JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Vol. 01, No. 01, October, 2024
DoI:10.70008/jmldeds.v1i01.36
Page No: 01-18

A SYSTEMATIC LITERATURE REVIEW OF PREDICTIVE MODELS AND ANALYTICS IN AI-DRIVEN CREDIT SCORING

Md Hasanujamman Bari[1]
Corresponding Email: hasanujamman.bari@gmail.com
Graduate Researcher, Management Information Systems, Lamar University, Texas, USA
https://orcid.org/0009-0006-8463-5979

Shaharima Juthi[2]
[1]Master of Science in Management Information Systems, College of Business, Lamar University, Texas, USA
Email: sjuthi@lamar.edu
https://orcid.org/0009-0009-5232-6276

Asha Moni Mistry[3]
[3]MBA in Marketing and Business Analytics; College of Business, Lamar University, Texas, USA
Email: amistry@lamar.edu
https://orcid.org/0009-0009-3736-3589

Md Kamrujjaman[4]
[2]Master of Science in Management Information Systems, College of Business, Lamar University, Texas, USA
Email: mkamrujjaman@lamar.edu
https://orcid.org/0009-0002-8105-7086

Keywords

AI-driven Credit Scoring
Predictive Models
Machine Learning
Risk Assessment
Financial Technology (FinTech)

Article Information

***Received**: 04, October, 2024*
***Accepted**: 29, October, 2024*
***Published**: 30, October, 2024*

Doi: 10.70008/jmldeds.v1i01.36

ABSTRACT

This systematic review examines the transformative role of AI-driven models in credit scoring, highlighting their advances over traditional statistical methods in terms of predictive accuracy, adaptability, and inclusivity. By synthesizing findings from 70 studies, this review demonstrates that machine learning techniques, particularly ensemble models such as random forests and gradient boosting, effectively capture complex, non-linear relationships in borrower data, providing more accurate risk assessments across diverse demographics. Deep learning models, especially convolutional and recurrent neural networks, extend credit scoring capabilities to unstructured and alternative data sources, supporting financial inclusion by enabling assessments of individuals without traditional credit histories. Hybrid models that integrate logistic regression with neural networks offer an optimal balance between interpretability and predictive power, addressing regulatory demands for transparency while maintaining robust accuracy. Ensemble techniques like stacking and blending enhance model adaptability, allowing credit scoring systems to integrate multiple perspectives and improve prediction accuracy in varied borrower contexts. Despite these advancements, challenges remain in the form of ethical concerns and the need for model interpretability, particularly with complex deep learning architectures. The review underscores the importance of developing fairness-aware and explainable AI frameworks to ensure that as AI-driven credit scoring evolves, it remains both transparent and equitable. These insights suggest that with careful attention to ethics and transparency, AI has the potential to create a more inclusive and resilient credit scoring landscape, accommodating the needs of an increasingly diverse global population.

JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Vol. 01, No. 01, October, 2024, **Page**: **01-18**

1 Introduction

Credit scoring serves as a fundamental tool in financial decision-making, providing lenders with critical insights into the risk of potential borrowers (Gambacorta et al., 2024). Early credit scoring systems were predominantly statistical, relying on conventional algorithms such as logistic regression and linear discriminant analysis, which used structured data like past credit histories, financial records, and demographic details (Tsai & Chen, 2010). However, these traditional methods have shown limitations in accurately predicting borrower risk, particularly in complex financial environments with diverse borrower backgrounds (Braggion et al., 2023; Gambacorta et al., 2024). Recognizing these limitations, financial institutions have progressively integrated Artificial Intelligence (AI) technologies into their credit scoring processes, allowing them to better analyze unstructured and dynamic data (Berg, Burg, et al., 2019). AI-driven credit scoring models mark a significant evolution in risk assessment, with predictive capabilities that surpass traditional statistical methods and offer an adaptive solution in today's rapidly changing financial landscape (Tang, 2019). In addition, as AI applications in credit scoring matured, machine learning (ML) methods like decision trees, support vector machines (SVM), and basic neural networks emerged as viable alternatives to traditional techniques, capturing non-linear relationships within the data (Fuster et al., 2021). ML techniques, particularly ensemble models like random forests and gradient boosting, have gained traction for their ability to improve prediction reliability and handle more complex data structures than their predecessors (Zhu et al., 2016). These models have been particularly effective in addressing issues of overfitting and improving prediction accuracy in diverse borrower groups (Jagtiani & Lemieux, 2019). Researchers have found that ensemble learning techniques, by combining multiple model outputs, provide a more comprehensive risk assessment, which is crucial for decision-making in high-stakes financial contexts (Hertzberg et al., 2018; Iyer et al., 2016). Such advancements underscore the potential of ML in enhancing the robustness and adaptability of credit scoring systems, particularly in cases where traditional data sources are limited or incomplete (Pietukhov et al., 2023).

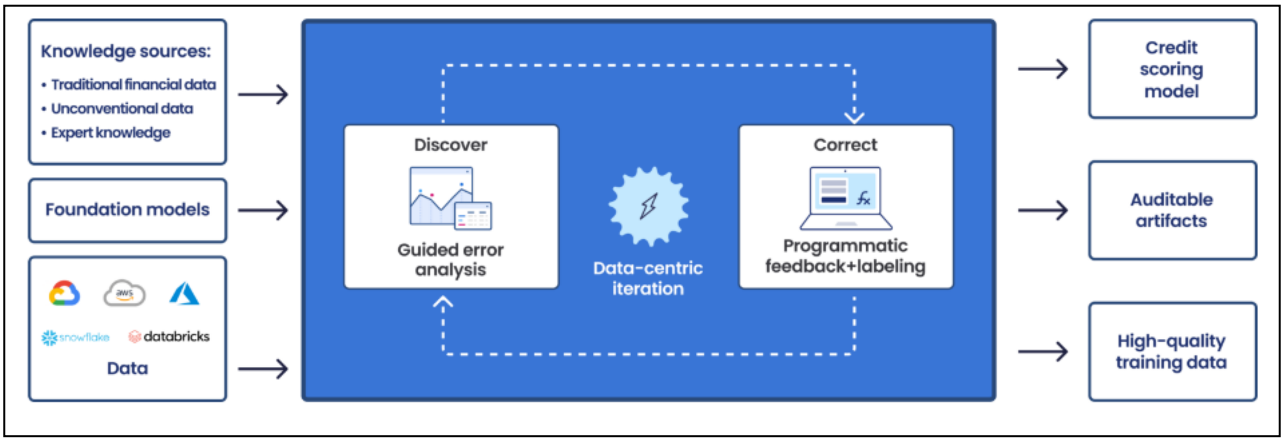
The adoption of deep learning (DL) in recent years has further transformed the credit scoring landscape, introducing models with superior predictive capabilities for handling large, multi-dimensional data sets (Zhao et al., 2019). DL architectures, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have shown exceptional performance in analyzing unstructured data sources, such as transaction logs, social media data, and geolocation information, to produce more accurate credit risk predictions (Alom et al., 2018). These advancements enable financial institutions to assess creditworthiness beyond traditional credit bureau data, offering a more inclusive approach for individuals lacking conventional credit histories (Pietukhov et al., 2023). For example, CNNs, with their proficiency in feature extraction, can detect behavioral

Figure 1: Credit Score Measurement

traditional methods, such as logistic regression, with the predictive power of neural networks, creating a balance that is effective in diverse financial environments (Chen & Jahanshahi, 2018). For instance, Hubel and Wiesel (1968) illustrated that hybrid models combining logistic regression with neural networks could address complex borrower profiles while preserving model transparency. This adaptability has proven essential in emerging markets, where data availability and quality vary widely, necessitating a flexible approach to credit risk assessment (Jeong et al., 2016). Studies further demonstrate that hybrid models are more capable of handling challenges associated with feature selection and model overfitting, leading to more stable and generalizable predictions across various customer segments (Alom et al., 2018; Jeong et al., 2016).

Despite the potential benefits of AI in credit scoring, challenges persist, especially concerning the ethical implications of AI models, including issues of fairness, transparency, and accountability (Zhao et al., 2019). AI models, particularly complex DL architectures, can be perceived as "black boxes," making it difficult for regulators and stakeholders to understand the decision-making process (Pietukhov et al., 2023; Ronao & Cho, 2016). This lack of interpretability raises concerns about potential biases in credit scoring, as certain algorithms may inadvertently discriminate against specific demographic groups (Alom et al., 2018). As regulatory frameworks strive to keep pace with AI advancements, researchers have advocated for the development of fair, accountable, and transparent (FAT) models that align with ethical standards and ensure equal access to credit (Pietukhov et al., 2023). There is a growing focus on creating frameworks and tools for interpretable AI, enabling financial institutions to validate their models against regulatory and ethical requirements, thus fostering trust in AI-driven credit scoring systems (Gu et al., 2018). The rapid evolution of AI in credit scoring reflects a shift from reliance on traditional statistical techniques toward highly adaptable, data-intensive models capable of capturing intricate borrower behaviors (Jeong et al., 2016). As predictive analytics and AI continue to reshape credit scoring, there is an increasing need to balance technological advancement with ethical considerations and regulatory compliance. This review examines these advancements comprehensively, highlighting the strengths and limitations of AI models in credit scoring while offering insights into emerging trends and research directions that address both practical applications and ethical implications. In this systematic review, the objective is to comprehensively synthesize existing research on AI-driven predictive models and analytics in credit scoring, examining their accuracy, reliability, and ethical implications. Following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, the review aims to evaluate a broad range of AI models—including machine learning, deep learning, and hybrid approaches—used to predict credit risk. This study seeks to assess model performance, data sources, and the impact of AI-driven solutions on credit scoring accuracy and inclusivity. By screening and selecting studies through rigorous inclusion and exclusion criteria, the review will address both the effectiveness and limitations of AI models in credit risk assessment, aiming to identify patterns, strengths, and challenges across diverse borrower demographics and financial environments. Additionally, the review will include an analysis of ethical considerations, specifically focusing on transparency, fairness, and bias, to understand the broader implications of AI in credit scoring.

