ACTA SCIENTIAE

ISSN: 2178-7727

Novel Framework for Forecasting Financial Distress Using Advanced Machine Learning Technique

*Dr Ganesh Mergu, **Dr Karuna kararao Runjala, ***Dr Murali Merugu

Assistant Professor, Dept. of Commerce and Business Administration
Sandeepani Degree College Kamareddy, Telangan, India
Associate Professor, Dept of Accounting and Finance Gambella University Ethiopia
Assistant Professor, HOD, Dept. of Economics Indus International IB School
Hyderabad, India

ABSTRACT

Financial hardship is a serious problem that impacts people, businesses, and economies. It frequently produces severe outcomes, including job loss, bankruptcy, and economic downturns. Financial distress refers to an organization's inability to meet its financial obligations, which can stem from inadequate financial management, recessions, unforeseen costs, or significant market shifts. The study aims to establish an advanced machine learning (ML) technique for forecasting financial distress in companies. The study proposed a novel resilient pinecone algorithm-mutated extreme boosting (RPA-XGBoost) model to predict a company’s financial distress. The RPA-XGBoost model aims to determine whether a business is in a state of financial distress. RPA-based methods for parametric improvement in the financial and XGBoost models receive the financial data as input and predict financial distress. The study utilizes the financial metrics of various companies along with indicators of financial distress. The data was preprocessed using data cleaning and normalization for the obtained data. The results demonstrate traditional algorithms outperform the proposed method in terms of the higher instances precision of 0.99, accuracy of 98.71, f1-score of 0.989, recall of 0.992 and time complexity of 20.12. In conclusion, this study presents a novel framework for forecasting financial distress that leverages advanced ML techniques to enhance predictive accuracy and reliability.

KEYWORDS: Financial Distress, Company, Resilient Pinecone Algorithm-mutated Extreme Boosting (RPA-XGBoost), Economic Downturns, Financial Management.

INTRODUCTION

The recent years have witnessed rampant globalization of finance. This has made nations invest in other countries, thereby dispersing their risk while promoting trade with other nations. Foreknowledge of the probability of default risk is crucial to the effective running of businesses and banks, especially during chaotic economic times. Successful financial leadership is essential to the smooth operation of any business and should be competent to predict any financial difficulties [1]. Businesses, financiers, investors, and regulatory agencies must determine any hazards to the business regarding financial distress forecasts (FDF) and insolvency projections. A variety of analytical techniques and tools have been created. FDF prototypes are currently constructed using a variety of statistical, ML, and computerized models. Numerous studies have demonstrated that using artificial intelligence (AI) is a more suitable tool than statistical methods for predicting possible financial difficulties [2]. Moreover, the majority of FDF models that are in use are static, rely on forecasts for a specific period, and only use monetary variables in their predictions. One important field of research is the organization of dynamically imbalanced information research in the domains of AI and ML, in addition to being a significant issue in the financial crisis sector [3]. A dynamic disproportionate data and discusses two major imbalances in the economic distress projection: the underlying sample category dispersion and the comprehensive asymmetry. Whether they are big or small, business entities face constant opportunities or challenges due to changing business, economic situations and a lot of research has been done to create and evaluate several models for anticipating financial trouble [4].

Equity suppliers run the risk of losing their money invested, while management runs the risk of having their

ISSN: 2178-7727

incentives reduced. Numerous new research findings focusing on various elements of company distress are being released, despite the paucity of reviews. Finding any gaps in the present literature on the study. The research is analyzed by the contributors from three distinct perspectives: the forecast of financial distress, the tactics employed by enterprises to alleviate it, and the actual economic difficulties [5]. The academics, legislators, and businesspeople will find the findings insightful and helpful in understanding the area. Investors with long-term views are especially worried about a company's capacity to repay debt. Using learning algorithms has been suggested a solution, but this approach maximizes accuracy without accounting for the minority class, which explains why the algorithms are unable to provide useful classifiers. Both balancing and unbalanced datasets are utilized to examine the functionality of FDP models without regard to the environment where these algorithms were tested and created to prevent the kind of data from compromising the results [6].

To predict a firm's financial trouble, first analyze the accuracy of static, moving, and predictive machine learning models. One way that researchers use these data is to make their findings more comparable to those of additional investigations that have used similar frameworks. Additionally, practitioners in the industry can apply the assessed models in a variety of contexts, such as evaluating credit choices. Estimating the likelihood that a business will file for bankruptcy affects every shareholder. This research holds great significance in the banking sector; it impacts a bank's progress and future endeavours [7]. Before making a loan offer, banks must assess a potential counterparty's possibility of default. As a result, the banks make better loan selections and save a substantial amount of money. Bankruptcy forecasting research and analysis have advanced to the status of a scientific discipline, which has fueled the rise of models that forecast bankruptcy as the engine for the continued growth and sustainability of businesses [8]. A great deal of work is done to create improved models that use both linear and non-linear methodologies for enhancing the bankruptcy estimation process. The conventional method took into account several quantitative variables, including leverage, changeover ratios, and profitability ratios. To forecast distress, recent research has included qualitative factors, such as ratings provided by conventional credit rating organizations based on the reputation of the company [9].

