Published on (https://bnm.md)

01.06.2025

Working Paper No. 1/2025: Predicting Bank Defaults with AI: Improvements over Statistical and Machine Learning Methods

Predicting Bank Defaults with AI: Improvements over Statistical and Machine Learning Methods [1]

Author: Alexandru Monahov (Banca Națională a Moldovei)

June 2025

Abstract

Accurately forecasting bankruptcies within the financial sector is an essential objective for prudential regulators tasked with maintaining financial stability. While Machine Learning techniques, in particular the more advanced ensemble methods, and neural networks have been proven to perform well in forecasting loan defaults, these methods have yet to be integrated into workflows for assessing risk and predicting the failure of financial institutions.

To identify the most effective approach to predicting the default of banks and NBFIs, this study investigates the performance of eight leading predictive modeling techniques of varying complexity – from traditional statistical models (Logistic and Linear Regression), to advanced Machine Learning methods in the field of classification (such as Random Forests, XGBoost and Support Vector Machines), against Large Language Models (LLMs), a rapidly growing area of Artificial Intelligence which has recently made notable improvements in its ability to process large quantities of unstructured textual and numeric data.

The paper develops a new workflow which uses LLMs to analyze the risk exposure of financial institutions and determine their probability of default. A new PD metric, that LLMs are capable of generating accurately, is created as the joint outcome of risk and profitability, whose impacts are separately estimated by the model. To further improve the analysis, the paper proposes a new financial performance indicator and adaptations for traditional ratios to enable their usage in both going concern and failure contexts.

The results of the study reveal that while traditional methods like regression models and Random Forests can provide very good predictive capabilities, the best performance is achieved with the Large Language Model, which significantly surpasses all other methods in the majority of evaluation metrics. The LLM's ability to capture complex patterns and contextual nuances within financial data results in superior predictive accuracy and robustness. This highlights the potential of incorporating advanced language-based modeling approaches into financial risk management systems, paving the way for more intelligent and adaptive frameworks that enhance decision-making and regulatory policy in the financial industry.

Keywords: default, bank, risk, financial institutions, AI, Large Language Models, regression, classification, Machine Learning, Random Forest, XGBoost, Neural Network.

JEL Classification: G21, G23, G33, C38, C45.

Important Notice: The views expressed in this paper are solely those of the author(s) and do not necessarily reflect the official position or involve the liability of the National Bank of Moldova. All rights reserved. The reproduction of this information is permitted exclusively for educational and non-commercial purposes, provided that the source is explicitly cited.

Tags

Working Paper No. 1/2025 [2]
Working paper [3]
Working paper No. 1 [4]
Predicting Bank Defaults with AI [5]
Improvements over Statistical [6]
Machine Learning Methods [7]
Machine Learning [8]
AI [9]
Predicting Bank Defaults [10]
Artificial Intelligence [11]
default [12]
bank [13]
risk [14]
financial institutions [15]
Large Language Models [16]
regression [17]
classification [18]
Random Forest [19]
XGBoost [20]
Neural Network [21]

Source URL:
https://bnm.md/en/content/working-paper-no-12025-predicting-bank-defaults-ai-improvements-over-statistical-and-0

Related links:

[1] https://bnm.md/files/Alexandru Monahov - Predicting Bank Defaults with AI_1.pdf [2] https://bnm.md/en/search?hashtags[0]=Working Paper No. 1/2025 [3] https://bnm.md/en/search?hashtags[0]=Working paper [4] https://bnm.md/en/search?hashtags[0]=Working paper No. 1 [5] https://bnm.md/en/search?hashtags[0]=Predicting Bank Defaults with AI [6] https://bnm.md/en/search?hashtags[0]=Improvements over Statistical [7] https://bnm.md/en/search?hashtags[0]=Machine Learning Methods [8] https://bnm.md/en/search?hashtags[0]=Machine Learning [9] https://bnm.md/en/search?hashtags[0]=AI [10] https://bnm.md/en/search?hashtags[0]=Predicting Bank Defaults [11] https://bnm.md/en/search?hashtags[0]=Artificial Intelligence [12] https://bnm.md/en/search?hashtags[0]=default [13] https://bnm.md/en/search?hashtags[0]=bank [14] https://bnm.md/en/search?hashtags[0]=risk [15] https://bnm.md/en/search?hashtags[0]=financial institutions [16] https://bnm.md/en/search?hashtags[0]=Large Language Models [17] https://bnm.md/en/search?hashtags[0]=regression [18] https://bnm.md/en/search?hashtags[0]=classification [19] https://bnm.md/en/search?hashtags[0]=Random Forest [20] https://bnm.md/en/search?hashtags[0]=XGBoost [21] https://bnm.md/en/search?hashtags[0]=Neural Network