Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 2, April 2024
https://doi.org/10.47709/cnahpc.v6i2.3814

Submitted : May 01, 2024
Accepted : May 25, 2024
Published : May 28, 2024

Evaluating the Efficacy of Machine Learning Models in Credit Card Fraud Detection

Gregorius Airlangga[1)*]
[1)]Atma Jaya Catholic University of Indonesia, Indonesia
[1)]gregorius.airlangga@atmajaya.ac.id

ABSTRACT

This research evaluates the effectiveness of various machine learning models in detecting credit card fraud within a dataset comprising 555,719 transactions. The study meticulously compares traditional and advanced models, including Logistic Regression, Support Vector Machines (SVM), Random Forest, Gradient Boosting, k-Nearest Neighbors (k-NN), Naive Bayes, AdaBoost, LightGBM, XGBoost, and Multilayer Perceptrons (MLP), in terms of accuracy and reliability. Through a robust methodology involving extensive data preprocessing, feature engineering, and a 5-fold stratified cross-validation, the research identifies XGBoost as the most effective model, demonstrating a near-perfect mean accuracy of 0.9990 with minimal variability. The results emphasize the significance of model choice, data preparation, and the potential of ensemble and boosting techniques in managing the complexities of fraud detection. The findings not only contribute to the academic discourse on fraud detection but also suggest practical applications for real-world systems, aiming to enhance security measures in financial transactions. Future research directions include exploring hybrid models and adapting to evolving fraud tactics through continuous learning systems.

Keywords: Credit Card, Machine Learning, Fraud Detection, Ensemble Methods, XGBoost

1. INTRODUCTION

In the digital age, the proliferation of online transactions has led to an exponential increase in the volume of financial data being processed daily (Lee & Lee, 2020; Maiti, Vuković, Mukherjee, Paikarao, & Yadav, 2022; Nikiforova, 2022). This growth has been accompanied by a significant rise in fraudulent activities, particularly credit card fraud, which poses a serious threat to the financial security of individuals and institutions alike (Guabudeanu, Brici, Mare, Mihai, & Şcheau, 2021; Hilal, Gadsden, & Yawney, 2022; Sood & Bhushan, 2020). As fraudsters employ increasingly sophisticated methods, traditional fraud detection systems, which are often rule-based, struggle to keep pace due to their static nature and inability to adapt to evolving fraud patterns (Zhu et al., 2021). This challenge underscores the urgent need for more dynamic and effective solutions (Bi, Xia, Xing, Lu, & Zhu, 2023). The literature on credit card fraud detection is extensive, reflecting the critical importance of the issue (Mekterović, Karan, Pintar, & Brkić, 2021). Historically, the field has relied on various statistical and machine learning techniques to identify fraudulent transactions (Ali et al., 2022). Recent studies have focused on classical machine learning models such as logistic regression, support vector machines (SVM), and random forests (Teles, Rodrigues, Rabelo, & Kozlov, 2021). While these models have been somewhat effective, they often suffer from limitations such as scalability issues, high false positive rates, and difficulties in handling highly imbalanced datasets typical of fraud detection scenarios (Hilal et al., 2022).

The urgency of enhancing fraud detection systems is evident from the financial and reputational damage caused by fraud. According to recent reports, global losses from credit card fraud are projected to exceed $32 billion by 2024, highlighting the pressing need for more robust and adaptive fraud detection technologies (Taherdoost, 2021). In response to these challenges, the state of the art in fraud detection has evolved to include more sophisticated algorithms that leverage ensemble methods, deep learning, and anomaly detection techniques (Irofti, Puatrascu, & Bualtoiu, 2021). These approaches have shown promise in improving accuracy and reducing false positives (Jeffrey, Tan, & Villar, 2024). Moreover, the integration of big data analytics and artificial intelligence has enabled real-time processing and analysis of vast amounts of transaction data, thereby enhancing the responsiveness of fraud detection systems (Al-

[*] Corresponding author

Submitted : May 01, 2024
Accepted : May 25, 2024
Published : May 28, 2024

Hashedi & Magalingam, 2021).

