

Machine Learning Models for Bankruptcy Prediction in Lebanon

Author: Elias Maroun Rhayem¹

Affiliation: CIRAME Research Center, Business School

Holy Spirit University of Kaslik, P.O Box 446, Jounieh, Lebanon

Abstract:

Predicting corporate insolvency is a critical component of financial risk management, particularly in economies experiencing severe macroeconomic shocks. Following the unprecedented 2019 financial collapse in Lebanon, traditional statistical tools for bankruptcy prediction have proven increasingly inadequate. This study investigates the applicability and necessary adaptation of ensemble machine learning (ML) architectures—specifically gradient-boosted algorithms like XGBoost and Random Forests—for assessing corporate bankruptcy risk within Lebanon's highly informal, dollarized economy. Utilizing a mixed-method approach, the research applies the Altman Emerging Market Z"-Score to a cross-section of Lebanese firms (2018–2023) while reviewing contemporary empirical ML literature. The findings identify a critical "Lebanese Paradox": numerous firms exhibiting highly distressed quantitative Z" scores successfully avoid insolvency by leveraging mechanisms invisible to traditional models, such as unrecorded hard-currency revenues and diaspora capital injections. Consequently, the study concludes that raw ML models trained on standard financial ratios will yield systematically high false-positive rates in Lebanon. To achieve the 82–95% accuracy observed in international studies, a Lebanon-calibrated ML model must inherently integrate macroeconomic stress variables, parallel exchange rates, and shadow economy proxies. Ultimately, while machine learning offers superior diagnostic capabilities for investors and creditors, its successful implementation requires a fundamental restructuring of input variables to accurately reflect the realities of crisis-stricken, unstructured economies.

Keywords:

Bankruptcy Prediction, Machine Learning, Lebanese Financial Crisis, Corporate Insolvency, Altman Z-Score, Ensemble Models, Shadow Economy, Financial Risk Management

Introduction: "The What & Why"

One of the most complex and important problems in applied financial analysis is predicting a company's insolvency before it happens. When a business fails, the harm is not just limited to its owners, employees lose their positions, creditors lose their money, suppliers lose their clients, furthermore, in economies where one industry dominates, entire communities may become unstable as a result, early and stable financial crisis prediction is beneficial for the public and investors, allowing them to safeguard their money, it also helps managers to try to solve bankruptcy before it becomes inevitable, as the times progress, new tools emerge for predictions, such as logistic regression, decision trees, random forests, neural networks and gradient boosted algorithms like XGboost, have been made possible because of the emergence of machine learning. These tools will always perform better than traditional predicational tools that dominated the markets for majority of the new age. The upside of these new ML techniques is that they can process larger datasets and spot mistakes that older models are not able to do, it can also modify and forecast logic in response to new data.

One of the most significant areas to test these applications of machine learning is Lebanon, the world bank has recognized Lebanon's financial crisis that started in 2019 one of the biggest catastrophic in recent history. From 2019 to 2024, the Lebanese lira lost almost 100% of its value, in 2023 inflation reached the extraordinary level of 220%. Lebanon was once considered a powerhouse, attracting international deposits, called the Switzerland of the middle east, with attractive interest rates and banking secrecy, the banking sector stopped becoming a source of credit and froze all depositors' access to their own funds. In addition, the GDP of Lebanon shrank by almost half (40%) between 2019 and 2024, by 2022 foreign investment fell from almost 3 billion dollars to half a billion only. These circumstances increase the danger of corporate insolvency in almost every economic sector.

Despite the country's serious corporate issues, the use of machine learning models for bankruptcy predictions in Lebanon is still not explored as it should be. Most of the studied machine learning prediction data sets are done in Europe or the US in economies that are stable and reporting is widely available, as well as strong standards for reporting, there are many challenges that call for adaptation of machine learning techniques in Lebanon, of these challenges are the dollarized cash transactions that avoid the banking system, non-existent financial disclosure system for non-listed businesses. The main problem this paper is addressing is : how can machine learning models be used to predict the risk of bankruptcy in the corporate world of Lebanon, and what modifications are we ought to take to adapt to the unique financial and economic conditions that Lebanon is going through. In contrast to traditional single-variable or linear statistical models, it is hypothesized that ensemble machine learning models—especially gradient-boosted algorithms—adapted to incorporate both financial ratio inputs and macroeconomic crisis indicators will provide more accurate bankruptcy risk assessments for Lebanese firms, which is especially useful for banks (Athari et al, 2023).

