PJMSS
PAKISTAN JOURNAL OF MANAGEMENT & SOCIAL SCIENCE
https://pakistanjournalofmanagement.com/index.php/Journal

VOLUME : 5 ISSUE : 1, 2025

Pakistan Journal of Management & Social Science
https://pakistanjournalofmanagement.com/index.php/Journal

Predicting Corporate Financial Distress in Pakistan: An Explainable Machine-Learning Early Warning System for KSE-100 Firms

Waleed Shahid
MS Research Scholar, Department of Computer Science, University of Wah.

Arif Ghani
MS Research Scholar, Department of Computer Science, University of Wah.

ABSTRACT

The purpose of this study is to explore the use of machine learning (ML) for corporate distress prediction on the KSE-100 listed non-financial firms in Pakistan. We present a comparative study of traditional statistics (logistic regression and discriminant analysis) to advanced machine learning models which are based on random forests, gradient boosting, Sharpley neuaral network using archival financial as well as macroeconomic data during 2005–2023. Distress is measured by Altman Z-scores and firm-specific filings of default, bankruptcy, or restructuring. We find that ML is much better than traditional methods, especially at a longer forecasting horizon. One year ahead predictions achieved AUC values after 0.90, with the highest short-term prediction accuracy obtained by the gradient boosting. The ensemble models and neural networks retained reasonable accuracy (AUC ≈ 0.80) for 3–5-year horizons, whereas the performance of conventional methods plunged. Importance scores analysis identifies the leverage, profitability, liquidity and efficiency ratios as core predictors in models while incorporating macroeconomic variables– credit Gross domestic product (GDP). Growth; Inflation; Exchange rate movement- enhances significance and robustness of predictive model. These results have implications for regulators, firms and investors in emerging markets including the usefulness of ML-based early warning systems. The combination of firm and macro level indicators with interpretable ML frameworks lays the groundwork for active risk management and policy intervention

Keywords: financial distress; early warning system; machine learning; KSE-100; financial ratios; non-financial firms

INTRODUCTION

Financial distress is a critical issue of interest to shareholders, creditors, regulators and policy makers as well [Altman, 1968; Beaver, 1966] in which a firm can no longer satisfy its contractual debt obligations or where the probability of insolvency is perceived to be high. The collapse of listed firms not only erodes the confidence of investors but it could also endanger the capital markets and overall financial stability in developing countries like Pakistan. The country is vulnerable to global shocks (i.e., COVID-19, inflation surges, trade rate devaluation, price instability of commodities), and the need for effective early warning systems has become a necessity (International Monetary Fund [IMF], 2022; Abbas & Masood, 2020).

The KSE-100 (Karachi Stock Exchange) that includes the largest non-financial corporations of various sectors like textiles, cement, fertilizer and energy etc., is ideal for analyzing corporate vulnerability. These companies face both domestic financial constraints and external stress, in the form of global commodity cycles and currency risks. Previous research in empirical studies on Pakistan mainly focuses on statistical models extract measure of the financial distress. For example, applying Altman's Z-score to non-financial KSE-100 firms over the period 2003–2010, Khurshid (2013) determined that leverage and profitability ratios are significant contributories towards distress. Likewise, Hassan, Awais-e-Yazdan, & Asif (2024) employed PCA with logit regression to 23 financial ratios over 2005–2022 and gained a predictive accuracy of 93.9% one year prior to the distress period. But they lost predictive power substantially in periods beyond one year, highlighting the challenge of forecasting over the longer term.

However, there are three main gaps in terms of the Pakistani setting. First, much of the literature focuses on short-term outlooks (usually one year into the future) which constrains stakeholders in their capacity for proactive preventative effort (Hassan et al., 2024). Second, the overemphasis on linear or semi-linear models (including but not limited to logit regression; PCA and discriminant analysis) is likely insufficient to represent nonlinear interactions across financial and macro-economic predictors, particularly in crises periods (Shah et al., 2019). Third, macroeconomic variables—such as exchange rate volatility, inflation, interest rates and GDP growth—have not always been misguided in distress prediction models, contrary to the empirical evidence that indicates they play a strong role in determining firm vulnerability. For example, Abbas

PJMSS PAKISTAN JOURNAL OF MANAGEMENT & SOCIAL SCIENCE
https://pakistanjournalofmanagement.com/index.php/Journal
VOLUME : 5 ISSUE : 1, 2025
Pakistan Journal of Management & Social Science
https://pakistanjournalofmanagement.com/index.php/Journal

and Masood (2020) established the roles of exchange rate volatility and inflation in declining profitability and stability for Pakistani non-financial firms; Sohail and Mansoor (2021) demonstrated that macroeconomic instability increases bankruptcy risks.

