WORLD COMET RESEARCH GROUP

International Journal of Artificial Intelligence & Digital Transformation
Vol. 6, No. 1, 2023
Doi: 10.67228/30713315/IJAIDT-2023PI6RS29
PP. 01-13

Original Article

AI-Based Credit Scoring Models for Smart Banking

Prof. Liam O'Connor

Faculty of Science and Technology, Durham College, Whitby, Canada.

Received: 01-02-2023 Revised: 01-21-2023 Accepted: 01-31-2023 Published: 02-03-2023

ABSTRACT

Credit scoring is vital in modern banking for loan decisions, risk management, and financial inclusion. Traditional models rely on limited, static financial data and struggle to adapt to changing conditions. This paper proposes an AI-based credit scoring approach that uses machine learning, deep learning, and hybrid models to improve accuracy and scalability. The framework integrates financial data with alternative data such as transaction behavior, digital footprints, and repayment patterns. Techniques like decision trees, random forests, SVMs, gradient boosting, and neural networks are analyzed and compared with traditional methods. It also emphasizes explainable AI (XAI) to ensure transparency, fairness, and regulatory compliance. Results show that AI models outperform conventional approaches in accuracy and risk prediction. The study highlights the importance of interpretability, bias reduction, and strong governance, concluding that AI-driven credit scoring enhances smart banking, customer experience, and financial inclusion.

KEYWORDS

Credit Scoring, Artificial Intelligence, Machine Learning, Smart Banking, Risk Assessment, Explainable Ai, Financial Technology.

ISSN: 3071-3315
PP: 01-13.

1. INTRODUCTION

1.1. Background

Borrower creditworthiness has been traditionally assessed using statistical methods like logistic regression and linear discriminant analysis in the credit scoring systems. These are estimates based on pre-calculated financial vehicles, such as income level, credit repayment history, credit employment stability and debt obligations as estimated by these techniques. This is possible because they cannot be mathematically complicated, they can be transparent and highly accepted by the regulation, as they are able to enable lenders to clearly understand the effect of individual factors on the credit decisions. Nevertheless, these conventional methods are also constrained by their use of linear presumptions and rigid model formulations which limit their capability to model nontrivial nonlinear association and interactions impacts between borrower characteristics. This causes them to usually be unable to adjust to the dynamic change in borrower behavior and change in financial environments. The recent explosion of digital banking has fundamentally altered the credit evaluation environment creating large quantities of diverse information using online banking, mobile banking applications, e-wallets, and other financial services. Customers are placing massive online trails that indicate real-time spending patterns, payment habits and financial habits. The systems of conventional credit scoring are poorly suited to work with and learn such large, heterogeneous datasets and as a result this wealth of information is not effectively mined. This loophole has underscored the necessity to develop superior, data-intensive credit scoring schemes, which will be in a position to utilize contemporary data feeds. Therefore, restrictions of conventional frameworks have inspired the desire to explore further artificial intelligence and machine learning methodologies to deliver more precise, flexible, and inclusive credit risk reports within the modern banking context.

