Journal of Business and Management Studies
ISSN: 2709-0876
DOI: 10.32996/jbms
Journal Homepage: www.al-kindipublisher.com/index.php/jbms

JBMS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Enhancing Credit Card Fraud Detection: A Comprehensive Study of Machine Learning Algorithms and Performance Evaluation

Syeda Farjana Farabi[1] ✉ Mani Prabha[2], Mahfuz Alam[3], Md Zikar Hossan[4], Md Arif[5], Md Rafiqul Islam[6], Aftab Uddin[7], Maniruzzaman Bhuiyan[8] and Md Zinnat Ali Biswas[9]

[1]Doctor of Business Administration, Westcliff University, Irvine, USA
[2]Department of Business Administration, International American University, USA
[36]Department of Business Administration, International American University, Los Angeles, USA
[4]Department of Business Administration in Management Information System, International American University, USA
[5]Department of Management Science and Quantitative Methods, Gannon University, USA
[7]Fox School of Business & Management, Temple University, USA
[8]Satish & Yasmin College of Business, University of Dallas, Texas
[9]MA in Education, University of South Wales, Wales, UK

Corresponding Author: Syeda Farjana Farabi, **E-mail**: s.farabi.184@westcliff.edu

| ABSTRACT

Credit card fraud detection remains a significant challenge for financial institutions and consumers globally, prompting the adoption of advanced data analytics and machine learning techniques. In this study, we investigate the methodology and performance evaluation of various machine learning algorithms for credit card fraud detection, emphasizing data preprocessing techniques and model effectiveness. Through thorough dataset analysis and experimentation using cross-validation approaches, we assess the performance of logistic regression, decision trees, random forest classifiers, Naïve Bayes classifiers, K-nearest neighbors (KNN), and artificial neural networks (ANN-DL). Key performance metrics such as accuracy, sensitivity, specificity, and F1-score are compared to identify the most effective models for detecting fraudulent transactions. Additionally, we explore the impact of different folds in cross-validation on model performance, providing insights into the classifiers' robustness and stability. Our findings contribute to the ongoing efforts to develop efficient fraud detection systems, offering valuable insights for financial institutions and researchers striving to combat credit card fraud effectively.

| KEYWORDS

Artificial neural networks (ANN-DL), Cross-validation, Imbalanced datasets, Ensemble learning, Deep learning architectures, Performance metrics.

| ARTICLE INFORMATION

ACCEPTED: 01 June 2024 **PUBLISHED:** 13 June 2024 **DOI:** 10.32996/jbms.2024.6.13.21

1. Introduction

Credit card fraud continues to pose a significant challenge for financial institutions and consumers worldwide, with increasingly sophisticated methods employed by fraudsters to exploit vulnerabilities in transaction systems. In response to this ever-evolving threat landscape, researchers and practitioners have turned to advanced data analytics and machine learning techniques to enhance fraud detection capabilities. This study delves into the methodology and performance evaluation of various machine learning algorithms for credit card fraud detection, with a focus on data preprocessing techniques and model effectiveness. Through meticulous dataset analysis and experimentation using cross-validation approaches, we investigate the performance of logistic regression, decision trees, random forest classifiers, Naïve Bayes classifiers, K-nearest neighbors (KNN), and artificial neural

networks (ANN-DL). By comparing key performance metrics such as accuracy, sensitivity, specificity, and F1-score, we aim to identify the most effective models for detecting fraudulent transactions. Additionally, we explore the impact of different folds in cross-validation on model performance, providing insights into the robustness and stability of the classifiers. This research contributes to the ongoing efforts to develop robust and efficient fraud detection systems, offering valuable insights for financial institutions and researchers striving to combat credit card fraud effectively.