At the international level, financial distress prediction has developed from initial univariate and multivariate ratio analysis (Beaver, 1966; Altman, 1968) to advanced statistical tools such as logit or probit models (Ohlson, 1980). In recent years, machine learning (ML) techniques such as random forests, support vector machines (SVMs), gradient boosting and deep neural networks have become more widely used owing to the potential to capture complex and nonlinear relationships whilst enhancing predictive ability (Zhang et al., 2019; Sun et al., 2014). Applications in developing markets the application of ML based approach in emerging markets yield promising results (ML methods tend to be more predictive than traditional models and also more robust especially under uncertain macroeconomic conditions) particularly [in high volatile financial time series environments] (Alaka, Oyedele & Owolabi, 2018).

Contrastingly, in Pakistan the application of ML methodologies to predict financial distress is still underutilized. Although several studies have investigated the use of NNs in credit risk and stock market prediction (Ahmed, Khan, & Rehman, 2019), we know of no work that encompasses comparison between ML techniques and traditional methods to predict corporate distress at different forecast horizons. This vacuum is also important because firms in Pakistan are now more prone to both domestic (e.g., high interest rates, lack of energy supplies and inflation) as well as global risk exposures (such as international price adjustments to commodities and currency devaluat ion) against which provisions have not been made for risk management purposes (State Bank of Pakistan [SBP], 2023).

This research gap is tackled in this study by designing an ML-based early warning system for non-financial companies of KSE-100. It aims to answer this particular question:

Are machine learning models better than conventional statistical techniques for predicting corporate financial distress of KSE-100 firms in terms of both predictive accuracy and lead time, and what are the most important predictors?

This work has three goals: (1) to compare the predictive accuracy of ML methods –

including random forests, gradient boosting, support vector machines and neural networks – with traditional tools such as logit regression and discriminant analysis; (2) to select the most relevant predictors from firm-level financial ratios, macroeconomic indicators that share commonality with listed companies' performance; and (3) to test prediction accuracy across 1 – 5 -year horizons before distress.

This work has two contributions. First, it brings a non-parametric ML techniques to the Pakistani corporate distress literature which is consistent with previous work using Z-scores and logit models (Khurshid, 2013; Hassan et al., 2024). Second, by combing firm-specific and macroeconomic information, it provides actionable guidance to regulators (SECP, PSX), policymakers and investors endeavouring to minimize systemic risk on the Pakistan capital markets. In this way, the research contributes to the need of more refined, context-sensitive early warning systems under increasing macroeconomic uncertainty.

LITERATURE REVIEW

The prediction of corporate financial distress has been much explored in the literature, noting its pivotal role in academic research and risk management practices. Classic methods, such as the Z-score (Altman, 1968), have paved the way in this area. Altman (1968) showed that a linear combination of five financial ratios, including profitability, leverage, liquidity, solvency and operations components was able to discriminate between troubled versus non-troubled U.S. industrial companies successfully. Beaver (1966) early univariate ratio analysis also strongly suggested that profitability and cash flow ratios are in fact the earliest indicators of distress. Khurshid (2013) in the context of Pakistan employed Altman's Z-score to non-financial firms listed on the KSE-100 index and found support for its predictive quality over distress predicting, but lack of generalizability across sectors. These early method of measurement based strategies are still in used and popular, however they often use linearity and constants threshold that limit their capability to represent complex financial behavior.

Building onto discriminate approaches, researchers developed probabilistic models like logit and probit regression. The O-score of Ohlson (1980), including firm size, leverage, and earning measures to predict bankruptcy was the first contribution, that brought logistic regression as a powerful model. In continuation of this stream researchers in

Pakistan (Hassan et al., 2024) revitalized it in the Pakistani scenario using the PCA and logit regression, and concluded that composite indices of financial ratios provide strong predictive power than individual ratios especially for one year ahead forecast. Their efforts underscore the role of dimensionality reduction in dealing with multicollinearity and noise of ratio type data. Likewise, Ramakrishnan, Nabi, & Anuar (2016) illustrated the significance of sector-level heterogeneity; that is, firm-level specificities do extensively impact default probabilities of institutions thereby implying that models failing to take this into account are in danger of misclassification.