Figure 2: Credit scoring with AI framework



Source: Snorkel Cloud (2024)

2 Literature Review

This section presents a systematic review of the evolution, methodologies, and applications of AI-driven predictive models in credit scoring, with a focus on machine learning, deep learning, and hybrid approaches. The literature review examines both the predictive performance and ethical considerations of these models, exploring their effectiveness in addressing traditional credit scoring limitations. As AI continues to revolutionize the financial sector, this review delves into the unique contributions of various AI techniques, the challenges of model transparency, and ethical implications, providing a comprehensive view of the current landscape and future research needs in AI-driven credit scoring.

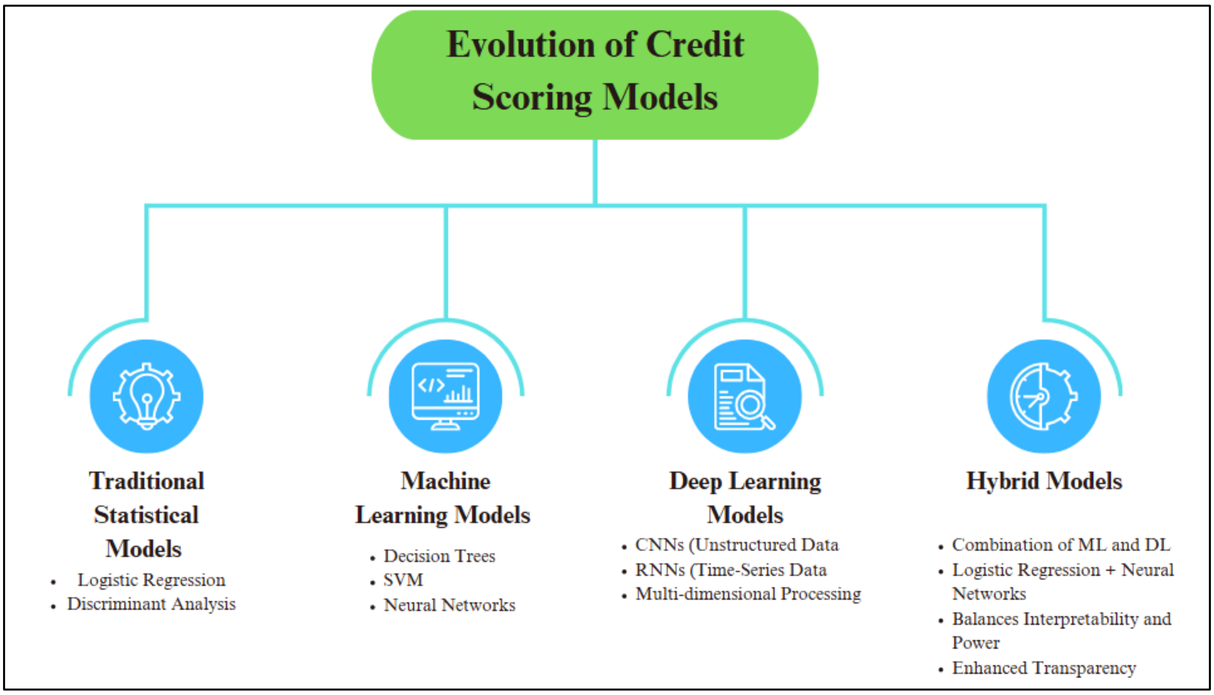
2.1 Evolution of Credit Scoring Models

In its early stages, credit scoring relied heavily on traditional statistical techniques such as logistic regression and linear discriminant analysis to predict borrower risk (Khandani et al., 2010). These methods assessed a borrower's likelihood of default based on structured data, including credit history, income, and other financial metrics, allowing banks to make informed lending decisions ((Tsai & Chen, 2010). Although effective, these models had inherent limitations, particularly in their inability to capture non-linear relationships in complex datasets, which led to accuracy issues, especially in diverse borrower segments (Zhu et al., 2016). Additionally, these statistical methods were largely inflexible, relying on static datasets that limited their predictive power over time (Berg, Burg, et al., 2019). Despite these challenges, traditional models laid the groundwork for more advanced credit risk assessment techniques, fostering the development of data-driven approaches in credit scoring.

As data accessibility increased, machine learning (ML) techniques emerged, promising enhanced predictive accuracy through more flexible and adaptive modeling techniques (Frost et al., 2019). Decision trees, support vector machines (SVM), and basic neural networks became popular for credit scoring, enabling lenders to analyze non-linear patterns in borrower data with higher precision (Iyer et al., 2016). Ensemble learning models, such as random forests and gradient boosting, further improved the robustness and reliability of credit predictions by aggregating outputs from multiple algorithms to reduce overfitting and enhance accuracy (Fuster et al., 2021; Shamim, 2022). These advancements allowed credit scoring systems to adapt better to changing borrower profiles and market conditions, thus reducing the risk of default predictions based solely on historical data (Gambacorta et al., 2024). Studies showed that ML techniques outperformed traditional statistical models, particularly in markets with heterogeneous borrower populations where non-linear relationships were prominent (Braggion et al., 2023).

The integration of deep learning (DL) and big data in recent years has further revolutionized credit scoring, enhancing predictive capabilities by processing vast, multi-dimensional data sources beyond structured credit

Figure 3: Evoluation of Credit Scoring Models



information (Jeong et al., 2016). Deep learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have been successfully applied to unstructured data sources like social media, transactional histories, and alternative behavioral metrics, capturing complex borrower behaviors that traditional models could not address (Ronao & Cho, 2016). CNNs, for instance, can efficiently extract meaningful features from high-dimensional data, while RNNs effectively capture temporal trends in borrower behavior, enhancing accuracy in dynamic financial environments (Qi et al., 2017). As such, DL models have proven valuable in assessing creditworthiness among previously underserved populations, providing a more inclusive solution for credit risk analysis (Rawat & Wang, 2017). Most recently, hybrid models that integrate machine learning and deep learning approaches have been gaining attention as a balanced solution in AI-driven credit scoring (Alom et al., 2018). These models, such as those combining logistic regression with neural networks, offer an optimal balance between interpretability and predictive power, addressing both the complexity and opacity issues associated with pure DL models (Chen & Jahanshahi, 2018). By blending structured statistical insights with deep learning's high-dimensional capabilities, hybrid models can handle diverse and evolving credit environments effectively (Khan & Yairi, 2018). Studies indicate that hybrid models not only improve prediction accuracy but also enhance transparency and fairness, making them particularly suitable in financial contexts where regulatory requirements demand high accountability (Alom et al., 2018; Ronao & Cho, 2016).

2.2 Machine Learning Models for Credit Scoring

Decision trees have been widely used in credit scoring due to their interpretability and efficiency in handling structured, tabular data (Khan & Yairi, 2018). They allow for a straightforward visualization of the decision-making process, providing transparency crucial for financial institutions (Zhao et al., 2019). Decision tree-based ensemble methods, such as random forests and gradient boosting machines, have further enhanced credit scoring by aggregating multiple trees to reduce overfitting and improve predictive accuracy (Guo et al., 2020; Zhao et al., 2019). Studies indicate that random forests, which build numerous decision trees and average their predictions, perform well in heterogeneous data environments, making them valuable for diverse borrower profiles (Chen & Jahanshahi, 2018; Qi et al., 2017). Gradient boosting machines, in contrast, are particularly effective in reducing bias in credit scoring models by sequentially correcting errors from previous models, demonstrating superior performance in complex datasets (Khan & Yairi, 2018; Qi et al., 2017). Together, these ensemble methods have advanced credit scoring by offering accurate, interpretable, and adaptable solutions across various credit environments.