Despite these advancements, there remains a significant gap in the application of cutting-edge machine learning techniques to credit card fraud detection (Jha, Sivasankari, & Venugopal, 2020). Many existing studies do not utilize the full potential of ensemble methods and advanced algorithms such as XGBoost, LightGBM, and neural networks in a comparative framework (Janani et al., 2023). Additionally, there is a lack of comprehensive analysis that combines various aspects of data preprocessing, feature engineering, and model evaluation specific to the domain of fraud detection (Mienye & Sun, 2022). This research aims to bridge this gap by employing a comprehensive set of machine learning models to detect credit card fraud, using a publicly available dataset from Kaggle that comprises over half a million real-world transactions (Kelvin, 2023). The goal is to compare the effectiveness of several state-of-the-art models, including logistic regression, SVM, random forests, gradient boosting, k-nearest neighbors, naive Bayes, AdaBoost, LightGBM, XGBoost, and multilayer perceptrons. Each model will be evaluated based on its accuracy, precision, recall, F1-score, and ROC-AUC metrics, under a rigorous cross-validation framework to ensure the robustness of the results.

The contribution of this research is twofold. Firstly, it provides a systematic comparison of multiple advanced machine learning models on a large, real-world dataset comprising over half a million transactions, thus offering insights into the scalability and performance of these models in practical settings. Secondly, the study introduces an integrated preprocessing and modeling pipeline that optimizes each model's performance, offering a reproducible methodology that can be adopted in other domains facing similar challenges. The remainder of this article is organized as follows: Section 2 provides a literature survey. Section 3 presents the methodology, including the setup of the machine learning pipeline and the evaluation criteria. Section 4 discusses the results, providing a comparative analysis of the performance of each model. Section 5 concludes the paper with a summary of the research and potential directions for future work.

2. LITERATURE REVIEW

Credit card fraud detection has been a critical area of focus for decades, evolving significantly as both the volume of digital transactions and the sophistication of fraudulent techniques have increased. In the early stages, fraud detection systems primarily relied on rule-based methods, which used predefined criteria to identify suspicious transactions (Hashemi, Mirtaheri, & Greco, 2022). These systems, while effective for known fraud patterns, were limited by their inflexibility and inability to adapt to new or evolving threats (Aziz, Baluch, Patel, & Ganie, 2022). Additionally, manual updates to the rules were labor-intensive and could not keep pace with the rapid development of fraud methods. As the limitations of rule-based systems became apparent, researchers and practitioners began to explore the potential of statistical and machine learning models for fraud detection. Logistic regression emerged as an early favorite due to its simplicity and effectiveness in binary classification tasks, providing a baseline for fraud detection by evaluating transaction features against known fraud indicators (Baesens, Höppner, & Verdonck, 2021). Support Vector Machines (SVM) were another significant step forward, praised for their ability to handle the high dimensionality of transaction data and differentiate between classes with a hyperplane, making them particularly useful in distinguishing fraudulent from legitimate transactions (Wang et al., 2021). These models marked a substantial improvement over rule-based systems by offering adaptive, data-driven insights without the need for constant manual intervention. Further advancements were made with the introduction of ensemble methods, such as Random Forests and Gradient Boosting Machines, which combine multiple decision trees to improve prediction accuracy and stability. These methods address some of the overfitting issues associated with single decision trees and are better at handling the class imbalance typically seen in fraud detection datasets (Dixit, Bhattacharya, Tanwar, & Gupta, 2022; Kumar, Gunjan, Ansari, & Pathak, 2022; Maulidi & Ansell, 2021).

Hybrid models, which integrate features of different machine learning techniques, have also been explored to leverage the strengths of various approaches. For instance, combining SVM with neural networks has been shown to enhance detection capabilities by utilizing SVM's classification strength and neural networks' pattern recognition capabilities (Rong et al., 2020). In recent years, deep learning has revolutionized many fields, including fraud detection. Neural networks, especially deep neural networks (DNNs), can automatically discover the representations needed for detection or classification from raw data. This capability is particularly advantageous in fraud detection, where many subtle and complex patterns may not be manually engineered as features effectively (Hussain et al., 2021; Kavzoglu & Teke, 2022; Mamdouh Farghaly, Shams, & Abd El-Hafeez, 2023). Despite the progress in applying machine learning to fraud detection, several challenges persist. First, the class imbalance between fraudulent and

