A Deep Learning-Based Framework for Accurate Bankruptcy Prediction

Ajay Sourashtriya[1], Shrivastava[2]

[1]MTech Scholar, [2]AP & Head, CSE
[1,2]NIIST, Bhopal
[1]ajsourashtriya04@gmail.com, [2]Anurag.shri08@gmail.com

ABSTRACT

Bankruptcy prediction is a critical challenge in financial risk management, as early identification of distressed firms can prevent severe economic consequences and aid stakeholders in making informed decisions. Traditional statistical models and traditional machine learning approaches have demonstrated limited accuracy in capturing the complex, non-linear relationships within financial datasets. To address these limitations, this work proposed a deep learning-based framework for accurate bankruptcy prediction. The proposed model leverages advanced neural network architectures to analyze multidimensional financial indicators, incorporating techniques such as feature normalization, dropout regularization, and adaptive optimization for improved performance and generalization. Experiments were conducted using publicly available bankruptcy datasets containing diverse financial ratios and company attributes. The model's effectiveness was evaluated using standard metrics including Accuracy, Precision, Recall, and AUC-ROC. Experiment results indicate that the proposed deep learning approach significantly outperforms traditional methods, achieving superior predictive accuracy and robustness. This work highlights the potential of deep learning as an intelligent solution for proactive bankruptcy risk assessment, contributing to improved financial decision-making and risk mitigation strategies.

Keywords: Bankruptcy, Bankruptcy prediction system Classification, machine learning techniques, Deep Learning, Credit Card Fraud Detection

1. INTRODUCTION

Bankruptcy prediction has long been a fundamental concern in finance, economics, and corporate management. The ability to anticipate financial distress at an early stage enables investors, creditors, policymakers, and management teams to take timely corrective actions and mitigate risks. Traditionally, statistical models such as logistic regression, discriminant analysis, and Altman's Z-score have been used to forecast bankruptcy by analyzing financial ratios and historical data. While these models offer interpretability, they are often constrained by their linear assumptions and limited capability to handle complex, high-dimensional financial data. With the advent of artificial intelligence, particularly deep learning, there has been a paradigm shift in predictive analytics. Deep learning models possess the capability to automatically learn intricate patterns and non-linear relationships from large datasets without extensive feature engineering. This makes them particularly well-suited for financial applications where data complexity and variability are high. Recent advances in deep learning offer superior capabilities in capturing intricate patterns from large datasets without manual feature engineering. In this work proposed a deep learning-based framework to improve prediction accuracy and robustness. By leveraging advanced neural architectures, the approach aims to enhance financial forecasting and support proactive corporate risk management strategies.

This paper proposed a deep learning-based framework aimed at enhancing the accuracy and robustness of bankruptcy prediction. By employing architectures such as multilayer perceptrons (MLP) and incorporating optimization techniques like dropout regularization and adaptive learning rate, the model is designed to effectively capture underlying financial distress signals. Using publicly available bankruptcy datasets with detailed financial attributes, the framework is trained and evaluated to determine its predictive effectiveness.

2. LITERATURE REVIEW

The literature review explores existing research on bankruptcy prediction models, comparing traditional statistical methods with advanced deep learning techniques. It highlights key findings, challenges, and emerging trends, emphasizing the growing role of AI in financial risk assessment.

In this research [1] proposed hybrid deep learning model through the use of convolutional neural network to enhance
Bankruptcy forecasting models. Authors address the high-dimensional data and imbalanced problems by introducing
feature selection strategically and Synthetic Minority Over-sampling Technique (SMOTE). In comparative evaluation, the performance of our model is over 81 %, which is better than that for Logistic Regression and Support Vector Machines. This leap in accuracy demonstrates the cutting edge unprecedented ability of our model to decrypt complex financial patterns and establishes a new precedent for deep learning applications in the nuanced field of financial analytics.