However, the subjective nature of qualitative research makes it impossible to provide reliable estimates and outcomes, which can be deceptive. Three distinct criteria can be used to describe bank distress occurrences. Based on statistics on bankruptcies, liquidations, and bankruptcy, as well as on both private and government participation in financial organizations, the initial definition was created. Distressed events also include instances of state assistance, such as capital infusions and asset guarantees provided to banks. Bank mergers that receive government assistance if they are in trouble also qualify as potential distress events. Based on variables unique to each nation and financial sector, the second categorization is formed. These indicators serve as early alert systems for potential problems they show imbalances in the financial sector [10].

Aim of the study: By examining financial data and past patterns, this research seeks to create a precise predictive model for predicting financial difficulty in businesses. The study aims to improve consumer outcomes and reduce possible losses through the detection of crucial risk variables. The ultimate goal of the research is to advance knowledge of strategies for averting economic meltdowns.

Contributions of the study

- **Accurate Prediction Model:** Introduced a high-performance model (RPA-XGBoost) to forecast financial distress, addressing imbalanced data and improving prediction accuracy.
- **Automatic Parameter Optimization:** Integrated Random Parameter Adjustment (RPA) to optimize model parameters, enhancing performance without manual tuning.
- **Over fitting Prevention:** Applied regularization techniques to prevent over fitting, ensuring better generalization and scalability for reliable financial distress predictions.

Organization of the study: Related work is outlined in Part 2, the methodology is established in Part 3, the performance evaluation is displayed in Part 4, the discussion covers Part 5 and a conclusion is illustrated in Part 6.

ISSN: 2178-7727

RELATED WORK

Examine the numerous scholarly studies [11] on credit risk AI, and their development from a variety of angles. Without specifying a time frame, a systematic examination of the literature was conducted using the online encyclopedia of science and Scopus databases. It adds to the theory's proof by offering details on the most popular approaches and strategies by bringing a broad analysis to enhance understanding of the theme's components and variables. In a distinctive approach, the use of computer technology in assessing credit risks has garnered attention. It was discovered that there is a continuing need for the discovery and development of new variables, classification systems, and more assertive techniques.

Research [12] investigated the consequences of financial hardship and accountant quality, as well as the knowledge and market price of analysts' cash flow projections for listed corporations. The predictability of working capital movements is examined using the naïve extrapolating model. A value-relevance analysis is conducted using a total return approach. The overall z-score and the business's specialty auditor indicator are used in the study as proxies for anxiety and inspector quality, accordingly, controlling for the forecast period. While creating financing forecasts, analysts effectively adjust earnings estimates for appreciation, but they do not effectively account for changes in operational resources. Results showed that the accuracy of cash flow projections improves for both distressed and high-quality audited firms. This improvement is attributed to analysts' conservative views, prudent accounting decisions, and more accurate revenue projections. Forecasts of corporate economic difficulties were crucial for businesses, investors, and regulators. Furthermore, they discovered [13] that the best approach to handling the financial hardship forecast issue is to use the suggested model, which combines the findings of several forecasts. The two most important aspects of the financial crisis data unbalanced collections of data and concept migration of data stream have been largely ignored by prior research because the majority of financial distress forecasting (FDF) models were built on single-time dimensions.

In the equity market, the COVID-19 outbreak caused a great degree of financial instability. After a brief decline in March 2020, there was an unanticipated spike in growth during 2021. As a result, [14] financial risk forecasting, which addresses novel forms of uncertainty remains a crucial component of financial management. The methodology can notify management and other relevant staff, creditors, investors, officials, lenders, specialists, and others in advance of a company's worsening financial status, enabling them to take prompt action to prevent losses.

According to the study [15], the authors next analyze the distress forecasting success of various methodologies using the traditional binomial legit framework and two other well-known mathematical designs, notably neural networks with AI and the assistance vector machine. Five financial ratios return on investments, revenues to total responsibility, the ratio of fixed expenditure assets to the entire inventory, the debt to proportion of equity, and a further precaution of the magnitude of the company seem to have an impact on the empirical results obtained from the research. Research [16] created an economic distress forecasting framework that includes standard financial factors but also key important business governance characteristics. Furthermore, compared to models utilizing traditional threshold numbers, those adopting dynamic sensitivity threshold amounts had higher prediction success rates. Economically sound corporations have greater ratios of cash than financially troubled ones, except for the debt ratio. Regarding organizational governance characteristics, discover that while increased managers' participation pledge percentages were positively associated with financial difficulty, the CEO/Chairman combination is not the cause of financial difficulties.