Literature Review: "The story so far "

The first works in bankruptcy prediction are the univariate studies of Beaver in the 1960s and the Z-Score of Altman in 1968 which used multiple discriminant analysis on five financial ratios and attained 72-90% accuracy with various prediction periods. The next "Z" of Altman, which is adjusted to non-manufacturing emerging market companies, having distress threshold of less than 1.1 and safe zone of more than 2.6 is the one that fits Lebanese mainly unlisted private sector. Nevertheless, a 2024 study by MDPI on Tunisian firms and a PRISMA-based review of 207 studies (Expert Systems with Applications) both revealed that ensemble ML models always outperform the Z-Score in the accuracy, F1-score, and AUC metrics, especially in times of macroeconomic shocks.

In the case of Lebanon, the most important methodological implications of international studies are: the introduction of macroeconomic stress factors (exchange rates, CPI, banking liquidity) and balancing strategies, such as SMOTE, to overcome the underrepresentation of bankrupt firms. A study of large datasets in the US reported XGBoost and random forest to reach an accuracy of around 95 percent compared to 57 percent in logistic regression. Lebanon-related issues add to the problem: the inability of the Commercial Code to introduce a modern mechanism of restructuring implies that most insolvent companies are still in business due to informal solutions, which creates critical data gaps and renders non-financial variables, such as the quality of governance, vulnerability of the sector, and diaspora financing, key contributions to any working model, even though in some cases, such as the Silicon valley Bank crisis, such prediction would be harder to expect (Beainy & Kamel, 2023).

Methodology: "The how we did it "

This research project uses mixed methodology that involves the use of structured literature review and an applied quantitative exercise. The design was influenced by two limitations: the lack of an available publicly accessible standardized financial database of Lebanese non-listed firms and the prevalence of informal reporting tendencies, such as the existence of multiple accounting records, use of dollarized cash transactions, which are more likely to conceal leverage and liquidity stress than the official statements. This research is a demonstration and adaptation exercise as opposed to a primary model-training research. The literature review included empirical research published in 2012-2025 through Google Scholar, SSRN, and ScienceDirect and complemented by World Bank reports, Banque du Liban publications, data on the ministry of finance, and reports of the sectoral associations.

The quantitative element used the Altman Emerging Market Z 2 formula - $Z_2 = 3.25 + 6.56(X_1) + 3.26(X_2) + 6.72(X_3) + 1.05(X_4)$ where X_1 = working capital/total assets, X_2 = retained earnings/total assets, X_3 = EBIT/total assets, and X_4 = book value of equity/total liabilities to Lebanese firm annual reports in 2018-2023 cross-section. All the values of the Lebanese pounds were converted into USD at the current market rates to allow inter-period comparability and represent the dollarized economic reality. The outcomes were measured by qualitative knowledge of the true operational position of each firm to determine predictive validity of the model in the Lebanese setting.

Results: "The facts "

The use of the Z" model on the sample showed that the grey-zone scores in 2018, which are indicative of the existing vulnerability of Lebanon, were replaced with the distress-zone scores in 2020, after the freezing of the financial system, the devaluation of the pound, and the Beirut port explosion. The most intense falls were observed in X2 (retained earnings/total assets) and X3 (EBIT/total assets) which were caused by currency translation losses and failed operating profitability. The model also rightly singled out domestically dependent sectors, including healthcare, retail, construction, that went into a decline or shut down.