This research [2] aims to compare the predictive performance of five models namely the Linear Discriminant Analysis (LDA), Logistic Regression (LR), Decision Trees (DT), Support Vector Machine (SVM) and Random Forest (RF) to forecast the bankruptcy of Tunisian companies. A Deep Neural Network (DNN) model is also applied to conduct a prediction performance comparison with other statistical and machine learning algorithms.

This review [3] examines the discriminatory power of an MLP in the context of bankruptcy prediction. The model was developed using a data set of Taiwanese firms composed of 95 financial ratios for the years 1999 to 2009. We compared different setups of four main parameters of MLPs: optimization algorithm, activation function, number of neurons, and number of layers. The goal was to find the parameter setup that achieve the best results in four evaluation metrics that we used: average accuracy, specificity, sensitivity, and precision.

In [4] author compare some traditional statistical methods for predicting financial distress to some more ‘‘unconventional’’ methods, such as decision tree classification, neural networks, and evolutionary computation techniques, using data collected from 200 Taiwan Stock Exchange Corporation (TSEC) listed companies. Empirical experiments were conducted using a total of 42 ratios including 33 financial, 8 non-financial and 1 combined macroeconomic index, using principle component analysis (PCA) to extract suitable variables.

Author [5], proposed a semi-parametric Cox survival analysis model and non-parametric CART decision trees have been applied to financial distress prediction and compared with each other as well as the most popular approaches. This analysis is done over a variety of cost ratios (Type I Error cost: Type II Error cost) and prediction

Intervals as these differ depending on the situation. The results show that decision trees and survival analysis models have good prediction accuracy that justifies their use and supports further investigation.

The proposed [6] algorithm is successfully applied in the bankruptcy prediction problem, where experiment data sets are originally from the UCI Machine Learning Repository. The simulation results show the superiority of proposed algorithm over the traditional SVM-based methods combined with genetic algorithm (GA) or the particle swarm optimization (PSO) algorithm alone.

In [7] researchers investigate the effect of sampling methods on the performance of quantitative bankruptcy prediction models on real highly imbalanced dataset. Seven sampling methods and five quantitative models are tested on two real highly imbalanced datasets. A comparison of model performance tested on random paired sample set and real imbalanced sample set is also conducted. The experimental results suggest that the proper sampling method in developing prediction models is mainly dependent on the number of bankruptcies in the training sample set. In this research, authors [8] propose the implementation of Jordan Recurrent Neural Networks (JRNN) to classify and predict corporate bankruptcy based on financial ratios. Feedback interconnection in JRNN enables to make the network keep important information well allowing the network to work more effectively. The result analysis showed that JRNN works very well in bankruptcy prediction with average success rate of 81.3785%. Neural Networks can process a tremendous amount of attribute factors; it results in over fitting frequently when more statistics is taken in. By using K-Nearest Neighbor and Random Forest, authors [9] obtain better results from different perspectives.

3. PROPOSED METHODOLOGY

The proposed model focuses on developing an accurate bankruptcy prediction framework using deep learning techniques while effectively addressing the common issue of class imbalance in financial datasets. The process begins with data collection, where publicly available bankruptcy datasets from Kaggle are utilized. These datasets typically contain financial ratios, company profiles, and a binary target variable indicating bankruptcy status (bankrupt or non-bankrupt). The next step involves data preprocessing, which includes handling missing values by applying mean or mode imputation for numerical and categorical attributes, respectively. To ensure uniform feature distribution, financial ratios are normalized using Min-Max scaling.

Since bankruptcy datasets are often highly imbalanced, with significantly fewer bankrupt cases compared to non-bankrupt cases, class balancing is a crucial step. To address this, the SMOTE (Synthetic Minority Over-sampling Technique) method is applied to generate synthetic samples for the minority class, thereby balancing the dataset and improving model learning.

