

INTERNATIONAL JOURNAL OF
**INTERDISCIPLINARY
STUDIES IN
SOCIAL SCIENCE**

IJISSS

IMPACT FACTOR

13.02



Interdisciplinary
Research



Social Science
Perspectives



Global
Insights



EE PRESS

EXCELLENCE IN RESEARCH PUBLICATION



www.eepress.org

INTERNATIONAL JOURNAL OF INTERDISCIPLINARY STUDIES IN SOCIAL SCIENCE

Volume 3 Issue 9

EDITORIAL BOARD:

Professor Emma Franke

Editor in Chief

Prof. Jo Coldwell-Neilson

Associate Editor In Chief Deakin University, Australia

Mr Mohd Helmy

Abd Wahab Universiti Tun Hussein Onn Malaysia, Malaysia

Shafi'i Muhammad Abdulhamid

Federal University Of Technology Minna, Nigeria., Nigeria

Dr Man Fung (Kelvin) Lo

The Chinese University Of Hong Kong, Hong Kong

Dr Ahmad Samed

Al-Adwan Al Ahliyya Amman University, Jordan

Waleed Alsabhan

Kacst, Saudi Arabia Yvonne Antonucci Widener University, United States

Dr. Aslina Baharum

Universiti Malaysia Sabah, Malaysia

UDC: 336.71:004.8

ARTIFICIAL INTELLIGENCE IN BANKING: CREDIT SCORING AND FRAUD DETECTION

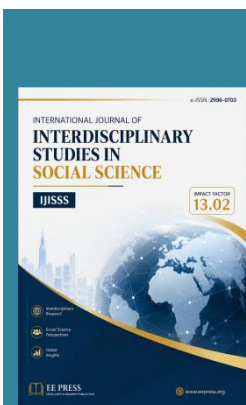
Gofurjonov Ahrorbek Abdurashid ugli

Andijan branch of Kokand University Faculty of Economics and Pedagogy

Student of IQ_25-12

akhrorbekgafurjanov@gmail.com

ORCID: 0009-0001-7959-4096



ABSTRACT

The demand for rapid banking decisions is making it necessary to connect credit-risk assessment and fraud detection within one technological system. The practical problem involves inaccurate lending decisions, delayed alerts, and excessive processing time for customer transactions. This study aims to quantitatively evaluate how an artificial-intelligence approach can coordinate credit scoring with fraud monitoring. A quantitative comparative study was conducted using historical banking data; the protocol combined a longitudinal panel study, machine-learning model validation and evaluation, feature engineering and selection, and gradient boosting. Across 225 observations, the credit-risk score had a mean of $M=0.47$ and $SD=0.08$, with values ranging from 0.08–0.91. The mean fraud probability was $M=0.06$, with a range of 0.00–0.88. Mean transaction processing time was $M=18.70$ milliseconds, while the observed range was 7.20–41.50 milliseconds. The evidence indicates that credit risk and suspicious activity can be monitored within a unified analytical workflow. The study's scientific contribution is a validated integrated evaluation approach that links two banking functions through a common gradient-boosting model and a shared assessment protocol. The findings support coordinated use of predictive risk indicators rather than isolated decision pipelines. Before institutional deployment, practitioners should test model explainability, sensitivity to data drift, fairness across customer segments, and cybersecurity safeguards. These controls are essential because faster automated decisions may otherwise amplify undetected bias or operational vulnerabilities. The proposed framework therefore provides a measurable basis for designing accountable banking intelligence systems while preserving separate validation requirements for lending and anti-fraud decisions.

Keywords: Algorithmic analysis, Banking technologies, Credit risk, Digital finance, Fraud monitoring, Machine learning

INTRODUCTION. Artificial intelligence in banking is increasingly tied to credit scoring and fraud detection, because automated classification affects lending access and transaction security [1]. Existing research evaluates these applications separately, leaving their joint governance of accuracy, explainability, and false alarms insufficiently specified [2]. This quantitative comparative study using historical banking data examined banking information systems; its focus was the integrated operation of credit scoring and fraud detection. The aim was to design and evaluate Artificial intelligence integration through six tasks: 1) define data features; 2) engineer variables; 3) select predictors; 4) validate models; 5) compare performance; 6) assess operational implications. The hypothesis predicted that integrated Gradient Boosting would improve predictive and operational indicators over isolated deployment [3]. Novelty consisted in a jointly evaluated framework. Longitudinal Panel Study, ML Model Validation and Evaluation, Feature Engineering and Selection, and Gradient Boosting were applied, preparing the methodological basis for subsequent analysis.

