International Journal of Advanced Research in Commerce, Management & Social Science (IJARCMSS)
(An International Quarterly Multidisciplinary Open Access Double Blind Peer Reviewed Refereed Journal)
ISSN :2581-7930 (Online), Impact Factor (Cosmos): 7.270 & PIF (I2OR): 5.010, Volume 09, No. 03(III), July-September, 2026, pp 72-80

A Review of Traditional and Machine Learning Approaches in Loan Default and Fraud Detection

Shama Rani*

Associate Professor of Commerce & Research Scholar of USM, KUK GCG, Palwal, Kurukshetra.

*Corresponding Author: shamagovtcollege@gmail.com

Citation: *Rani, S. (2026). A Review of Traditional and Machine Learning Approaches in Loan Default and Fraud Detection. International Journal of Advanced Research in Commerce, Management &Amp; Social Science, 09(03(III)), 72–80.*

Abstract

With all the growing number of financial transactions and lending activities carried out digitally, there is a minimal need for precision and adapting approaches in loan defaults prediction and fraud detection. This study aims to compare and contrast the traditional statistical tools and Machine-learning tools for financial risk assessment. The review focuses on the following techniques: expert-based lending practices, credit scorecards, discriminant analysis, logistic regression, rule-based fraud detection, and advanced machine-learning techniques such as decision trees, random forests, support vector machines, artificial neural networks as well as ensemble learning. Because they are intuitive, easily understood, legitimated, and user-friendly, traditional methods are still being widely practiced. They, however, rely on specific assumptions, linearity and reliance on manual feature engineering, which hampers the ability to model complex borrower behavior as well as changes in fraud patterns. Machine-learning methods offer more advanced features that encompass the ability to detect nonlinear relationships, handle big and complex financial information, and reveal underlying patterns in the risks. However, they are only effective if they can be correctly evaluated, are explainable, have quality data and represent the features. The study underscores the importance of considering other dimensions beyond predictive accuracy in the assessment of financial-risk models, such as calibration, robustness, fairness, model interpretability and economic impact. The review finds that hybrid approaches combining the seemingly clear statistical models with predictive predictive power of the machine-learning algorithms exist and look promising in reliable and sustainable credit risk management and fraud detection.

Keywords: *Credit Risk Assessment, Loan Default Prediction, Financial Fraud Detection, Machine Learning, Credit Scoring, Ensemble Learning.*

Introduction

Credit-risk assessment and fraud detection are integral parts of financial decision-making. Credit-risk assessment is the estimation of the likelihood that the borrower will default under the contractual repayment obligations, and fraud detection is the identification of fraud or theft, or intentional deception, unauthorized activity, manipulation of the identity or strategically misrepresented financial behaviour. Both problems are conceptually different and are described as classification problems. Loan default can be caused by an actual financial hardship, the borrower's mismanagement, or bad economic news, and fraud is when a person deliberately defaults on a loan in order to secure some type of financial gain.

Traditionally, banks used expert judgment, borrower interviews and financial statements, and collateral, repayment history and institutional lending policies. The growth of consumer credit then led to the growth of statistical credit-scoring systems that were able to handle large numbers of consumer credit applications and process them uniformly. Credit scoring became a formal way of sorting out borrowers in relatively high-risk and low-risk categories by observing borrower characteristics and their past performance (Hand & Henley, 1997; Thomas, 2000).

A similar phase took place regarding fraud detection. Initial systems were mainly based on rules, thresholds in transactions, internal demarcation and the experiential knowledge of investigators. While these systems could detect known instances of fraud, the inflexible nature of these systems made them unable to adapt to new fraud attacks. The ideas of statistical learning, data mining and machine learning then enabled financial institutions to detect far more complicated, unpredicted and evolving risk trends than they could otherwise have seen (Bolton & Hand, 2002; Ngai et al., 2011; Abdallah et al., 2016).

The real question is not if machine learning should entirely supplant traditional methods of doing things, however. Instead, decision-making reliability for each methodological family needs to be established under different circumstances: those that lead to reliable and interpretable decisions; those that lead to economically useful decisions; and those that lead to institutionally defensible decisions. This review discusses how traditional and machine-learning methods have evolved, the underlying assumptions they make, their predictive capabilities, their level of interpretability, adaptability, and limitations when it comes to operation.