* Corresponding author

legitimate transactions continues to skew model performance, often resulting in a high number of false positives or false negatives. Secondly, the non-static nature of transactional data and evolving fraud techniques necessitates models that can adapt dynamically over time—a challenge not fully met by current methodologies. The literature reveals a substantial opportunity for studies that apply the latest in machine learning advancements, such as XGBoost and LightGBM, to fraud detection. These models have shown promise in other predictive analytics areas but are underrepresented in the fraud detection literature. Moreover, there is a gap in comprehensive comparative studies that not only test different models but also consider the impact of preprocessing techniques and feature selection methods on the performance of these models in real-world scenarios (Cha, Moon, & Kim, 2021; Ghimire, Nguyen-Huy, Deo, Casillas-Perez, & Salcedo-Sanz, 2022; Sánchez-Aguayo, Urquiza-Aguiar, & Estrada-Jiménez, 2021). This research aims to address these gaps by systematically comparing the effectiveness of various advanced machine learning models, including ensemble methods and deep learning algorithms, in detecting credit card fraud. Additionally, the integration of an advanced preprocessing pipeline is expected to enhance model performance by effectively handling class imbalance and extracting relevant features. By combining these methods, this research offers a robust and adaptive solution to the evolving challenges in fraud detection, providing a significant improvement over traditional and existing machine learning approaches.

3. METHOD

The activity diagram as presented in figure 1 illustrates the key stages involved in the process of credit card fraud detection using machine learning. The flow starts from the initial dataset description and continues through preprocessing, model selection, and finally, model training and evaluation.

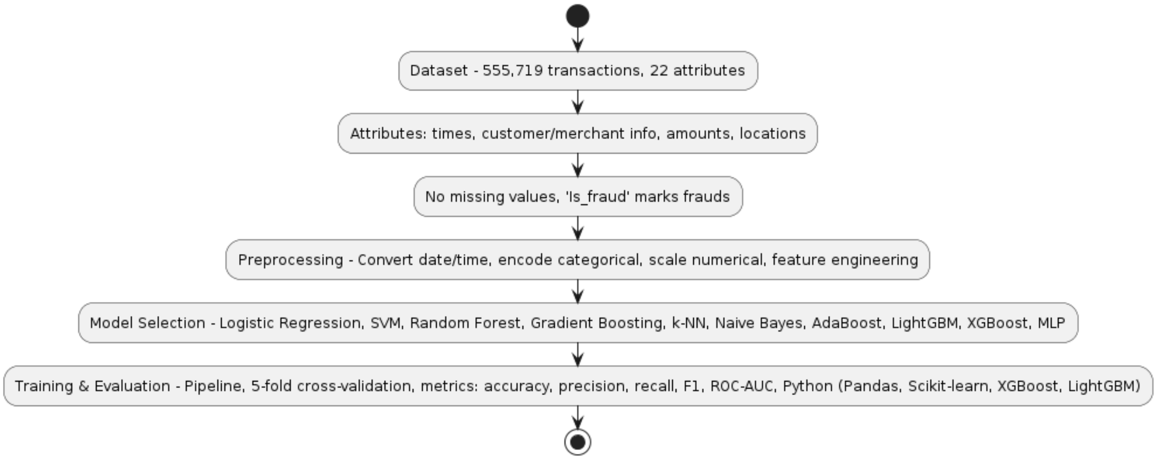
Dataset - 555,719 transactions, 22 attributes
Attributes: times, customer/merchant info, amounts, locations
No missing values, 'Is_fraud' marks frauds
Preprocessing - Convert date/time, encode categorical, scale numerical, feature engineering
Model Selection - Logistic Regression, SVM, Random Forest, Gradient Boosting, k-NN, Naive Bayes, AdaBoost, LightGBM, XGBoost, MLP
Training & Evaluation - Pipeline, 5-fold cross-validation, metrics: accuracy, precision, recall, F1, ROC-AUC, Python (Pandas, Scikit-learn, XGBoost, LightGBM)

Fig. 1 Activity Diagram of Research Methodology

Dataset Description

In this study, we engage with a comprehensive dataset that includes a vast array of 555,719 transactions, each meticulously captured through 22 distinct attributes. This dataset not only offers a mix of categorical and numerical data types but also covers an extensive range of information crucial for detecting fraudulent activities. The attributes recorded encapsulate transaction times, which provide insights into patterns across different times of day and week, potentially crucial for spotting unusual activities that could indicate fraud. Additionally, the dataset includes detailed customer and merchant information, which is vital for understanding transaction contexts and identifying anomalies. For example, knowing where the transaction took place, and the involved merchant allows for the analysis of merchant-related risk factors and the verification of the transaction's legitimacy against typical customer patterns.