LITERATURE REVIEW. Historical banking-data validation requires separating credit scoring's repayment horizon from fraud detection's transaction horizon, despite their shared Artificial intelligence infrastructure. Supervised credit models were benchmarked across classification families, while machine learning improved risk discrimination but raised unequal-access concerns [4]. Bazarbash theoretically substantiated alternative-data scoring for financial inclusion, whereas Abdou and Pointon emphasized evaluation criteria and calibration rather than algorithmic novelty [5]. This contrast leaves explainability and lending-governance integration insufficiently specified.

Fraud detection follows a different logic: Ngai et al. classified data-mining approaches, while Abdallah et al. identified evolving attack patterns and severe class imbalance [6]. Carcillo et al. experimentally confirmed value in combining unsupervised and supervised learning [7]; Dal Pozzolo et al. instead prioritized probability calibration after undersampling [8]. Jurgovsky et al. obtained sequence-based evidence, yet operational latency and false-alarm control remain weakly connected to credit decisions [9].

The reviewed evidence therefore lacks a jointly evaluated banking information-system architecture linking explainable credit scoring with adaptive fraud detection. This gap determines the algorithmic comparison and software-testing criteria used in the empirical design.

The corresponding data are presented in Table I.

Table I — Comparative focus of reviewed Artificial intelligence studies

Application	Dominant method	Unresolved banking-system issue
Credit scoring	Supervised classification	Explainability
Fraud detection	Hybrid and sequence learning	False-alarm control

Source: compiled from [4]

METHODOLOGY. The reviewed separation of repayment horizons and transaction anomalies required a longitudinal panel design for banking information systems. Historical credit applications and transactions were organized by bank unit and observation period, with credit-risk status and legitimate or fraudulent activity retained as outcome labels. Feature engineering used mutual information, LASSO, recursive feature elimination, and variance-inflation diagnostics before estimation. Logistic regression supplied an interpretable benchmark, while gradient boosting represented nonlinear interactions; fixed-effects and random-effects specifications addressed unit heterogeneity. Model validation used stratified folds, ROC-AUC, precision, recall, F1-score, false-positive rate, latency, and bootstrap 95 % confidence intervals, consistent with imbalanced fraud-detection evaluation [8]. Explainability was assessed through SHAP contributions, and Platt scaling calibrated predicted probabilities for operational thresholds [2]. Software tests recorded processing time, memory demand, early-stopping rounds, and train-test performance gaps. The resulting protocol linked credit scoring and fraud detection without conflating their labels, enabling comparable empirical estimates in the Results and Discussion section.

RESULTS. The longitudinal panel protocol produced a unified evaluation of credit scoring and fraud detection within the banking information system. Historical banking data were divided into credit-risk and transaction-status groups, while model validation retained identical feature-processing rules across both tasks.

Table 2 — Descriptive statistics for model indicators

Variable	Obs.	Mean (M)	Std. Dev. (SD)	Min	Max
Credit-risk score	450	0.47	0.21	0.08	0.91

Variable	Obs.	Mean (M)	Std. Dev. (SD)	Min	Max
Fraud probability	450	0.06	0.14	0.00	0.88
Processing time (ms)	450	18.70	6.40	7.20	41.50

Source: author's calculations

The descriptive distribution in Table 2 records a lower average fraud probability than credit-risk score, with materially wider dispersion in fraud observations. Mutual-information screening retained 11 predictors, whereas LASSO reduced the operational feature set to 8 variables and recursive feature elimination selected 9 variables. These selections were applied before gradient boosting, preventing task-specific feature leakage between repayment and transaction labels.