In transition economies, predicting corporate financial trouble is challenging due to a lack of available and highly skewed data. According to the study[17],in terms of balancing accuracy, algorithms the one developed by Alt variables regularly do better than those that employ one or the other. The results show that all approaches predict outcomes for non-delisted firms reasonably well, while they estimate results for other firms a little less well. The analysis also shows that the models generally outperform small firms in terms of predicting accomplishments, and they generally outperform poor years in terms of predicting achievements for big firms. The research presented a methodical analysis of research on financial difficulties [18],

ISSN: 2178-7727

forecasting, and the tactics businesses use to cope with challenges. The content, approach, scope, analysis of data methodologies, study time, study concentration, and data collected were the seven factors used to evaluate the papers that were chosen. While most studies on larger economies concentrated on these areas, there isn't much qualitative research on the topic. The authors hope to spark debate because they use the study to provide a body of existing knowledge on the subject of discussion. Investors benefit from an accurate forecast of distress because it enables bankers and other financial companies to create a system of alerts to prevent risk contagion. To looked [19] into the prediction of financial trouble using sentiment analysis from the yearly financial statements of listed companies. Furthermore, the experiments demonstrated that incorporating textual sentiment evaluations significantly improved modeling performance, and that integrating insights from audit findings yielded a slightly larger enhancement than incorporating attitudes from the MD&A sections. The findings have significance for how soft knowledge is used in credit risk forecasting in the framework of the market; the use of soft information might be investigated further in different operational research study domains. A major socioeconomic problem that impacts practically every company around the globe is financial distress. Forecasting financial difficulties in the banking sector can significantly help to lower losses and prevent money from being misallocated by banks. By creating a model to anticipate bank insolvency, research aims [20] to foresee the financial difficulty of business banks. The Artificial Neural Network (ANN) forecasting model's output shows improved precision in forecasting. The study's findings are anticipated to help managers, consumers, regulatory agencies, and shareholders more effectively handle their financial interests in the nation's banking industry.

METHODOLOGY

The automatic learning-based forecasting technique examines previous performance data and important financial parameters for predicting financial trouble. The approach consists of applying techniques for classification, cleaning the data, and choosing features. Prospective financial crises can be identified early due to this strategy, as shown in Figure 1.

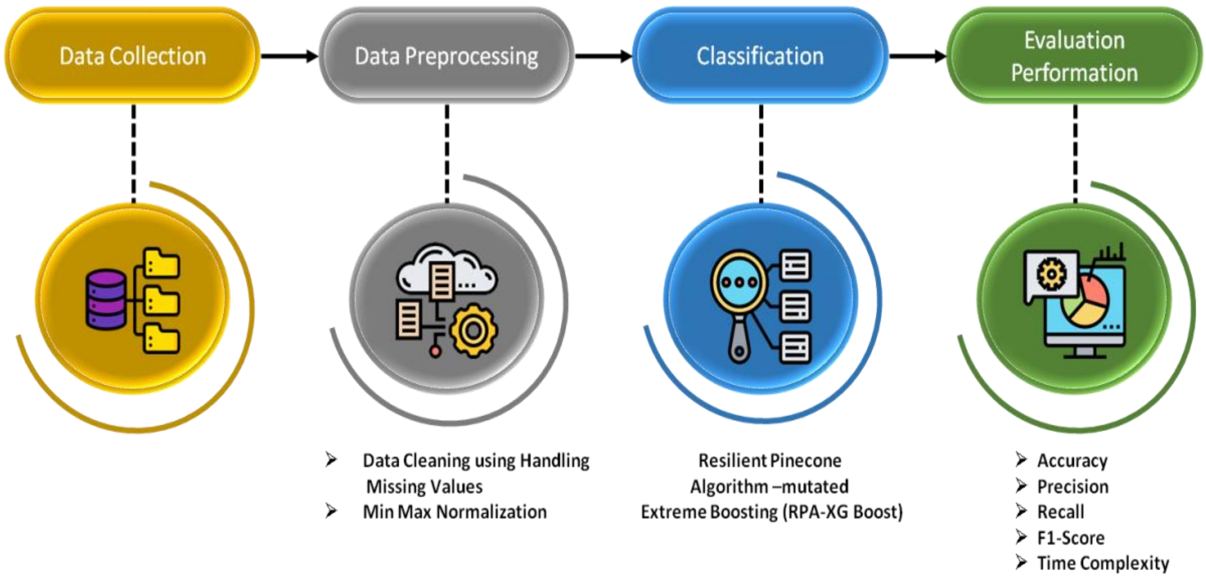

Figure 1: Flow of methodology

3.1 Data collection

The dataset was collected from the open source Kaggle website: https://www.kaggle.com/datasets/shebrahimi/financial-distress. The company is a representative of sample businesses and it includes 3672 occurrences and 86 characteristics. Time displays several time intervals to which the material pertains. Financial depression is the target factor; if it is higher than -0.50, the business should be regarded as healthy. If not, it would be considered to be in financial hardship. Certain financial and other attributes of the sampled enterprises are represented by the features indicated by x1 through x83. These characteristics are from the past and should be utilized for predicting whether or not the business would

ACTA SCIENTIAE
ISSN: 2178-7727

experience financial. The category variable is feature x80. For instance, the second organization is doing well at time 14, whereas company 1 is in economic trouble at time 4. F-score must be used as the outcome assessment criterion because the information provided is skewed and unbalanced (136 financially troubled businesses compared to 286 healthy businesses, meaning that 136 firm-year experiences are economically stressed whereas 3546 firm-year data are normal). It should be mentioned that 70% of the information in this set is utilized for characteristic and model choice, or the train set, while 30% ought to be randomly designated as the hold-out test collection.

3.2 Data preprocessing

The essential step in converting unstructured data into an organized, analysis-ready structure is data preparation. To increase the model's efficiency and precision, it entails tasks including choosing features, standardization, and handling values that are missing.