The proposed deep learning model is designed with a multi-layer architecture. It consists of an input layer corresponding to the number of selected financial features, followed by multiple dense hidden layers with ReLU activation and dropout layers to reduce overfitting. The output layer contains a single neuron with a Sigmoid activation function for binary classification. The model is optimized using the Adam optimizer with a learning rate scheduler for efficient convergence, and training is performed on 80% of the dataset while reserving 20% for testing. Finally, model evaluation is carried out using performance metrics such as Accuracy, Precision, Recall, F1-score, and AUC-ROC, with particular emphasis on Recall and F1-score, as these metrics are critical in detecting bankrupt companies accurately.

DATASET: The data were collected from the Taiwan Economic Journal for the years 1999 to 2009. Company bankruptcy was defined based on the business regulations of the Taiwan Stock Exchange [21]. The dataset used in this study was sourced from Kaggle's Company Bankruptcy Prediction repository. It contains 6,819 records with 96 financial indicators, including liquidity, profitability, solvency, and efficiency ratios. The target variable is binary, indicating whether a company is bankrupt or not. The dataset is highly imbalanced, requiring class balancing techniques for accurate modeling.

4. RESULT ANALYSIS

The proposed deep learning model was evaluated using a publicly available bankruptcy dataset from Kaggle, which includes various financial ratios and company details. The dataset was split into 80% for training and 20% for testing. To address class imbalance, SMOTE was applied, ensuring equal representation of bankrupt and non-bankrupt companies. The model's performance was compared with baseline machine learning models, including Logistic Regression, Random Forest, and Support Vector Machine (SVM). The evaluation metrics used were Accuracy, Precision, Recall, F1-score, and AUC-ROC, as these metrics provide a comprehensive understanding of predictive power, especially in imbalanced scenarios.

Table 4.1: Performance Comparison of the model

Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**
Logistic Regression	85.4	80.2	75.8	77.9
Random Forest	89.6	87.1	84	85.5
SVM	88.3	85.6	81.9	83.7

Proposed DL Model	94.2	92.8	91.5	92.1

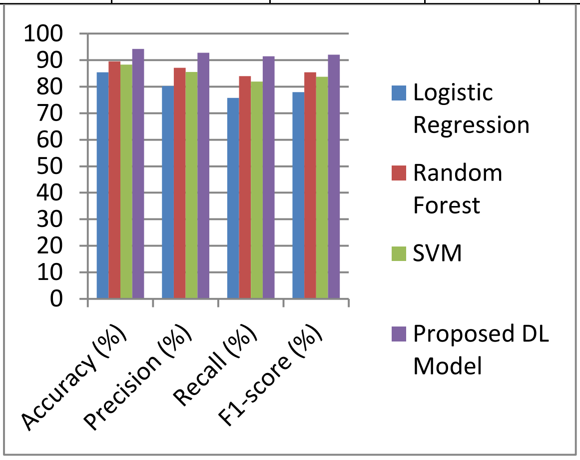

Figure 4.1: Performance comparison graph of proposed model

CONCLUSION

In this research proposed a deep learning-based framework for accurate bankruptcy prediction, addressing the critical issue of class imbalance in financial datasets through SMOTE. By leveraging advanced neural network architectures with dropout regularization and adaptive optimization techniques, the proposed model achieved superior predictive performance compared to traditional machine learning models, reaching an accuracy of 94.2% and an AUC-ROC score of 0.96. These results demonstrate the model's effectiveness in detecting financially distressed firms, which is essential for mitigating risks and supporting informed decision-making by investors, creditors, and policymakers. The work underscores the transformative potential of deep learning in financial risk assessment by improving accuracy, robustness, and sensitivity toward minority classes. However, real-world scenarios often involve dynamic financial environments, requiring continuous model adaptation.

Future work will focus on integrating time-series data to capture temporal patterns in financial health and exploring hybrid deep learning architectures combining CNN and LSTM for enhanced predictive capability. Additionally, implementing explainable AI (XAI) techniques will improve transparency and trustworthiness, facilitating adoption in regulatory and financial decision-making processes. Deployment as an interactive web-based platform is also envisioned for practical application.