1.2. Emergence of AI in Smart Banking

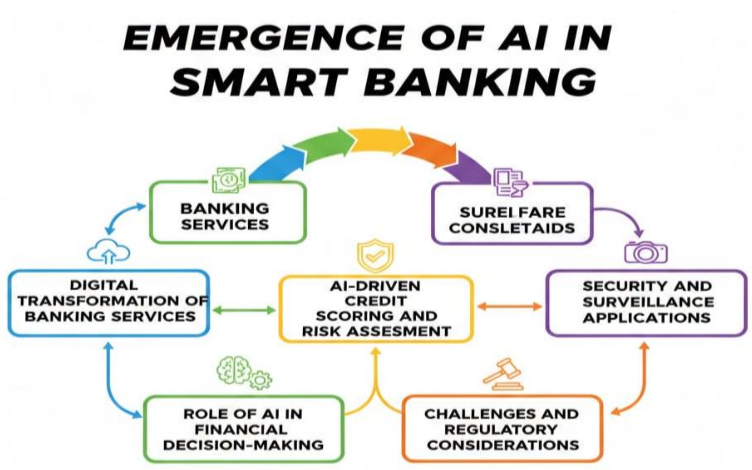

Fig 1 - Emergence of AI in Smart Banking

1.2.1. Digital Transformation of Banking Services

The digital transformation in the banking industry has been brought about by the development of information technology, cloud-computing, and mobile technology. Digital communication platforms like internet banking, mobile applications and automated service kiosks are gradually substituting physical offices via the traditional branch-based services. This change has led to the creation of huge amounts of real-time customer information, such as transaction and spending history and logs of customer interaction. New opportunities have been presented by the availability of this data to allow banks to use artificial intelligence (AI) to better their decision-making, personalization, and operational effectiveness in different banking operations.

1.2.2. Role of AI in Financial Decision-Making

AI has become an important facilitator to smart banking systems by making decisions intelligently. Data analytics and machine learning methods enable banks to deal with voluminous and complicated data volumes and discover additional patterns that are obscured using standardized methods. AI models have the capability to learn through historical data to make correct predictions and adjust to evolving customer behavior in fields like credit scoring, fraud detection and risk control. This will make financial institutions react more promptly to the market forces and demands of customers resulting in better accuracy and consistency in the financial decision making.

1.2.3. AI-Driven Credit Scoring and Risk Assessment

Credit scoring and risk assessment is one of the most influential AI tools in smart banking. In contrast to traditional models, AI-based credit scoring systems can use other types of data including digital payment behavior, frequency of transaction, and regularity of spending. AI models can offer deeper insights into the creditworthiness of the borrowers by extracting nonlinear associations between features and other intricate interactions. This method enhances predictive performance as well as financial inclusion since it allows banks to evaluate their customers with little or no traditional credit history.

1.2.4. Challenges and Regulatory Considerations

Although it has some benefits, the application of AI in smart banking presents certain difficulties with the transparency, privacy of data, and regulatory compliance. The AI models should be in line with the financial regulations that ensure that the decisions made are explainable and fair. This compels the increased focus on explainable AI frameworks and ethical AI practice to facilitate making intelligent banking systems as trustworthy, accountable, and within the scope of regulatory surveillance to be possible.

1.3. AI-Based Credit Scoring Models for Smart Banking

AI-driven credit score models have become a cigar of an element of smart banking given that the model allows the course to assess creditworthiness of borrowers more accurately, adaptively and inclusively. In contrast to the historical statistical models which are based on select financial markers and linearity, AI-based architectures can utilize the machine learning and deep learning algorithms to analyze vast, high dimensional data produced by the compatibility of digital banking ecosystem. These data sets are not just traditional credit bureau data but also other alternative data sets like transaction patterns, consistency of payment, use of mobile banking and online tracks. The AI-built models will also enable the financial institutions to make wise lending decisions by incorporating various data streams to offer a comprehensive insight into the risk profile of the borrowers in real time. Common machine learning methods like decision trees, random forests, gradient boosting and support vector machines have become very popular in AI-based credit scoring since they are able to reflect nonlinear correlations and more sophisticated interactions among features. Deep learning models take this one step further as they learn the hierarchical representations of borrower behavior based on raw data hence performing exceptionally in extensive and ever-changing banking setting. The advanced models enhance the predictive accuracies, diminish the default risks, and increase the efficiency in their operations since the procedures of credit evaluation are automated. Moreover, AI-driven credit scoring will contribute to encompassing a wider range of individuals through banking due to its ability to work with customers with a limited or no credit history, which is essential in new markets and underbanked areas. Along with their advantages, AI-based credit scoring models also present the problems of interpretability, fairness, and compliance with regulators. The complexity models can also be black-box making it hard to explain individual credit-taking, which is a major requirement in regulated financial systems. To mitigate these issues, explainable AI methods and hybrid modelling are also finding their way to credit scoring systems. These approaches are designed to satisfy high predictive accuracy and transparency and accountability, as AI-assisted credit scoring models should be reliable, ethical, and taboo to be used in a sustainable manner in intelligent banking setup.

