INTEGRATION OF FINANCIAL RISK ASSESSMENT METHODS USING ARTIFICIAL INTELLIGENCE AND BIG DATA

Rakhimov K.
Doctor of Technical Sciences, Professor (PhD).
Fergana State University.
Akhmedova E'zozkhon Ergasheva
1st-year Master's Student in the Program of Applied Mathematics,
Fergana State University, Fergana, Uzbekistan.
E-mail: ezozxon@gmail.com
https://doi.org/10.5281/zenodo.20199733

ARTICLE INFO

Qabul qilindi: 11-may 2026 yil
Ma'qullandi: 13- may 2026 yil
Nashr qilindi: 15- may 2026 yil

KEY WORDS

artificial intelligence; big data; financial risks; machine learning; credit risk; forecastin.

ABSTRACT

It has been established that the integration of artificial intelligence (AI) methods and big data technologies makes it possible to increase the accuracy of financial risk forecasting. The effectiveness of using neural network models for analyzing credit risk and market volatility has been evaluated. It was revealed that the use of hybrid machine learning algorithms reduces the likelihood of errors in the classification of problem assets. The main approaches to building early warning systems based on the analysis of unstructured data are described. The possibilities of adapting the developed methods to the conditions of the financial market of Uzbekistan are considered.

The modern financial environment is characterized by high volatility, an increasing complexity of risk structure, and exponential growth in data volumes. Traditional statistical models, such as Value-at-Risk (VaR) and CreditMetrics, are often unable to promptly account for nonlinear relationships among risk factors and to process unstructured data (news feeds, social media posts, internal documents). This leads to delays in risk identification and underestimation of the real threat

The modern financial environment is characterized by high volatility, increasing complexity of risk structures, and exponential growth in data volumes. Traditional statistical models, such as Value-at-Risk (VaR) and CreditMetrics, are often unable to promptly account for nonlinear relationships among risk factors and to process unstructured data (news feeds, social media posts, internal documents). This leads to delays in risk identification and underestimation of the real threat

In Uzbekistan, where the banking sector is actively undergoing digital transformation, the problem of adequate risk assessment becomes particularly acute. The growth of lending to individuals and legal entities, the introduction of remote banking services, and the increase in transaction activity generate data volumes that traditional econometric approaches can no longer analyse efficiently. In this regard, the integration of artificial intelligence (AI) methods and Big Data technologies into financial risk management systems is a pressing task.

Aim of this work is to substantiate and experimentally verify a hybrid approach to financial risk assessment that combines classical econometric methods with machine learning algorithms for processing large volumes of structured and unstructured data.

Research objectives:

1. Develop an architecture for collecting and preprocessing data from internal and external sources.
2. Build ensemble models (XGBoost, Random Forest) for credit default prediction and compare them with baseline logistic regression.
3. Apply recurrent neural networks (LSTM) to analyse market volatility.
4. Test natural language processing (NLP) methods for market sentiment evaluation based on news feeds.
5. Evaluate the effectiveness of the proposed methods on real data from a commercial bank of Uzbekistan for the period 2022–2024.

Overall System Architecture.The proposed architecture comprises three sequential stages: (1) data collection and preprocessing; (2) predictive model construction; (3) verification and interpretation of results. Large-scale data processing was carried out using distributed computing on the Apache Spark platform (4-node cluster, 128 GB RAM).

Data Sources

Two classes of sources were used:- Internal databases of a commercial bank (anonymised transactions of 150 thousand clients, payment histories over 3 years, delinquencies, collateral data). External sources – daily news feeds (Uzbek and international financial media, official press releases of the Central Bank of the Republic of Uzbekistan), macroeconomic indicators (inflation, key interest rate, business activity index).Total volume of processed structured records – 5.2 million rows. Unstructured text corpus – 48 thousand news messages.

Machine Learning Models.For credit risk, three models were implemented:

Logistic regression (baseline).

Random Forest – 200 trees, max_depth=15.

Gradient boosting (XGBoost) – 300 iterations, learning_rate=0.05, L1/L2 regularisation.

The target variable was default event (overdue more than 90 days within 12 months after observation). The feature set included 72 indicators: socio-demographic, transactional (average cheque value, transaction frequency), behavioural (overdraft usage, password change frequency), as well as derived features – moving averages, payment-to-income ratio.For market risk, a Long Short-Term Memory (LSTM) network with two hidden layers of 64 neurons each and dropout=0.2 was used. The input vector consisted of the last 20 observations of the volatility index of the Uzbek stock exchange and news sentiment (see below).

Unstructured Data Processing (NLP)

News texts were normalised (lemmatisation, stop-word removal) using the `uzbek-nlp-toolkit` (an experimental model for the Uzbek language). For sentiment analysis, a pre-trained RuBERT model was fine-tuned on a labelled corpus of 3000 news articles (three classes: positive, neutral, negative). The final sentiment index (ranging from –1 to +1) was aggregated by day and supplied as an additional feature to LSTM and XGBoost. Evaluation Metrics Quality of credit risk classification was assessed using Accuracy, Recall, F1-score, and the area under the ROC curve (AUC-ROC). For market risk, the Mean Absolute Error (MAE) between predicted and actual volatility was used. An economic metric – the reduction in the non-performing loan (NPL) ratio in the test portfolio – was also applied.