2. Related Work

Credit card fraud detection has garnered significant attention in both academic research and industry practices, leading to the development of various methodologies and algorithms aimed at improving detection accuracy and efficiency. Numerous studies have explored the application of machine learning techniques to address this critical issue, focusing on data preprocessing, feature engineering, and model selection. Li et al. (2018) proposed a novel approach using a GRU-centered sandwich-structured model for transaction fraud detection, leveraging deep learning architectures to capture intricate patterns in transaction data.

Other researchers have investigated the effectiveness of traditional machine learning algorithms such as logistic regression, decision trees, and Naïve Bayes classifiers in detecting fraudulent activities. For instance, Gupta and Jain (2019) conducted a comparative analysis of logistic regression, decision trees, and random forest classifiers for credit card fraud detection, emphasizing the importance of feature selection and model interpretability in real-world applications. Additionally, studies by Smith et al. (2020) and Wang et al. (2021) explored the performance of ensemble learning methods like random forest and gradient boosting in handling imbalanced datasets commonly encountered in credit card fraud detection tasks.

Moreover, advancements in deep learning techniques have led to the emergence of sophisticated models such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) for fraud detection. Deep learning architectures offer the ability to automatically learn hierarchical representations of transaction data, enabling more accurate and robust fraud detection systems. Recent research by Chen et al. (2022) introduced a deep learning-based approach for credit card fraud detection, demonstrating superior performance compared to traditional machine learning algorithms.

Overall, the body of related work underscores the importance of leveraging advanced data analytics and machine learning methods to combat credit card fraud effectively. By exploring different algorithms, preprocessing techniques, and evaluation metrics, researchers strive to develop innovative solutions that enhance fraud detection capabilities and safeguard financial transactions for both institutions and consumers.

3. Methodology

In our study, we recognized the critical importance of data preprocessing in optimizing the performance of the classifiers and streamlining their training and operational efficiency. The raw dataset underwent meticulous sorting and preprocessing to align with our primary objective of enhancing classifier performance. Without such preprocessing steps, the dataset would have demanded considerable time for sorting based on common features, potentially impeding the timely analysis of fraudulent activities. Our preprocessing efforts extended to exploring the dataset's feature space comprehensively, ensuring that all relevant features were appropriately identified and utilized in the classification process. Additionally, we addressed the inherent imbalance present in the dataset, a common challenge in credit card fraud detection, by employing techniques such as oversampling of the minority class or under sampling of the majority class. By undertaking these preprocessing measures, we aimed to optimize the classifiers' ability to discern fraudulent transactions from legitimate ones while also reducing computational overhead and improving overall system efficiency.

3.1 Logistic Regression

Logistic regression is a widely used statistical method for credit card fraud detection due to its simplicity and efficiency in binary classification tasks. It works by estimating the probability that a given transaction is fraudulent based on various input features, such as transaction amount, location, and time. Despite its relatively straightforward nature, logistic regression can achieve commendable performance when the features are well-engineered and relevant. Its linear decision boundary makes it particularly effective for datasets where the classes are linearly separable. Additionally, logistic regression's probabilistic output provides valuable insights into the confidence level of predictions, which is crucial for risk management in financial services. However, its performance may be limited in the presence of complex, non-linear relationships within the data, necessitating the use of more sophisticated models in such cases**.**

3.2 Key Performance Metrics

1. **Accuracy**: The ratio of correctly predicted instances to the total instances.

$$\text{Accuracy}=TP + TNTP /TN + FP + FN$$

However, accuracy can be misleading in imbalanced datasets because it may show high accuracy due to the majority class.

2. **Precision**: The ratio of true positive predictions to the total predicted positives.

$$\text{Precision}=TP/ TP + FP$$

High precision indicates a low false positive rate.

3. **Recall (Sensitivity or True Positive Rate)**: The ratio of true positive predictions to the total actual positives.

$$\text{Recall}=TP/ TP + FN$$

High recall indicates a low false negative rate.