2. LITERATURE SURVEY

2.1. Traditional Credit Scoring Approaches

The initial studies on credit scoring largely depended on classical statistical methods including but not limited to logistic regression, probit models and linear discriminant analysis to gain a rough measure of the likelihood of a borrower default. These methods modelled the connection among the borrower characteristics in the form of income, job status, repayment track record, and loan balances and the credit hazard utilizing functional forms which were established as predetermined. The most widely used was logistic regression because it was easy to implement and use mathematically because it was a simple method that gave the lender an opportunity to place an interpretable weight on each of the risk factors. Further, these models were preferable to regulatory bodies due to their transparency and auditability. Nonetheless, the large body of empirical research emphasized the intrinsic limitations, in particular, they are solely based on linear assumptions and unpredictability of the predictors. Because credit data increasingly became large and more complicated, the traditional models could not represent nonlinear trends, effects of interactions and latent correlations, and the predictive performance of the model tended to diminish, particularly with the borrowers who had thin or atypical credit histories.

2.2. Machine Learning-Based Credit Scoring

The advent of machine learning (ML) algorithms was an enormous result that changed the approach to credit risk modeling by allowing complex patterns to be learned by using data without making any direct assumptions regarding the relationship among variables. The models based on

ISSN: 3071-3315
PP: 01-13.

decision trees quickly received popularity because of its intuitive design and the possibility of nonlinear modeling of decision boundaries. Random forests also enhanced performance in that it aids in aggregating many trees to decrease variance and increase generalization. The applications of support vectors machines (SVMs) showed significantly good results in terms of classification in high-dimensionality feature spaces and in situations where kernel functions are used to model the nonlinear separability. Ensemble learning methods based on gradient boosting machines (GBM), XGBoost, and LightGBM have been more recently demonstrated to have a higher predictive accuracy due to the ability to address model errors and give attention to hard-to-classify cases. Many comparative analyses express that the accuracy, recall, and area under the ROC curve (AUC) of the models which are based on ML are greater than the ones based on conventional statistics. Such performance improvement, however, is usually accompanied by lower interpretability, making it difficult to apply to a controlled financial setting.

2.3. Deep Learning and Hybrid Models

The availability of computational power and data have enabled the solution of credit scoring problems using the methods of deep learning. Multilayer perceptrons and recurrent neural networks are types of deep neural networks (DNNs), which have been shown to be strong in modeling non-linear and multiplex relationships between borrower attributes. These models have especially been shown to be useful in handling large scale data sets which involve alternative data like transaction logs, behavioural patterns as well as digital footprints. Although deep learning models have a strong predictive power, they are often criticized as being able to make predictions without understanding the mechanisms behind them and being computationally expensive. In order to solve this problem, recent studies have suggested hybrid modeling solutions using either traditional statistical models with machine learning or deep learning. As an example, baseline risk estimation or feature selection can be available through the use of statistical models, and predictions can be refined through a neural network. These hybrid frameworks are designed to balance between predictive performance, robustness, and interpretability and are better optimized to fit the real world credit assessment systems.

2.4. Explainable AI in Credit Scoring

Explainable Artificial Intelligence (XAI) is a new research direction that would play a crucial role in credit scoring because more and more regulators are subject to greater criticism and a more visible decision-making process. Laws that govern finance like fair lending laws demand the lender to give justifications that are clear on why they have granted or denied credit. This has led to the development of post-hoc methods of explanation that are able to explain complex machine learning and deep learning models. SHAP (Shapley Additive Explanations) uses cooperative game theory in order to estimate the role of each feature in a prediction, providing both local and global interpretability. In a similar fashion, the Local Interpretable Model-agnostic Explanations (LIME) classifies complex models with a simple surrogate model around single predictions. Empirical research findings suggest that XAI techniques possess the ability to promote trust, accountability, and regulatory compliance, but do not negatively affect predictive performance. This has led to an

ISSN: 3071-3315
PP: 01-13.

expandable but not uncritically yet, explainable is being characterized as a necessary condition rather than as a supplementary feature of the application of AI-driven credit scoring systems.