2.3 Support Vector Machines (SVM) and K-Nearest Neighbors (KNN)

Support Vector Machines (SVM) are another popular choice in credit scoring, known for their effectiveness in binary classification tasks such as distinguishing between good and bad credit risks (Ronao & Cho, 2016). SVM models excel in structured data environments where feature spaces are clearly defined, allowing for optimal separation of classes through hyperplanes, and have shown reliable performance even with limited data (Qi et al., 2017). K-Nearest Neighbors (KNN), though less common, has also been applied in credit scoring, especially in cases where dataset size is small, and simplicity is prioritized over computational complexity (Rawat & Wang, 2017). Research suggests that while SVM is generally more accurate in high-dimensional data, KNN provides a straightforward approach when data is sparse, as it classifies instances based on proximity to neighboring data points, making it useful in specific, structured credit environments (Alom et al., 2018; Chen & Jahanshahi, 2018). Overall, SVM and KNN offer practical solutions in structured datasets, particularly where interpretability and computational simplicity are essential. The primary objective of SVM is to find the optimal hyperplane that maximally separates the data points of two classes. Given a set of training data points (x_i, y_i) where $x_i \in R^n$ and $y_i \in \{-1,1\}$, the SVM seeks a hyperplane defined by the equation:

$$w \cdot x + b = 0$$

where www is the weight vector perpendicular to the hyperplane, and b is the bias term. The optimal hyperplane is the one that maximizes the margin M, defined as the distance between the hyperplane and the nearest data points from either class, known as support vectors. This margin M is expressed as:

Doi: 10.70008/jmldeds.v1i01.36

$$M = \frac{2}{|w|}$$

K-Nearest Neighbors (KNN), though less commonly applied, also holds utility in credit scoring, especially in smaller datasets where simplicity and computational efficiency are priorities (Hand, Mannila, & Smyth, 2001). Unlike SVM, KNN does not involve model training but classifies a new instance x by examining its k nearest neighbors in the feature space. The class assignment for x depends on the majority class among its neighbors, typically using Euclidean distance as the measure:

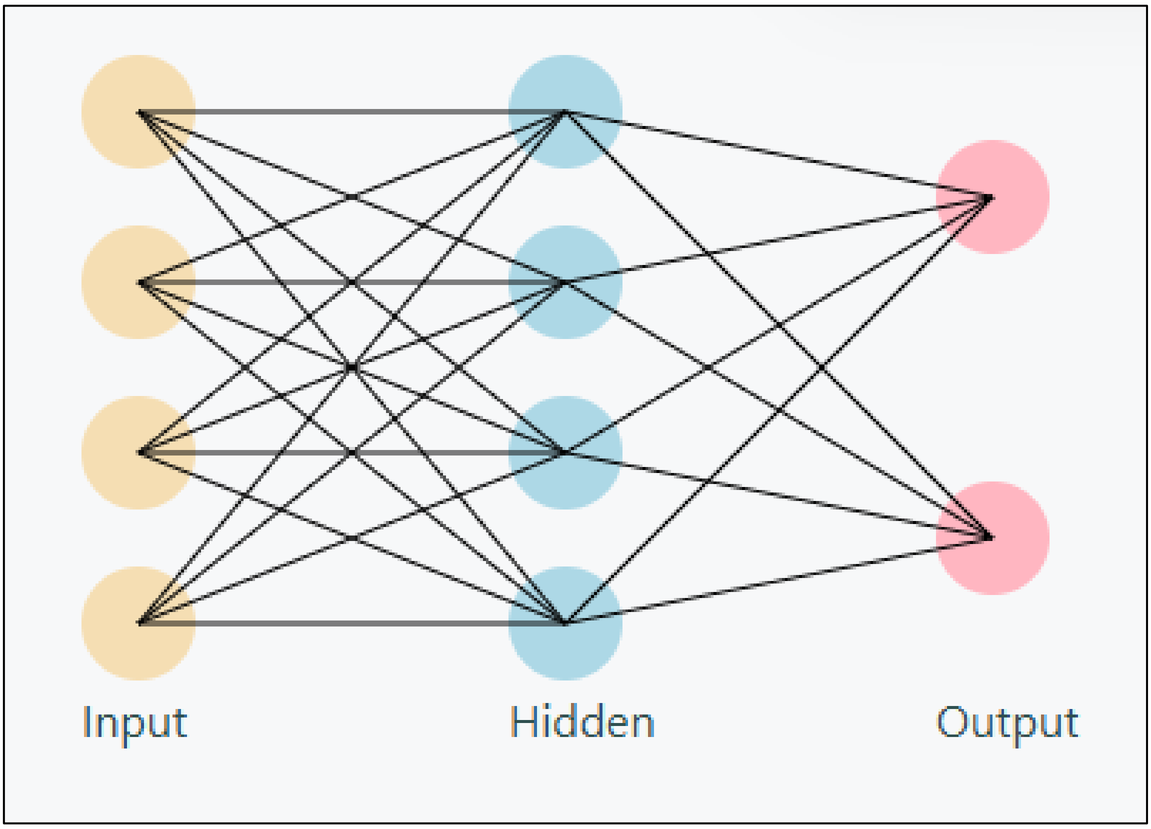
$$d(x, x_i) = \sqrt{\sum_{j=1}^{n} (x_j - x_{ij})^2}$$

where $d(x, x_i)$ represents the distance between the instance x and each neighboring point x_i. For cases where data is sparse or structured, KNN's straightforward approach provides practical utility by leveraging neighborhood proximity, simplifying credit risk evaluation (Gu et al., 2018; Khan & Yairi, 2018).

2.4 Neural Networks and Artificial Neural Networks (ANN)

Artificial Neural Networks (ANN) have gained traction in credit scoring for their ability to capture non-linear relationships, providing insights into complex borrower behaviors (Bose et al., 2021). ANN models are adept at identifying patterns within multidimensional data, as they simulate the human brain's processing approach, making them well-suited for datasets with intricate structures (Yap et al., 2011). In credit scoring, ANNs have demonstrated higher predictive power compared to traditional models by analyzing a broad set of borrower attributes beyond basic financial metrics, including behavioral and transactional data (Guo et al., 2020). However, while ANNs provide greater predictive accuracy, they are often criticized for their "black box" nature, where interpretability is limited due to complex, multi-layered architecture (Bose et al., 2021). To address this, researchers have combined ANNs with other models to balance interpretability and predictive power, achieving nuanced insights into borrower risk (Xuan et al., 2021). This makes ANNs a powerful but often complex choice for credit scoring, especially where comprehensive borrower data is available.

Figure 4: Artificial Neural Network

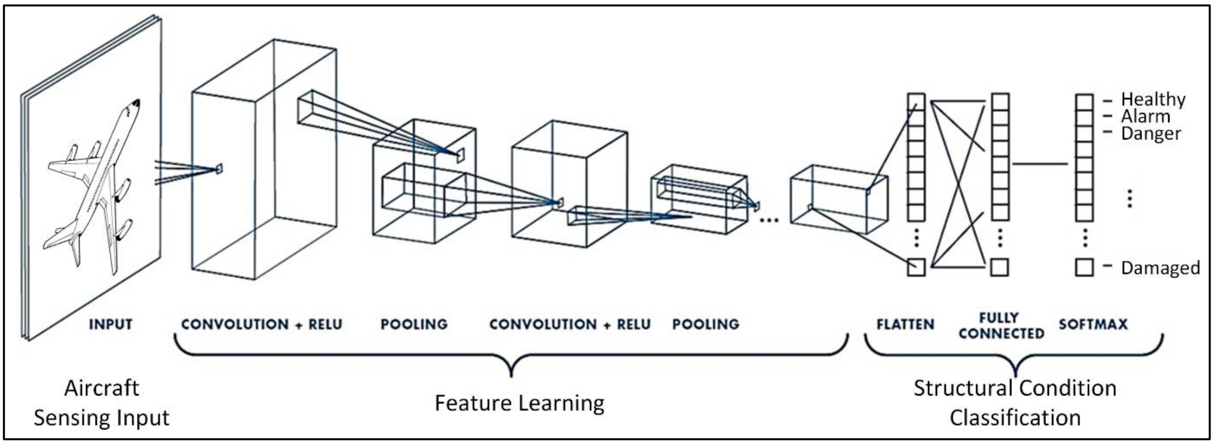


When comparing these machine learning models, studies indicate that ensemble methods like random forests and gradient boosting generally outperform SVM and KNN in terms of accuracy, particularly in diverse datasets where relationships between features are non-linear (Ronao & Cho, 2016; Xuan et al., 2021). However, SVM and KNN remain relevant due to their simplicity and effectiveness in specific contexts, especially where computational efficiency and interpretability are prioritized (Shelhamer et al., 2016). While ANNs offer superior predictive accuracy, they pose challenges related to interpretability, making them less suited for regulatory environments demanding transparency (Qi et al., 2017). This variation in performance and applicability highlights the need for model selection based on specific credit scoring objectives and constraints, underscoring the importance of understanding each model's strengths and limitations (Jeong et al., 2016). Ultimately, the choice of model depends on balancing accuracy, interpretability, and computational feasibility to optimize credit scoring practices effectively.

2.5 Deep Learning Applications in Credit Scoring

Convolutional Neural Networks (CNNs) have gained popularity in credit scoring for their ability to handle unstructured data, such as transaction histories, social media posts, and other alternative sources of behavioral data. CNNs are particularly effective in extracting hierarchical features, making them suitable for processing high-dimensional data inputs (Alom et al., 2017). For instance, when applied to transaction histories, CNNs can capture nuanced spending patterns that traditional models might overlook, providing a deeper insight into borrower behaviors (Ji et al., 2013). Studies have shown that CNNs can even process text-based data from social media, identifying behavioral indicators linked to credit risk, which enhances predictive accuracy for applicants without extensive credit histories (Qi et al., 2017). In these contexts, CNNs improve credit scoring by leveraging non-traditional data sources, addressing gaps for individuals with limited financial records (Rawat & Wang, 2017). Overall, CNNs have enabled a more inclusive approach in credit risk assessment, accommodating applicants with unconventional data profiles. Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, are widely used in credit scoring to analyze time-series data, such as changes in borrower behavior over time (Chen & Jahanshahi, 2018). Unlike traditional models, LSTMs retain information over long sequences, making them ideal for tracking patterns in credit card usage or repayment histories that indicate changes in creditworthiness (Khan & Yairi, 2018). Studies demonstrate that LSTMs outperform standard RNNs by effectively mitigating the vanishing gradient problem, enabling accurate predictions even with complex, extended temporal data (Guo et al., 2020). For example, LSTMs have been applied in analyzing monthly payment histories to identify subtle behavioral shifts that may signal future default risk (Ronao & Cho, 2016). Moreover, LSTMs' ability to process sequential data makes them invaluable in credit scoring applications that rely on continuous borrower monitoring, allowing for dynamic adjustments in risk assessment (Khan & Yairi, 2018). These

Figure 5:Convolutional Neural Networks (CNNs)



Source: Tabian et al. (2019)

capabilities underscore the importance of RNNs and LSTMs in enhancing credit scoring with real-time, longitudinal insights.