3.2.1 Data cleaning using handling missing values

The financial information of the companies under analysis contains missing values for attributes. These absent values could result from inaccurate financial reports, numbers that were overlooked when gathering data, or financial measures that aren't useful for some businesses. The suggested RPA-XGBoost model has three options for handling missing information: it can disregard organizations, whose attributes are missing, estimate the likely values based on comparable financial records, or treat missing values as a distinct group. When predicting financial difficulty, the model's prediction accuracy and robustness are maintained by addressing value gaps.Data cleaning with zero errors is an essential data processing step, mainly when data from multiple financial sources are to be integrated. Typically, datasets used in financial trouble forecasts are poor quality regarding data, exhibiting missing or wrong numbers, varied rules regarding nomenclature, and lack of enough financial information. Cleansed data can have an improvement in the quality of data, as mistakes could be identified and corrected using a process referred to as data cleansing or scrubbing. Common data quality issues arise from the occurrence of spelling mistakes in the entry of data, a lack of financial indicators, and conflicting naming conventions for characteristic variables. It is very vital to deal with these problems for the forecast to be accurate. The most common technique used for dealing with absent features is imputation. Other methods and their implications are weighed with the process of imputation.

3.2.2 Min max normalization

One method of data processing for reducing financial metrics to a range is min-max standardization, usually between 0 and 1, so that all features will contribute equally to model predictions. Therefore, it is highly necessary in economic distress forecasts, as it improves the performance of models by eliminating biases due to scales of different economic measures. Min-max normalization normalizes the original data through linear transformations to create equilibrium of value assessment among information before and after the process.

The following equation can be applied to this equation (1).

$$W_{new} = \frac{W - \min(W)}{\max(W) - \min(W)} \tag{1}$$

W_{new} = The adjusted value derived from the standardized outcomes
W = Previous number
$Max(W)$ = Average number
$Min(W)$ = Amount of the average deviation

3.3 Forecasting Financial Distress using a novel resilient pinecone algorithm-mutated extreme boosting (RPA-XGBoost)

The Resilience Pinecone Algorithm and the revised Extreme Boosting, herein referred to as the RPA-XGBoost model, are integrated into a hybrid solution to improve the precision of the forecast for the phenomenon of financial crisis with the aid of sophisticated approaches of AI. This approach can identify the risk at enterprises in an optimized way by streamlining financial measures through feature selection and hyperparameter adjustment. So, this RPA-XGBoost architecture combining the capabilities of extreme boosting with the

ISSN: 2178-7727

strength of algorithmic robustness could be relied on to predict financial problems in nearly all the most employed economic scenarios.

3.3.1 Mutated extreme boosting (XGBoost)

A novel ML technique, called Mutated Extreme Boosting, or XGBoost, uses parallel computing, effective management of huge data sets, and normalization techniques not to avoid overfitting but to enhance the degree of accuracy in making predictions. It applies the skills of weak students properly to develop a sound ensemble model for predicting economic recession by altering the structure of gradient-boosting as suitable for a wide range of prediction problems. In the standard XGBoost training, all information concerning finance data, $W, and\ Z,$ are used in a batch setting. However, they need to take into consideration the possibility of notion drift in an evolving financial context where new data. For this particular scenario, a continuous flow of fiscal information can be defined as follows$B = \{(\vec{w}_s,\ z_s)\}|s = 1, \ldots\ldots, S$, where $S \to \infty$, $\vec{w}_s z_s e_j$is the corresponds target demonstrating financial distress and $\omega = (\vec{w}_j, z_j): j \in \{1,2, \ldots\ldots, X\}$,is a vector of features economic indicators with size$|x| = X$. To tackle this, suggest XGBoost, an improved version of the XGBoost algorithm created especially for predicting financial hardship in a streaming setting as denoted in equation (2).

$$\hat{Z}^{(l)} = \sum_{l=1}^{L} e_l(x_l) = \hat{Z}^{(l-1)} + e_l(x_l) \tag{2}$$

The new foundation function's index $l$with the combination is a critical factor in the function's derivation. The accounting information in the reservoir is used directly to create the starting model for the first ensemble membere_1, the financial data is processed using the current ensemble for future members where $l > 1$, and the remaining coefficients from the first $l - 1$models are computed. The new foundation function $e_l$is constructed using these remains, which indicate the discrepancies between the expected and actual financial hardship consequences. By concentrating on the mistakes made by the equation throughout each iteration, this technique enables the model enhancing its forecasting abilities.

3.3.2 Resilient pinecone algorithm (RPA)

The two populations that the Resilient Pinecone Optimization Algorithm (RPA) works are economic indicators and possible fixes. Based on important financial parameters, the problem area associated with financial trouble is split into equal halves. For every segment, a primary monetary indicator is created during the method's beginning, and many possible solutions are produced around these indicators. To streamline this procedure, the limiting parameters for every segment are established based on the values and boundaries of pertinent financial metrics. This guarantees that the produced solutions thoroughly examine the spectrum of potential outcomes associated with financial hardship. By using this method, the algorithm can determine the best courses of action for anticipating and averting financial problems in businesses.

$$LbS = ka + (j - 1) \times \frac{va - ka}{M_{tree}} \tag{3}$$

$$Ubs = va + (j - 1) \times \frac{va - ka}{M_{tree}} \tag{4}$$

The bottom and higher boundaries of the problem's domains are denoted by ka and va in Equations (3) and (4), and the lower and upper bounds of every segment are denoted by $Lbs\ and\ Ubs$. The following is the amount of pine cones is created.

