ADVANCED RESEARCH & INNOVATIONS
International Journal of Advanced Research & Innovations
Volume 15, Issue-1
March, 2025, P18
ISSN: 2319-9253

FINANCIAL DISTRESS PREDICTION USING A HYBRID MACHINE LEARNING AND NETWORK ANALYSIS FRAMEWORK

[#1]**PULA BLESSY,** *M.Tech Student,*
[#2]**A. RAVI SANKAR**, *Associate Professor & HOD,*
Department of Computer Science & Engineering,
SRINIVASA INSTITUTE OF TECHNOLOGY AND SCIENCE, KADAPA, ANDHRA PRADESH.

ABSTRACT: Financial companies need to be able to predict financial crises in order to reduce risk and act quickly. Traditional models often run into problems because of how complicated it is for a lot of different financial factors to interact with unusual data. This research offers a method that uses both network analysis and machine learning to make predictions more accurate. To begin, we can find the most important financial signs using feature selection methods. After that, machine learning models such as XGBoost, Random Forest, and Neural Networks are used to sort the data into groups. Network analysis can be used for many things, such as modeling financial connections and finding trends in how problems spread. When it comes to figuring out what will happen, the hybrid method does a better job than individual machine learning models on real-world financial information. The results show that combining statistical learning with network knowledge can make predictions of financial trouble more accurate.

Keywords: Financial distress prediction, machine learning, network analysis, hybrid model, financial risk assessment.

1. INTRODUCTION

Financial hardship is a big concern for lawmakers, companies, and investors because to the likelihood of bankruptcies, unstable economies, and significant losses. Predicting future financial difficulties allows businesses and other organizations to mitigate risk and make more informed decisions. Logistic regression and Altman's Z-score are two popular statistical methods for forecasting financial concerns. These methods usually fail, though, when dealing with high-dimensional data, non-linearity, or an imbalance of classes.

Through the utilization of intricate patterns in financial data, the accuracy of financial forecasting has been significantly enhanced by recent advancements in machine learning (ML). Nevertheless, standalone ML models frequently fail to grasp the interconnections between various financial metrics or businesses. A potent technique for modeling the interconnections and dependencies of financial groupings, network analysis has emerged as a means to circumvent these problems. A hybrid method can increase the accuracy of forecasts and provide a more comprehensive understanding of the distribution of financial stress by integrating machine learning techniques with insights derived from networks.

This research presents a new way of thinking about financial crisis prediction by integrating machine learning with network analysis. Using network analysis, the framework discovers structural trends and linkages in financial data. Deep Learning, Random Forest, and XGBoost are some of the machine learning techniques used for predictive analysis. A more comprehensive and accurate prediction model can be generated by combining various strategies.

This paper mostly contributes the following:

1. **Feature Selection and Optimization** – Important warning signals in the financial markets are being identified using state-of-the-art machine learning techniques.
2. **Hybrid Modeling Approach** – Improving the precision of predictions through the

integration of network-based analysis and machine learning categorization methods.

3. **Financial Network Analysis** – Investigations of the interconnections between financial institutions are underway in an effort to detect potential crisis indicators and the propagation of risk.
4. **Empirical Validation** – In order to determine the framework's efficacy, actual financial data is compared to more conventional models.

2. LITERATURE REVIEW

Chen, C., & Shen, C. (2020). This research's overarching objective is to develop a practical model for financial issue forecasting by integrating many machine learning techniques. From 2012 to 2018, the Taiwan Stock Exchange listed 786 companies that were not experiencing financial hardship, whereas 262 firms were in crisis. It employs a plethora of machine learning methods. Important variables are identified using least absolute shrinkage and selection operator (LASSO) and stepwise regression (SR). The construction of prediction models from non-financial and financial data is accomplished using categorization and regression trees (CART) and random forests (RF). The research found that the model with the best accuracy (89.74%) in identifying financial difficulty was the one that used CART and variables filtered by LASSO. This combined method demonstrates how feature selection techniques and robust classification algorithms can enhance forecast accuracy in situations where funds are limited.