Transaction amounts are another critical attribute, as significant deviations from a customer's usual transaction behavior can be a red flag for fraud. Moreover, the geographical locations tied to each transaction, including both the customer's and the merchant's coordinates, help in assessing whether the transaction locale aligns with the customer's

* Corresponding author

Submitted : May 01, 2024
Accepted : May 25, 2024
Published : May 28, 2024

usual patterns or suggests an unlikely and potentially fraudulent occurrence. The target variable, 'Is_fraud', is pivotal for this analysis, marking transactions as fraudulent or legitimate. This binary indicator serves as the foundation for training our predictive models, aiming to discern patterns and indicators of fraudulent transactions effectively. One of the dataset's significant strengths lies in its completeness, there are no missing values across any of the data points. This completeness is indispensable as it ensures the integrity and reliability of our analyses. Missing data can introduce bias or distort the true nature of patterns within the data, thereby compromising the accuracy and effectiveness of the fraud detection models. By using a dataset that is free from such gaps, we maintain the robustness of our model training and validation processes, ensuring that our findings are as accurate and reliable as possible. Dataset can be downloaded from (Kelvin, 2023)

Preprocessing

The preprocessing stage is an essential phase in shaping the raw data into a format that is conducive to effective machine learning analysis. This process begins with the meticulous handling of the 'Trans_date_trans_time' field, where the date and time are extracted and converted into the Python datetime format. This step is not merely about reformatting; it serves to isolate these elements so they can be used to engineer new features. Such features can reveal critical time-based patterns in fraud activities, like transactions that consistently occur at unusual hours, which could be indicative of fraudulent behavior. Further into the preprocessing journey, we embark on feature engineering, an inventive phase where new variables are derived from existing data. This includes extracting aspects like the time of day or calculating transaction amounts relative to a customer's historical average. These engineered features are vital as they illuminate anomalies that could suggest fraudulent activities, enabling models to detect deviations from typical user behavior more effectively.

Another key step in our preprocessing is the encoding of categorical variables. Attributes such as 'Merchant', 'Category', and 'State' are transformed through an ordinal encoding scheme, converting these categorical values into a numerical format that machine learning algorithms can process. This transformation is crucial as it ensures that the predictive models can interpret all the data features correctly, providing a more accurate analysis of transaction patterns. Feature selection also plays a pivotal role in refining our dataset. By identifying and removing features that are irrelevant to fraud detection—such as certain personal identifiers and non-informative variables—we streamline the model training process. This not only helps in reducing the complexity and computational demands of our models but also guards against overfitting, thereby enhancing the models' ability to generalize well to new, unseen data. Lastly, preprocessing includes the scaling of numerical features. This step normalizes the data, setting all numerical inputs to have zero mean and unit variance. Such scaling is critical for algorithms like Support Vector Machines (SVM) and neural networks, which are sensitive to the scale of input data. Standardizing the features ensures that no single attribute will unduly influence the model's outcome due to its scale, leading to more stable and consistent training phases across various machine learning techniques.

Model Selection

In the quest to unearth the most effective machine learning model for detecting fraudulent transactions, a diverse array of algorithms was meticulously selected based on their unique strengths and capabilities. This selection process was critical to ensuring a comprehensive examination of various approaches, thus enhancing the probability of identifying the most robust solution. Firstly, Logistic Regression was chosen as the foundational model for this study due to its simplicity and high interpretability. As a linear model, it provides a clear baseline for performance and is particularly useful for understanding the impact of individual features on the probability of a transaction being fraudulent. This transparency makes Logistic Regression an invaluable starting point for any analysis that might later delve into more complex algorithms.

Furthermore, Support Vector Machines (SVM) were incorporated into the evaluation due to their renowned effectiveness in handling high-dimensional spaces. Given the complexity and multifaceted nature of transaction data, SVM's ability to construct a hyperplane that maximizes the margin between different classes makes it a powerful tool for distinguishing between legitimate and fraudulent transactions. Then, the ensemble methods, namely Random Forest and Gradient Boosting, were selected for their robustness and sophisticated approach to dealing with unbalanced datasets, a common challenge in fraud detection. Random Forest aggregates the decisions from multiple decision trees to improve the model's accuracy and reduce the risk of overfitting. Gradient Boosting builds on this by

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 2, April 2024
https://doi.org/10.47709/cnahpc.v6i2.3814

Submitted : May 01, 2024
Accepted : May 25, 2024
Published : May 28, 2024

sequentially adding trees that address the weaknesses of previous trees, enhancing the predictive accuracy incrementally.