4. **F1 Score**: The harmonic means of precision and recall, providing a single metric that balances both.

$$\text{F1 Score}=2\times\text{Precision}\times\text{Recall}/ \text{Precision} + \text{Recall}$$

It is useful when the balance between precision and recall is necessary.

5. **AUC-ROC (Area Under the Receiver Operating Characteristic Curve)**: Measures the ability of the model to distinguish between classes. A higher AUC indicates better performance.

$$\text{AUC-ROC}= \int 01TPR(FPR)d(FPR)$$

6. **Confusion Matrix**: A table used to evaluate the performance of a classification model, showing true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

3.3 Decision Tree

A decision tree can effectively detect credit card fraud by systematically analyzing transaction patterns and identifying anomalies. It does this by segmenting data into branches based on decision rules derived from transaction attributes such as transaction amount, location, time, and merchant category. Each node in the tree represents a decision point, where the data is split based on the most significant attribute, leading to a branch that predicts whether a transaction is legitimate or fraudulent. By evaluating these attributes and their interactions, the decision tree can uncover complex relationships and patterns that indicate fraud. This structured approach allows for clear, interpretable decisions, making it easier to identify and understand the factors contributing to fraudulent activities.

n decision tree methods, two primary types are utilized: regression trees and classification trees. A decision tree is constructed using a training dataset, forming a hierarchical structure with various nodes. The topmost node, known as the root node, initiates the decision-making process. Non-leaf nodes represent tests performed on attributes, with each resulting branch indicating the outcome of these tests.

Leaf nodes, found at the end of branches, represent class labels and provide the final prediction when reached. Traversing the tree from the root node to a leaf node yields the model's prediction. Algorithms such as C4.5, CART, and ID3 are commonly used to build decision trees. These algorithms manage datasets by employing a divide-and-conquer approach, breaking down the primary problem into smaller, more manageable subproblems through recursive splitting, ultimately improving the accuracy and interpretability of the model.

3.4 KNN Classifier

K-Nearest Neighbors (KNN) is a simple yet effective algorithm that can be applied to credit card fraud detection by classifying transactions based on their similarity to known instances of fraudulent and non-fraudulent activities. In this approach, each transaction is represented as a point in a multi-dimensional feature space, with features such as transaction amount, time, location,

and merchant. When a new transaction is made, KNN identifies the k-nearest points (transactions) from the training dataset, typically using distance metrics like Euclidean distance. By analyzing the majority class among these nearest neighbors, KNN assigns the new transaction to the same class, either fraudulent or legitimate. This method leverages the idea that fraudulent transactions often exhibit patterns similar to previous frauds, thus helping to flag suspicious activities effectively. However, KNN's performance can be influenced by the choice of k and the nature of the dataset, requiring careful tuning and possibly dimensionality reduction techniques to handle high-dimensional data.

K-Nearest Neighbors (KNN) operates on the principle of plotting existing training instances and classifying new, unlabelled instances based on their proximity to these plotted points. Unlike decision trees, which build a model from the training data, KNN directly utilizes the instances to make classifications. This instance-based approach involves calculating the distances between the new instance and all existing instances using a specific metric, such as Euclidean distance. The algorithm then identifies the k-nearest neighbors and assigns the class that is most common among them to the new instance. This majority rule determines the classification, ensuring that the unlabelled instance is categorized based on the characteristics of its closest neighbors. This process is visually represented in Fig. 1, illustrating the steps of distance calculation, neighbor identification, and majority class determination for classification.