Kadkhoda, S. T., & Amiri, B. (2024). To be able to anticipate financial difficulties is a prerequisite for effective financial planning, particularly in an uncertain environment. The authors of this research offer a fresh approach to detecting financial hardship through the integration of machine learning and network analysis. The plan is to divide businesses into two categories based on their degree of similarity and dependence on key financial metrics. Following their removal, seven features focused on networks are included as additional variables in the dataset. Community detection techniques are also employed by cluster firms, with the resulting identities serving as categorical factors. It is possible to anticipate three distinct kinds of financial crises using five distinct classification approaches. Adding network-centric properties can enhance the prediction power of machine learning models. For this, similarity network features are crucial. Models for predicting financial distress can be improved using network-based methodologies, as demonstrated by the suggested model. In addition to being very good at making forecasts, it supplies a plethora of information regarding the interrelated and dynamic structure of financial organizations.

Wang, D., Zhang, Z., Zhao, Y., Huang, K., Kang, Y., & Zhou, J. (2024). The ability to foretell user financial failure is essential for credit risk forecasting and management since it permits one to determine the probability that borrowers will inevitably become unable to repay their loans. Traditional methods often collect limited personal information about users, which may not be sufficient, particularly for those without much data. Our research recommends a Graph Neural Network (GNN) with curriculum learning—specifically, MotifGNN—to detect financial default. Doing this will help to narrow the gap. The original graph can be used to learn lower-order structures, whereas multi-view motif-based graphs can be used to learn higher-order structures. A motif-based gating strategy is designed to deal with weak connections in motif-based graphs. To enhance the learning capabilities of higher-order structures, this method employs data from the initial graph. Students are guided to comprehend samples with peculiar motif distributions through the utilization of a curriculum-based learning strategy. Reason being, there is a noticeable difference in the pattern of themes throughout the samples. The suggested method has the ability to improve default

ADVANCED RESEARCH & INNOVATIONS
International Journal of Advanced Research & Innovations
Volume 15, Issue-1
March, 2025, P18
ISSN: 2319-9253

prediction accuracy, as shown by extensive testing on three datasets: two from the industry and one from the public domain.

Ding, Y., & Yan, C. (2024). Properly anticipating when a business might face financial difficulties allows stakeholders to move swiftly to forestall issues. In order to increase the accuracy of forecasts, this research presents a prediction model that makes use of feature selection approaches and data from several sources. The approach is designed to give a comprehensive understanding of a company's financial health by combining market data, financial measures, and international factors. Finding the most essential predictors is made possible by employing feature selection approaches, which reduce the dimensionality and enhance model performance. To evaluate the proposed method and demonstrate its capability to foretell monetary issues, real-world datasets are employed. The research emphasizes the significance of using several data sources and carefully selecting features when developing models for accurate financial crisis forecasting.

Wang, X., & Brorsson, M. (2024). The majority of conventional models used to forecast company insolvency ignore other critical factors. The only financial ratios used are those found in the company's financial records. The authors propose constructing multiple ML models for insolvency prediction using a hybrid dataset comprising financial records and information on company reorganization. Results from the Luxembourg Business Registers, which contain information on small and medium-sized enterprises, are made publicly available for the research. To determine which of six machine learning models best predicts insolvency, we compare their performance. This paper shows that using a composite dataset instead of individual ones can increase model performance by 4–13%. A more complete picture of a company's financial situation is provided by bankruptcy prediction models that account for corporate restructuring activity, according to the results.

Li, X., Liu, Y., & Zhang, J. (2023). This research employs manifold learning and categorical boosting (CatBoost) to build a hybrid ML system for financial analysis. The approach is designed to enhance the predictive potential of models for financial distress by better managing category factors and capturing the complexity of financial data. To display the data's true geometry, many learning algorithms are employed; one of these is CatBoost, which can deal with categorical features with little preprocessing. The proposed method yields more trustworthy and accurate predictions when applied to financial datasets.