3. METHODOLOGY

3.1. Textual Description of AI-Based Credit Scoring Workflow

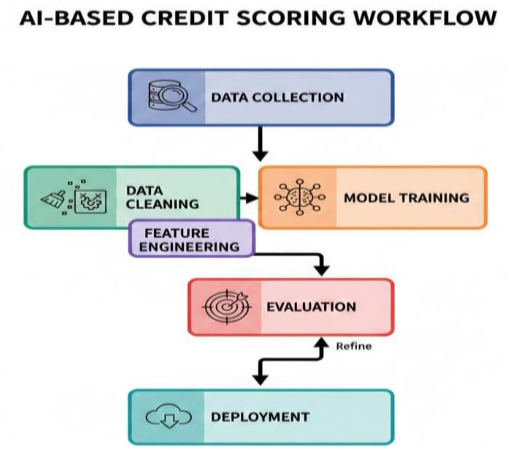

Fig 2 - Textual Description of AI-Based Credit Scoring Workflow

3.1.1. Data Collection

The process of credit scoring based on AI starts with the process of data collection whereby data that is structured and unstructured are collected including information on credit bureaus, banking transactional records, loan application forms and other sources of data including digital payment histories and behavioral data are gathered. Relevance, diversity, and quality of the gathered information are essential in dictating the performance of credit rating model. The thorough information gathering would mean that the risk profiles of borrowers are properly represented, and it minimizes the bias that would be the result of incomplete or obsolete data.

3.1.2. Data Cleaning

Data cleaning aims at enhancing data quality by eliminating problematic outcomes through the absence of values, duplicate entries, outliers, and inconsistent data types. In credit data, missing data can occur as a result of incomplete applications or reporting late and outliers can be a result of data entry or abnormal financial behavior. Normalization, noise filtering and imputation techniques are used to facilitate data consistency and reliability. Data cleaning also leads to stability in the model and ensuring that there is no pattern of misleading credit-related decisions that would be made.

3.1.3. Feature Engineering

The feature engineering process is used to convert raw data into some useful input variables, which more accurately reflects the creditworthiness of the borrowers. The development of financial ratios, aggregation of transaction behaviors and encoding of categorical variables as well as defining the most appropriate features in statistical or machine learning approaches are part of this step. Good features facilitate the process of finding concealed features by models and enhance prediction efficiency. With AI-driven credit scoring, mythical knowledge and automated features of selection are usually combined together in feature engineering to address the dilemma between interpretability and accuracy.

3.1.4. Model Training

The processed data during machine learning or deep learning algorithms are exposed to default or repayment trends during model training. Logistic regression, decision trees, gradient boosting or neural networks are algorithms that are trained with historical labeled data. The training process consists in a process that aims at optimization of the parameters of the model in a manner that minimizes prediction errors and prevents overfitting. When the model is trained appropriately, the system will generalize effectively on nonobserved data on the borrowers.

3.1.5. Evaluation

Model evaluation verifies the performance and reliability of the training credit scoring model with use of test or validation data. Other frequent measures of evaluation are accuracy and precision, recall, F1-score, and the area under the ROC curve (AUC). Besides predictive performance, metrics such as fairness, bias, and explainability are also inherently looked into in an effort to ascertain a regulatory compliance. The close assessment will help to discover the possible flaws and facilitate the reasonable choice of the model.