Results

Comparison of Credit Risk Models-all models were trained on a sample of 120 thousand clients (2022–2023) and tested on 30 thousand clients (2024). The results are presented in Table 1

Table 1 – Comparison of efficiency of credit risk assessment models

Model	Accuracy	Recall	F1-мера	AUC-ROC
Logistic Regression	0,82	0,71	0,76	0,78
Random Forest	0,89	0,85	0,87	0,91
XGBoost	0,91	0,88	0,89	0,93

Below is the continuation of the English translation from the previous scientific text, starting where the previous translation ended.

XGBoost demonstrated a 12% increase in AUC-ROC compared to logistic regression (absolute values: 0.93 vs. 0.78). Recall increased from 0.71 to 0.88, meaning that an additional 17% of problematic borrowers were identified at the same false positive rate.

Market Risk Analysis Results

The LSTM model, trained on daily data from 2021 to 2023 (730 points), predicted next-day volatility for 2024. The baseline benchmark was the GARCH(1,1) model. The MAE for LSTM was 0.0132 compared to 0.0217 for GARCH, i.e. accuracy improved by 39%. The inclusion of the sentiment index reduced MAE by another 8% (to 0.0121), confirming the significance of unstructured data.

Impact on Portfolio Risks

On the test segment of the credit portfolio (30 thousand clients), applying XGBoost with a classification threshold of 0.7 made it possible to build an early warning scoring system. Over the 12 months of 2024, the share of non-performing loans (NPL > 90 days) in the portfolio managed by the new model was 4.2%, compared to 6.8% in the control group (where classical logistic regression was used). Thus, the reduction in the problem loan ratio reached 38.2% (relative).

Expected Loss Formula with AI-Adjusted Probability of Default

In the risk management framework, expected loss (EL) is calculated as:

$$EL = \sum_{i=1}^{n} PD_{AI}^{(i)} \times LGD^{(i)} \times EAD^{(i)}$$

The obtained results are consistent with international studies, where ensemble methods demonstrate superiority over linear models when dealing with a large number of features and nonlinearities [4]. However, under the conditions of Uzbekistan, an important specific feature was the necessity to adapt NLP tools for the Uzbek and Russian languages. The fine-tuned RuBERT model used in this study showed acceptable accuracy (F1 = 0.84 on the test corpus), but full coverage requires the creation of specialised linguistic resources.

It was also found that adding behavioural features (password change frequency, use of night-time transactions) increased the AUC-ROC of XGBoost by 4 percentage points compared to the model using only socio-demographic and standard transactional data. This indirectly indicates that the digital behaviour of clients in Uzbekistan has specific patterns different from those described in European studies.

Limitations of the study

- The experiment was conducted on data from a single commercial bank, which may limit the generalisability of the conclusions.
- The testing horizon (1 year) does not cover a full economic cycle; during a systemic crisis, model performance may change.
- Time lags in collecting external news data sometimes reached 2–3 days, which reduced the quality of LSTM forecasts in real-time mode.

Practical significance

The practical significance of the proposed hybrid approach is confirmed by the reduction in the NPL ratio. Moreover, the XGBoost model provided transparency at the level of feature importance (SHAP values), allowing loan officers to explain rejections to borrowers.

Further research will be directed toward the implementation of Explainable Artificial Intelligence (XAI) methods to comply with the requirements of the regulator – the Central Bank of the Republic of Uzbekistan, which plans to introduce transparency standards for automated scoring systems in 2025.

Integration of artificial intelligence and Big Data methods into the financial risk assessment system makes it possible to significantly improve forecast accuracy and model adaptability to changing market conditions.

The hybrid approach, combining structured (transactions, payment history) and unstructured (news background) data processing, has proven its effectiveness on local data from the Uzbek banking sector.

The XGBoost gradient boosting model outperformed logistic regression by 12% in AUC-ROC, reducing the share of non-performing loans in the portfolio by 38% over one year.

The application of LSTM with the inclusion of a news sentiment index reduced market volatility forecast error by 39% relative to GARCH.

For further improvement of quality and trust in AI-based systems, it is necessary to develop Uzbek-language NLP models and implement Explainable Artificial Intelligence (XAI) methods.

The obtained results may be useful for commercial banks and regulators when building early warning systems for financial risks in the digital economy of Uzbekistan.

References:

1. Abdullaev A., Tursunov B. Application of Machine Learning in Credit Scoring for Commercial Banks of Uzbekistan // International Journal of Banking and Finance. – 2023. – Vol. 8. – No. 2. – P. 45-58.

2. Karimov Sh.Kh. Digital Transformation of the Banking System of Uzbekistan: Risks and Prospects // Economics and Innovative Technologies. – Tashkent, 2024. – No. 1. – P. 112-120.

3. Niyazmetov D.A. Big Data in the Financial Sector: Methodology of Analysis and Processing // Bulletin of Tashkent State University of Economics. – 2023. – No. 4. – P. 78-85.

4. Goodfellow I., Bengio Y., Courville A. Deep Learning. – Cambridge: MIT Press, 2023. – 800 p.

5. Rakhmatov A.A., Yusupov B.M. Use of Neural Networks for Forecasting Financial Stability // Abstracts of the Republican Scientific and Practical Conference "Digital Economy: Challenges and Solutions". – Tashkent: NUz, 2024. – P. 210-212.

6. Mirzaev D.I. Credit Risk Assessment Using Machine Learning Methods: Master's Thesis in Economics. – Tashkent: TSUE, 2024. – 120 p.
7. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. – San Francisco, 2016. – P. 785-794.
8. Hochreiter S., Schmidhuber J. Long Short-Term Memory // Neural Computation. – 1997. – Vol. 9, No. 8. – P. 1735-1780.