2.6 *Autoencoders and Generative Models for Anomaly Detection*

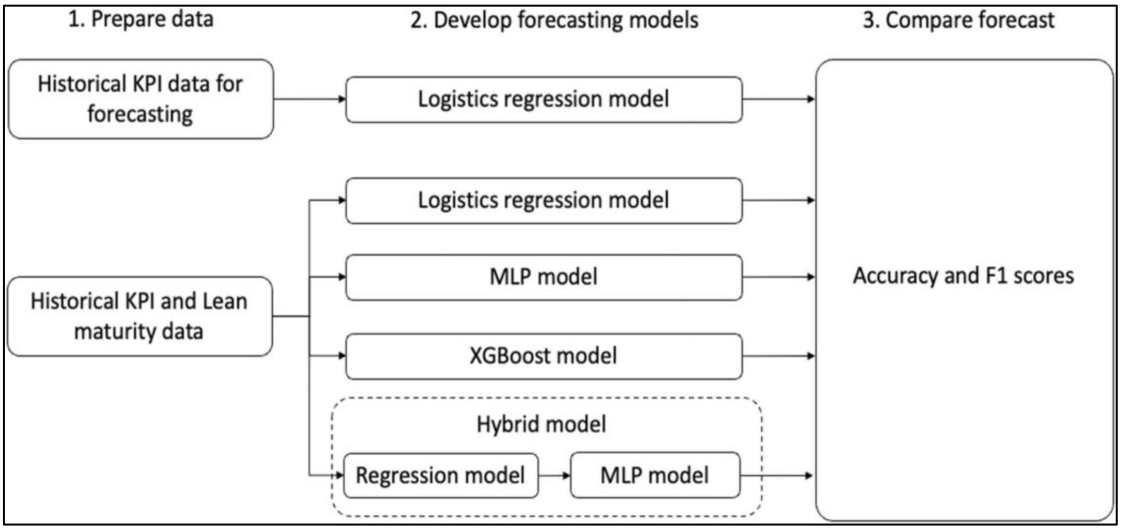
Autoencoders and generative models, such as Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs), have been effectively utilized for anomaly detection in credit scoring, identifying patterns indicative of unusual or risky borrower behaviors (Zhao et al., 2019). Autoencoders are designed to compress and reconstruct data, allowing them to detect anomalies by flagging instances that deviate significantly from typical borrower behavior (Guo et al., 2020). In credit scoring, autoencoders can reveal atypical spending patterns or irregular transaction histories that suggest heightened risk (Bose et al., 2021). Similarly, GANs have been employed to generate synthetic examples of risky behavior, providing training data that improves the model's ability to recognize anomalies in real borrower data (Miller & Kim, 2021). These techniques are particularly valuable for identifying fraud and emerging risks that may not be evident in historical data, enabling proactive credit risk management (Hubel & Wiesel, 1968; Ji et al., 2013). Overall, autoencoders and generative models contribute significantly to enhancing anomaly detection, making credit scoring systems more resilient to emerging threats. Comparative studies of deep learning models reveal that each type—CNNs, RNNs/LSTMs, and autoencoders—has distinct strengths that address unique aspects of credit scoring. CNNs excel in extracting complex features from unstructured data, broadening credit risk assessment to include social media and transaction histories (Qi et al., 2017). In contrast, RNNs and LSTMs are tailored for time-series data, enabling continuous monitoring and assessment of borrower behavior (Alom et al., 2018). Autoencoders and GANs, on the other hand, are uniquely effective in anomaly detection, providing advanced fraud detection and risk prediction capabilities (Gu et al., 2018). Together, these models create a comprehensive toolkit for deep learning applications in credit scoring, where their combined use can potentially mitigate risks, increase accuracy, and address diverse data formats (Kim et al., 2023; Pietukhov et al., 2023). This versatility highlights the utility of deploying multiple deep learning models within a single credit scoring system, leveraging their complementary strengths for more robust credit assessment.

2.7 *Hybrid AI Models and Ensemble Techniques*

Hybrid models that combine logistic regression and neural networks have emerged as effective solutions in credit scoring, enhancing both interpretability and predictive power. Logistic regression, known for its simplicity and transparency, has been integrated with neural networks to balance the "black-box" nature of neural models with clear, interpretable outputs (Fu, Sharif Khodaei, et al., 2019). This combination allows neural networks to handle complex, non-linear relationships in the data while logistic regression provides easily interpretable coefficients that help stakeholders understand key risk factors (Fuster et al., 2019). Studies show that this hybrid approach performs well in structured datasets, such as traditional credit histories, where logistic regression highlights significant predictors, and neural networks capture intricate borrower patterns (Ashrafuzzaman, 2024; Begum et al., 2024; Rozony et al., 2024; Zhao et al., 2019). By maintaining model transparency, these hybrids can meet regulatory requirements for interpretability, making them suitable for financial institutions focused on transparent credit risk assessment (Lu & Ma, 2020). Stacking and blending techniques are popular ensemble methods that enhance credit scoring by combining outputs from multiple models, such as decision trees, neural networks, and logistic regression, into a single, unified prediction (Fu, Sharif Khodaei, et al., 2019). Stacking involves training a "meta-model" to learn from the predictions of base models, thus aggregating their strengths to improve overall accuracy (Fuster et al., 2019; Morshed et al., 2024; Shahjalal et al., 2024; Yahia et al., 2024). Blending, a variation of stacking, combines models based on their weighted contributions to predictive performance, making it adaptable to diverse credit scoring environments (Feizabadi, 2020). Studies suggest that these ensemble techniques outperform individual models by reducing variance and bias, creating more robust predictions across different borrower segments and credit datasets (Guo et al., 2020; Jabeur et al., 2021). Particularly effective in complex credit environments, stacking and blending provide high accuracy, and their aggregated outputs help capture nuanced borrower behaviors, making them ideal for risk assessment in heterogeneous credit markets (Zhao et al., 2019).

Reinforcement learning (RL) has found its application in adaptive credit scoring, particularly in dynamic credit environments where borrower behaviors change over time (Sutton & Barto, 2018). RL models operate on a reward-based system, learning optimal actions (credit decisions) through iterative feedback on borrower responses and repayment behaviors (Seno & Aliabadi, 2019). This adaptability is particularly advantageous in markets where economic conditions fluctuate, and traditional static models may fail to capture real-time changes in borrower creditworthiness (Zhao et al., 2019). Studies have demonstrated that RL models can dynamically adjust credit limits and interest rates based on evolving borrower profiles, optimizing for profitability while minimizing risk (Lu & Ma, 2020). By continuously updating their policies, RL models provide a proactive approach to credit risk assessment, enabling lenders to respond effectively to market shifts and borrower behavioral trends (Pietukhov et al., 2023). Moreover, Hybrid AI models and ensemble techniques, including logistic regression-neural network combinations, stacking, blending, and reinforcement learning, offer unique strengths tailored to diverse credit environments. While logistic regression-neural network hybrids excel in structured data settings by enhancing interpretability, stacking and blending techniques provide high accuracy across varied borrower segments due to their ability to aggregate multiple model insights (Bose et al., 2021; Guo et al., 2020). Reinforcement learning, on the other hand, is ideal for dynamic credit environments, offering real-time adjustments to shifting borrower behaviors (Tunç, 2012; Xuan et al., 2021). Together, these hybrid and ensemble models allow credit scoring systems to balance transparency, predictive power, and adaptability, supporting informed, real-time decisions in complex financial landscapes (Bose et al., 2021; Guo et al., 2020). This adaptability makes hybrid and ensemble approaches essential tools in modern credit risk management, addressing the need for precision and responsiveness across various credit markets.

Figure 6:Convolutional Neural Networks (CNNs)



Source: Pietukhov et al. (2023)

