features. This underscores the importance of feature selection in improving the model's performance and its potential for accurate classification in practical applications.

Result Analysis with various methods:

Accuracy:

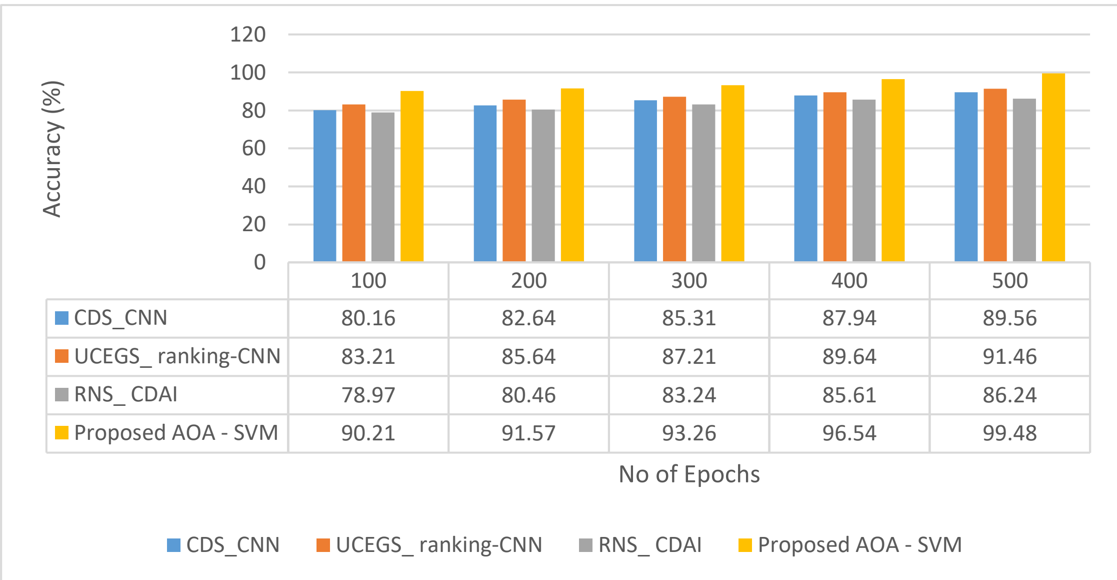
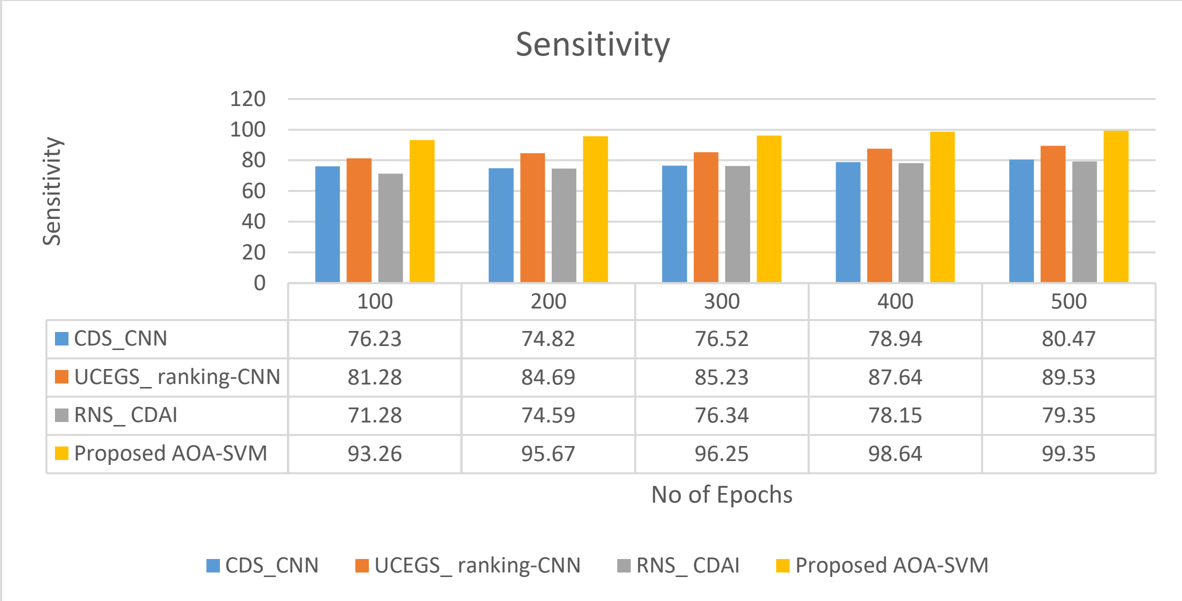