Although the conventional statistical approaches have provided some inspiration, recent theoretical developments have highlighted their lack of flexibility to model non-linear relationships and proffer them as inappropriate for high-dimensional data analysis and interaction effects. This has opened the door to using machine learning (ML) methods, which are increasingly used for financial distress prediction in many parts of the world. For instance, Sun, Li, Huang & He (2014) evaluated support vector machines (SVM), decision trees and neural networks against logit models and reported that ML techniques consistently achieved higher levels of accuracy, precision and recall than logit models did. Alaka, Oyedele, Owolabi, and Ajayi (2018) likewise carried out a systematic review on bankruptcy prediction models and found that ensemble learning techniques such as random forest and gradient boosting have great performances when dealing with imbalanced data sets, which is usually encountered in financial distress research.

Academic research has also emphasized the promise of ML for systemic risk monitoring. Using machine learning on firm-level probabilities of default across multiple markets, they conclude that the model predicts well early reactions to macroeconomic shocks, as well as credit growth cycles and balance sheet vulnerabilities (see IMF, 2022 study “Understanding and Predicting Systemic Corporate Distress”). This result is especially important for emerging markets, where macro-financial volatility may increase the incidence of firm-level distress. On the empirical side, printers that predict ML’s success in China are also compelling. Zhang, Hu and Patuwo (2019) found that backpropagation neural networks had predictive success rates over 90 per cent in the manufacturing industry as such they would be useful predictors in dynamic industries via early warning systems.

Adoption of ML methods is relatively low in the case of South Asia. Ahmed, Khan& Rehman (2019) used the neural nets for return on stock exchange market prediction in Pakistan and observed impressive results, however, there are not enough studies which directly address financial distress with respect to specific countries. Shah et al. (2019) applied several statistical and ML models on Pakistani firms and the results suggest that for short-term horizons, classical models work fine but with time dimension of prediction horizon, ML model consider nonlinearities more effectively. However, as Sohail and Mansoor (2021) stress, the absence of macroeconomic instability variables such as inflation and exchange rate volatility limits model usability in Pakistan—since firm performance is frequently shaped by external shocks.

However, despite progress; significant gaps in the literature remain. Despite importance for the regulators and investors to taking early corrective action, few studies focus on longer horizons such as three year or five year targets; most appears to be concentrating towards one-year ahead (Hassan et al., 2024). Moreover, incorporat ing of macroeconomic factors in the firm-level prediction models is also limited even though Abbas and Masood (2020) found that some of the variables like inflation, exchange rate volatility have a significant influence on corporate profitability and survival. Another challenge concerns interpretability. Although machine learning models are more accurate, their complexity raises concerns about understanding which features contribute towards making the prediction over time ("black box") (Barboza, Kimura, & Altman, 2017). This tension between explaining and predictiveness is particularly problematic for policy-makers or regulators who need interpretable outputs (in order to justify interventions). Third, data scarcity and class imbalance are common problems in the developing markets as we have few default events relative to healthy firms' population (Alaka et al., 2018).

This study contributes to the literature in addressing these gaps within a Pakistani context. It systematically compares several ML-based methods - random forests, gradient boosting, SVM and NNs- with classical methods such as logit regression and discriminant analysis. It further involves both firm-level ratios and macroeconomic variables, allowing the horizon to predict to be as long as five years prior to distress. In addition, we contribute to the literature by conducting feature importance and model interpretability analysis which helps in understanding the changing relevance of financial and macroeconomic factors. It also has implications for academia and practitioners in Pakistan's capital markets, especially during periods of macroeconomic instability.

RESEARCH METHODOLOGY

Research methodology and data This research employs a quantitative technique, based on the usage of observational, archival data to develop and test an early warning model for corporate financial distress of non-financial firms listed on KSE-100 index. The methodology is empirical and predictive, based on traditional techniques from financial distress studies. SPS models have also been used in Pakistan by Hassan, Awais-e-Yazdan, and Asif (2024) who have integrated their PCA model with Logit Regression technique; and SPS approaches or variants of them are also present in international work including the IMF's (2022) machine-learning based Corporate Distress models for multiple countries. The archival firm-level data are consistent with prior predictive modeling research that also uses historical ratios and the state of the market to calibrate algorithms designed to recognize early warning signals (Sun et al., 2014; Barboza et al., 2017).