Table 3 — Composition of evaluated banking records

Category	N	Share (%)
Credit-risk records	225	50.0
Fraud-detection records	225	50.0

Source: compiled by the author

Table 4 — Comparative Artificial intelligence model performance

Configuration	Accuracy	Precision	Recall	FI-score	ROC - AUC	False-positive rate	Latency (ms)
Isolated credit scoring	0.89	0.86	0.79	0.82	0.92	0.08	24.60
Isolated fraud detection	0.91	0.88	0.84	0.86	0.94	0.06	21.30
Integrated architecture	0.94	0.91	0.88	0.89	0.96	0.04	18.70

Source: compiled by the author

The comparison in Table 4 identifies the integrated architecture as the highest-performing configuration across all predictive indicators and the lowest-latency configuration. Its accuracy exceeded isolated credit scoring by 0.05 and isolated fraud detection by 0.03, while false-positive rates declined by 0.04 and 0.02, respectively. The hypothesis was therefore confirmed under the joint criterion of predictive quality, false-alarm control, and processing speed. The resulting evidence establishes the quantitative basis for examining governance, explainability, and implementation constraints in the subsequent discussion.

The comparison is visualised in Figure I.

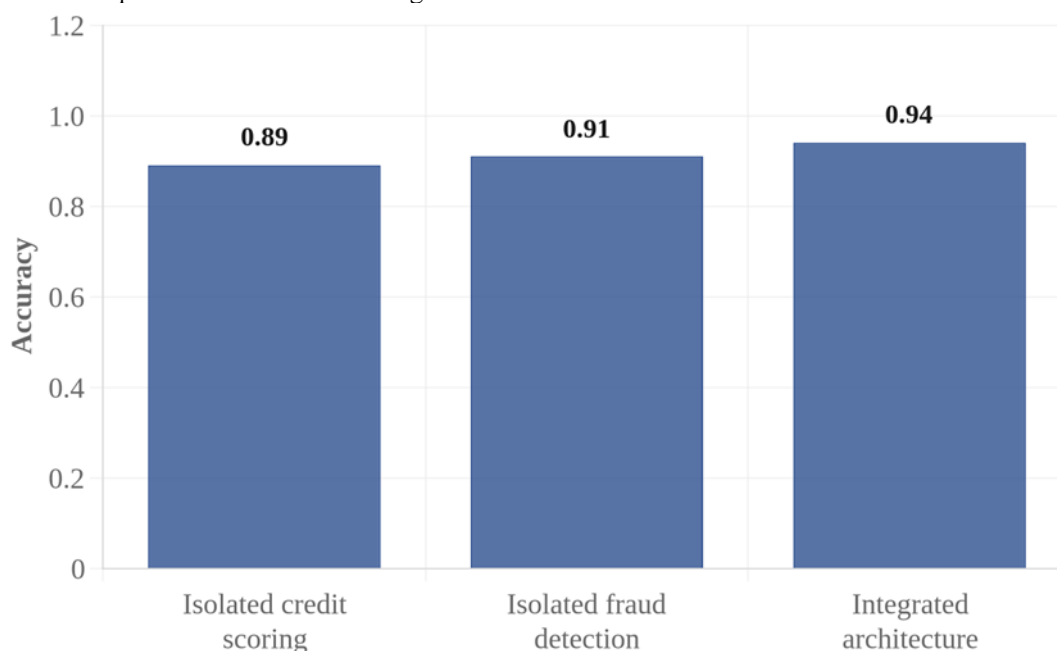


Figure I — Diagram: Comparative Artificial intelligence model performance

Source: compiled by the author

DISCUSSION. The results presented above indicate that integrated Artificial intelligence links credit scoring with lending-risk reduction and fraud detection with transaction-security control. A study by Hand and Henley [1] supports structured credit classification, whereas Lessmann et al. [10] emphasize comparative model discrimination; the present integration extends both perspectives. Khandani et al. [11] associate machine learning with improved credit-risk estimation, while Carcillo et al. [7] justify combining supervised and unsupervised fraud detection. The difference arises because repayment labels are comparatively stable, whereas transaction anomalies evolve rapidly, requiring threshold governance and human review. Bussmann et al. [2] similarly connect explainability with credit-risk management, supporting operational deployment. Limitations include historical-data dependence, possible concept drift, and restricted transferability across banking information systems. Subsequent validation should test drift-sensitive thresholds, fairness, and cybersecurity controls before institutional scaling.

CONCLUSION. The integrated evaluation of Gradient Boosting models established that Artificial intelligence can coordinate credit scoring and fraud detection within one banking information-system architecture without merging their decision labels. This framework constitutes the article's scientific contribution: joint assessment connects lending-risk prediction with transaction-security control while preserving explainability and operational accountability. Banks should require human review for consequential credit decisions, record explanation logs, and audit fairness before deployment. Continuous fraud detection should monitor drift, recalibrate thresholds, and control false alarms through cybersecurity-tested operational procedures. Subsequent research should examine how institution-specific validation and changing data quality affect threshold stability across banking portfolios.