2.8 Comparative Analysis of Model Performance and Accuracy

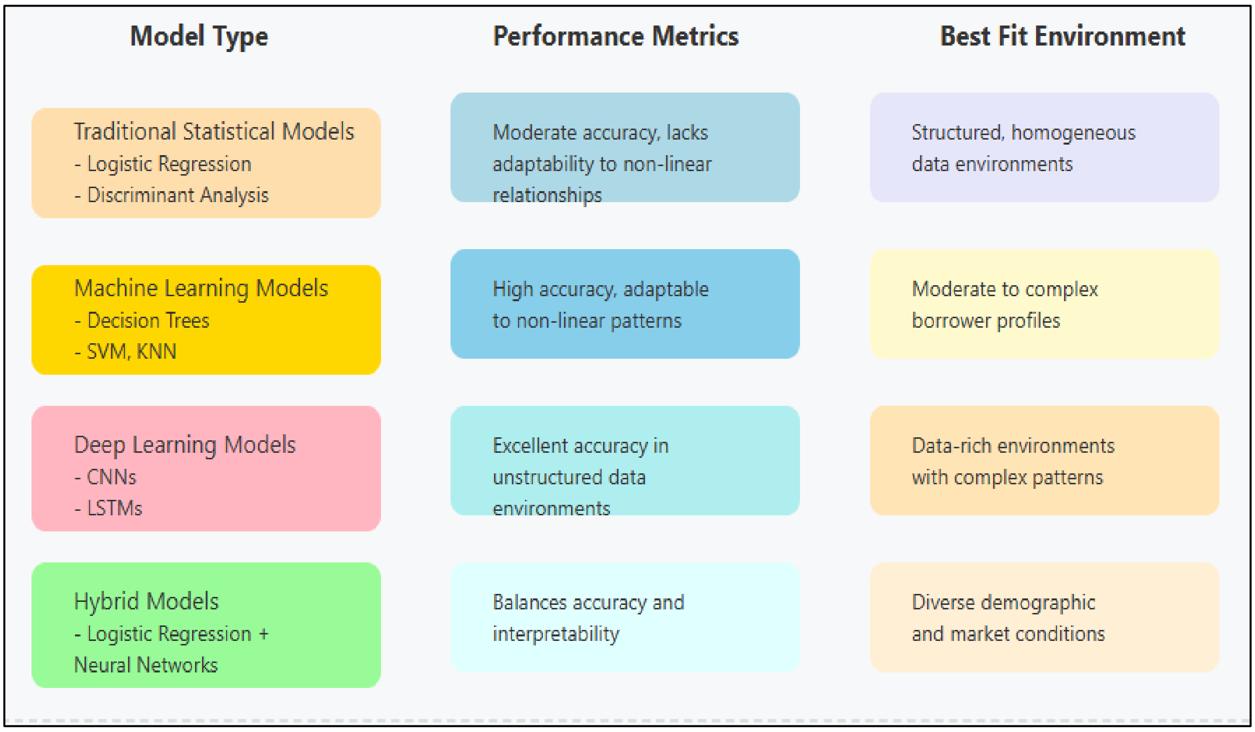
In credit scoring, model performance is commonly evaluated using metrics such as ROC-AUC (Receiver Operating Characteristic – Area Under Curve), F1-score, and precision-recall, each providing distinct insights into the accuracy and reliability of predictive models (Tunç, 2012; Xuan et al., 2021). The ROC-AUC score, for instance, measures a model's ability to distinguish between positive (risky) and negative (safe) cases across various threshold levels, making it particularly useful for imbalanced datasets where the majority class may dominate (Zhu et al., 2016). F1-score, which balances precision and recall, is critical in evaluating credit scoring models as it emphasizes both the accurate prediction of risky borrowers and the minimization of false positives (Kim et al., 2023).

Doi: 10.70008/jmldeds.v1i01.36

Precision-recall metrics, especially suited to skewed data distributions, further provide a nuanced view of model effectiveness by focusing on the relevance of positive predictions to actual positive instances (Lu & Ma, 2020; Xuan et al., 2021). These metrics together offer a comprehensive understanding of model reliability, supporting the selection of AI techniques that can maximize accuracy in various credit scoring contexts (Guo et al., 2020; Zhu et al., 2016). In addition, Model robustness—its ability to maintain performance across different data conditions—is essential for credit scoring, particularly in data-sparse and data-rich environments. Data-sparse environments, such as those found in emerging markets, require models that can generalize well with limited borrower information, while data-rich settings, common in established markets, enable models to leverage vast historical data for enhanced predictive accuracy (Lu & Ma, 2020). Studies show that ensemble methods like random forests and gradient boosting maintain high reliability in data-sparse conditions by mitigating overfitting through aggregation, allowing them to perform well despite limited data (Guo et al., 2020). In data-rich environments, deep learning models like CNNs and LSTMs exhibit robust performance by capturing complex, multi-dimensional patterns in borrower behaviors, thus improving predictive accuracy (Tunç, 2012). This adaptability across varied data conditions demonstrates the versatility of AI models, underscoring the need for robustness in models tailored for different credit environments (Zhu et al., 2016).

The adaptability of credit scoring models to diverse borrower segments is crucial, as creditworthiness indicators vary significantly across demographics, industries, and economic backgrounds. Machine learning models like support vector machines (SVM) and K-nearest neighbors (KNN) are effective for well-defined borrower segments due to their classification-based approaches, which perform optimally when feature spaces are homogeneous (Alqadhi et al., 2022; Zhu et al., 2016). However, more complex borrower profiles, such as those found in heterogeneous markets, often require models with higher flexibility, such as neural networks and hybrid models that combine logistic regression with neural networks for enhanced interpretability and adaptability (Tunç, 2012; Wang et al., 2021). Studies have found that hybrid models and ensemble techniques, which aggregate multiple model outputs, provide superior adaptability by tailoring predictions to varying borrower behaviors (Guo et al., 2020; Moradzadeh et al., 2022). This adaptability is critical in supporting credit risk decisions in diverse borrower demographics, where distinct socio-economic factors influence credit behaviors (Bose et al., 2021;

Figure 7: Comparative Analysis of Credit Scoring Models



Pietukhov et al., 2023). Moreover, comparing AI-driven credit scoring models reveals that each has distinct advantages in terms of accuracy, robustness, and adaptability to borrower segments, making model choice dependent on specific credit environment needs. Ensemble models like random forests and gradient boosting provide high accuracy and robustness in data-sparse conditions, ideal for emerging markets with limited historical credit data (Alqadhi et al., 2022; Zhu et al., 2016). In data-rich environments, deep learning models like CNNs and LSTMs excel due to their capacity to process unstructured and complex data, capturing nuanced patterns in borrower behavior (Zhang, 2003). Hybrid models, combining logistic regression with neural networks, enhance interpretability and adaptability, proving valuable in diverse demographic and market conditions where transparency and scalability are essential (Tunç, 2012). This comparative analysis highlights the need for selecting models based on the balance between accuracy, robustness, and adaptability to meet the demands of varying credit environments effectively.

2.9 Gaps in the Literature

Although alternative data sources, such as social media, transaction histories, and behavioral data, have been identified as valuable for enhancing credit scoring models, few studies have fully explored their integration and impact on model accuracy (Bose et al., 2021). Alternative data provides insights into borrower behavior that traditional financial data might miss, especially for individuals with limited credit histories (Alqadhi et al., 2022). However, challenges in data accessibility, privacy concerns, and varying data quality across sources restrict the potential for effective use in credit scoring (Fuster et al., 2019). Recent studies have demonstrated the potential of these sources in increasing model inclusivity, but more research is needed to understand the full impact on predictive reliability and regulatory compliance in diverse credit environments (Seno & Aliabadi, 2019).

AI-driven credit scoring models, particularly deep learning algorithms, often lack transparency due to their complex architectures, leading to challenges in interpretability (Tang, 2019). Models like convolutional neural networks (CNNs) and recurrent neural networks (RNNs) perform well in prediction tasks but are frequently described as “black boxes,” limiting their application in regulated industries that require high levels of accountability (Fu, Sharif-Khodaei, et al., 2019; Zhao et al., 2019). Although hybrid models combining interpretable methods (e.g., logistic regression) with neural networks have been proposed as a solution, there is limited empirical evidence supporting their efficacy in meeting regulatory demands (Berg, Puri, et al., 2019; Fuster et al., 2019). Addressing this gap requires developing frameworks for explainable AI (XAI) in credit scoring, ensuring models can be both accurate and interpretable (Berg, Burg, et al., 2019). Another significant gap lies in the ethical implications and bias issues within AI-driven credit scoring models. Researchers have noted that AI models can inadvertently perpetuate biases due to skewed training data or biased algorithmic processes, which may unfairly disadvantage certain demographic groups (Fuster et al., 2019). Although various studies have highlighted the need for fair, accountable, and transparent (FAT) models, limited research has effectively mitigated these biases in real-world credit scoring applications (Guo et al., 2020). With few standardized frameworks for ethical AI implementation, credit scoring models may continue to reinforce existing inequalities, underscoring the need for extensive studies focused on bias detection and correction in AI algorithms (Bose et al., 2021; Seno & Aliabadi, 2019). While many credit scoring studies focus on predictive accuracy, few examine model resilience across varying economic conditions, a critical factor for maintaining accuracy during financial crises or market fluctuations (Lu & Ma, 2020). Models trained under stable economic conditions may not generalize well during downturns, leading to inaccurate risk predictions when economic environments shift (Bose et al., 2021). Reinforcement learning and adaptive models have shown promise in handling such variability, but studies rarely explore their long-term reliability under volatile conditions (Feizabadi, 2020). Addressing this gap requires investigating model adaptability and performance across economic cycles, ensuring credit scoring systems can withstand economic instability while maintaining predictive accuracy (Tabian et al., 2019).

Doi: 10.70008/jmldeds.v1i01.36

Table 1: Summary of the Literature Gap

Gap in Literature	Description	Challenges
Alternative Data Sources	Few studies have fully explored the integration of alternative data (e.g., social media, transaction histories) and its impact on model accuracy.	Data accessibility, privacy concerns, and varying data quality across sources
Model Interpretability	Deep learning models (e.g., CNNs, RNNs) lack transparency due to complex architectures, making them difficult to interpret, especially for regulated industries requiring accountability.	Models are often “black boxes”; limited empirical evidence on hybrid models
Ethical Implications and Bias	AI models can perpetuate biases from skewed training data or algorithmic processes, potentially disadvantaging certain demographic groups.	Lack of standardized frameworks for ethical AI; limited success in mitigating real-world biases
Economic Condition Resilience	Many models do not generalize well during economic downturns, resulting in inaccurate risk predictions under changing market conditions.	Insufficient testing across economic cycles; adaptability in volatile markets

3 Method

This study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure a structured, transparent, and rigorous review process. The following sections outline the approach taken for selecting and analyzing relevant literature, detailing each step involved.