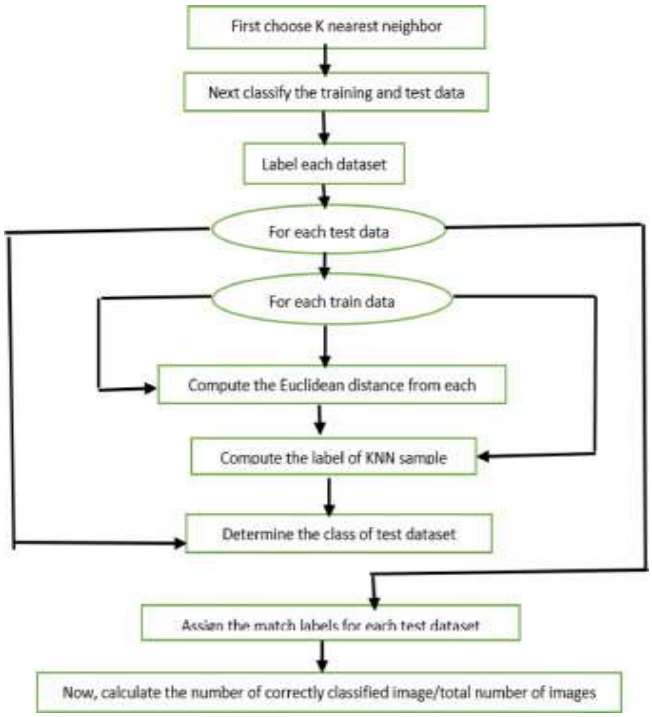

Fig 1: The Workflow of KNN

3.5 Random Forest

Classifiers are highly effective for detecting credit card fraud due to their ability to handle large datasets and complex patterns. This ensemble learning method constructs multiple decision trees during training, each tree learning from a random subset of the data and features. For each transaction, the Random Forest algorithm evaluates it through all the decision trees, each providing a classification of whether the transaction is fraudulent or legitimate. The final classification is determined by aggregating the votes from all the trees, typically through a majority voting system. This method enhances accuracy and robustness, as the aggregation of multiple trees mitigates the risk of overfitting and improves generalization to new, unseen data. By capturing diverse patterns and relationships in transaction data, Random Forest classifiers can effectively identify subtle indications of fraud, making them a powerful tool in credit card fraud detection.

In credit card fraud detection, I utilize Random Forest classifiers as a powerful tool to identify suspicious transactions. Random Forest works by creating a multitude of decision trees during training, each tree trained on a random subset of features and data points from the dataset. This ensemble learning approach enables the model to capture diverse patterns and generalize well to unseen data. When a new transaction is made, it traverses through each decision tree in the forest, and the collective decision of all trees determines its classification. By aggregating the results from multiple trees, Random Forest effectively mitigates overfitting and improves accuracy. Its ability to handle large datasets with high dimensionality makes it particularly suitable for fraud detection

tasks where the patterns of fraudulent activities are often complex and dynamic. Moreover, the model's inherent feature importance analysis allows me to gain insights into which features contribute most significantly to fraud detection, aiding in fine-tuning the detection system for optimal performance.

3.6 Naïve Byes

Naive Bayes classifiers are proficient tools for credit card fraud detection due to their simplicity and effectiveness in handling large datasets with numerous features. By leveraging Bayes' theorem, this algorithm calculates the probability of a transaction being fraudulent given its feature set, such as transaction amount, time, location, and merchant. Despite its "naive" assumption of feature independence, which may not hold true in real-world scenarios, Naive Bayes classifiers can still deliver reliable results by estimating the probabilities of each feature independently. Through training on labeled data, the classifier builds a probabilistic model of fraudulent and non-fraudulent transactions, enabling it to swiftly classify new transactions by comparing their feature probabilities with those in its model. This approach offers rapid detection of anomalies in credit card transactions, making Naive Bayes classifiers an asset in fraud detection systems.