3.1.6. Deployment

Deployment entails adopting the validated credit scoring model into operational use of the model in financial systems. The model is integrated into loan approval processes, decision support systems or risk management platforms. In deployment, constant observation is necessary to identify the data drift, performance deterioration, or some new bias. Regular retraining and remodeling would keep the AI-based credit score system relevant, legal, and in line with the changing levels of financial and regulatory settings.

3.2. Data Preprocessing

Preprocessing is an important task in AI-based credit scoring, and the quality of the input data directly determines the output of the model and its dependability. The credit data is usually riddled with missing values as there are not all loan applications that are submitted late or even system errors. Such absent records are capable of skewing learning trends in case they are not addressed; consequently, suitable methods of handling missing data including mean or median values of quantitative variables, mode replacement of categorical values or model-based methods of imputation are used. Excessive missing information on records is, in other instances, cleared out to ensure the overall integrity of the data. In conjunction with the missing value treatment, preprocessing also solves inconsistency of data formats, where all the variables will have identical, standardized representations to feed to machine learning algorithms. Another fundamental preprocessing task is normalization, especially when the features are considered to differ in value (i.e. income, loan amount, and credit utilization ratio). The features with larger numbers may prevail in the learning process without normalization, and affect the model results. This is usually accomplished by minmax scaling in which each one, x i, of the feature values is converted to a normalized value by calculating the difference between the minimum value and the maximum value of a particular feature and dividing the outcome by the difference between the maximum and minimum values of the feature. Simply stated this is a repackaged version of the process which takes all feature values and brings them up to a standard range, usually the 0 to 1 range so that the model can treat all the input variables in a more equal manner when they are learning. The outlier detection in the credit score is also valuable in that the extreme figures can either be caused by errors in entering the data, fraud, or infrequent but valid financial circumstances. Z-score analysis, the use of interquartile range (IQR) are just a few of the available statistical methods to evaluate anomalous data value. Outliers can be capped, transformed or deleted once identified based on the nature and the effect they have. Good preprocessing increases the consistency of data, minimizes noises and makes AI-based credit scoring models more robust which eventually results in more accurate and fair credit risk evaluation.

3.3. Feature Engineering

Engineering of features significantly contributes to AI-abilities in credit scoring by converting raw financial and behavioral information into informative variables that have major effects as far as representing the creditworthiness of the borrowers. Dimensional reduction is one of the key aims of feature engineering to create the most efficient, interpretable, and generalisable models using the most predictive information. Credit data may have many correlated or redundant variables, including the existence of two or more variables based on income, repayment history, or credit utilization. The method of feature selection is therefore used in order to select and preserve the most useful attributes. Correlation analysis is often taken as the first feature selection technique to assess the strength and direction of the linear relationship among the input variables and the target outcome, e.g. loan default or repayment. Features that have a high correlation can cause multicollinearity, and this can be detrimental to the stability of a model, especially a linear model. Based on the correlation coefficients, one can eliminate the redundant features and focus on the ones that have a high correlation with credit risk. The feature space is simplified and computational efficiency is increased with only substantial loss of predictive power. Besides correlation analysis, mutual information has been extensively used to learn all the nonlinear and linear dependencies between features and the target variable. The degree to which the particular feature is providing essential information that predicts the outcome, irrespective of the shape of the relationship, is given by the mutual information. This is quite handy with complicated credit data in which the behavior of the borrowers can display a nonlinear trend. Attributes with a greater mutual information value are deemed as more informative and they are used during training the models. By identifying the features well and then employing the engineering technique, a credit scoring model can be more accurate, robust and interpretable so that financial institutions can be applied to make sound and fair lending decisions and decrease overfitting as well as enhance scalability.

ISSN: 3071-3315
PP: 01-13.