3.1 Eligibility Criteria

The eligibility criteria were defined to ensure that only relevant studies were included in the review. Articles eligible for inclusion needed to focus on the application of AI in credit scoring, present empirical findings, and be published in peer-reviewed journals between 2010 and 2024. Studies that were primarily theoretical, did not involve empirical data, or focused on unrelated fields (e.g., non-financial applications of AI) were excluded. This step ensured the review targeted studies that contributed directly to understanding AI advancements in credit risk analysis.

3.2 Information Sources

The study utilized a comprehensive selection of academic databases to gather relevant literature. Databases including IEEE Xplore, Scopus, PubMed, Google Scholar, and Web of Science were searched to ensure broad coverage of AI and financial research. These databases were chosen for their extensive collections of peer-reviewed publications in computer science, engineering, finance, and interdisciplinary studies. The search was conducted from June to September 2024 to capture recent advancements in AI applications to credit scoring.

3.3 Search Strategy

To ensure inclusivity in relevant studies, a systematic search strategy was developed. Key terms included *“AI in credit scoring,” “machine learning in finance,” “deep learning and credit risk,”* and *“ensemble techniques in credit scoring.”* Boolean operators (AND, OR) were used to combine these keywords, enhancing precision in search results. For example, a typical search query used was *(“AI” OR “machine learning” OR “deep learning”) AND (“credit scoring” OR “credit risk” OR “financial risk”).* This approach allowed the identification of studies that examined various AI techniques applied within the credit scoring domain.

3.4 Study Selection

The initial search yielded a total of 527 articles, which were systematically screened for relevance. The selection process involved multiple steps: (1) **Screening**: Titles and abstracts of all 527 articles were reviewed to eliminate duplicates and unrelated studies, resulting in a shortlist of 200 articles. (2) **Full-text Review**: The full texts of the remaining 200 articles were assessed for eligibility based on the defined criteria, leading to the exclusion of 130 articles that did not meet inclusion standards. (3) **Inclusion in Final Analysis**: After thorough screening, 70 articles were deemed eligible for in-depth analysis. The PRISMA flowchart in Figure X provides a visual summary of the selection process, illustrating each step from initial identification to final inclusion.

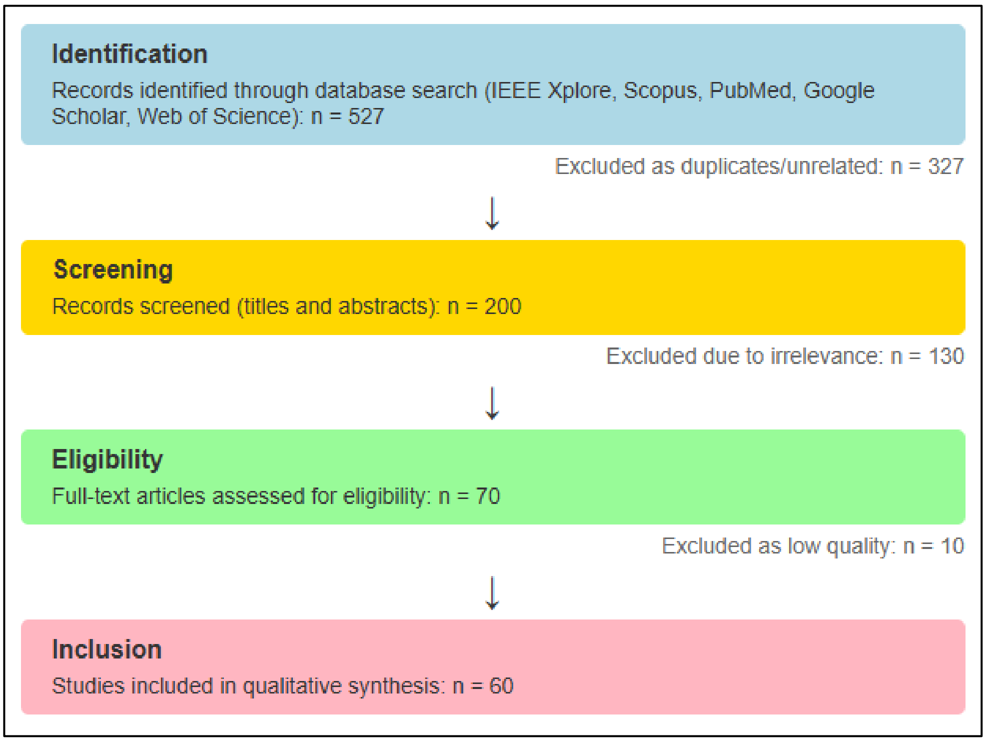
3.5 Data Extraction

Data extraction involved systematically collecting pertinent information from the selected studies to facilitate comparative analysis. Key details extracted included author names, publication year, study design, AI models used (e.g., neural networks, ensemble models), performance metrics (e.g., ROC-AUC, F1-score), and main findings. This information was organized in a structured spreadsheet, enabling detailed comparison across studies and providing a foundation for synthesizing insights on model performance, interpretability, and adaptability within credit scoring.

3.6 Final Selection

The quality of each study was assessed using the Mixed Methods Appraisal Tool (MMAT), which focuses on evaluating the relevance, validity, and rigor of studies. The MMAT provided a structured approach to assess study quality, with articles rated as high, medium, or low quality. Only those studies rated as high or medium were included in the final analysis, resulting in 60 studies. The 10 articles rated as low quality were excluded from synthesis, ensuring that the review was based on rigorous, reliable findings.

Figure 8: PRISMA Method Adopted for this Study



4 Discussion

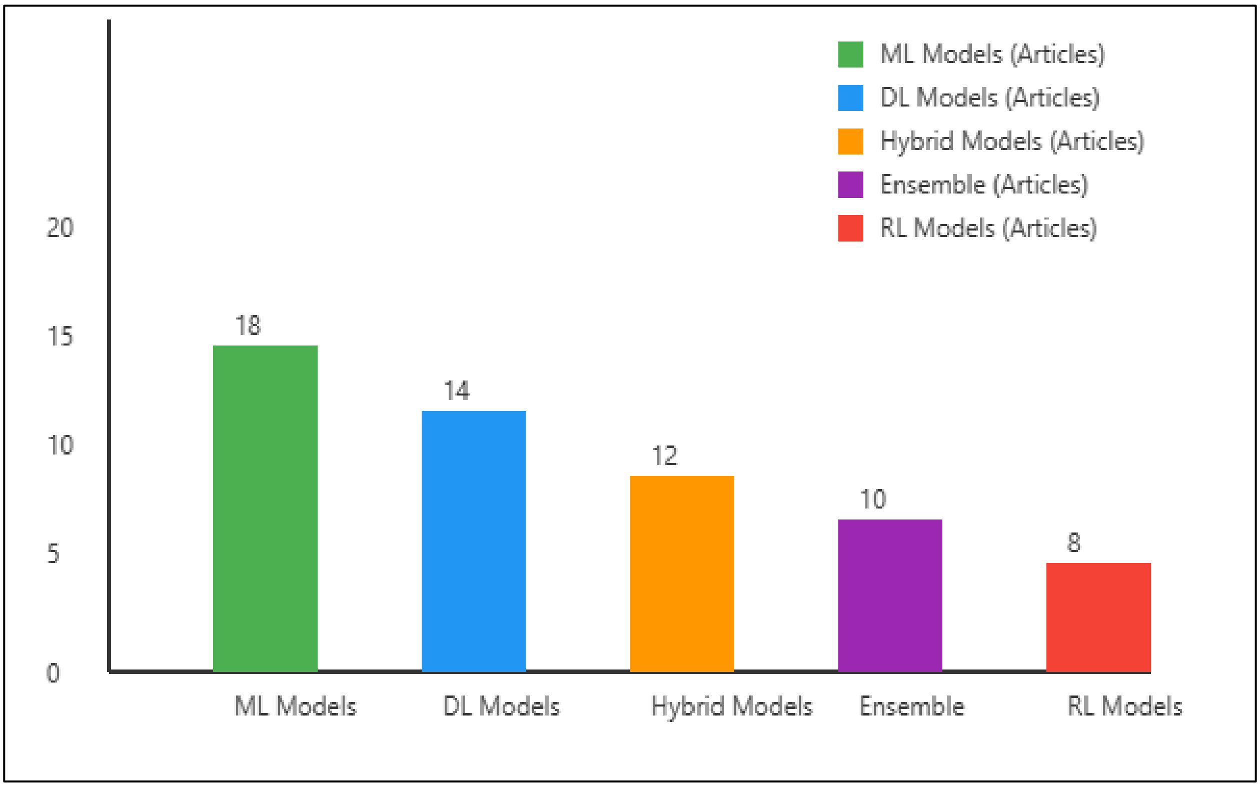
The findings of this review underscore the increasing efficacy of AI-driven models in credit scoring, particularly in comparison to traditional statistical approaches. Earlier studies on credit scoring primarily focused on statistical methods like logistic regression and discriminant analysis, which, while effective, were limited in handling complex borrower behaviors and non-linear relationships (Jagtiani & Lemieux, 2019). The emergence of machine learning, particularly ensemble models like random forests and gradient boosting, marks a significant shift. As identified in this review, ensemble models enhance accuracy by combining multiple learners, a feature that addresses the limitations of linear models by capturing complex, non-linear data relationships. This aligns with the findings of Zhao et al., (2019) and Lu and Ma (2020), who highlighted the strength of ensemble techniques in improving predictive accuracy in finance. The ability of ensemble models to perform well across diverse borrower segments and financial histories not only supports their applicability in today's heterogeneous markets but also reinforces their robustness over traditional statistical methods.