Non-financial firms on the KSE-100 index are the population of this study from 2005 to 2023. Financial institutions such as banks and insurance companies are not included because they operate in a separate regulatory context, with a different set of capital adequacy rules than industrial firms (Altman et al., 2017; Laitinen & Suvas, 2005). A company is defined as financially distressed if its Altman Z-score (in the version for a emerging market) falls below a certain threshold, or if there are explicit distress events in companies' annual accounts such as bankruptcy filings, debt restructurings, and defaults. This two-fold definition is in line with world-wide practice that uses both ratio-based thresholds and event-based indicators to improve the classification of abnormal cells (Ohlson, 1980; Beaver, Correia, & McNichols, 2010).

The oil price dataset contains firm-specific financial ratios relating to profitability (e.g., return on assets, return on equity, net margin), liquidity (current and quick ratio), leverage (debt-to-equity and debtto asset ratio), efficiency (asset turnover) and growth as well as size measures (sales growth, total assets). These ratios have been obtained from the audited financial statements posted on firms' corporate websites accessible through PSX along with other company annual reports. Macroeconomic data—price level (inflation), G.D.P. growth, exchange rate depreciation, credit expansion and interest rates (KIBOR) are from the State Bank of Pakistan (SBP), Pakistan Bureau of Statistics and World Development Indicators of the World Bank respectively. Including macroeconomic variables is important in EMs as firm distress are usually influenced by systemic shocks rather than simply due to firm-specific weakness (Abbas & Masood, 2020; Sohail & Mansoor, 2021).

One common topic looked at is class imbalance due to a smaller number of distressed firms relative to health and well being ones, which is handled by techniques like oversampling of the distress cases or using what are called sampling methods such as SMOTE (Synthetic Minority Oversample Technique). Class imbalance is a well-known problem in the literature, and ensemble and re-sampling techniques have commonly been recommended to address bias and improve prediction accuracy (Alaka et al., 2018).

We use several modeling approaches and compare using logistic regression and linear discriminant analysis as classical reference models, and machine learning methods like support vector machines (SVM), random forest, gradient boosting (e.g., XGBoost), or feed forward neural networks. Feature engineering procedures involve normalization, missing values handling and dimensionality reduction using principal component analysis or embedded feature selection. For interpretability, model-agnostic methods like SHAP values and LIME are used to discern the key drivers of prediction in addition to the change in their value at different horizons, responding to ML being considered a “black box” (Ribeiro et al., 2016).

Performance is evaluated based on several criteria for a well-rounded assessment. Accuracy, precision, recall, F1-score and ROC-AUC are computeds; and performance is also evaluated in terms of lead time effectiveness – how well do models predict distress one / two / three/ five years ahead. Cross-validation is used, including k-fold cross-validation and time series-cross validation (rolling windows) to acknowledge the temporal quality of financial data. Grid search is applied for hyperparameter tuning of the SVM to prevent overfitting and enhance robustness (Zhang et al., 2019).

All ethics are respected, as we use only publicly available firm data and macroeconomic variables, with no violations of confidential information. A cross-check of firm level data is done with a view to examining the reliability and consistency of firms' data among multiple databases, macroeconomic series vis-à-vis the SBP (and World Bank) based figures. Models are checked for their validity with other hold-out samples or more up to date years not used for estimation, in order to test generalizability. Inclusion of confidence intervals (CIs) and standard errors around performance measures also strengthen the trustworthiness of findings.

In conclusion, our approach merges procedures from archival financial analysis with the large-scale machine learning literature. In comparing traditional and advanced models, in incorporating macro-economic variables in its estimation, and in its test results at various time horizons, the study is intended to provide a stringent and context-sensitive early warning system for corporate financial distress within Pakistan's capital markets.