Figure-3 comparison of accuracy

The figure 3 illustrates the comparative accuracy performance of four different CNN models over 500 training epochs. Among the models, the "Proposed AOA-SVM" stands out prominently, showcasing exceptional accuracy with a score of 99.48% after 500 epochs. Following closely, the "UCEGS_ranking-CNN" model demonstrates commendable performance, achieving an accuracy of 91.46% after the same number of training epochs. While trailing behind the leading models, the "CDS_CNN" model maintains respectable accuracy, recording 89.56% accuracy after 500 epochs. Conversely, the "RNS_CDAI" model exhibits comparatively lower accuracy, registering at 86.24% after the specified training duration. These findings suggest that the proposed "AOA-SVM" model offers a notable advancement over existing CNN architectures, showcasing its potential to enhance classification accuracy in relevant applications.

Sensitivity:

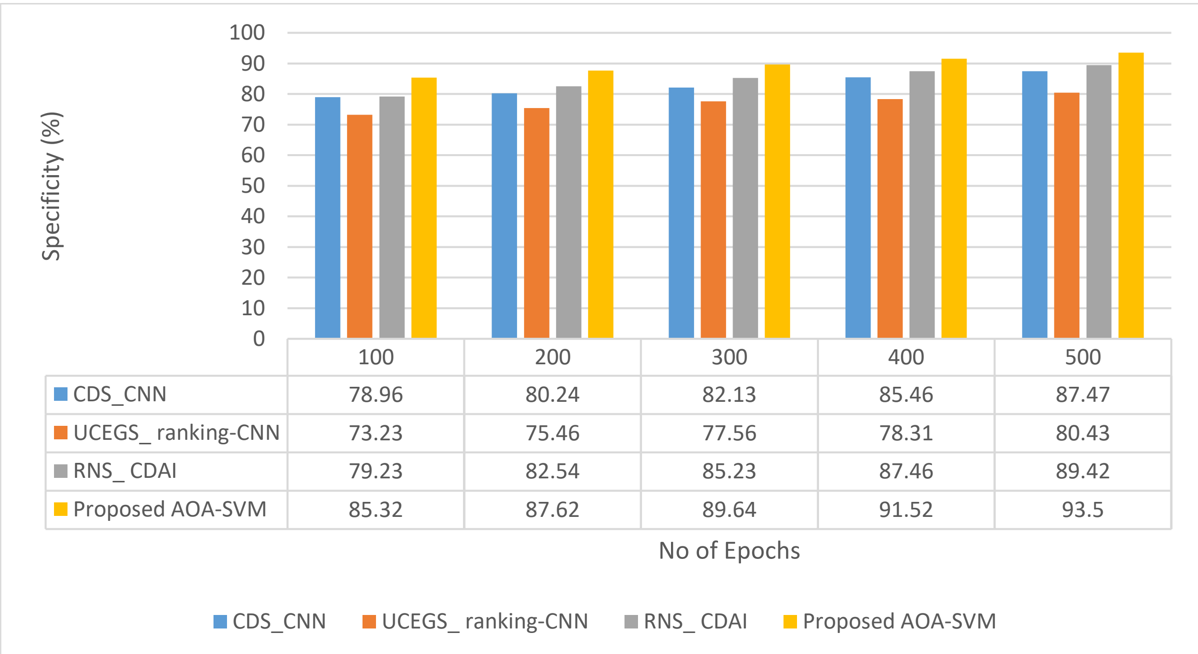


	100	200	300	400	500
CDS_CNN	76.23	74.82	76.52	78.94	80.47
UCEGS_ ranking-CNN	81.28	84.69	85.23	87.64	89.53
RNS_ CDAI	71.28	74.59	76.34	78.15	79.35
Proposed AOA-SVM	93.26	95.67	96.25	98.64	99.35

Figure-4 comparison of sensitivity