3.4. Model Training

The core aspect of AI-based credit scoring is model training, whereby supervised learning algorithms get trained on historical and labeled data that allows the prediction of whether a borrower will default or not on a loan, or has made the correct loan payments. The data of each training example is a feature vector, denoted by X, a value describing the financial, demographic and behavioral traits of the applicant, and an associated class value of either default or non-default. The training process aims to learn a mapping between input features and credit outcomes so that the model can be able to obtain the likelihood of default on new unseen applicants. In probabilistic credit scoring models like logistic regression, a linear combination of the input features is calculated and the nonlinear transformation to compute the probability of default defined as P (D = 1 X) is applied. To be more precise, every element within X is multiplied by a weight that is posted next to it and together they are added up with a bias factor b. This is weighted sum indicates the aggregate effect of all the borrower characteristics on credit risk. The output is then subjected to a sigmoid function which transforms the output into a probability value ranging between one and zero. In plain words, the sigmoid function will narrow the raw score of the model into a meaningful probability that will give an indication of how a borrower is likely to default. In the training process, the weights and the bias of every model are minimally optimized based on iterative algorithms, such as gradient descent, to reduce a predefined loss function, which is usually the log-loss. It is a process of optimization such that the probabilities that are predicted are close to the reality of credit. The use of proper model training, regularization method and validation strategies will prevent overfitting and can improve the generalization ability of the credit scoring system resulting in the provision of reliable and data-driven lending.

3.5. Evaluation Metrics

The evaluation metrics are required to analyze the level of effectiveness, consistency, and practicability of AI-based credit scoring models. As credit risk prediction is generally a binary classification problem default and non-default, various performance measures are employed to propose various facets of model behavior. Accuracy is the ratio of the correct instances divided by all the predictions and is a general picture of the model accuracy. It can, however, be misleading in datasets where the proportion of creditors which default is often minor, and the proportion of non-defaulters is greater; in such cases, accuracy will be valid, but will not be a very neutral measure. Precision and recall give more sample details of model activity, especially in the default category, which occupies more importance to lenders. Precision is the ratio of number of borrowers identified as defaulters of the total number of borrowers identified in the model to identify a false comparison and non-rejection of credit. Recall on the other hand is the proportion of the number of actual defaulters the model correctly identified, or the success of the model in identifying risky borrowers. The high level of recall is particularly critical in credit risk management where the consequences of missing the need to identify the potential defaulters may lead to huge losses. The F1-score is an amalgamation of both accuracy and recall into a single score by computing the harmonic mean of the two which offers a good balance in cases where the two are in trade-off with each other. Such a measure is especially practical where both false positives and false negatives have high costs. Also,

ISSN: 3071-3315
PP: 01-13.

the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) assesses how effective the model is at discriminating between the defaulters and non defaulters on the various classification points. Higher the value of AUC, there is greater discriminative power. A combination of these assessment metrics offers a holistic system of choosing and certifying effective credit score models.

4. RESULTS AND DISCUSSION

4.1. Performance Analysis

Table 1: Performance Analysis

Model	Accuracy (%)	AUC-ROC
Logistic Regression	78.5	0.81
SVM	84.2	0.87
Random Forest	89.6	0.92

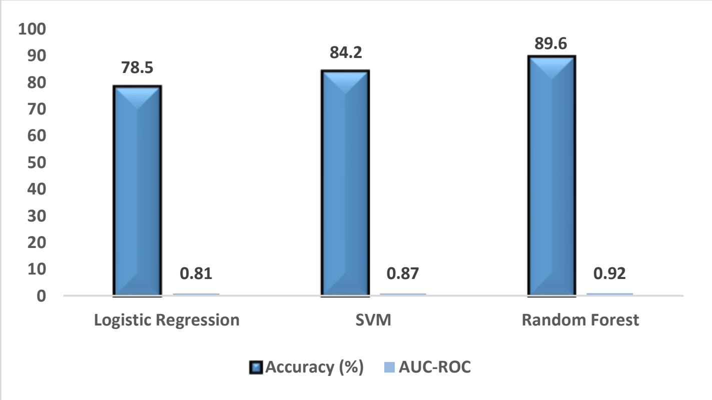

Fig 3 - Performance Analysis

4.1.1. Logistic Regression

The logistic regression can be used as a benchmark model used to compare performances since it is widely used in the traditional credit scoring system. As indicated in the experimental results, the accuracy of logistic regression was 78.5% with an AUC-ROC value of 81 and this implies that there was a reasonable predictive power to classify the defaulters and non-defaulters. Although this model is easy to understand, interpret, and acceptable by regulators, its worse performance can be adopted by the fact that it has a linear decision edge and is not effective in creating complex nonlinear relationships between the borrower attributes. Such constraints minimize its performance on high-dimensional credit data of the modern world.