Rehabilitation of Expert Based Credit Assessment and Traditional Lending Practices

For many years prior to the current automated systems, a substantial amount of the credit decisions relied on the experience of the lending officer. Their interviews, documentary proof, job security, source of funds, other debts, securities and financial statements, and past relationships with the financial institutions assessed borrowers. There were a number of aspects of this information that could be classified according to the Five Cs of Credit: character, capacity, capital, collateral and conditions.

Character refers to how the borrower is perceived in regards to willingness and reliability to repay. Capacity was defined as the borrower's capacity to have sufficient earnings or cash flow to repay the debt. Capital “was the borrower's wealth and contribution. All the collateral was recoverable in the event of default, and all conditions were about the use of the loan and the general economy. These factors led to a focus for comprehensive evaluation, as they included both quantitative financial data and qualitative information about the context.

However, there was significant variation in expert assessment. The same piece of evidence may be interpreted differently by different lending officers or judgment may be impaired by incomplete information, cognitive bias, institutional culture or perhaps inconsistent risk tolerance. Manual assessment was also cumbersome and scaling the process to consumer-credit applications proved challenging. Thomas (2000) explained that as consumer loans grew in number, there was a need for systematic approaches capable of making consistent predictions of repayment risk when assessed across a broad spectrum of consumers, and Anderson (2007) explained how credit scoring was an important mechanism for making automated, repeatable decisions based on information about borrowers.

It is widely accepted, however, that expert judgement shouldn't be discarded completely! A structured data-driven approach often fails to capture exceptional conditions, contextual risk, and/or inconsistencies in the documentation that experienced credit officers may see and consider. The weakness comes when subjective judgment is in the absence of documented, explicit or validated criteria, or accountability. As a result, many new lending platforms have a human element within their system, for on the border-line, high-dollar or unusual loans, where they avoid the subtleties of automated lending but get a closer look.

Predicting loan default using traditional statistical methods

- **Discriminant Analysis**

One of the earliest Multivariate Statistical Methods applied to separately classify financially distressed and financially stable entities was Discriminant analysis. The approach used is taken that returns a linear combination of predictors that best separates the predetermined groups. Altman’s (1968) z score model was one of the models used to identify ratios that can be used together to predict

bankruptcy. One of the significant contributions of this research was the application of multivariate techniques which demonstrated that a more systematic approach can be provided for the classification of financial-distress, as opposed to only considering individual financial ratios.

Discriminant analysis is bankrupt mathematically, yet it gives a simple equation that can be conveyed fairly easily, something that can and should produce a "scoring function. It is appropriate when relatively stable group distributions are found, and when relationships between predictors and outcomes can be characterized by linear boundaries. But the classical linear discriminant analysis (LDA) requires that the data is multivariate normally distributed with equal covariance matrices for each group. Often the assumptions are not valid for financial data because variables can be correlated, heterogeneous, skewed, or even extreme. The procedure also only captures nonlinear links and interactions between borrower characteristics to a limited degree (Hand & Henley, 1997; Thomas et al., 2002).

Therefore, the use of discriminant analysis still has historically and methodologically critical roles, but might not be sufficient when using large, heterogeneous and behaviourally complex data sets, particularly in areas of lending data.

- **Logistic Regression**

Logistic regression has proved to be one of the most popular methods used in credit scoring as it directly estimates the probability of a binary outcome, e.g. default versus nondefault, as a result of the algorithmic procedure. Logistic regression has the advantage over linear discriminant analysis in that there is no requirement that the predictor variables be multivariate normally distributed. It can also be given coefficients with log odds and odds ratios interpretations, which gives analysts the opportunity to explore the nature or direction and strength of the associations between borrower characteristics and the hazard of default.

Hand and Henley (1997) have noted that logistic regression is one of the most important statistical credit scoring methods and Thomas et al. (2002) point out that logistic regression plays a key role in developing usable credit scorecards. Anderson (2007) also pointed to its ability to support financial institutions that have clear models, relationships that can be documented, probabilities to be forecasted and valid lending decisions. It remains an important bench mark for more complex algorithms to be judged against, and logistic regression continues to be used.