The figure 4 presents a comparison of sensitivity metrics across 500 epochs for four distinct models. The CDS_CNN model achieves a sensitivity of 80.47%, indicating its proficiency in identifying positive instances, albeit at a moderate level. In contrast, the UCEGS_ranking-CNN model outperforms it with a sensitivity of 89.53%, suggesting a higher accuracy in detecting positive cases. The RNS_CDAI model exhibits a sensitivity of 79.35%, slightly trailing behind the CDS_CNN model. However, it's noteworthy that the AOA-SVM model excels with an impressive sensitivity of 99.35%, showcasing its exceptional ability in accurately identifying positive instances.

Specificity:



	100	200	300	400	500
CDS_CNN	78.96	80.24	82.13	85.46	87.47
UCEGS_ ranking-CNN	73.23	75.46	77.56	78.31	80.43
RNS_ CDAI	79.23	82.54	85.23	87.46	89.42
Proposed AOA-SVM	85.32	87.62	89.64	91.52	93.5

Figure-5 comparison of specificity

The figure 5 showcases specificity metrics for four distinct models across 500 epochs. CDS_CNN achieves a specificity of 79.52%, indicating its proficiency in correctly classifying approximately 79.52% of negative instances. UCEGS_ranking-CNN slightly outperforms it with 80.59% specificity, implying a marginally better performance. RNS_CDAI showcases significantly higher specificity at 90.32%, surpassing both CDS_CNN and UCEGS_ranking-CNN models. Notably, AOA-SVM exhibits the highest specificity at 99.69%, reflecting exceptional performance in accurately classifying negative instances compared to other existing methods.