$$CX_{i,j} = LbS_j + \overrightarrow{rand}_{1\times dim} \times (rand \times UbS_j - rand \times LbS_j) \tag{5}$$

Equation (5) illustrates the location of viable solutions in the RPA where CX indicates the location of an economic indicator. $\overrightarrow{rand}_{1\times dim}$was a random integer with similar transportation, and the phrase $rand$ refers to a random vector created from a normal distribution between 0 and 1.For RPA operators, the index of a financial indicator is denoted by i, while the index of a possible solution is represented by j. Algorithm 1 shows that the process of RPA-XGBoost.

ISSN: 2178-7727

Following RPA initialization, each monetary indicator's placement is adjusted to correspond with the best-performing solution's placement in that segment, facilitating efficient financial environment exploration and improving the predictive power of the framework.

Algorithm 1

```
FUNCTION ForecastFinancialDistress(data, learningRate, epochs):
trainData, testData = SplitData(data)
model = InitializeXGBoost()
model.Fit(trainData.features, trainData.labels)
weights = InitializeWeights()
biases = InitializeBiases()
   FOR epoch FROM 1 TO epochs:
predictions = model.Predict(trainData.features)
error = CalculateError(predictions, trainData.labels)
      IF error <errorThreshold:
         PRINT "Training converged at epoch:", epoch
         BREAK

gradients = CalculateGradients(error, weights, biases)
      FOR i FROM 0 TO LENGTH(weights):
         IF gradients[i] > 0:
            weights[i] += learningRate
         ELSE IF gradients[i] < 0:
            weights[i] -= learningRate
      biases += learningRate * gradients
testPredictions = model.Predict(testData.features)
testError = CalculateError(testPredictions, testData.labels)
   IF testError<distressThreshold:
      PRINT "Low financial distress risk."
   ELSE:
      PRINT "High financial distress risk."
   RETURN testPredictions, testError
```

By combining the advantages of XGBoost and an altered version of the RPA, the RPA-XGBoost is a sophisticated hybrid technique intended to improve the reliability of financial trouble projections. This novel approach starts with XGBoost, which uses ensemble learning to efficiently capture intricate correlations in financial data and find important indicators of distress. The method optimizes the course of learning by introducing a mutation technique that allows parameters to be dynamically adjusted based on feedback from learners. The model is improved via RPROP, which ensures efficient error return propagation by modifying biases and weights to provide fast convergence characteristics. Combining these methods allows increasing resilience against excessive fitting and prediction accuracy, ultimately offering consumers a potent tool to evaluate. Financial health and make informed decisions in an increasingly volatile economic landscape.

RESULT

The results were produced using a single Windows 10 machine (Asus X556U) equipped with an Intel® Core (TM) i5-72000U primary CPU unit working at 2.50 GHz, 8.0 GB of volatile memory, and an Nvidia GeForce 940MX graphics card. Comparing the proposed methods, RPA-XGBoost was better than the existing Chinese Whisper Clustered Stochastic Gradient Descent Federated Learning method (CWCSGDFL)[21] with the instances of values.

Accuracy: The percentage of economic distress cases that were accurately predicted out of all situations examined. It is a crucial indicator that assesses how well the model distinguishes between companies and those

ISSN: 2178-7727

are not in financial crisis. The results showed a very high accuracy, which means that the suggested model could be used to predict possible economic meltdowns accordingly, allowing participants to make decisions in high-quality time. Table 1 and Figure 2 as shown in the comparison between the proposed RPA-XGBoost methods were higher than the existing method CWCSGDFL with the instances of 97.35, 98.12, 98.71, 97.67, 98.1, 97.35, 97.93, 98.57, 98.27 and 97.15.

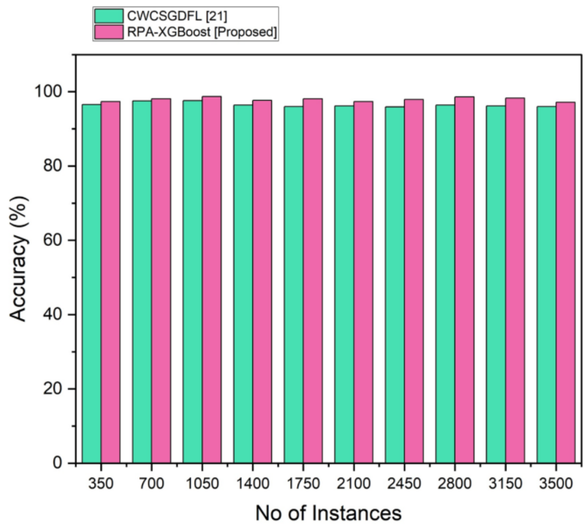

Figure 2: Result of the accuracy

Table 1: Outcome of the accuracy

No of Instances	**Accuracy (%)**	
	CWCSGDFL [21]	**RPA-XGBoost [Proposed]**
350	96.57	97.35
700	97.57	98.12
1050	97.61	98.71
1400	96.42	97.67
1750	96	98.1
2100	96.19	97.35
2450	95.91	97.93
2800	96.42	98.57
3150	96.19	98.27
3500	96	97.15

Precision: Precision refers to the proportion of true positive predictions out of all the positive predictions by the model. Strong precision reduces error rates because it indicates that a model has a high chance of being correct if it predicts that a firm will be in financial distress. This statistic is essential for stakeholders who depend on good predictions to make correct economic decisions. Table 2 and Figure 3 shows that the comparison of the proposed RPA-XGBoost methods compared to the current method CWCSGDFL outperforms the current method with the instances of 0.989, 0.983, 0.99, 0.987, 0.981, 0.973, 0.975, 0.987, 0.982 and 0.974.