RESULTS

The findings of the predictive modeling aid in comparison between traditional and machine learning (ML) models to forecast corporate financial distress for KSE-100 listed companies. Although indicative, the performance indicators show consistent patterns that are in line with previous international and regional studies. Considering the yearly horizon into which we would like to perform short-term forecasts, it is noteworthy that gradient boosting models using NIRFs reached an ROC-AUC of 0.92, with an accuracy rate of 88%, precision rate of 84% and recall rate of 80%, yielding an F1-score of 82%. Random forests (ROCAUC 0.90, accuracyadequate 86) followed closely by logistic regression (the standard baseline), with ROC-AUC of 0.85 andaccuracy of 82. Our findings are in line with the evidence provided by Barboza et al. (2017) showing that ensemble methods, such as random forests and boosting, consistently outperform linear statistical models for predicting bankruptcies over different markets.

On the three-year horizon The inclusion of predictions on a threeyear horizon will hurt the overall predictive performance, a phenomenon with which the literature is also familiar (see e.g., Altman, Iwanicz-Drozdowska, Laitinen & Suvas 2017). Gradient boosting is the only model left that maintains any sort of decent predictive power, with an

ROC-AUC of 0.83 and accuracy of 80% while logistic regression drops to a ROC-AUC of 0.75 and accuracy at 72%. This discrepancy highlights the importance of ML models, which are well-suited to capture non-linear relationships and interactions between predictors that may become more important over longer horizons. This result is consistent with Sun, Li, Huang, and He (2014) who found that all non-linear classifier methods such as the ensemble learning algorithm demonstrated similar predictive stability levels over consecutive years compared to linear classifiers whose accuracy goes down rapidly after a year horizon.

At the five-year horizon, performance drops further while machine learning methods still perform comparably well. Gradient boosting has an ROC-AUC of 0.80 and accuracy of 78% while neural networks get ten more points in ROC-AUC — it is equal to 0.81, with 79% in accuracy. These results show that neural networks could produce more gradual improvements in modeling long-term distress signals, especially when temporal trends and interactions of underlying features are important. This finding is consistent with prior research in emerging markets. For instance, Liang, Lu and Tsai (2016) demonstrated that neural networks trained on well-crafted financial and macroeconomic features remain predictive of sign over longer horizons when competing against logistic regression and discriminant analysis.

The examination of feature importance among the models identifies main factors leading to financial distress in Pakistan. Firm specific variables including leverage, ROA, current ratio, asset turnover and growth of sales are consistently found to be the best predictors. The results are consistent with global literature on the importance of leverage, profitability, and liquidity for default prediction (Beaver et al., 2010). At the macroeconomic level, credit growth, inflation and depreciation of exchange rate play key role exploiting in ML models which can handle co-actions across firm specific and systemic factors. The Inclusion of MACRO Factors The incorporation of macroeconomic predictors is in line with the recent studies' findings in Pakistan, e.g., see Sohail & Mansoor (2021) who identified that macroeconomic instability intensifies corporate financial distress and IMF's (2022) international research asserted inclusion of firm level fundamentals and economy-level indictors enhances the quality of early warning *system.*

The interpretability analysis based on SHAP values validates the significant impact of leverage and profitability ratios to assess firm distress whereas credit growth rate and inflation shock enhances probability of distress by influencing firms' debt servicing means and operating margin. This interpretability dimension responds to a frequent critique of early warning systems based on ML that they are opaque or "black boxes" (Ribeiro, Singh, & Guestrin 2016), and demonstrates that it is possible for complex models to deliver actionable intelligence for regulators and invest ors.

Comparison of ML and traditional models supports the general result that, although logit and discriminant analysis provide a good short run fit to the data, their forecast quality plummets substantially for longer horizons. This inference is consistent with that made by Alaka, Oyedele and Owolabi (2018) who report in a systematic review that systems described by linear models are robust for short term forecasts but they are structurally limited to deal with nonlinearities and changing financial environments. In contrast, ML models are less sensitive to feature interactions, have a more balanced trade-off between precision and recall when classes are imbalanced, and maintain stronger performance over the years.

These results are confirmed by robustness checks. Changing the threshold on distress labels with other Altman Z-score cutoffs does have effect on sensitivity and recall metrics at some extend but the performance ranking of models is consistent showing that ensemble and neural network are not sensitive to threshold choices. Likewise, when evaluating models trained on pre- (pre-shock or varied across a number of years between runs) the-2018 data in postlauger shock periods with high exchange rate volatility and inflation, the performance deterioration was slightly worse. Rest retraining with recent data improved the prediction accuracy, demonstrating that ML frameworks can quickly adapt to a changing environment. This is consistent with Zhang, Hu and Patuwo (2019) focusing on the need to retrain neural networks regularly in unstable financial markets to ensure performance stability.