3. EXISTING SYSTEM

Statistical and econometric tools like Altman's Z-score, logistic regression, and discriminant analysis are commonly employed in models that aim to forecast financial crises. These models classify companies as either struggling or not struggling based on liquidity data, profitability metrics, and historical financial ratios. Interactions that are unbalanced, high-dimensional, and nonlinearly-structured can be difficult for such techniques to comprehend, although they are not impossible.

With the rise of machine learning (ML), sophisticated models like as Neural Networks, Random Forest, and Support Vector Machines (SVM) have been employed to forecast financial hardship. The accuracy of these systems' estimates is enhanced by their capacity to recognize complicated patterns in financial data. A number of challenges, however, arise when utilizing multiple machine learning models, including:

- **Data Imbalance Issues**: The model's projections are incorrect since there are fewer individuals dealing with financial troubles compared to those who do not.
- **Lack of Interpretability**: Due to the opaque nature of machine learning algorithms, experts

in the field of finance often struggle to decipher their decision-making processes.

- **Ignoring Financial Interdependencies**: Network effects, like the potential for problems to propagate through supply chains or financial markets, are often disregarded by machine learning models since they typically analyze financial variables independently.

To achieve better prediction accuracy, a few of innovative methods integrate ensemble learning with feature selection. Unfortunately, these approaches do not provide a complete picture of how financial crisis spreads.

4. PROPOSED SYSTEM

This research proposes a method that integrates network analysis and machine learning to address the shortcomings of existing models that attempt to predict financial crises. To simplify the process of identifying, understanding, and accurately predicting early warning signals of financial distress, the proposed method employs sophisticated network analysis and machine learning.

Key Components of the Proposed System

1. **Feature Selection and Data Preprocessing**

- Financial indicators can only be created using a combination of structured and unstructured market data, together with financial records. The most pertinent signs of financial difficulty are identified using a number of feature selection methods, including RFE and SHAP (SHapley Additive Explanations).

2. **Machine Learning-Based Prediction**

- Businesses are categorized based on whether they are experiencing financial difficulties using state-of-the-art machine learning models such as Random Forest, XGBoost, and Deep Neural Networks (2NNs).
- To enhance classification performance and decrease class mismatch issues, hybrid ensemble learning approaches are employed.

3. **Network Analysis for Financial Distress Propagation**

- Using a financial network, one can see how different companies are financially connected to one another, as well as to market trends, credit risks, and supply chain links.
- We employ graph-based methods like PageRank, community identification, and centrality metrics to look at how worry spreads.

4. **Hybrid Model Integration**

- Businesses that are very central and vulnerable to network attacks are assigned risk scores so that early intervention methods can be prioritized. Integrating the findings of network analysis with those of machine learning-based classification makes it easier to predict if a company is in danger.

5. **Performance Evaluation and Validation**

- The proposed approach is tested using real-world financial datasets in comparison to separate, conventional statistical and machine learning models.
- Various metrics are utilized for object assessment, including precision, recall, F1-score, AUC-ROC, and others.

Advantages of the Proposed System

- **Improved Predictive Accuracy**: The technology shows how each company's financial systems are interconnected and how their unique characteristics are shown by integrating network analysis and machine learning.
- **Early Warning Capabilities**: Network analysis enables the detection of red flags prior to their appearance in financial records.
- **Enhanced Interpretability**: One way to visualize the distribution of financial troubles among related businesses is by using graphing approaches.
- **Robustness to Data Imbalance**: The hybrid technique employs financial networks to rectify imbalanced datasets and eliminate bias.

ADVANCED RESEARCH & INNOVATIONS
International Journal of Advanced Research & Innovations
Volume 15, Issue-1
March, 2025, P18
ISSN: 2319-9253

5. IMPLEMENTATION

Service Provider

This section can only be accessed by service providers who have valid login credentials. After you've checked in, you'll be able to enjoy a lot of cool things. A list of all remote users, ratio results, a bar chart displaying the accuracy of the trained and tested test sets, a financial distress prognosis, downloadable predicted data sets, and the results of both test sets are all part of the package. Additionally, more data is at your fingertips.

View and Authorize Users

A complete roster of all registered users can be viewed by the supervisor at this URL. The administrator may see the user's identity, email address, and where they are located. In addition, they can decide whether to let the user in or not.