REFERENCES

1. Akshat Gaurav et. al. Bankruptcy forecasting in enterprises and its security using hybrid deep learning models, Cyber Security and Applications Volume 3, December 2025, https://doi.org/10.1016/j.csa.2024.100070
2. Manel Hamdi et al "Artificial Intelligence Techniques for Bankruptcy Prediction of Tunisian Companies: An Application of Machine Learning and Deep Learning-Based Models" 2024, 17, 132. https://doi.org/10.3390/jrfm17040132, https://www.mdpi.com/journal/jrfm
3. Raffael F¨orch Brenes et al "An intelligent bankruptcy prediction model using a multilayer perceptron" https://doi.org/10.1016/j.iswa.2022.200136 Elsevier 2022

4. Mu-Yen Chen "Bankruptcy prediction in firms with statistical and intelligent techniques and a comparison of evolutionary computation approaches" Elsevier 2011
5. Adrian Gepp and Kuldeep Kumar∗ "Predicting Financial Distress: A Comparison of Survival Analysis and Decision Tree Techniques" Elsevier 2015
6. Yang Lu,1 Nianyin Zeng,2 Xiaohui Liu,3,4 and Shujuan Yi1 "A New Hybrid Algorithm for Bankruptcy Prediction Using Switching Particle Swarm Optimization and Support Vector Machines" Hindawi 2015
7. Ligang Zhou "Performance of corporate bankruptcy prediction models on imbalanced dataset: The effect of sampling methods" Elsevier 2012
8. Lingga Hardinata1 , Budi Warsito1 , Suparti1 Bankruptcy prediction based on financial ratios using Jordan Recurrent Neural Networks: a case study in Polish companies IOP Conf. Series: Journal of Physics: Conf. Series 1025 (2018) 012098 doi :10.1088/1742-6596/1025/1/012098
9. Wenhao Zhang Machine Learning Approaches to Predicting Company Bankruptcy Journal of Financial Risk Management,2017,6,364-http://www.scirp.org/journal/jfrm ISSN Online: 2167-9541 ISSN Print: 2167-9533
10. Björn mattsson & olof steinert corporate bankruptcy prediction using machine learning techniques department of economics university of gothenburg school of business economics and law,2017
11. Duaa Alrasheed1 , Dongsheng Che1 Improving Bankruptcy Prediction Using Oversampling and Feature Selection Techniques Int'l Conf. Artificial Intelligence | ICAI'18 |
12. Jacky C. K. Chow "analysis of financial credit risk using machine learning" Aston University Birmingham, United Kingdom April 2017
13. Kalyan Nagaraj and Amulyashree Sridhar "a predictive system for detection of bankruptcy using machine learning techniques" International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.5, No.1, January 2015
14. M. Krivko, "A hybrid model for plastic card fraud detection systems," Expert Systems with Applications, vol. 37, no. 8, pp. 6070–6076, Aug. 2010.
15. Benson Edwin Raj, A. Annie Portia, "Analysis on Credit Card Fraud Detection Methods", IEEE International Conference on Computer, Communication and Electrical Technology – ICCCET2011, 978-1-4244-9394-4/11, 2011 IEEE.
16. David Opitz and Richard Maclin, "Popular Ensemble Methods: An Empirical Study", Journal of artificial intelligence research 169-198, 1999.
17. L. Breiman, "Bagging predictors," Machine Learning, vol. 24, no. 2, pp. 123–140, 1996.
18. Freund, Y., & Schapire, R. (1996). Experiments with a new boosting algorithm. In Proceedings of the thirteenth international conference on machine learning, Bari, Italy (pp. 148–156).
19. Wolpert, D. H. (1992). Stacked generalization. Neural Networks, 5(2), 241–259.
20. Masoumeh Zareapoor, Pourya Shamsolmolia, "Application of Credit Card Fraud Detection: Based on Bagging Ensemble Classifier", International Conference on Intelligent Computing, Communication & Convergence, (ICCC 2015), Elsevier, Procedia Computer Science 48 (2015) 679 – 685.
21. https://www.kaggle.com/datasets/fedesoriano/company-bankruptcy-prediction