ISSN: 2178-7727

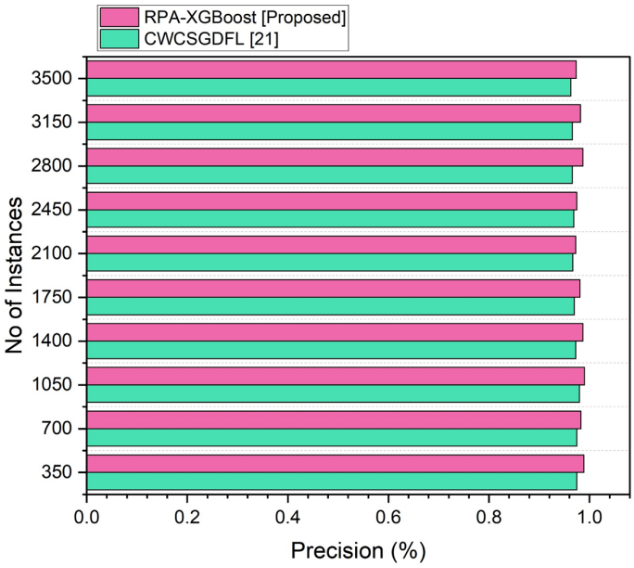

Figure 3: Result of the precision

Table 2: Outcome of the precision

No of Instances	Precision (%)	
	CWCSGDFL [21]	RPA-XGBoost [Proposed]
350	0.975	0.989
700	0.975	0.983
1050	0.98	0.99
1400	0.973	0.987
1750	0.97	0.981
2100	0.967	0.973
2450	0.969	0.975
2800	0.966	0.987
3150	0.966	0.982
3500	0.963	0.974

Recall: In the case of financial distress predictions, recall measures the extent to which the model is able to identify those firms that are really stressed. It measures how well the model captures stressed entities and is calculated as the number of true positives relative to all real positives. The higher a firm's recall, the more at-risk firms the model can identify for prompt action and preventive measures. Table 3 and Figure 4 shows that the comparison of the proposed RPA-XGBoost methods, it was higher compared to existing CWCSGDFL with the instances of0.989, 0.992, 0.992, 0.988, 0.989, 0.987, 0.99, 0.989, 0.991 and 0.982.

ISSN: 2178-7727

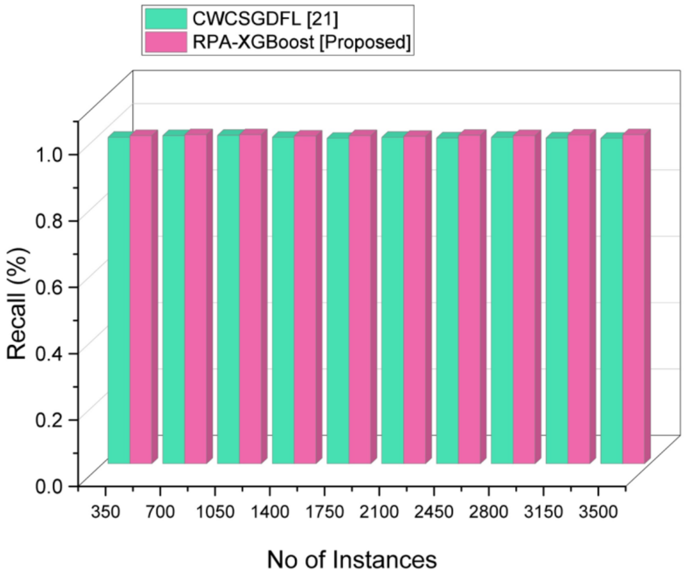

Figure 4: Result of the recall

Table 3: Outcome of the recall

No of Instances	Recall (%)	
	CWCSGDFL [21]	RPA-XGBoost [Proposed]
350	0.984	0.989
700	0.989	0.992
1050	0.99	0.992
1400	0.984	0.988
1750	0.981	0.989
2100	0.984	0.987
2450	0.982	0.99
2800	0.984	0.989
3150	0.982	0.991
3500	0.981	0.992

F1-score: An F1-score is a performance parameter to assess the prediction model for financial trouble forecasting by merging precision and recall into one statistic. It does represent the harmonic average of recall and accuracy, thus providing a fair assessment of the model's efficacy. A high F1-score in this case study means that the model minimizes false positives and false negatives but is efficient in selecting at-risk firms of economic collapse. Table 4 and Figure 5 as shown in the comparison between the proposed developed RPA-XGBoost methods perform better than the existing method CWCSGDFL with the instances of 0.983, 0.989, 0.989, 0.983, 0.981, 0.986, 0.984, 0.987, 0.985 and 0.988.