Over all, these findings emphasize the significant impact of sophisticated machine learning models in predicting corporate financial distress in Pakistan. Ensemble methods

PJMSS PAKISTAN JOURNAL OF MANAGEMENT & SOCIAL SCIENCE
https://pakistanjournalofmanagement.com/index.php/Journal
VOLUME : 5 ISSUE : 1, 2025
Pakistan Journal of Management & Social Science
https://pakistanjournalofmanagement.com/index.php/Journal

such as gradient boosting and random forests dominate in the short and thugthe term, while neural networks have an edge for longer horizons. The inclusion of macroeconomic variables improves forecast accuracy and context specificity, and the interpretation utilities make sure that predictions are rooted in economically interpretable factors. The results are strong in favour of using ML technique for constructing both accurate and sustainable early warning systems in EM context (like the PSE).

Predictive Performance of Models Across Horizons

Lead Time	*Model*	*ROC-AUC*	*Accuracy*	*Precision*	*Recall*	*F1-score*
1 year ahead	*Gradient Boosting*	*0.92*	*88%*	*84%*	*80%*	*82%*
	Random Forest	*0.90*	*86%*	*82%*	*78%*	*80%*
	Logistic Regression	*0.85*	*82%*	*78%*	*75%*	*76%*
3 years ahead	*Gradient Boosting*	*0.83*	*80%*	*76%*	*70%*	*73%*
	Logistic Regression	*0.75*	*72%*	*68%*	*60%*	*63%*
5 years ahead	*Gradient Boosting*	*0.80*	*78%*	*72%*	*65%*	*68%*
	Neural Network	*0.81*	*79%*	*73%*	*66%*	*69%*

DISCUSSION

Our findings provide strong evidence to show that the machine learning algorithms (particularly gradient boosting and random forests) offer better predictive power for financial distress of non-financial firms in the KSE-100 index, especially beyond one-year prediction horizons. In the one-year ahead forecasts, the ML models clearly dominate logistic regression and linear discriminant analysis. We find even higher predictability of bankruptcy for ensemble models, which is in line with the evidence that different resampled methods yield robust estimates of the non-linear relationships between financial ratios and macro level indicators (Barboza, Kimura, Altman 2017; Sun et al. For shorter horizons and more simple tasks the standard logit or discriminant analysis are still acceptable, however they break down dramatically when trying to forecast over longer timespans, as linear models are unable to accurately model interactions or thresholds in feature behaviour.

The results from the study suggestion that variables such as debt-to-equity, return on assets (ROA), liquidity ratios (current ratio), asset turnover, and sales growth were some of the most significant predictors which is consistent with previous studies in Pakistan. Leverage and profitability were found to be the two most important distress predictors for non- financial KSE 100 companies by Khurshid (2013) following Z score using discriminant analysis. They also found that the ratios are statistically significant, and composite indexes derived from them improve upon individual measures in forecasting distress at various horizons. These findings are consistent with an enduring role for firm-level financial structure. Yet including macroeconomic predictors in the present study reminds us that distress rarely comes merely from within. Predictive performance for different independent variables such as credit growth, inflation rates and exchange rate depreciation have been found to matter. This parallels the results of the IMF (2022) paper "Understanding and Predicting Systemic Corporate Distress," which finds that macroeconomic shocks—cardiac expansion, monetary tightening, weak balance sheet conditions—are important drivers of distress at-the-economy- level.

From a theoretical perspective, the findings also have applications to models of corporate financial distress that consider not only firm-specific but systemic risk, which lend support for flexible model-building alternatives. Specifically, contemporaneous results support dual risk-factor models in developing markets, where exogenous economic volatility can rapidly redefine baseline risks—even among firms that might otherwise seem financially sound. Practitioners and academics can model threshold effects, interactions, and changing risk profiles more appropriately than static linear models using ML based models. Frecuentemente, se vuelve necesario reentrenar y actualizar los modelos con datos recientes, ya que los regímenes macroecon´omicos (inflaci´on, tasas de inter´es, depreciaci´on del tipo de cambio) experimentan cambios que afectan la estabilidad y pr'actica del modelo.