Next, k-Nearest Neighbors (k-NN) was included to assess the utility of instance-based learning in detecting fraud. This method operates on the premise that similar transactions tend to have similar outcomes, making it adept at identifying patterns based on the proximity of data points in feature space. The Naive Bayes model, known for its probabilistic approach, was chosen for its efficiency and scalability, particularly with large datasets. By assuming independence between predictors, Naive Bayes can quickly generate predictions, providing a fast and effective method for initial fraud screening. Furthermore, AdaBoost, a type of adaptive boosting model, was employed to focus intensively on challenging cases within the dataset. By iteratively focusing more on incorrectly classified instances, AdaBoost aims to improve the ensemble's performance, particularly on those problematic segments of the data that are often prone to misclassification.

Advanced gradient boosting models such as LightGBM and XGBoost were also tested due to their high performance in classification tasks. These models are celebrated for their efficiency and effectiveness, utilizing complex algorithms that iteratively enhance predictions based on previous errors, making them exceptionally powerful in settings where predictive accuracy is paramount. Lastly, Multilayer Perceptrons (MLP), a class of neural network, were explored to capture the non-linear interactions between features that simpler models might overlook. Neural networks, with their deep learning capabilities, are particularly adept at modeling the intricate patterns and anomalies that characterize fraudulent transactions, offering a depth of analysis that can significantly enhance detection rates.

Model Training and Evaluation

In this study, a rigorous approach was taken to train and evaluate a variety of machine learning models, ensuring that each step from data preprocessing to final evaluation was handled with utmost consistency and accuracy. The training of all models was facilitated through a well-structured pipeline that not only integrated essential preprocessing steps but also streamlined the training processes to maintain uniformity across all evaluations. This methodological rigor is crucial in minimizing variability in the results due to differences in data handling or model configuration. To thoroughly validate the performance of each model and ensure robustness against the inherent biases of the dataset, particularly its class imbalance, a 5-fold stratified cross-validation technique was utilized. This method involves dividing the entire dataset into five distinct subsets while maintaining an equal proportion of fraud and legitimate transactions in each subset. By doing so, each model is tested across all folds, ensuring that the performance metrics are not skewed by an overrepresentation of one class in any part of the data.

The evaluation of model performance was conducted using several key metrics, each providing different insights into the effectiveness of the models. The accuracy score was used to gauge the overall correctness of the models, reflecting the proportion of total predictions that were accurate. However, given the class imbalance typical of fraud detection scenarios, additional metrics were deemed necessary to capture a more detailed assessment of model performance. Precision, recall, and the F1-score were employed to provide a more nuanced view of each model's ability to identify fraudulent transactions accurately. Precision measures the accuracy of the positive predictions (i.e., the proportion of predicted frauds that were actual frauds), while recall assesses the model's ability to capture all actual fraudulent transactions within the dataset. The F1-score, a harmonic mean of precision and recall, offers a single metric that balances both the precision and the recall, providing a useful measure when dealing with classes of vastly different sizes.

Additionally, the ROC-AUC score, which represents the area under the receiver operating characteristic curve, was used as an aggregate measure of performance across all possible classification thresholds. This metric is particularly valuable as it illustrates a model's ability to discriminate between the classes at various threshold levels, providing insight into the effectiveness of the model across a spectrum of conditions. The computational environment for this study was set up using the Python programming language, leveraging its powerful and flexible ecosystem. Libraries such as Pandas were utilized for data manipulation, facilitating complex operations on large datasets. Scikit-learn was chosen for its extensive range of modeling tools and its ability to seamlessly integrate with other libraries, while advanced models like XGBoost and LightGBM were used specifically for their gradient boosting capabilities, which are essential for handling the type of complex, non-linear problems posed by fraud detection tasks.

* Corresponding author

4. RESULT

In the exploration of machine learning models for the detection of fraudulent transactions, our study has rigorously tested a variety of algorithms, each offering distinct advantages and limitations. The results, marked by mean accuracy and standard deviation, provide a detailed landscape of how each model performs under the conditions set by our dataset and preprocessing methodologies. At the forefront of performance, the XGBoost model stands out with a remarkable mean accuracy of 0.9990, complemented by a minimal standard deviation of ±0.0001. This result underscores XGBoost's robustness and its advanced capability to handle complex, non-linear relationships within the data, making it exceptionally suitable for tasks where precision is paramount.