Although its potential drawback is that it can only model simple data, its ability to assume that the relationship between each continuous predictor and the log-odds of the outcome is linear is the standard specification. Transformations, splines, interactions tasks, bining and Weight of Evidence coding are examples of steps that can be done for the purpose of introducing nonlinearity, but are deliberate feature engineering steps. When many interactions have not been identified in the data, or the surfaces of response are particularly irregular, or the relationship among the behaviours changes rapidly, logistic regression can fail badly.

Although logistic regression has its intrinsic limitations, it can still be useful because in the regulated-lending environment, predictive performance is not the only measure. Other considerations are: suitability for governance, interpretability, coefficient stability and the ability to explain adverse lending decisions, as well as probability calibration. Measuring a modest improvement in discrimination in a complex model may not be worth adopting below such a cost of a lack of transparency or on-going reliability in the model.

- **Traditional Credit Scorecards**

Credit scorecards are sets of values that represent borrower traits and are added together to estimate the risk of the borrower. Income, number of years at a job, prior delinquency, debt, credit utilization, home stability, and payment history are examples of variables that can be organized into categories and score given different values to each. The results are then totalled and compared to a cutoff to approve, reject or refer the application for further review.

Thomas et al. (2002) included the concept of scorecards as systematic methods of Credit and Behavioural Scoring while Siddiqi (2006) introduced operational steps for the development of scorecards such as variable binning, weight-of-evidence transformation, variable screening, variable estimation and validation, scaling and implementation. Older scorecards were useful at boosting consistency, with similar wishing applicants receiving similar scores. They also enabled the lending institutions to record the connection among the characteristics of the applicants and the estimated risks.

The scorecards are not necessarily constant every year, however. A sound scorecard can be undermined by shifts in economic conditions, consumer behaviour, and lending policies, as well as through changes in the type of people that apply for loans, and shifts in the nature of the data collected. Other growth and change effects could result in model deterioration as well, such as population drift, approval policies, and new borrower populations. The scorecards need to be monitored, calibrated, redeveloped and validated periodically.

In particular, traditional scorecard methods rely on a large degree on manually engineered features, and a given risk relationship. Such dependence can enhance the transparency but can also prevent the identification of complex patterns which were not foreseen at the time of the model construction.

Traditional Approaches to Financial Fraud Detection

- **Rule-Based Fraud Detection**

Rule-based systems look for transactions or applications which match certain conditions that indicate a potentially suspicious pattern. They could be an amount above a threshold, a number of repeated loan application sessions made by the same device, a mismatch in the declared and verified income, adding multiple identities to the same address, a customer's information changed too frequently, unusual repayment patterns or a transaction coming from an area with high-risk ratings.

Rule-based systems are appealing since they're simple, straightforward, and simple to align with organizational guidelines. They can respond instantly if a definite suspicious condition is met. They are therefore of special value for identifying fraud schemes that have been detected in the past and for imposing mandatory compliance measures.

But normally it is restricted to be static as per the rules unless experts change it. A fraudster could only be able to make adjustments to their transaction to either fall below a threshold or adjust their behaviour to not comply with a known fraud rule. As the number of rules increases, the system could get harder to maintain, and also create conflicts of alert triggers. Random false-positive rates in a rule-based system can also occur if normal but odd customer action meets suspicious-activity criteria.

As noted by Bolton & Hand (2002), identifying fraud is challenging because fraudulent activity is adaptive and relatively uncommon. However, Ngai et al. (2011) and Abdallah et al. (2016) demonstrated that the approaches that are built upon fixed rules are not effective in detecting financial fraud because of the changing data, imbalance, and heterogeneous characteristics.

- **Traditional statistical and anomaly-detection methods**

Fraud-detection systems that are based on statistics rate the "normal" behavior of customers and flag observations that stray significantly from the norm. Regional techniques include regression, probability models, distance based measures, clustering and peer-group analysis, time-series monitoring and statistical process-control procedures.

These techniques are superior to simple rules because they allow for judgements about the suspicious activity in relation to the behavior that is the subject of the suspicious activity or in relation to similar customers. For instance, a transaction could be a significant price for the overall population but a very peculiar one for other accounts. This may be detected by statistical profiling. However, there are some challenges for statistical anomaly detection. Not every odd transaction is fraudulent, but a deftly engineered fraudulent transaction can exhibit the characteristics of a legitimate one. Labels can also be delayed, incomplete or incorrect. Furthermore, the behaviour of normal customers and fraud strategies evolve over time, resulting in concept drift. The findings by Phua et al. (2010), Ngai et al. (2011) and Hilal et al. (2022) have shown that there is a growing need for adaptive supervised, unsupervised, semi-supervised and hybrid fraud detection techniques that can learn as the transaction patterns change, in order to defeat fraudsters.