Nonetheless, the findings reveal a Lebanese paradox: the companies with highly distressed Z" scores still managed to stay in business using the mechanisms unseen by ratio-based models. The export-oriented producers and technology companies had dollar revenues and the diaspora capital completely out of the official Lebanese financial statements, which leads to high false positive rates in the model. This establishes that any strict ML model of Lebanon should have non-financial inputs, such as hard-currency revenue share, diaspora presence in the form of investors, sectoral coefficient of dollarization, in addition to the financial ratio inputs that XGBoost and random forest models focus on when international research is conducted, with ensemble models attaining 82-95% accuracy in prediction.

Discussion and conclusion: "The so what ? "

The key conclusion is subtle: ensemble ML models provide an actual better predictive ability compared to classical models without reconsidering the input variables and data infrastructure, which is impossible in Lebanon. A Lebanon-calibrated model has three design characteristics: hard-currency revenue share as an input characteristic; time-varying characteristics (macroeconomic stress variables, parallel exchange rate, monthly CPI, bank liquidity); and a training dataset which represents informal economic activity using alternative proxies (records of electricity consumption, social security contributions, customs export declarations). In the absence of such adaptations, any model will be systematically wrong in classifying the most resilient Lebanese firms as distressed.

Application is also limited by the institutional environment. The insolvency regime in Lebanon based on Articles 459-668 of the Commercial Code does not provide a modern restructuring mechanism and the eight IMF-conditioned reform actions that were agreed in April 2022 are yet to be implemented as of 2024. ML predictions of bankruptcy in Lebanon are, therefore, diagnostic tools to creditors and investors, and not inputs to a formal regulatory response. There are three priorities that research should have in the future: building a proprietary dataset that incorporates official financials with other data in the sample of Lebanese SMEs; creating and back testing a Lebanese-specific XGBoost or random forest model; and collaborating with Lebanese banking organizations on early-warning system implementation to mitigate the social cost of future corporate failure.

References

- Athari, S. A., Saliba, C., Khalife, D., & Salameh-Ayanian, M. (2023). The role of country governance in achieving the banking sector's sustainability in vulnerable environments: New insight from emerging economies. *Sustainability*, 15(13), 10538.
- 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. (1983). *Corporate financial distress: A complete guide to predicting, avoiding, and dealing with bankruptcy*. John Wiley & Sons.
- Altman, E. I., Hartzell, J., & Peck, M. (1995). *Emerging markets corporate bonds: A scoring system*. Salomon Brothers.
- Altman, E. I. (2018). Applications of distress prediction models: What have we learned after 50 years from the Z-score models? *International Journal of Financial Studies*, 6(3), 70.
- Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, 83, 405–417.
- Beainy, R. H., & Kamel, J. K. (2023). Surviving the Post Silicon Valley Bank Crisis. *International Journal of Membrane Science and Technology*, 10(3), 775-785.
- Dasilas, A., & Rigani, A. (2024). Machine learning techniques in bankruptcy prediction: A systematic literature review. *Expert Systems with Applications*, 255, 124761.
- Doha Institute for Graduate Studies. (2024). Navigating Lebanon's financial collapse: Dollarization, banking dysfunction, and the road to recovery. ACRPS Policy Paper.
- Mestiri, S., & Hamdi, M. (2024). Artificial intelligence techniques for bankruptcy prediction of Tunisian companies. *Journal of Risk and Financial Management*, 17(4), 132.
- Rashid, F., Khan, R. A., & Qureshi, I. H. (2024). A comprehensive review of the Altman Z-score model across industries. SSRN Working Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5044057
- Shetty, S., Musa, M., & Brédart, X. (2022). Bankruptcy prediction using machine learning techniques. *Journal of Risk and Financial Management*, 15(1), 35.
- United States Department of State. (2024). 2024 investment climate statements: Lebanon. U.S. Department of State.
- Wikipedia. (2025). Lebanese liquidity crisis. https://en.wikipedia.org/wiki/Lebanese_liquidity_crisis
- World Bank. (2024). Lebanon macro poverty outlook. World Bank Group. <https://thedocs.worldbank.org>
- World Bank. (2020). Lebanon economic monitor: The deliberate depression. World Bank MENA Region, Fall 2020.