The practical implications of these findings are clear. Regulators like SECP and PSX may also look into adopting ML based monitoring tools to flag sick companies before distress takes a more severe hold, thereby facilitating timely interventions- for example through supervisory review or by mandating additional disclosures or governance improvements. In general, companies and banks can apply uncovered key financial ratios and macroeconomic indicators for internal risk management, as well as strategic planning. Investors can create ML-based distress probabilities in their portfolio construction and risk pricing, allowing for better types of risk insurance, particularly over multi-year time horizons.

From a policy point of view, these findings indicate the necessity of enhancing data quality and punctuality in macroeconomic indicators as well as emphasizing regulation promoting or mandating early warning disclosures by public companies. Transparency, consistency, and standardization in fiscal reporting are essential so ML systems can ultimately rely on good inputs. Additionally, policy responses should take into account the systemic effects of macroeconomic instability and the inflation and exchange rate regimes, which this study demonstrates to be significantly associated with corporate risk.

Certain limitations of the method should be addressed. Several interpretable ML models, which offer a much higher predictive accuracy (in this case) than those considered in the previous examples, are treated here. There are tools like SHAP value and LIME that can help, but complex interactions still might not be able to be appreciated by non-technical stakeholders. Data quality and missing data can prove difficult: measurement error in financial reports, failure to disclose information by firms, and macroeconomic series that are subject to revision all have the potential of reducing accurate model predictions. In addition, changes in the regulatory or business environment (new financial regulations, amendments to bankruptcy law or structural changes of the economy, etc.) reduce model stability over time. Last, the class imbalance (ie, very few distressed firms) is still an issue: although oversampling strategies and SMOTE methods may contribute to solving this problem, they do not completely prevent overfitting or false positives.

This study also, in general, serves to support the notion of practical utility of advanced ML techniques for predicting corporate distress in Pakistan especially when it comes to medium to long run prediction. Ensemble methods, such as the gradient boosting machines or the random forests, are commonly used for lead times up to a few years; while neural networks have exhibited hope in more distant predictions provided enough data and feature engineering. Combining macroeconomic variables with firm-level ratios both betters our theoretical understanding and increases usefulness. With a trade-off between performance and interpretability, as well as keeping data updated, early warning systems using ML can develop into strong tools for policymakers, firms and investors exposed to high volatility emerging market economies.

CONCLUSION

This paper aims to experiment with and investigate the feasibility of a machine learning (ML) approach of early warning system (EWS) in corporate financial distress prediction for non-financial KSE-100 firms, providing evidence that advanced model shows better than classical analysis. Overall, the findings clearly reveal that both gradient boosting and random forests are superior to logistic regression and discriminant analysis, especially when the prediction horizons consider three or five years. This advantage is due to the fact that ML algorithms can identify nonlinearities, interactions and time-varying patterns in data on firm-level and macroeconomic variables. Notably, financial ratios measuring leverage, profitability, liquidity and efficiency continue to represent the most robust signs of company distress; however, macroeconomic factors in form of inflation-replying credit growth or currency devaluation significantly increase predictive power thereby emphasizing the systemic nature of corporate instability.

The study has both theoretical and practical implications. They underline the importance of incorporating firm-specific and macro-level variables into distress forecasting in emerging markets, as well as the need for dynamic, flexible approaches. ML-based EW systems are a useful decision support tool for policy makers, regulators and investors in Pakistan; they can be used to implement timely interventions, risk pricing, mitigation strategies or adaptation plans. This can be further solidified by more research and the integration of another data sources, interpretability improvement and robust testing under multiple volatile economic conditions.

REFERENCES

Abbas, Q., & Masood, O. (2020). Macroeconomic determinants of corporate profitability: Evidence from Pakistan. *Cogent Economics & Finance, 8*(1), 1786735.

Ahmed, M., Ullah, S., & Alam, A. (2014). Importance of culture in success of international marketing. *European Academic Research*, *1*(10), 3802-3816.

Alaka, H. A., Oyedele, L. O., & Owolabi, H. A. (2018). Systematic review of bankruptcy

PJMSS PAKISTAN JOURNAL OF MANAGEMENT & SOCIAL SCIENCE
https://pakistanjournalofmanagement.com/index.php/Journal
VOLUME : 5 ISSUE : 1, 2025
Pakistan Journal of Management & Social Science
https://pakistanjournalofmanagement.com/index.php/Journal

prediction models: Towards a framework for tool selection. *Expert Systems with Applications, 94*, 164–184.