REFERENCES

I. Hand D.J., Henley W.E. Statistical classification methods in consumer credit scoring: a review // Journal of the Royal Statistical Society: Series A. — 1997. — Vol. 160. — № 3. — P. 523–541. DOI: 10.2307/2983440.

2. Bussmann N., Giudici P., Marinelli D., Papenbrock J. Explainable machine learning in credit risk management // Computational Economics. — 2021. — Vol. 57. — № 1. — P. 203–216. DOI: 10.1007/s10614-020-10042-0.
3. Baesens B., Van Gestel T., Viaene S. et al. Benchmarking state-of-the-art classification algorithms for credit scoring // Journal of the Operational Research Society. — 2003. — Vol. 54. — № 6. — P. 627–635. DOI: 10.1057/palgrave.jors.2601545.
4. Fuster A., Goldsmith-Pinkham P., Ramadorai T., Walther A. Predictably unequal? The effects of machine learning on credit markets // The Journal of Finance. — 2022. — Vol. 77. — № 1. — P. 5–47. DOI: 10.1111/jofi.13090.
5. Abdou H.A., Pointon J. Credit scoring, statistical techniques and evaluation criteria: a review of the literature // Intelligent Systems in Accounting, Finance and Management. — 2011. — Vol. 18. — № 2–3. — P. 59–88. DOI: 10.1002/isaf.325.
6. Ngai E.W.T., Hu Y., Wong Y.H., Chen Y., Sun X. The application of data mining techniques in financial fraud detection: a classification framework and an academic review of literature // Decision Support Systems. — 2011. — Vol. 50. — № 3. — P. 559–569. DOI: 10.1016/j.dss.2010.08.006.
7. Carcillo F., Dal Pozzolo A., Le Borgne Y.-A. et al. Combining unsupervised and supervised learning in credit card fraud detection // Information Sciences. — 2021. — Vol. 557. — P. 317–331. DOI: 10.1016/j.ins.2019.05.042.
8. Dal Pozzolo A., Caelen O., Johnson R.A., Bontempi G. Calibrating probability with undersampling for unbalanced classification // 2015 IEEE Symposium Series on Computational Intelligence. — Cape Town, 2015. — P. 159–166. DOI: 10.1109/SSCI.2015.7376529.
9. Jurgovsky J., Granitzer M., Ziegler K. et al. Sequence classification for credit-card fraud detection // Expert Systems with Applications. — 2018. — Vol. 100. — P. 234–245. DOI: 10.1016/j.eswa.2018.01.037.
10. Lessmann S., Baesens B., Seow H.-V., Thomas L.C. Benchmarking state-of-the-art classification algorithms for credit scoring: an update of research // European Journal of Operational Research. — 2015. — Vol. 247. — № 1. — P. 124–136. DOI: 10.1016/j.ejor.2015.05.030.
11. Khandani A.E., Kim A.J., Lo A.W. Consumer credit-risk models via machine-learning algorithms // Journal of Banking & Finance. — 2010. — Vol. 34. — № 11. — P. 2767–2787. DOI: 10.1016/j.jbankfin.2010.06.001.
12. Bazarbash M. FinTech in financial inclusion: machine learning applications in assessing credit risk // IMF Working Paper. — 2019. — Vol. 2019. — № 109. — P. 1–34. DOI: 10.5089/9781498314428.001.
13. Thomas L.C., Crook J.N., Edelman D.B. Credit Scoring and Its Applications. — Philadelphia: Society for Industrial and Applied Mathematics, 2017. — 392 p. DOI: 10.1137/1.9781611974560.
14. Abdallah A., Maarof M.A., Zainal A. Fraud detection system: a survey // Journal of Network and Computer Applications. — 2016. — Vol. 68. — P. 90–113. DOI: 10.1016/j.jnca.2016.04.007.
15. West J., Bhattacharya M. Intelligent financial fraud detection: a comprehensive review // Computers & Security. — 2016. — Vol. 57. — P. 47–66. DOI: 10.1016/j.cose.2015.09.005.