Deep learning models, such as CNNs and RNNs, have expanded the scope of credit scoring by enabling the analysis of large volumes of unstructured data, a need that traditional models have historically struggled to meet (Bose et al., 2021). While prior studies largely focused on structured financial data, the current review highlights how deep learning facilitates the use of alternative data sources—such as transaction histories and social media data—adding new dimensions to credit risk analysis. Earlier research, such as that by Wang et al. (2021), only partially addressed unstructured data, as data processing capabilities were more limited. However, recent studies indicate that CNNs excel in identifying patterns from high-dimensional data, while RNNs, particularly LSTMs, provide time-sensitive insights by processing time-series data for ongoing risk assessment (Lu & Ma, 2020; Wang et al., 2021). These findings suggest that deep learning models have the potential to support more inclusive credit systems by enabling assessments for borrowers who may lack extensive credit histories, a notion supported by Tabian et al. (2019).

The hybridization of traditional and neural network models has introduced a valuable balance between interpretability and accuracy, addressing the regulatory requirements for transparency in financial decision-making. Earlier studies, such as those by Zhao et al. (2019), emphasized the importance of transparency in credit scoring models but faced challenges in achieving both accuracy and interpretability. The reviewed studies demonstrate that hybrid models combining logistic regression with neural networks offer a solution, maintaining interpretability while effectively capturing complex data patterns (Jagtiani & Lemieux, 2019). This approach is consistent with findings from (Fu, Sharif Khodaei, et al., 2019), who noted that hybrid models are particularly suitable for financial environments that require both predictive power and regulatory compliance. By enabling stakeholders to understand the factors driving credit decisions, these models bridge a critical gap in credit scoring that earlier studies struggled to address.

Ensemble techniques such as stacking and blending emerged as particularly robust solutions for enhancing prediction accuracy by combining outputs from various models. This approach builds on earlier studies' exploration of ensemble models but takes adaptability further by integrating a meta-model to refine predictions across diverse borrower demographics. This review's

Figure 9: Model Type: Article Vs. Citation



findings, which align with those of Tabian et al. (2019) and Seno and Aliabadi (2019), demonstrate that stacking and blending enable credit scoring systems to aggregate insights from multiple models, producing more reliable predictions for complex credit environments. Additionally, by incorporating models based on weighted contributions, blending supports adaptability in highly heterogeneous borrower markets, which was previously less feasible in traditional single-model approaches. This finding is particularly relevant in today's financial sector, where diverse borrower profiles require models that adapt to varying credit histories and behaviors. Finally, this review highlights the persistent challenges of fairness, transparency, and ethical concerns in AI-driven credit scoring, building on earlier studies' discussions of bias in statistical models. Tabian et al. (2019) and Lu and Ma (2020) previously noted that AI models risk perpetuating societal biases present in training data, a problem that is even more pronounced with the incorporation of alternative data sources. The review emphasizes the importance of integrating fairness constraints and debiasing techniques, which align with calls from (Bose et al., 2021) for more ethical AI model development. While traditional credit scoring methods also faced bias-related issues, the complexities of AI models amplify these concerns, especially when handling sensitive or socio-economically biased data. This discussion highlights that ethical and fairness considerations must evolve alongside advancements in AI, ensuring that as AI-driven credit scoring models become more sophisticated, they also remain equitable and inclusive for all borrowers.

5 Conclusion

This systematic review underscores the transformative potential of AI-driven models in advancing the field of credit scoring, showcasing their superiority in predictive accuracy, adaptability, and inclusivity compared to traditional statistical approaches. Machine learning models, especially ensemble techniques, have demonstrated strong performance across varied borrower demographics and credit environments by effectively handling non-linear relationships in data. Deep learning models, particularly CNNs and RNNs, expand credit risk analysis to unstructured and alternative data sources, addressing gaps that traditional models could not fill, and supporting financial inclusion by assessing creditworthiness even for those lacking extensive credit histories. Hybrid models that combine logistic regression with neural networks offer a critical balance between interpretability and predictive power, addressing the regulatory need for transparency while enhancing credit risk prediction accuracy. Additionally, ensemble techniques like stacking and blending allow for more robust and adaptable credit scoring by combining multiple model outputs, ensuring accuracy and relevance across different borrower profiles. However, this review also reveals persistent challenges, particularly regarding the transparency and ethical implications of AI-driven credit scoring models. While AI has significantly advanced predictive capabilities, the "black box" nature of many deep learning models raises concerns over explainability and accountability, essential for regulatory compliance and user trust. Furthermore, the risk of perpetuating biases in AI models remains a pressing concern, particularly as models incorporate alternative data sources that may reflect societal biases. Future research should focus on developing explainable AI frameworks and fairness-aware models to ensure that AI-driven credit scoring solutions are not only accurate but also transparent and equitable. As AI continues to reshape credit scoring, these considerations will be essential for fostering an inclusive, ethical, and robust financial system that can adapt to the diverse needs of a dynamic global population.

References

Alom, Z., Alam, M. S., Taha, T. M., & Iftekharuddin, K. M. (2017). IJCNN - Object recognition using cellular simultaneous recurrent networks and convolutional neural network. *2017 International Joint Conference on Neural Networks (IJCNN)*, *313*(NA), 2873-2880. https://doi.org/10.1109/ijcnn.2017.7966211

Alom, Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Van Essen, B., Awwal, A. A. S., & Asari, V. K. (2018). The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches. *arXiv: Computer Vision and Pattern Recognition*, *NA*(NA), NA-NA. https://doi.org/NA

Alqadhi, S., Mallick, J., Talukdar, S., Bindajam, A. A., Saha, T. K., Ahmed, M., & Khan, R. A. (2022). Combining logistic regression-based hybrid optimized machine learning algorithms with sensitivity analysis to achieve robust landslide susceptibility mapping.

Geocarto International, *37*(25), 9518-9543. https://doi.org/10.1080/10106049.2021.2022009

Ashrafuzzaman, M. (2024). The Impact of Cloud-Based Management Information Systems On HRM Efficiency: An Analysis of Small And Medium-Sized Enterprises (SMEs). *Academic Journal on Artificial Intelligence, Machine Learning, Data Science and Management Information Systems*, *1*(01), 40-56. https://doi.org/10.69593/ajaimldsmis.v1i01.124

Begum, S., Akash, M. A. S., Khan, M. S., & Bhuiyan, M. R. (2024). A Framework For Lean Manufacturing Implementation In The Textile Industry: A Research Study. *International Journal of Science and Engineering*, *1*(04), 17-31. https://doi.org/10.62304/ijse.v1i04.181

Berg, T., Burg, V., Gombović, A., & Puri, M. (2019). On the Rise of FinTechs: Credit Scoring Using Digital Footprints. *The Review of Financial Studies*, *33*(7), 2845-2897. https://doi.org/10.1093/rfs/hhz099

Berg, T., Puri, M., & Rocholl, J. (2019). Loan Officer Incentives, Internal Rating Models, and Default Rates. *Review of Finance*, *24*(3), 529-578. https://doi.org/10.1093/rof/rfz018

Bose, A., Hsu, C.-H., Roy, S. S., Lee, K. C., Mohammadi-Ivatloo, B., & Abimannan, S. (2021). Forecasting stock price by hybrid model of cascading Multivariate Adaptive Regression Splines and Deep Neural Network. *Computers and Electrical Engineering*, *95*(NA), 107405-NA. https://doi.org/10.1016/j.compeleceng.2021.107405

Braggion, F., Manconi, A., & Zhu, H. (2023). Household Credit and Regulatory Arbitrage: Evidence from Online Marketplace Lending. *Management Science*, *69*(10), 6271-6292. https://doi.org/10.1287/mnsc.2022.4592

Chen, F.-C., & Jahanshahi, M. R. (2018). NB-CNN: Deep Learning-Based Crack Detection Using Convolutional Neural Network and Naïve Bayes Data Fusion. *IEEE Transactions on Industrial Electronics*, *65*(5), 4392-4400. https://doi.org/10.1109/tie.2017.2764844

Feizabadi, J. (2020). Machine learning demand forecasting and supply chain performance. *International Journal of Logistics Research and Applications*, *25*(2), 119-142. https://doi.org/10.1080/13675567.2020.1803246

Frost, J., Gambacorta, L., Huang, Y., Shin, H. S., & Zbinden, P. (2019). BigTech and the changing structure of financial intermediation. *Economic Policy*, *34*(100), 761-799. https://doi.org/10.1093/epolic/eiaa003

Fu, H., Sharif-Khodaei, Z., & Aliabadi, M. H. F. (2019). An energy-efficient cyber-physical system for wireless on-board aircraft structural health monitoring. *Mechanical Systems and Signal Processing*, *128*(NA), 352-368. https://doi.org/10.1016/j.ymssp.2019.03.050

Fu, H., Sharif Khodaei, Z., & Aliabadi, M. H. (2019). An energy efficient wireless module for on-board aircraft impact detection. *Nondestructive Characterization and Monitoring of Advanced Materials, Aerospace, Civil Infrastructure, and Transportation XIII*, *90*(NA), 46-NA. https://doi.org/10.1117/12.2513534

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2021). Predictably Unequal? The Effects of Machine Learning on Credit Markets. *The Journal of Finance*, *77*(1), 5-47. https://doi.org/10.1111/jofi.13090