Machine-Learning Approaches to Loan-Default Prediction

- **Decision Trees**

Decision Trees are classification procedures that recursively partition the space of the predictors into more and more homogeneous groups of borrowers. Internal nodes denote decision conditions; branches denote the decision conditions outcomes and terminal nodes denote predicted classes or probabilities.

Breiman et al. (1984), introduced systematically Classification and Regression Trees, which provided a nonparametric approach to model nonlinear relationships, interactions, mixed variable type and threshold effects. Baesens et al. (2003) nicely showed in the context of credit scoring the importance of tree-based models and other advanced classification models on several real-world data sets.

The interpretability of decision trees is that they can be represented as decision rules. They are able to automatically detect interactions without assumptions of normality and linearity. But individual trees are susceptible to slight variations in the training data. Too deep tree might result in overfitting and too aggressively pruned trees might lead to loss of information. For this reason, individual decision trees are often being used as interpretable baseline models or as part of ensemble models.

- **Random Forest**

Random forest is an ensemble of many decision trees, created from bootstrapped samples and randomly selected subsets of predictors. They make their predictions and the totals are added together. Breiman (2001) demonstrated that a collection of randomized trees can yield a model with lower instability and lower variance from individual decision trees, and maintain the ability of the individual trees to model nonlinear relationships and interactions.

Random forests are able to deal with high dimensional data, mixed types of predictors, missing-value strategies, nonlinear effects and complex interactions. They also offer variable-importance measures that may help to interpret the models. Ensemble models have often been shown to be as suitable as or better than single statistical classifiers in some datasets (Lessmann et al.,2015; Dastile et al.,2020).

The main drawbacks are that the transparency is decreased, and that some basic importance measures may be unstable or biased. It is hard to interpret a forest with hundreds of trees directly like a logistic-regression equation or a small decision tree. Post hoc explanation techniques can be useful to understand the results, but they are not as interpretable as an inherently interpretable model.

- **Support Vector Machines**

Support vector machines find the decision boundary that keeps the classes as far apart as possible. They can work in the kernel space to map data into higher dimensional space and create nonlinear classification boundaries. This is appropriate for credit data where it is not possible to distinguish between default and non-default borrowers with a linear function.

Baesens et al. (2003) found support-vector-machine performance to be competitive in credit-scoring data sets and Louzada et al. (2016) concluded that the support vector machines are one of the most important families of methods in the credit-classification research.

In complex and high dimensional setups, SVMs can work well but require proper selection of the kernel function, regularization, and scaling of the features as well as hyperparameter tuning. They also create less intuitive explanations than logistic regression or a decision tree with few rules. The output of the SVM is not directly calibrated probability of default, although probability estimates can be obtained by calibration.

- **Artificial Neural Networks**

Artificial neural networks represent relationships using a series of interconnected computational units in layers. Using hidden layers, neural networks can learn complex nonlinear functions and interactions that cannot be expressed a priori by analysts.

In his study, West (2000) considered a range of neural network architectures for solving the credit scoring problem and showed that neural network models could compete with respect to classification performance. Baesens et al., (2003) also looked at neural methods in the list of advanced methods for credit-risk classification.

Neural networks can be useful for identifying the complex and distributed patterns in large data sets. But, in general, they need more computing power, an appropriate architecture choice, regularisation, hyperparameter tuning and large amounts of training data. It's also very hard for their internal reasoning to permeate into regulators, customers and credit officers. So, don't take it for granted that deep learning will automatically outperform simpler models, especially when credit datasets are relatively small, highly structured or dominated by tabular variables.

- **Boosting and Ensemble Learning**

Boosting builds a series of models that give increasing attention to observations that were incorrectly classified by preceding learners. Nonlinear relationships, complex interactions and heterogeneous predictor effects can be modelled by gradient boosting, AdaBoost, XGBoost and LightGBM and related methods.