4.1.2. Support Vector Machine (SVM)

SVM model showed better results compared to logistic regression with 84.2 percent accuracy and 87 AUC-ROC. SVMs can be used to model nonlinear relations with a high-dimensional feature space which they can deal with effectively. This ability enables the model to be more differentiated between default and non-default classes than linear methods. Nevertheless, tuning SVMs can be sensitive and computationally impractical with very large credit data which is likely to restrict their scalability in real-time credit decision systems.

4.1.3. Random Forest

In the models that we have tested, the random forest classifier presented the best results since it had an accuracy of 89.6% and an AUC-ROC of 92. Its ensemble learning structure, which integrates a number of decision trees, is the main reason of the improved results because of mitigating variability and enhancing generalization. Random forests have the advantage of desensitizing a complex nonlinear interaction and are resistant to noise and outliers of credit data. These attributes are very with credit scoring solutions where predictive it is essential and stability is required which is why they are becoming more and more prevalent in current AI-based credit risk scoring models.

4.2. Discussion

The experimental findings reveal clearly that AI based credit scoring model is far more successful than the traditional statistical methods in the complexity patterns inherent in borrower behavior. Improved machine learning algorithms like random forests and gradient boosting algorithms are good at helping to detect nonlinear relationships and complex feature interactions that are not easily modeled by other methods. The models are designed to capitalize on the improvement of predictive accuracy and robustness through an ensemble learning strategy and are suited well to the current credit environment, which has large, diverse and high-dimensional datasets. The witnessed high changes in the values of AUC-ROC demonstrates high discriminative power, and hence, is capable of increasing monetary institutions differentiating high and low risk borrowers, leading to lower default rates at a profitable lending activity. The disadvantage of lower interpretability is a significant issue despite these benefits to performance that make it harder to implement AI models in real-world credit decision systems. Most of the successful models are black boxes and it is a challenge to the stakeholders to have insight into the relationship between particular borrower characteristics and credit. This disadvantage of transparency creates issues with regard to the regulatory compliance especially in those regions that demand explanations of credit grants or declines be clearly made by the lenders. Moreover, the lack of interpretability will decrease the confidence of the regulators and even customers, which may plunge the reception of AI-based credit scoring solutions. Explainable AI (XAI) frameworks have become a more critical aspect to resolve these issues. Methods like SHAP and LIME give significant insights into the predictions made by a model since the computation of feature contributions in the global and individual levels are made. A balance can be found between prediction accuracy and transparency by integrating AI models with high performance with powerful explanation mechanisms to help financial institutions achieve it. These hybrid solutions facilitate equitable, responsible and legal credit decision-making, and explainability is especially relevant in the context of sustainable implementation of AI-based credit scorecards.

5. CONCLUSION

This paper has provided a critical and systematised discussion on AI-based credit score models in the field of smart banking systems, and its significance in the present-day financial ecosystems is increasing. The AI-based credit scoring models can substantially increase the accuracy, consistency, and scalability of credit risk assessment by using more refined machine learning

ISSN: 3071-3315
PP: 01-13.