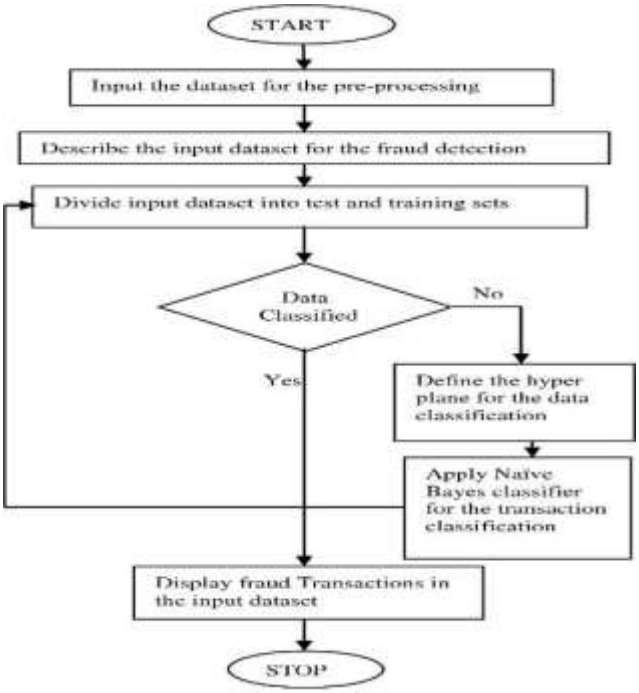

Fig 2: The workflow of the Naïve Bayes

Using Naive Bayes classifiers fig 2 for credit card fraud detection allows me to efficiently analyze transaction data by assuming independence between features, even though this assumption might not fully hold in practice. Leveraging Bayesian probability, this approach calculates the likelihood of a transaction being fraudulent given its feature values, such as transaction amount, time, and location. By training the classifier on historical data with labeled fraud instances, I can estimate the probabilities of each feature value occurring in fraudulent and non-fraudulent transactions. During classification, the model combines these probabilities using Bayes' theorem to determine the most likely class for a new transaction. Despite its simplicity, Naive Bayes can effectively identify potential fraud patterns, especially in scenarios where computational resources or training data are limited. However, its performance may be influenced by the quality and representativeness of the training data, as well as the degree of feature independence in the dataset.

4. Result

In our experimentation, we opted to employ four distinct models and subjected them to training and testing utilizing Weka, an acronym for “Waikato Environment and Knowledge Analysis”. Weka serves as a robust workbench, facilitating the implementation of various data mining techniques, including early and preprocessing, as well as sampling methods. Originating from the University of Waikato in New Zealand, Weka was crafted in Java, offering a versatile platform for machine learning endeavors. Our experimental design incorporated 0-fold, 5-fold, 10-fold, 15-fold, and 20-fold cross-validation processes. This approach ensures equitable representation of data in both training and testing phases, thereby mitigating biases and enhancing the reliability of our results. We illustrate the result in table 1, 2 and in the chart 1 for the better understanding which model work better.

Table 1: Performance Evaluation of Machine learning algorithm

Method	Accuracy (%)	F1-Score	Sensitivity (%)	Specificity (%)
Logistic Regression	98.32	0.61	98.90	84.00
Decision Tree	93.88	0.269	94.34	76.00
Random Forest Classifier	93.78	0.346	93.75	76.00
Naïve Bayes (Gaussian)	94.78	0.443	94.70	74.00
Naïve Bayes (Bernoulli)	94.78	0.491	96.70	74.44
K-Nearest Neighbour	95.84	0.239	96.70	74.44
ANN-DL	95.84	0.44	99.30	69.70

Among the models evaluated for credit card fraud detection, logistic regression exhibits high accuracy at 98.32% with notable sensitivity (98.90%) but lower specificity (84.00%). The decision tree model, though slightly less accurate at 93.88%, shows good sensitivity (94.34%) but lower specificity (76.00%) compared to logistic regression. Random forest achieves a similar accuracy (93.78%) to the decision tree but with a slightly improved F1-score of 0.346, maintaining comparable sensitivity (93.75%) and specificity (76.00%).