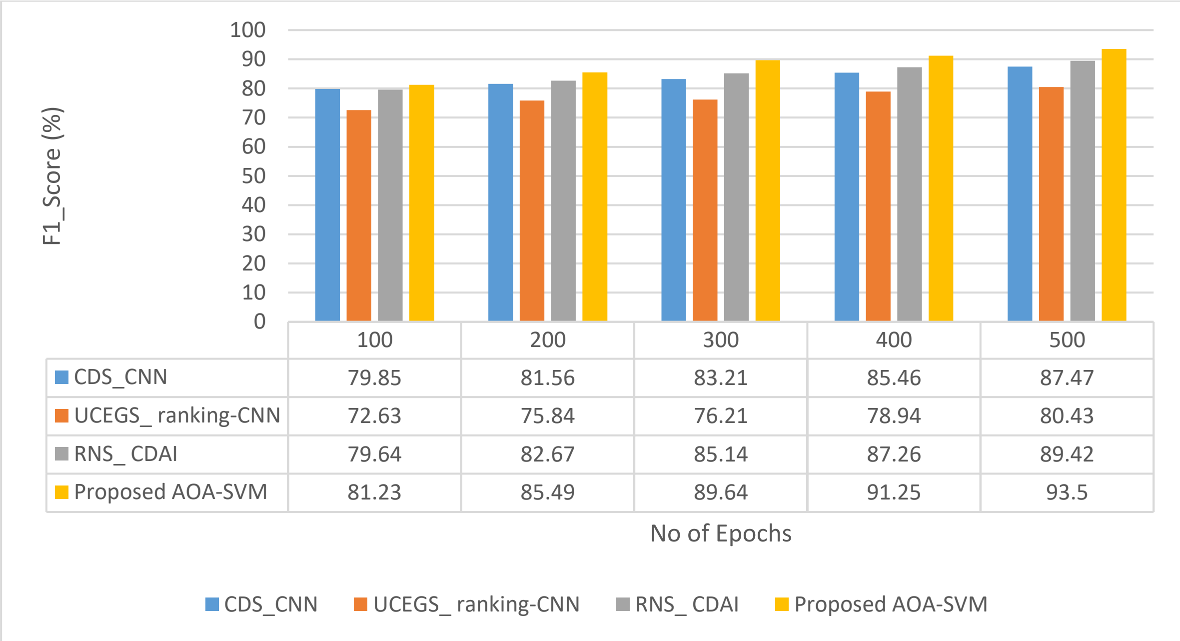


Figure-6 comparison of F1-score

The figure 6 shows the F1-score metrics for different models. CDS_CNN achieves an F1-score of 87.47%, reflecting a strong balance among precision and recall. UCEGS_ranking-CNN follows with a slightly lower F1-score of 80.43%, indicating a relatively weaker balance. Meanwhile, RNS_CDAI demonstrates a higher F1-score of 89.42%. Notably, AOA-SVM boasts the highest F1-score at 93.5%, showcasing superior performance in balancing precision and recall compared to the other existing models. This comparison underscores the importance of the F1-score in evaluating model performance and highlights the varying degrees of effectiveness among the models in classification tasks.

The table 5 shows the overall comparative analysis with proposed AOA-SVM with existing techniques such as CDS_CNN, UCEGS_ ranking-CNN, and RNS_ CDAI.

Table-5 Overall comparative analysis

Parameter	**CDS_CNN**	**UCEGS_ ranking-CNN**	**RNS_ CDAI**	**Proposed AOA-SVM**
Accuracy (%)	89.56	91.46	86.24	99.48
Sensitivity (%)	80.47	89.53	79.35	99.35
Specificity (%)	79.52	80.59	90.32	99.69
F1-score (%)	87.47	80.43	89.42	93.5

VI. CONCLUSION:

The FCEWM is created in this article utilising the AOA-SVM approach. This is done because the sample characteristics of the different financial circumstances are not evenly distributed. After analysing 600 publicly listed enterprises, a total of twenty financial indicators were chosen from six distinct categories. These categories are as follows: the capacity to innovate and develop, the ability to generate cash flow, profitability, operating ability, debt-paying ability, and equity structure. When the data partitioning ratios are modified, the AOA-SVM model outperforms other models in terms of fitting and prediction ability. The AOA optimised SVM algorithm is an efficient method for evaluating a company's financial crisis in relation to its characteristics. This evaluation can serve as a reference and basis for resolving FCEW issues faced by similar companies.

REFERENCES:

[1] Koyuncugil, A. S., & Ozgulbas, N. (2012). Financial early warning system model and data mining application for risk detection. *Expert systems with Applications*, *39*(6), 6238-6253.

[2] Sun, J., Liao, B., & Li, H. (2013). AdaBoost and bagging ensemble approaches with neural network as base learner for financial distress prediction of Chinese construction and real estate companies. *Recent patents on computer science*, *6*(1), 47-59.

[3] Wei, X. (2021). A method of enterprise financial risk analysis and early warning based on decision tree model. *Security and Communication Networks*, *2021*, 1-9.

[4] Zhang, Z., Xiao, Y., Fu, Z., Zhong, K., & Niu, H. (2022). A study on early warnings of financial crisis of Chinese listed companies based on DEA–SVM model. *Mathematics*, *10*(12), 2142.