methods and adding additional and more behavioral data. The AI models have proven to be more effective at identifying and representing complex, nonline relationships and subtle behavioural trends, which are likely signals of creditworthiness in a borrower compared to the traditional statistical methods of evaluation. Such abilities allow financial institutions to make better lending decisions, minimize the risk of default, and enhance operational effectiveness via accelerated and more automated credit reviews. Organizations that transition to AI-based credit scoring systems have these benefits, but specific issues also arise, especially on the front of transparency, interpretation, and equity. Most of the top performers are black boxes and it is hard to justify a specific credit decision to the regulators and the customers. This interpretive deficiency brings up issues pertaining to regulation compliance, responsibility and trust particularly in the setting of stringent fair lending and consumer protection policies. Moreover, the application of massive and heterogeneous data sets creates a risk of hereditary biases, which, careless not to recognize them and address them, can inevitably result in a discriminatory outcome. To mitigate the risks that AI use in the financial decision-making process may entail, it is important to address such issues to make its adoption ethical and responsible. These limitations can be addressed by an avenue of overcoming them that is provided by the integration of explainable AI methods. SHAP and LIME, among the methods, can offer valuable information about how the model behaves, emphasizing the value of each feature with regard to credit decisions, thus, enhancing a better understanding of transparency as well as regulatory needs. In the future, new paradigms like federated learning should be discussed since this would allow collaborative model training without violating data privacy. The bias detection and mitigation strategies should also be given more significance, along with the creation of the regulatory-compliant AI schemes. This will play a critical role in the pursuit of long-term, credible, and ethical implementation of smart credit scoring system in intelligent banking systems.

REFERENCES

[1] D. Durand, Risk Elements in Consumer Installment Financing, New York, NY, USA: National Bureau of Economic Research, 1941.
[2] E. I. Altman, “Financial ratios, discriminant analysis and the prediction of corporate bankruptcy,” J. Finance, vol. 23, no. 4, pp. 589–609, 1968.
[3] D. B. Thomas, J. N. Crook, and L. C. Edelman, Credit Scoring and Its Applications, 2nd ed., Philadelphia, PA, USA: SIAM, 2017.
[4] S. Hand and W. E. Henley, “Statistical classification methods in consumer credit scoring: A review,” J. R. Stat. Soc. A, vol. 160, no. 3, pp. 523–541, 1997.
[5] L. Breiman, “Random forests,” Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.
[6] C. Cortes and V. Vapnik, “Support-vector networks,” Mach. Learn., vol. 20, no. 3, pp. 273–297, 1995.
[7] J. H. Friedman, “Greedy function approximation: A gradient boosting machine,” Ann. Stat., vol. 29, no. 5, pp. 1189–1232, 2001.
[8] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining, San Francisco, CA, USA, 2016, pp. 785–794.
[9] Y. Lessmann, B. Baesens, H. V. Seow, and L. C. Thomas, “Benchmarking state-of-the-art classification algorithms for credit scoring,” Eur. J. Oper. Res., vol. 247, no. 1, pp. 124–136, 2015.
[10] G. F. Harris, M. P. He, and S. Zhang, “Deep learning in credit scoring: Opportunities and challenges,” IEEE Access, vol. 8, pp. 102900–102915, 2020.
[11] S. Wang, Y. Zhang, and Z. Zhang, “Hybrid credit scoring model based on logistic regression and neural networks,” Expert Syst. Appl., vol. 38, no. 3, pp. 2434–2440, 2011.

[12] R. K. Jain and P. Singh, “A hybrid machine learning framework for credit risk assessment,” IEEE Trans. Comput. Soc. Syst., vol. 7, no. 5, pp. 1325–1336, 2020.
[13] M. T. Ribeiro, S. Singh, and C. Guestrin, “Why should I trust you? Explaining the predictions of any classifier,” in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining, 2016, pp. 1135–1144.
[14] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in Proc. 31st Int. Conf. Neural Inf. Process. Syst., Long Beach, CA, USA, 2017, pp. 4765–4774.
[15] B. Baesens, R. Setiono, C. Mues, and J. Vanthienen, “Using neural network rule extraction and decision tables for credit-risk evaluation,” Manage. Sci., vol. 49, no. 3, pp. 312–329, 2003.