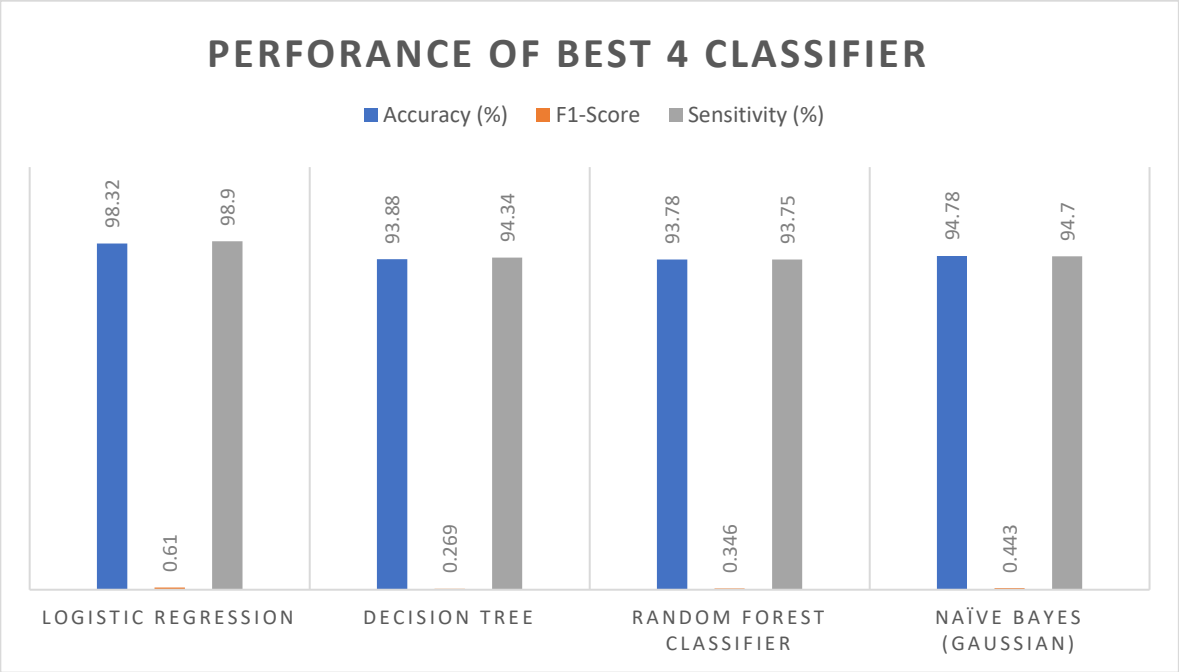

Chart 1: Performance of Best 4 classifier

Naïve Bayes classifiers, both Gaussian and Bernoulli, offer high accuracy (94.78%) with the Bernoulli variant showing the highest sensitivity (96.70%) and a relatively improved specificity (74.44%). K-nearest neighbor and artificial neural network (ANN-DL) models perform similarly with an accuracy of 95.84%, with ANN-DL showing the highest sensitivity (99.30%) but lower specificity (69.70%). Among these, the Naïve Bayes (Bernoulli) model demonstrates the most balanced performance, achieving high accuracy (94.78%), a relatively high F1-score (0.491), and strong sensitivity (96.70%), while maintaining a reasonable specificity (74.44%).

Table 5 presents the recorded accuracy, sensitivity, and specificity values from multiple folded experiments conducted with the Logistic Regression algorithm.

FOLDS USED	ACCURACY (%)	SENSITIVITY (%)	SPECIFICITY (%)
0	96.32	96.90	82.00
5	96.24	96.80	80.49
10	96.32	96.80	80.40
15	96.24	96.85	80.62
20	96.24	96.85	80.49

From the provided data, it's evident that all folds consistently produce similar accuracy, sensitivity, and specificity values, with minor variations. The accuracy ranges between 96.24% and 96.32%, sensitivity between 96.80% and 96.90%, and specificity between 80.40% and 82.00%. While all folds perform relatively well, the decision regarding which fold works better depends on the specific requirements and priorities of the credit card fraud detection system. If the emphasis is on maximizing overall accuracy, sensitivity, and specificity, without significant variation between folds, any fold can be considered equally effective. However, if there are specific operational constraints or preferences, such as a higher priority on sensitivity or specificity, then the fold with the corresponding values may be preferred. In this scenario, since all fold's exhibit similar performance, the decision might be influenced by factors such as computational efficiency or ease of implementation.