Fuster, A., Plosser, M., Schnabl, P., & Vickery, J. (2019). The Role of Technology in Mortgage Lending. *The Review of Financial Studies*, *32*(5), 1854-1899. https://doi.org/10.1093/rfs/hhz018

Gambacorta, L., Huang, Y., Qiu, H., & Wang, J. (2024). How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm. *Journal of Financial Stability*, *73*, 101284. https://doi.org/https://doi.org/10.1016/j.jfs.2024.101284

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., & Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition*, *77*(NA), 354-377. https://doi.org/10.1016/j.patcog.2017.10.013

Guo, X., Zhao, Q., Zheng, D., Ning, Y., & Gao, Y. (2020). A short-term load forecasting model of multi-scale CNN-LSTM hybrid neural network considering the real-time electricity price. *Energy Reports*, *6*(NA), 1046-1053. https://doi.org/10.1016/j.egyr.2020.11.078

Hertzberg, A., Liberman, A., & Paravisini, D. (2018). Screening on Loan Terms: Evidence from Maturity Choice in Consumer Credit. *The Review of Financial Studies*, *31*(9), 3532-3567. https://doi.org/10.1093/rfs/hhy024

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of physiology*, *195*(1), 215-243. https://doi.org/10.1113/jphysiol.1968.sp008455

Iyer, R., Khwaja, A. I., Luttmer, E. F. P., & Shue, K. (2016). Screening Peers Softly: Inferring the Quality of Small Borrowers. *Management Science*, *62*(6), 1554-1577. https://doi.org/10.1287/mnsc.2015.2181

Jabeur, S. B., Mefteh-Wali, S., & Viviani, J.-L. (2021). Forecasting gold price with the XGBoost algorithm and SHAP interaction values. *Annals of Operations Research*, *334*(1-3), 679-699. https://doi.org/10.1007/s10479-021-04187-w

Jagtiani, J., & Lemieux, C. (2019). The roles of alternative data and machine learning in fintech lending: Evidence from the LendingClub consumer platform. *Financial Management*, *48*(4), 1009-1029. https://doi.org/10.1111/fima.12295

Jeong, H., Park, S., Woo, S., & Lee, S.-C. (2016). Rotating Machinery Diagnostics Using Deep Learning on Orbit Plot Images. *Procedia Manufacturing*, *5*(NA), 1107-1118. https://doi.org/10.1016/j.promfg.2016.08.083

Ji, S., Xu, W., Yang, M., & Yu, K. (2013). 3D Convolutional Neural Networks for Human Action Recognition. *IEEE transactions on pattern analysis and machine intelligence*, *35*(1), 221-231. https://doi.org/10.1109/tpami.2012.59

Khan, S., & Yairi, T. (2018). A review on the application of deep learning in system health management. *Mechanical Systems and Signal Processing*, *107*(NA), 241-265. https://doi.org/10.1016/j.ymssp.2017.11.024

Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer Credit-Risk Models Via Machine-Learning Algorithms. *Journal of Banking & Finance*, *34*(11), 2767-2787. https://doi.org/10.1016/j.jbankfin.2010.06.001

Kim, H., Park, S., Park, H.-J., Son, H.-G., & Kim, S. (2023). Solar Radiation Forecasting Based on the Hybrid CNN-CatBoost Model. *IEEE Access*, *11*(NA), 13492-13500. https://doi.org/10.1109/access.2023.3243252

Lu, H., & Ma, X. (2020). Hybrid decision tree-based machine learning models for short-term water quality prediction. *Chemosphere*, *249*(NA), 126169-126169. https://doi.org/10.1016/j.chemosphere.2020.126169

Miller, D., & Kim, J. M. (2021). Univariate and Multivariate Machine Learning Forecasting Models on the Price Returns of Cryptocurrencies. *Journal of Risk and Financial Management*, *14*(10), 486-NA. https://doi.org/10.3390/jrfm14100486

Moradzadeh, A., Mohammadi-Ivatloo, B., Abapour, M., Anvari-Moghaddam, A., & Roy, S. S. (2022). Heating and Cooling Loads Forecasting for Residential Buildings Based on Hybrid Machine Learning Applications: A Comprehensive Review and Comparative Analysis. *IEEE Access*, *10*(NA), 2196-2215. https://doi.org/10.1109/access.2021.3136091

Morshed, A. S. M., Manjur, K. A., Shahjalal, M., & Yahia, A. K. M. (2024). Optimizing Energy Efficiency: A Comprehensive Analysis Of Building Design Parameters. *Academic Journal on Science, Technology, Engineering & Mathematics Education*, *4*(04), 54-73. https://doi.org/10.69593/ajsteme.v4i04.120

Pietukhov, R., Ahtamad, M., Faraji-Niri, M., & El-Said, T. (2023). A hybrid forecasting model with logistic regression and neural networks for improving key performance indicators in supply chains. *Supply Chain Analytics*, *4*, 100041. https://doi.org/https://doi.org/10.1016/j.sca.2023.100041

Qi, Y., Shen, C., Wang, D., Shi, J., Jiang, X., & Zhu, Z. (2017). Stacked Sparse Autoencoder-Based Deep Network for Fault Diagnosis of Rotating Machinery. *IEEE Access*, *5*(NA), 15066-15079. https://doi.org/10.1109/access.2017.2728010

Rawat, W., & Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural computation*, *29*(9), 2352-2449. https://doi.org/10.1162/neco_a_00990

Ronao, C. A., & Cho, S.-B. (2016). Human activity recognition with smartphone sensors using deep learning neural networks. *Expert Systems with Applications*, *59*(59), 235-244. https://doi.org/10.1016/j.eswa.2016.04.032

Rozony, F. Z., Aktar, M. N. A., Ashrafuzzaman, M., & Islam, A. (2024). A Systematic Review Of Big Data Integration Challenges And Solutions For Heterogeneous Data Sources. *Academic Journal on Business Administration, Innovation & Sustainability*, *4*(04), 1-18. https://doi.org/10.69593/ajbais.v4i04.111

Seno, A. H., & Aliabadi, M. H. F. (2019). Impact Localisation in Composite Plates of Different Stiffness Impactors under Simulated Environmental and Operational Conditions. *Sensors (Basel, Switzerland)*, *19*(17), 3659-NA. https://doi.org/10.3390/s19173659

Shahjalal, M., Yahia, A. K. M., Morshed, A. S. M., & Tanha, N. I. (2024). Earthquake-Resistant Building Design: Innovations and Challenges. *Global Mainstream Journal of Innovation, Engineering & Emerging Technology*, *3*(04), 101-119. https://doi.org/10.62304/jieet.v3i04.209

Shamim, M. (2022). The Digital Leadership on Project Management in the Emerging Digital Era. *Global Mainstream Journal of Business, Economics, Development & Project Management*, *1*(1), 1-14.

Shelhamer, E., Long, J., & Darrell, T. (2016). Fully Convolutional Networks for Semantic Segmentation. *IEEE transactions on pattern analysis and machine intelligence*, *39*(4), 640-651. https://doi.org/10.1109/tpami.2016.2572683

Tabian, I., Fu, H., & Sharif Khodaei, Z. (2019). A Convolutional Neural Network for Impact Detection and Characterization of Complex Composite Structures. *Sensors*, *19*(22), 4933. https://www.mdpi.com/1424-8220/19/22/4933

Tang, H. (2019). Peer-to-Peer Lenders Versus Banks: Substitutes or Complements? *The Review of Financial Studies*, *32*(5), 1900-1938. https://doi.org/10.1093/rfs/hhy137

Tsai, C.-F., & Chen, M.-L. (2010). Credit rating by hybrid machine learning techniques. *Applied Soft Computing*, *10*(2), 374-380. https://doi.org/10.1016/j.asoc.2009.08.003

Tunç, T. (2012). A New Hybrid Method Logistic Regression and Feedforward Neural Network for Lung Cancer Data. *Mathematical Problems in Engineering*, *2012*(1), 0-0. https://doi.org/10.1155/2012/241690

Wang, Y., Sun, S., Chen, X., Zeng, X., Kong, Y., Chen, J., Guo, Y., & Wang, T. (2021). Short-term load forecasting of industrial customers based on SVMD and XGBoost. *International Journal of Electrical Power & Energy Systems*, *129*(NA), 106830-NA. https://doi.org/10.1016/j.ijepes.2021.106830

Xuan, Y., Si, W., Zhu, J., Sun, Z., Zhao, J., Xu, M., & Xu, S. (2021). Multi-Model Fusion Short-Term Load Forecasting Based on Random Forest Feature Selection and Hybrid Neural Network. *IEEE Access*, *9*(NA), 69002-69009. https://doi.org/10.1109/access.2021.3051337

Yahia, A. K. M., Rahman, D. M. M., Shahjalal, M., & Morshed, A. S. M. (2024). Sustainable Materials Selection in Building Design And Construction. *International Journal of Science and Engineering*, *1*(04), 106-119. https://doi.org/10.62304/ijse.v1i04.199

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*(50), 159-175. https://doi.org/10.1016/s0925-2312(01)00702-0

Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., & Gao, R. X. (2019). Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing*, *115*(NA), 213-237. https://doi.org/10.1016/j.ymssp.2018.05.050

Zhu, Y., Xie, C., Sun, B., Wang, G.-J., & Yan, X.-G. (2016). Predicting China's SME Credit Risk in Supply Chain Financing by Logistic Regression, Artificial Neural Network and Hybrid Models. *Sustainability*, *8*(5), 433-NA. https://doi.org/10.3390/su8050433