Lessmann et al. (2015) tested and compared many classification methods and found that sophisticated ensemble methods often exhibited excellent credit-scoring performance. Dastile et al. (2020) also found that ensemble learning had become an important research direction in credit-scoring. The results in this paper do not necessarily show that one algorithm is always better than the other, but rather that ensemble techniques are often good choices when subjected to the proper validation process.

If the model complexity, learning rate, trees depth and number of iterations are not properly controlled, it will lead to the problem of boosting overfitting. Even if their ranking is good, they can give poor estimates of probabilities. Hyperparameter tuning should therefore be done on the training data and probability calibration should be assessed independently from discrimination.

Machine Learning for Financial Fraud Detection

- **Supervised Classification**

In supervised fraud-detection models, transactions are associated with a known fraudulent and/or legitimate category, and the models are trained based on that. This can be accomplished with logistic regression, decision trees, random forests, SVC, neural networks as well as boosting algorithms.

When historical fraud labels are correct and a good representation of future fraud, supervised models can produce good performance. The datasets, however, are typically highly imbalanced with fraudulent transactions being a minority of the transactions. Overall, then, a model can be very well accurate by effectively classifying virtually all transactions as valid. In the field of fraud research, it is therefore necessary to measure three or more than 3 measures other than just accuracy, namely precision, recall, F1-score, precision–recall area, detection rate, false-positive rate, and financial cost.

Feature engineering is intertwined, especially. It was proven by Correa Bahnsen et al. (2016) that transaction characteristics exploiting periodic and aggregated transaction patterns could be used to enhance fraud detection. Customer-level, merchant-level, device-level, temporal and network information can then bolster transaction-level variables.

- **Unsupervised and Semi-Supervised Anomaly Detection**

Without complete fraud labels, it is possible to use unsupervised methods to detect suspicious observations. Some approaches that are relevant are clustering, isolation methods, autoencoders, 1-class support vector machines, density estimation, and distance-based detection of outliers.

The techniques are useful for situations where new types of fraud are not nameable. They can recognize transactions that vary significantly from what they may anticipate customers or peer group to behave. The main drawback is that it isn't necessarily fraudulent behaviour if it's anomalous. These could also be considered abnormal responses if someone is traveling for legitimate reasons, making large purchases, facing emergencies, or changing their lifestyle as they are new customers.

A possible approach to semi-supervised methods is to merge the advantages of labelled and unlabelled data. For instance, Carcillo et al. (2021) showed the added value of adding unsupervised outlier scores to supervised classification. Hilal et al. (2022) also noted semi-supervised and unsupervised anomaly detection as key areas for utilizing a small number of labels and new types of fraud.

- **Sequential, Hybrid, and Cost-Sensitive Models**

Tracking fraudulent behaviour can only be discovered when multiple transactions are viewed as a series and not an individual transaction. Order and the timing transactions can be represented using recurrent neural networks, long short-term memory networks, hidden Markov models, or temporal aggregation methods. Jurgovsky et al. (2018) studied credit-card fraud detection as a sequential classification problem and demonstrated the benefits of sequential information over the static classification techniques.

Hybrid systems integrate expert rules, supervised classifiers, anomaly scores, network relationships and human investigation. The operational advantages are that rules can impose enforceable controls, supervised systems can detect patterns of previously known fraud, and unsupervised systems can detect patterns of new fraud.

There is also unequal importance given to classification errors so that a cost-sensitive notion of learning is of importance. A false negative can result in an immediate financial loss while a false positive can cause the inconvenience to a valid customer, send up investigation costs, slow the transaction process, create a loss of customer confidence. Money involved in each transaction is different and missing a fraudulent transaction can have a big impact versus another. Hence, the actual economic costs should be taken into account to maximize fraud models rather than setting an assumption on the equality of the importance of errors.

Comparisons of Traditional and Machine-Learning Approaches

The differences between traditional & machine learning methods should be measured by some additional metrics other than predictive accuracy.