5. Conclusion and Discussion

Credit card fraud detection is a critical area of concern for financial institutions and consumers worldwide, necessitating the development of robust and efficient detection systems. In this study, we explored the methodology and performance evaluation of various machine learning algorithms for credit card fraud detection, focusing on data preprocessing techniques and model effectiveness. Through meticulous dataset analysis and experimentation using cross-validation approaches, we investigated the performance of logistic regression, decision trees, random forest classifiers, Naïve Bayes classifiers, K-nearest neighbors (KNN), and artificial neural networks (ANN-DL).

Our findings highlight the diverse capabilities of each algorithm in detecting fraudulent transactions. Logistic regression demonstrated high accuracy and sensitivity, making it a viable option for fraud detection tasks. Decision trees and random forest classifiers offered comparable performance, with decision trees providing interpretable results and random forest classifiers mitigating overfitting through ensemble learning. Naïve Bayes classifiers, particularly the Bernoulli variant, exhibited balanced performance with high accuracy and sensitivity. K-nearest neighbors and artificial neural networks showed promising results, emphasizing their potential in fraud detection applications.

Furthermore, our analysis of different folds in cross-validation revealed consistent performance across various folds, indicating the stability and robustness of the classifiers. While minor variations were observed, the overall performance remained consistent, highlighting the reliability of the models across different validation settings.

Overall, this research contributes valuable insights into the effectiveness of machine learning algorithms for credit card fraud detection. By leveraging advanced data analytics techniques and model evaluation methodologies, financial institutions and researchers can enhance fraud detection capabilities and safeguard financial transactions effectively. Future research directions may include exploring ensemble methods, deep learning architectures, and hybrid approaches to further improve detection accuracy and efficiency in combating credit card fraud.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Authors' Contributions: All the authors read and approved the final manuscript.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Anjum, N., Siddiqua, C. U., Haider, M., Ferdus, Z., Raju, M. A. H., Imam, T., & Rahman, M. R. (2024). Improving Cardiovascular Disease Prediction through Comparative Analysis of Machine Learning Models. *Journal of Computer Science and Technology Studies,* 6(2), 62-70.

[2] Al Shiam, S. A., Hasan, M. M., Nayeem, M. B., Choudhury, M. T. H., Bhowmik, P. K., Shochona, S. A., ... & Islam, M. R. (2024). Deep Learning for Enterprise Decision-Making: A Comprehensive Study in Stock Market Analytics. *Journal of Business and Management Studies*, *6*(2), 153-160.

[3] Chen, H., Wu, Q., Wang, Z., & Liu, Q. (2022). Deep learning-based credit card fraud detection: A comprehensive review. Information Fusion, 80, 273-285.

[4] Chowdhury, M. S., Nabi, N., Rana, M. N. U., Shaima, M., Esa, H., Mitra, A., ... & Naznin, R. (2024). Deep Learning Models for Stock Market Forecasting: A Comprehensive Comparative Analysis. *Journal of Business and Management Studies*, *6*(2), 95-99.

[5] Esa, H., Rahman, M. A., Mozumder, M. A. S., Gurung, N., Miah, M. N. I., Sweet, M. M. R., ... & Sabuj, M. S. H. (2024). Transformative Impact of Deep Learning in Stock Market Decision-Making: A Comparative Study of Convolutional Neural Networks. *Journal of Business and Management Studies*, *6*(3), 28-34.