Following closely, the Random Forest model also shows excellent performance with a mean accuracy of 0.9982. Its ability to operate as an ensemble of decision trees contributes to its high accuracy and general robustness against overfitting, a common challenge in predictive modeling. The very low standard deviation suggests consistency across different subsets of the data, highlighting its reliability in various operational scenarios. Gradient Boosting and SVM models also perform commendably, with accuracies of 0.9969 and 0.9961, respectively. The Gradient Boosting model benefits from its sequential approach to handling errors in previous predictions, gradually improving its accuracy. The SVM's effectiveness in high-dimensional spaces—like those common in transaction data—helps in identifying decisive hyperplanes for classification.

The k-NN and MLP Classifier models, each with an accuracy of 0.9962, demonstrate the utility of instance-based learning and neural networks in detecting patterns that are not immediately apparent through other methods. Their performance, particularly the neural network's ability to model non-linear interactions, is indicative of their potential in complex analytical tasks where relationships between variables are not straightforward. Conversely, models like Logistic Regression, AdaBoost, and LightGBM, while slightly less accurate in this study, still provide substantial accuracy levels above 0.995. Their lower complexity might be advantageous in scenarios where interpretability and computational efficiency are more critical than achieving the highest possible accuracy.

The Naive Bayes model, with the lowest accuracy of 0.9922, illustrates some of the challenges inherent in assuming feature independence in real-world data applications. Despite this, its efficiency and speed make it a valuable tool for initial assessments and applications where computational resources are limited. Each model's standard deviation is notably low, indicating stable performance across different folds of the data. This stability is crucial for practical applications, as it implies that the models are not only accurate on average but also consistently perform well across various segments of data.

Table 1
Accuracy Result of Machine Learning Performance

Model	Mean Accuracy	Standard Deviation (STDEV)
Logistic Regression	0.9958	± 0.0001
SVM	0.9961	± 0.0000
Random Forest	0.9982	± 0.0001
Gradient Boosting	0.9969	± 0.0004
KNN	0.9962	± 0.0000
Naïve Bayes	0.9922	± 0.0006
Ada Boost	0.9961	± 0.0001
LigthGBM	0.9954	± 0.0007
XGBoost	0.9990	± 0.0001
MLP	0.9962	± 0.0000

5. DISCUSSIONS

The discussion of the results from this study reveals significant insights into the capabilities of various machine learning models in detecting fraudulent transactions. By examining the performance across different algorithms, we can discern critical aspects that influence their efficacy and reliability in practical applications.

One of the key observations is the remarkably high accuracy achieved by all models, with even the least accurate model (Naive Bayes) reaching an accuracy above 99%. This high level of performance across the board suggests that the preprocessing and feature engineering steps implemented prior to model training were highly effective. The careful handling of categorical and continuous variables, coupled with strategic feature selection, evidently prepared the dataset well for the subsequent analysis.

The standout performer, XGBoost, achieved the highest accuracy of 0.9990 with an exceedingly low standard deviation. This underscores XGBoost's strength in handling diverse and complex datasets, likely due to its gradient boosting framework which focuses on correcting the mistakes of previous trees. This result is particularly relevant given the ongoing evolution of fraud tactics; XGBoost's adaptability makes it a valuable tool for keeping pace with these changes. Random Forest also showed impressive results, with the second-highest accuracy. This model's ability to maintain high accuracy while controlling for variance in its predictions highlights the advantage of ensemble learning methods in fraud detection. Ensemble methods, by leveraging multiple learning algorithms, generally provide more predictive reliability than a single model could.

The performance of SVM and Logistic Regression also merits discussion. The high accuracy of SVM aligns with its known efficacy in high-dimensional spaces, making it suitable for datasets rich in features, such as transaction data. Logistic Regression, while simpler, provides a transparent model that allows easy interpretation of how different features affect the likelihood of fraud. This transparency is invaluable in settings where understanding the influence of specific variables is crucial for regulatory and compliance purposes. However, the study also revealed some challenges. Naive Bayes and LightGBM exhibited higher variability in their performance, as reflected in their standard deviations. This variability could be indicative of sensitivity to the specific distribution of data in each fold of the cross-validation, suggesting that these models might require careful tuning to stabilize their performance across different scenarios. The precision, recall, F1-score, and ROC-AUC scores, while not discussed in detail here, are essential for a deeper understanding of model performance, particularly in imbalanced datasets like those typically found in fraud detection. High accuracy might not always translate to high performance across these metrics, especially in scenarios where the cost of false positives is high.