Comparison of Traditional and Machine-Learning Approaches for Loan Default and Fraud Detection

Evaluation Criteria	Traditional Statistical Approaches	Machine-Learning Approaches
Methodology	Based on predefined statistical assumptions, expert knowledge, and manually designed rules	Data-driven approaches that automatically learn complex patterns from historical data
Common Techniques	Discriminant analysis, logistic regression, credit scorecards, rule-based systems	Decision trees, random forests, support vector machines, neural networks, boosting algorithms
Relationship Modelling	Primarily effective for linear and predefined relationships between variables	Capable of modelling nonlinear relationships and complex interactions among variables
Feature Engineering	Requires manual selection, transformation, and predefined risk indicators	Can identify complex feature interactions and support automated feature learning
Interpretability	High interpretability; decision-making process and coefficients are easily explained	Generally lower interpretability, especially for complex models such as deep neural networks
Predictive Capability	Provides reliable predictions for structured and stable financial environments	Often achieves higher predictive performance in large-scale and complex datasets
Data Requirements	Performs effectively with relatively smaller structured datasets	Usually requires larger datasets for training complex models effectively
Handling Complex Patterns	Limited ability to detect hidden patterns and nonlinear borrower behaviour	Strong capability to identify hidden relationships, behavioural patterns, and emerging risks
Adaptability to Changing Behaviour	Requires periodic manual updating and model redevelopment	Can be retrained and adapted to changing financial and fraud patterns
Fraud Detection Capability	Effective for predefined fraud rules and known suspicious activities	More effective for detecting unknown fraud patterns and complex anomalies
Class Imbalance Handling	Often requires additional statistical techniques for rare-event prediction	Provides advanced methods such as resampling, cost-sensitive learning, and anomaly detection
Computational Complexity	Generally low computational requirement and easier implementation	Higher computational requirements due to model training and optimization
Regulatory Acceptance	Widely accepted due to transparency and explainability	Requires additional validation, explainability, and governance frameworks
Model Stability	Relatively stable when underlying assumptions remain valid	May be sensitive to data changes, hyperparameter selection, and model drift
Validation Requirements	Mainly based on statistical significance, accuracy, and calibration	Requires comprehensive evaluation including accuracy, robustness, fairness, explainability, and generalization
Operational Application	Suitable for traditional banking systems requiring transparent decisions	Suitable for modern digital lending, fintech platforms, and real-time fraud monitoring
Major Limitation	Limited flexibility and inability to capture complex financial behaviour	Reduced transparency and challenges in interpretation and regulatory compliance
Future Potential	Useful as benchmark models and components of hybrid systems	Strong potential when combined with explainable AI and hybrid intelligent frameworks

Comparative Analysis & Discussion

Literature review shows interest has recently shifted from using the traditional type of statistical tools to machine learning (ML) tools to evaluate financial risk. Given their simplicity, ease of use, interpretability and acceptance by regulators, the traditional classification models (discriminant function analyses or logistic regression and credit scorecards) remain relevant**.** The models provide explainable decision-making processes that can be applied in areas where explainability and governance are required, while maintaining good predictive power.

But the traditional methods of approximation consider linear relationship and the variables and features are preprogrammed. However, they may not perform as well with larger amounts of financial data as fraud patterns and behaviours are more complex there and, therefore, non-linear, and may change more rapidly.

Machines learning technique including Decision Trees/Rand Forrest/SVMs/NNs/Ensemble Modeling provides an additional flexibility in identifying hidden patterns and modelling complicated relationships. Ensemble techniques like random forests and boosting algorithms have demonstrated outstanding predictive performance in the field of credit-risk, where these algorithms are capable to handle the high and heterogeneous dimensions. Similarly, machine learning based fraud detection approaches correct for the lack of flexibility of static rule-based ones in order to accurately detect new fraudulent behaviours.

Machine-learning models offer many advantages in the way of prediction, but there also are problems with model interpretability, computational complexity, model data requirements, model stability, and regulatory acceptance. So, while predictive accuracy may be the first evaluation criterion, others such as calibration, robustness, fairness, explainability and an impact assessment on the economy should also be emphasized for the evaluation of models.

The literature, as a whole, indicates that none of the modelling methods is always optimum. Although machine learning models can perform better than traditional statistical models in predicting outcomes in complex financial environments, traditional statistical models also have their merits as baseline models and for interpretation of the results. Hybrid solutions that combine a system that is statistically transparent with one based on machine learning, we believe, can be used in the future as a solution for managing financial risks.