[5] Nylén-Forthun, E., Møller, M., & Abrahamsen, N. G. B. (2022). *Financial Distress Prediction Using Machine Learning and XAI: Developing an Early Warning Model for Listed Nordic Corporations* (Master's thesis, NTNU).

[6] Yang, H., Li, E., Cai, Y. F., Li, J., & Yuan, G. X. (2021). The extraction of early warning features for predicting financial distress based on XGBoost model and shap framework. *International Journal of Financial Engineering*, *8*(03), 2141004.

[7] Li, S., Shi, W., Wang, J., & Zhou, H. (2021). A deep learning-based approach to constructing a domain sentiment lexicon: a case study in financial distress prediction. *Information Processing & Management*, *58*(5), 102673.

[8] Sun, X., & Lei, Y. (2021). Research on financial early warning of mining listed companies based on BP neural network model. *Resources Policy*, *73*, 102223.

[9] Beltman, J., Osterrieder, J., & Machado, M. (2024). Predicting Retail Customers' Distress: Early Warning Systems and Machine Learning Applications. *Available at SSRN 4730470*.

[10] Shen, F., Liu, Y., Wang, R., & Zhou, W. (2020). A dynamic financial distress forecast model with multiple forecast results under unbalanced data environment. *Knowledge-Based Systems*, *192*, 105365.

[11] Sun, J., Li, H., Fujita, H., Fu, B., & Ai, W. (2020). Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting. *Information Fusion*, *54*, 128-144.

[12] Sun, J., Fujita, H., Zheng, Y., & Ai, W. (2021). Multi-class financial distress prediction based on support vector machines integrated with the decomposition and fusion methods. *Information Sciences*, *559*, 153-170.

[13] Du, G., Liu, Z., & Lu, H. (2021). Application of innovative risk early warning mode under big data technology in Internet credit financial risk assessment. *Journal of Computational and Applied Mathematics*, *386*, 113260.

[14] Samitas, A., Kampouris, E., & Kenourgios, D. (2020). Machine learning as an early warning system to predict financial crisis. *International Review of Financial Analysis*, *71*, 101507.

[15] Craja, P., Kim, A., & Lessmann, S. (2020). Deep learning for detecting financial statement fraud. *Decision Support Systems*, *139*, 113421.

[16] Banerjee Chattapadhyay, D., Putta, J., & Rao P, R. M. (2021). Risk identification, assessments, and prediction for mega construction projects: A risk prediction paradigm based on cross analytical-machine learning model. *Buildings*, *11*(4), 172.

[17] Bluwstein, K., Buckmann, M., Joseph, A., Kapadia, S., & Şimşek, Ö. (2023). Credit growth, the yield curve and financial crisis prediction: Evidence from a machine learning approach. *Journal of International Economics*, *145*, 103773.

[18] Fang, L., Yao, Y., & Liang, X. (2023). New binary archimedes optimization algorithm and its application. *Expert Systems With Applications*, *230*, 120639.

[19] Ding, G., Wang, W., Liu, H., & Tu, L. (2023). Defect of Archimedes optimization algorithm and its verification. *Soft Computing*, *27*(2), 701-722.

[20] Wang, Y. J., Chen, M. C., & Ku, C. S. (2023). An improved archimedes optimization algorithm (IAOA). *J. Netw. Intell*, *8*(3), 693-709.

[21] Quan, Z., & Pu, L. (2023). An improved accurate classification method for online education resources based on support vector machine (SVM): Algorithm and experiment. *Education and information technologies*, *28*(7), 8097-8111.

[22] Cengiz, A. V. C. I., Budak, M., Yağmur, N., & Balçik, F. (2023). Comparison between random forest and support vector machine algorithms for LULC classification. *International Journal of Engineering and Geosciences*, *8*(1), 1-10.

[23] Wang, H., Li, G., & Wang, Z. (2023). Fast SVM classifier for large-scale classification problems. *Information Sciences*, *642*, 119136.