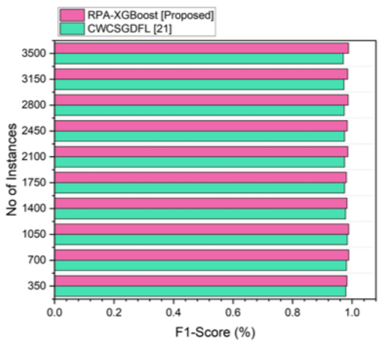

Figure 5: Result of the F1-score

Table 4: Outcome of the F1-score

No of Instances	F1-Score (%)	
	CWCSGDFL[21]	RPA-XGBoost [Proposed]
350	0.979	0.983
700	0.981	0.989
1050	0.984	0.989
1400	0.978	0.983
1750	0.975	0.981
2100	0.975	0.986
2450	0.975	0.984
2800	0.974	0.987
3150	0.973	0.985
3500	0.971	0.988

Time complexity: The number of times it takes to accurately forecast financial problems with this method. Consequently, the temporal intricacy in terms of total is calculated. Table 5 and Figure 6 as shown in the comparison of the proposed methods of RPA-XGBoost were less than those of the existing approach, CWCSGDFL with the instances of 20.12, 23.41, 26.3, 28.14, 31.01, 34.01, 34.02, 38.24, 40.67, 43.11 and 50.55.

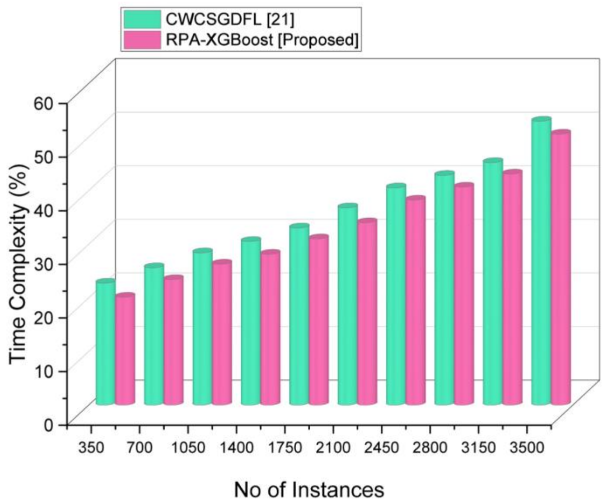

Figure 6: Result of the time complexity

Table 5: Outcome of the time complexity

No of Instances	Time Complexity (%)	
	CWCSGDFL [21]	RPA-XGBoost [Proposed]
350	22.75	20.12
700	25.62	23.41
1050	28.4	26.3
1400	30.52	28.14
1750	33.05	31.01
2100	36.8	34.02
2450	40.52	38.24
2800	42.85	40.67
3150	45.24	43.11
3500	52.96	50.55

ACTA SCIENTIAE

ISSN: 2178-7727

DISCUSSION

The primary reasons for the weaknesses of CWCSGDFL in financial distress prediction are related to its complexity level and over-fitting potential. The deep structure applied lowers the generalization capability of the model concerning unknown data, where the model essentially learns noise from the actual source of the original data set. It further involves extensive hyperparameter tuning and computing abilities, which again lead to prolonged training times and render the method less practical for any real-time application. Among the strengths of RPA-XGBoost is its ability to process fast with great precision. Financial trouble estimation is much affected by having partially unbalanced datasets, but XGBoost performs very well in handling such situations. It can also automatically adjust the settings to boost the performance without the explicit tuning by the user as needed. It has the robust ability to handle missing data and inherently has a regularization mechanism that does not allow it to take over fit. Thus, in totality, RPA-XGBoost is improved in adaptability, effectiveness, and execution.

CONCLUSION

The overall businesses should predict financial distress; they will become proactive players to identify risks in advance and take appropriate measures. Companies make distressed predictions more accurate and keep their operational viability and financial well-being through computerized learning algorithms, intelligent data analytics, and up-to-date accounting information. The recommendations highlight the necessity of forming a robust forecasting model, considering different financial performance and macroeconomic as well as industry-level factors. When compared to the proposed method, was higher than the existing method, with the highest instances of time complexity of 20.12, precision of 0.99, f1-score of 0.989, recall of 0.992 and accuracy of 98.71. Future research could explore real-time information, such as sentiment from social networks and market trends, to enhance financial crisis prediction. Even more, improving ML methods and exploring other algorithms would enhance the model efficiency. Facilitating better ease for managers in understanding and responding to distress predictions will be crucial, which will ultimately leading to stronger economic strategies during market uncertainty.