Final Remarks and Outlook for the Future

Though studied models suffer from drawbacks, the comparative analysis shows that these traditional methods like discriminant analysis and logistic regression and credit scorecards still play a vital role in financial risk management due to their transparency and regulations requirements and interpretability. However, because of some of the simplistic assumptions made, and their poor performance in capturing very nonlinear relationships, they are not very effective in a dynamic context within financial markets. Other machine-learning techniques, such as ensemble learning, neural networks and complex classification algorithms provide better abilities to detect hidden patterns, deal with large volume and high dimensionality data and track shifts in patterns of loan defaults and frauds. If implemented, it will still need to overcome issues of explainability, quality of data, model bias, computer overhead and regulatory requirements. Moving towards the future, more studies are required to come up with hybrid intelligent systems that integrates the capability of traditional methods with the capabilities of machine learning methods. For instance, explainable AI, graph-based learning, real-time fraud analysis, federated learning, and adaptive models are all key emerging trends that will help more reliable, fair and sustainable credit-risk and fraud-detection systems.

References

1. Abdallah, A., Maarof, M. A., & Zainal, A. (2016). Fraud detection system: A survey. *Journal of Network and Computer Applications, 68*, 90–113. doi:10.1016/j.jnca.2016.04.007.
2. Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance, 23*(4), 589–609. doi:10.1111/j.1540-6261.1968.tb00843.x.
3. Anderson, R. (2007). *The credit scoring toolkit: Theory and practice for retail credit risk management and decision automation*. Oxford University Press.

4. Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J. A. K., & Vanthienen, J. (2003). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society, 54*(6), 627–635. doi:10.1057/palgrave.jors.2601545.

5. Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical Science, 17*(3), 235–255. doi:10.1214/ss/1042727940.

6. Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. doi:10.1023/A:1010933404324.

7. Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Wadsworth International Group.

8. Carcillo, F., Le Borgne, Y.-A., Caelen, O., Kessaci, Y., Oblé, F., & Bontempi, G. (2021). Combining unsupervised and supervised learning in credit card fraud detection. *Information Sciences, 557*, 317–331. doi:10.1016/j.ins.2019.05.042.

9. Correa Bahnsen, A., Aouada, D., Stojanovic, A., & Ottersten, B. (2016). Feature engineering strategies for credit card fraud detection. *Expert Systems with Applications, 51*, 134–142. doi:10.1016/j.eswa.2015.12.030.

10. Dastile, X., Celik, T., & Potsane, M. (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing, 91*, Article 106263. doi:10.1016/j.asoc.2020.106263.

11. Hand, D. J., & Henley, W. E. (1997). Statistical classification methods in consumer credit scoring: A review. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 160*(3), 523–541. doi:10.1111/j.1467-985X.1997.00078.x.

12. Hilal, W., Gadsden, S. A., & Yawney, J. (2022). Financial fraud: A review of anomaly detection techniques and recent advances. *Expert Systems with Applications, 193*, Article 116429. doi:10.1016/j.eswa.2021.116429.

13. Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P.-E., He-Guelton, L., & Caelen, O. (2018). Sequence classification for credit-card fraud detection. *Expert Systems with Applications, 100*, 234–245. doi:10.1016/j.eswa.2018.01.037.

14. Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research, 247*(1), 124–136. doi:10.1016/j.ejor.2015.05.030.

15. Louzada, F., Ara, A., & Fernandes, G. B. (2016). Classification methods applied to credit scoring: Systematic review and overall comparison. *Surveys in Operations Research and Management Science, 21*(2), 117–134. doi:10.1016/j.sorms.2016.10.001.

16. Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems, 50*(3), 559–569. doi:10.1016/j.dss.2010.08.006.

17. Phua, C., Lee, V., Smith, K., & Gayler, R. (2010). *A comprehensive survey of data mining-based fraud detection research* [Preprint]. arXiv:1009.6119.

18. Siddiqi, N. (2006). *Credit risk scorecards: Developing and implementing intelligent credit scoring*. John Wiley & Sons.

19. Thomas, L. C. (2000). A survey of credit and behavioural scoring: Forecasting financial risk of lending to consumers. *International Journal of Forecasting, 16*(2), 149–172. doi:10.1016/S0169-2070(00)00034-0.

20. Thomas, L. C., Edelman, D. B., & Crook, J. N. (2002). *Credit scoring and its applications*. Society for Industrial and Applied Mathematics.

21. West, D. (2000). Neural network credit scoring models. *Computers & Operations Research, 27*(11–12), 1131–1152. doi:10.1016/S0305-0548(99)00149-5.

❑❑❑