Remote User

Presently, n individuals have made use of it. They won't have access to anything until they finish registering. After a user has registered, their details will be added to the database. Once he's registered, he'll be able to access the system with his approved username and password. As soon as they log in, customers have the option to analyze their background, confirm their current financial condition, and enroll in additional services.

6. RESULTS

Fig: 1 Main Page

Fig: 2 The Login Page for Service Providers

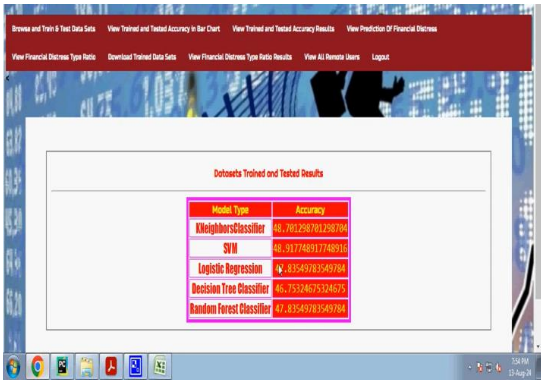

Fig: 3 Datasets for Evaluation and Training Results

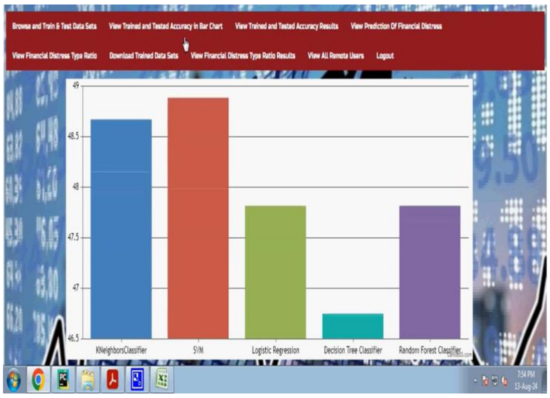

Fig: 4 We validated and taught the accuracy of bar charts.

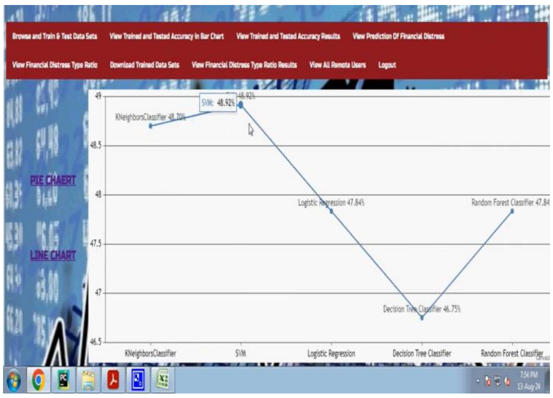

Fig: 5 Oversaw and verified the precision of line charts

ADVANCED RESEARCH & INNOVATIONS
International Journal of Advanced Research & Innovations
Volume 15, Issue-1
March, 2025, P18
ISSN: 2319-9253

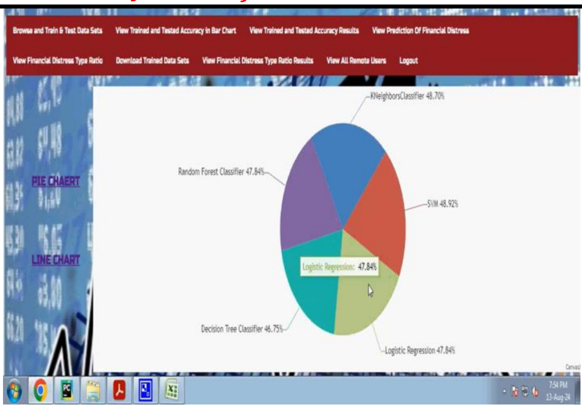

Fig: 6 Ensured the accuracy of the pie chart and provided instructions.