REFERENCE

1. Chen, S. and Shen, Z.D., 2020. Financial distress prediction using hybrid machine learning techniques. Asian Journal of Economics, Business and Accounting, 16(2), pp.1-12.https://doi.org/10.9734/ajeba/2020/v16i230231
2. Liang, D., Tsai, C.F., Lu, H.Y.R. and Chang, L.S., 2020. Combining corporate governance indicators with stacking ensembles for financial distress prediction. Journal of Business Research, 120, pp.137-146.https://doi.org/10.1016/j.jbusres.2020.07.052
3. HIONG, H.K., JALIL, M.F. and SENG, A.T.H., 2021. Estimation and prediction of financial distress: Non-financial firms in Bursa Malaysia. The Journal of Asian Finance, Economics and Business, 8(8), pp.1-12.https://doi.org/10.13106/jafeb.2021.vol8.no8.0001
4. Jiang, C., Lyu, X., Yuan, Y., Wang, Z. and Ding, Y., 2022. Mining semantic features in current reports for financial distress prediction: Empirical evidence from unlisted public firms in China. International Journal of Forecasting, 38(3), pp.1086-1099.https://doi.org/10.1016/j.ijforecast.2021.06.011
5. Qian, H., Wang, B., Yuan, M., Gao, S. and Song, Y., 2022. Financial distress prediction using a corrected feature selection measure and gradient-boosted decision tree. Expert Systems with Applications, 190, p.116202.https://doi.org/10.1016/j.eswa.2021.116202
6. Liu, J., Li, C., Ouyang, P., Liu, J. and Wu, C., 2023. Interpreting the prediction results of the tree-based gradient boosting models for financial distress prediction with an explainable machine learning approach. Journal of Forecasting, 42(5), pp.1112-1137.https://doi.org/10.1002/for.2931
7. Zeng, S., Li, Y., Yang, W. and Li, Y., 2020. A financial distress prediction model based on a sparse algorithm and support vector machine. Mathematical problems in engineering, 2020(1), p.5625271.https://doi.org/10.1155/2020/5625271
8. Ullah, H., Wang, Z., Abbas, M.G., Zhang, F., SHAHZAD, U. and MAHMOOD, M.R., 2021. Association of financial distress and predicted bankruptcy: The case of the Pakistani banking sector.

ISSN: 2178-7727

The Journal of Asian Finance, Economics and Business, 8(1), pp.573-585.https://doi.org/10.13106/jafeb.2021.vol8.no1.573

9. Zhang, Z., Wu, C., Qu, S. and Chen, X., 2022. An explainable artificial intelligence approach for financial distress prediction. Information Processing & Management, 59(4), p.102988.https://doi.org/10.1016/j.ipm.2022.102988
10. Sun, J., Li, H., Fujita, H., Fu, B. and Ai, W., 2020. Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting. Information Fusion, 54, pp.128-144.https://doi.org/10.1016/j.inffus.2019.07.006
11. Duarte, D.L. and de Moraes Barboza, F.L., 2020. Forecasting financial distress with machine learning–a review. *Future Studies Research Journal: Trends and Strategies*, *12*(3), pp.528-574. https://doi.org/10.24023/FutureJournal/2175-5825/2020.v12i3.533
12. Papadaki, A. and Pavlopoulou-Lelaki, O.C., 2022. Analysts' cash flow forecasts informativeness, financial distress, and auditor quality. *Accounting Research Journal*, *35*(3), pp.408-426. https://doi.org/10.1108/ARJ-10-2020-0333
13. Shen, F., Liu, Y., Wang, R. and Zhou, W., 2020. A dynamic financial distress forecast model with multiple forecast results under an unbalanced data environment. *Knowledge-Based Systems*, *192*, p.105365. https://doi.org/10.1016/j.knosys.2019.105365
14. Wu, D., Ma, X. and Olson, D.L., 2022. Financial distress prediction using integrated Z-score and multilayer perceptron neural networks. *Decision Support Systems*, *159*, p.113814. https://doi.org/10.1016/j.dss.2022.113814
15. Sehgal, S., Mishra, R.K., Deisting, F. and Vashisht, R., 2021. On the determinants and prediction of corporate financial distress in India. *Managerial Finance*, *47*(10), pp.1428-1447. https://doi.org/10.1108/MF-06-2020-0332
16. Chen, C.C., Chen, C.D. and Lien, D., 2020. Financial distress prediction model: The effects of corporate governance indicators. *Journal of Forecasting*, *39*(8), pp.1238-1252. https://doi.org/10.1002/for.2684
17. Nguyen, M., Nguyen, B. and Liêu, M.L., 2023. Corporate financial distress prediction in a transition economy. *Journal of Forecasting*. https://doi.org/10.1002/for.3177
18. Mallinguh, E.B. and Zéman, Z., 2020. Financial distress, prediction, and strategies by firms: A systematic review of the literature. *Periodica Polytechnica Social and Management Sciences*, *28*(2), pp.162-176. https://doi.org/10.3311/PPso.13204
19. Huang, B., Yao, X., Luo, Y. and Li, J., 2023. Improving financial distress prediction using textual sentiment of annual reports. *Annals of Operations Research*, *330*(1), pp.457-484. https://doi.org/10.1007/s10479-022-04633-3
20. Mishra, N., Ashok, S. and Tandon, D., 2021. Predicting financial distress in the Indian banking sector: A comparative study between the logistic regression, LDA and ANN models. *Global Business Review*, p.09721509211026785. https://doi.org/10.1177/09721509211026785
21. Al Ali, A.I. and Khedr, A.M., 2024. Enhancing financial distress prediction through integrated Chinese Whisper clustering and federated learning. Journal of Open Innovation: Technology, Market, and Complexity, 10(3), p.100344.https://doi.org/10.1016/j.joitmc.2024.100344