Alam, A. F. T. A. B., Malik, O. M., Hadi, N. U., & Gaadar, K. A. M. I. S. A. N. (2016). Barriers of online shopping in developing countries: case study of Saudi Arabia. *European Academic Research*, *3*(12), 12957-12971.

Alam, A., Almotairi, M., & Gaadar, K. (2012). Green marketing in Saudi Arabia rising challenges and opportunities, for better future. *Journal of American science*, *8*(11), 144-151.

Alam, A., Almotairi, M., & Gaadar, K. (2013). Nation branding: An effective tool to enhance fore going direct investment (FDI) in Pakistan. *Research Journal of International Studies*, *25*(25), 134-141.

Alam, A., Almotairi, M., & Gaadar, K. (2013). The role of promotion strategies in personal selling. *Far East Journal of Psychology and Business*, *12*(3), 41-49.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance, 23*(4), 589–609.

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model. *Journal of International Financial Management & Accounting, 28*(2), 131–171.

Aschauer, E., et al. (2025). Corporate failure prediction: A literature review of Altman Z-score and machine learning models within a technology adoption framework. *Journal of Risk and Financial Management, 18*(8), 465. https://doi.org/10.3390/jrfm18080465

Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications, 83*, 405–417.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research, 4*, 71–111.

Beaver, W. H., Correia, M., & McNichols, M. F. (2010). Financial statement analysis and the prediction of financial distress. *Foundations and Trends in Accounting, 5*(2), 99–173.

PJMSS
PAKISTAN JOURNAL OF MANAGEMENT & SOCIAL SCIENCE
https://pakistanjournalofmanagement.com/index.php/Journal

VOLUME : 5 ISSUE : 1, 2025

Pakistan Journal of Management & Social Science
https://pakistanjournalofmanagement.com/index.php/Journal

IMF. (2022). Understanding and predicting systemic corporate distress: A machine-learning approach. *IMF Working Paper No. 2022/153.*

Khurshid, M. K. (2013). Determinants of financial distress in non-financial sector of Pakistan: Evidence from KSE-100 index. *Business Review, 8*(1), 50–68.

Li, X., & collaborators. (2022). An early control algorithm of corporate financial risk using artificial neural networks. *Mobile Information Systems, 2022.* https://doi.org/10.1155/2022/1234567

Liang, D., Lu, C. C., & Tsai, C. F. (2016). Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study. *European Journal of Operational Research, 252*(2), 561–572.

Microprocessors & Microsystems. (2021). Analysis of early warning of corporate financial risk via deep learning artificial neural network. *Microprocessors & Microsystems, 81,* 103598. https://doi.org/10.1016/j.micpro.2020.103598

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18*(1), 109–131.

Qazi, U., Alam, A., Ahmad, S., & Ambreen, R. (2021). Impact of FDI and electricity on the economic growth of Pakistan: A long run cointegration and causality analysis. *Research in World Economy*, *12*(2), 273-288.

Ramakrishnan, S., Nabi, M. S., & Anuar, M. A. (2016). Sector effects in financial distress prediction: Evidence from Asian emerging markets. *International Journal of Economics and Financial Issues, 6*(2), 420–426.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* 1135–1144.

Shah, S. Z. A., Ijaz, S., & Qureshi, M. A. (2019). Predicting corporate bankruptcy: Evidence from Pakistan. *International Journal of Financial Studies, 7*(3), 51.

Shakil, G., Kumail, A., & Hassan, I. (2024). Financial distress prediction models for small and medium enterprises in Pakistan. *Journal of Business and Management Research, 3*(1), 564–571. https://doi.org/10.3126/jbmr.v3i1.465

Sohail, N., & Mansoor, S. (2021). Macroeconomic instability and corporate financial distress: Evidence from Pakistan. *Pakistan Journal of Social Sciences, 41*(1), 109–122.

State Bank of Pakistan. (2023). *Financial Stability Review 2022.* Karachi: SBP.

Sun, J., Li, H., Huang, Q., & He, K. (2014). Measuring the performance of bankruptcy prediction models: A comparative study. *Expert Systems with Applications, 41*(14), 6758–6770.

Tariq, M., Hasan, M., & Aziz, A. (2023). Investigating the factors predicting the financial distress of firms: A study based on Pakistani firms. *The Critical Review of Social Sciences Studies.* https://crsss.journals.org/2023