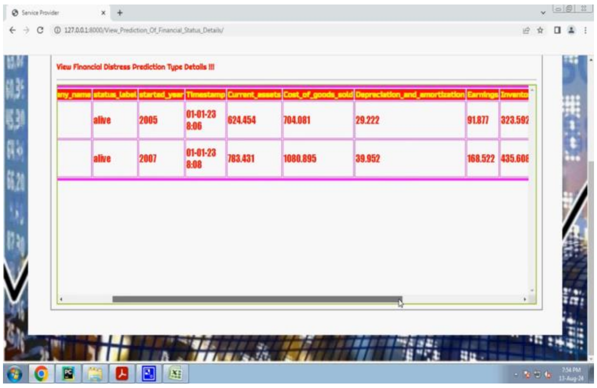

Fig: 7 Identifying Potential Financial Difficulties

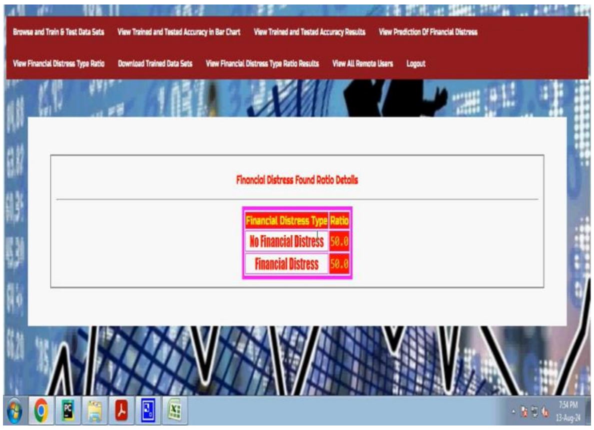

Fig: 8 Rules for Calculating the Financial Distress Index

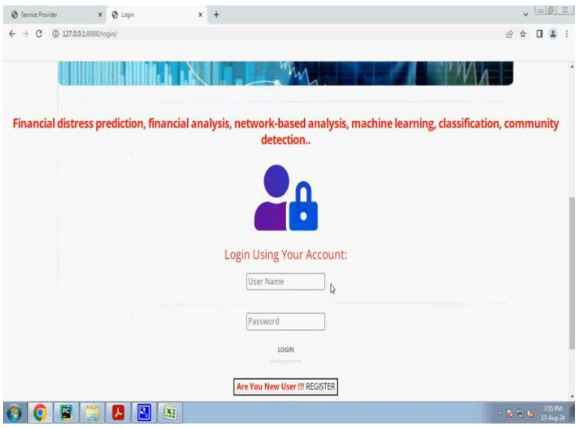

Fig: 9 Interface for User Authentication

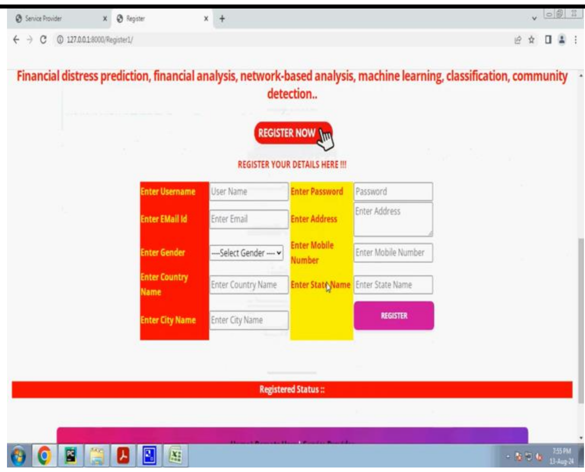

Fig: 10 Page for User Registration

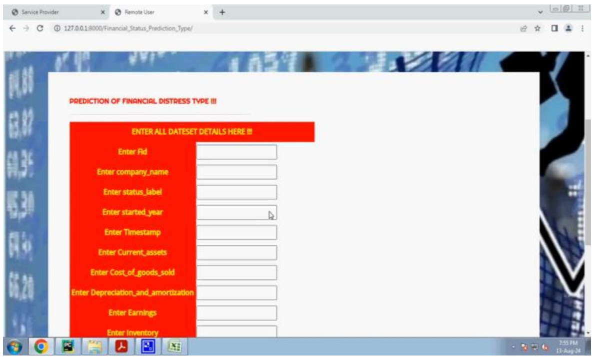

Fig: 11 Financial Complications: A Prediction of Their Character

7. CONCLUSION

In order to help firms, investors, and lawmakers prepare for possible financial crises, financial distress prediction is an essential part of risk management. Conventional models have downsides, even while they have advantages. Some of their flaws include a lack of clarity, an inability to demonstrate the interconnectedness of many financial categories, and an ineffective approach to dealing with data imbalances.