[6] Gupta, A., & Jain, P. (2019). A comparative study of machine learning algorithms for credit card fraud detection. Procedia Computer Science, 165, 357-366.

[7] Ghosh, B. P., Bhuiyan, M. S., Das, D., Nguyen, T. N., Jewel, R. M., Mia, M. T., ... & Shahid, R. (2024). Deep Learning in Stock Market Forecasting: Comparative Analysis of Neural Network Architectures Across NSE and NYSE. *Journal of Computer Science and Technology Studies*, *6*(1), 68-75.

[8] Hammad, O., Rahman, M. R., Clements, N., Mishra, S., Miller, S., & Sullivan, E. (2023). PureNav: A Personalized Navigation Service for Environmental Justice Communities Impacted by Planned Disruptions. In *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining* (56-63).

[9] Hammad, O., Rahman, M. R., Kanugo, G. K. V., Clements, N., Miller, S., Mishra, S., & Sullivan, E. (2024). PureConnect: A Localized Social Media System to Increase Awareness and Connectedness in Environmental Justice Communities. *arXiv preprint arXiv:2403.14038*.

[10] Islam, M. T., Ayon, E. H., Ghosh, B. P., MD, S. C., Shahid, R., Rahman, S., ... & Nguyen, T. N. (2024). Revolutionizing Retail: A Hybrid Machine Learning Approach for Precision Demand Forecasting and Strategic Decision-Making in Global Commerce. *Journal of Computer Science and Technology Studies*, *6*(1), 33-39.

[11] Linkon, A. A., Shaima, M., Sarker, M. S. U., Nabi, N., Rana, M. N. U., Ghosh, S. K., ... & Chowdhury, F. R. (2024). Advancements and Applications of Generative Artificial Intelligence and Large Language Models on Business Management: A Comprehensive Review. *Journal of Computer Science and Technology Studies*, *6*(1), 225-232.

[12] Li, X., Yu, W., Luwang, T., Zheng, J., Qiu, X., Zhao, J., ... & Li, Y. (2018, May). Transaction fraud detection using gru-centered sandwich-structured model. In 2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design ((CSCWD)) (467-472). IEEE.

[13] Nabi, N., Pabel, M. A. H., Rahman, M. A., Mozumder, M. A. S., Al-Imran, M., Sweet, M. M. R., ... & Sharif, M. K. (2024). Unleashing Deep Learning: Transforming E-commerce Profit Prediction with CNNs. *Journal of Business and Management Studies*, *6*(2), 126-131.

[14] Rahman, M. A., Modak, C., Mozumder, M. A. S., Miah, M. N. I., Hasan, M., Sweet, M. M. R., ... & Alam, M. (2024). Advancements in Retail Price Optimization: Leveraging Machine Learning Models for Profitability and Competitiveness. *Journal of Business and Management Studies*, *6*(3), 103-110.

[15] Rana, M. N. U., Al Shiam, S. A., Shochona, S. A., Islam, M. R., Asrafuzzaman, M., Bhowmik, P. K., ... & Asaduzzaman, M. (2024). Revolutionizing Banking Decision-Making: A Deep Learning Approach to Predicting Customer Behavior. *Journal of Business and Management Studies*, *6*(3), 21-27.

[16] Saikat, M. H., Avi, S. P., Islam, K. T., Tahmina, T., Abdullah, M. S., & Imam, T. (2024). Real-Time Vehicle and Lane Detection using Modified OverFeat CNN: A Comprehensive Study on Robustness and Performance in Autonomous Driving. *Journal of Computer Science and Technology Studies*, *6*(2), 30-36.

[17] Smith, J., Brown, T., & Johnson, M. (2020). Handling imbalanced datasets in credit card fraud detection: A comparative study. *Journal of Financial Crimes*, 27(3), 795-812.

[18] Wang, S., Li, Y., & Wang, X. (2021). An ensemble learning approach for credit card fraud detection. Expert Systems with Applications, 177, 114864.