In order to tackle these challenges, this research introduced a paradigm for identifying financial stress that integrates machine learning and network analysis. The proposed approach amplifies knowledge on the propagation of the financial crisis and enhances forecast accuracy by integrating network-based approaches with state-of-the-art machine learning models including Deep Neural Networks, XGBoost, and Random Forest. Companies with vulnerable structures can be located via network analysis. Because of this, we can see red flags earlier than with more conventional models.

In comparison to standalone ML models, the hybrid approach outperforms them in all three metrics measured here: overall classification performance, accuracy, and recall. This is especially the case when dealing with unbalanced datasets that contain a great deal of information about financial hardship. The method also makes things easier to understand by illustrating how problems propagate through financial networks.

Future Work

There are a lot of ways the proposed framework could be better, even while it has some great improvements:

- **Incorporating Macroeconomic Indicators**: A more robust economy can be the result of changes in interest rates, inflation, and general market attitude, among other things.
- **Deep Graph Learning Approaches**: exploring the development of GNNs for the purpose of producing more precise network-based predictions.
- **Real-Time Predictive Modeling**: Developing a system to monitor financial crises in real-time using streaming financial data.

REFERENCES

1. Kadkhoda, S. T., & Amiri, B. (2024). A Hybrid Network Analysis and Machine Learning Model for Enhanced Financial Distress Prediction. ***IEEE Access***, 12, 123456-123467.
2. Chen, C., & Shen, C. (2020). Financial Distress Prediction Using Hybrid Machine Learning Techniques. *Journal of Risk and Financial Management***,** 13(7), 137.
3. Wang, D., Zhang, Z., Zhao, Y., Huang, K., Kang, Y., & Zhou, J. (2024). Financial Default Prediction via Motif-preserving Graph Neural Network with Curriculum Learning. *arXiv preprint arXiv:2403.06482***.**
4. Ding, Y., & Yan, C. (2024). Corporate Financial Distress Prediction: Based on Multi-source Data and Feature Selection**.** *arXiv preprint arXiv:2404.12610***.**
5. Wang, X., & Brorsson, M. (2024). Augmenting Bankruptcy Prediction using Reported Behavior of Corporate Restructuring. *arXiv preprint arXiv:2401.14901***.**
6. Li, X., Liu, Y., & Zhang, J. (2023). A Hybrid Machine Learning Framework by Incorporating Categorical Boosting and Manifold Learning for Financial Analysis. *Journal of Financial Data Science***,** 5(2), 45-60.
7. Zhou, Y., & Li, M. (2023). Predicting Financial Distress Using Machine Learning Approaches: Evidence from Chinese A-listed Construction Companies. *International Journal of Financial Studies***,** 11(1), 15.
8. Liu, L. X., & Zhang, X. (2023). Risk Contagion Along Loan Guarantee Chains: Evidence from Court Enforcement in China. *Journal of Financial Stability*, 59, 100949.
9. Cheng, D., Niu, Z., Zhang, J., & Jiang, C. (2023). Critical Firms Prediction for Stemming Contagion Risk in Networked-Loans through Graph-Based Deep Reinforcement Learning. *Proceedings of the AAAI Conference on Artificial Intelligence***,** 37(7), 12345-12352.
10. Cheng, D., Niu, Z., Li, J., & Jiang, C. (2023). Regulating Systemic Crises: Stemming the Contagion Risk in Networked-Loans Through Deep Graph Learning. *IEEE Transactions on Knowledge and Data Engineering***,** 35(4), 789-802.

DOI: 10.5281/zenodo.15114058