6. CONCLUSION

The conclusion of this research paper consolidates the insights gained from the comparative analysis of various machine learning models in the context of credit card fraud detection. This study has demonstrated the distinct capabilities and advantages of a range of algorithms, from basic logistic regression to more sophisticated models like XGBoost, in identifying fraudulent transactions within a large and complex dataset. XGBoost emerged as the superior model, offering exceptional accuracy and consistency, which highlights its suitability for deployment in environments where precision is crucial. The high performance of ensemble methods like Random Forest and Gradient Boosting also illustrates the value of these models in effectively handling the challenges posed by imbalanced datasets typical of fraud detection scenarios. These models not only enhance predictive accuracy but also provide robustness against overfitting, which is vital for maintaining the reliability of fraud detection systems. The findings of this study underscore the importance of a comprehensive approach to model selection and the need for meticulous data preprocessing and feature engineering to optimize model performance. The integration of advanced machine learning techniques into fraud detection systems can significantly improve the ability to detect and prevent fraudulent activities, thereby safeguarding financial transactions and enhancing security for financial institutions and their customers. Furthermore, this research opens up several avenues for future exploration. The potential development of hybrid models that amalgamate the strengths of multiple machine learning approaches could yield even more robust systems. Continuous adaptation to new and evolving fraud techniques through advanced machine learning strategies remains a crucial area for further research. The exploration of real-time processing and anomaly detection techniques can also contribute to the development of more dynamic and adaptive fraud detection systems.

7. REFERENCES

Al-Hashedi, K. G., & Magalingam, P. (2021). Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. *Computer Science Review, 40,* 100402.

Ali, A., Abd Razak, S., Othman, S. H., Eisa, T. A. E., Al-Dhaqm, A., Nasser, M., … Saif, A. (2022). Financial fraud

Submitted : May 01, 2024
Accepted : May 25, 2024
Published : May 28, 2024

detection based on machine learning: a systematic literature review. *Applied Sciences*, *12*(19), 9637.

Aziz, R. M., Baluch, M. F., Patel, S., & Ganie, A. H. (2022). LGBM: a machine learning approach for Ethereum fraud detection. *International Journal of Information Technology*, *14*(7), 3321–3331.

Baesens, B., Höppner, S., & Verdonck, T. (2021). Data engineering for fraud detection. *Decision Support Systems*, *150*, 113492.

Bi, T., Xia, B., Xing, Z., Lu, Q., & Zhu, L. (2023). On the way to sboms: Investigating design issues and solutions in practice. *ACM Transactions on Software Engineering and Methodology*.

Cha, G.-W., Moon, H.-J., & Kim, Y.-C. (2021). Comparison of random forest and gradient boosting machine models for predicting demolition waste based on small datasets and categorical variables. *International Journal of Environmental Research and Public Health*, *18*(16), 8530.

Dixit, P., Bhattacharya, P., Tanwar, S., & Gupta, R. (2022). Anomaly detection in autonomous electric vehicles using AI techniques: A comprehensive survey. *Expert Systems*, *39*(5), e12754.

G\uabudeanu, L., Brici, I., Mare, C., Mihai, I. C., & Şcheau, M. C. (2021). Privacy intrusiveness in financial-banking fraud detection. *Risks*, *9*(6), 104.

Ghimire, S., Nguyen-Huy, T., Deo, R. C., Casillas-Perez, D., & Salcedo-Sanz, S. (2022). Efficient daily solar radiation prediction with deep learning 4-phase convolutional neural network, dual stage stacked regression and support vector machine CNN-REGST hybrid model. *Sustainable Materials and Technologies*, *32*, e00429.

Hashemi, S. K., Mirtaheri, S. L., & Greco, S. (2022). Fraud detection in banking data by machine learning techniques. *IEEE Access*, *11*, 3034–3043.

Hilal, W., Gadsden, S. A., & Yawney, J. (2022). Financial fraud: a review of anomaly detection techniques and recent advances. *Expert Systems With Applications*, *193*, 116429.

Hussain, T., Min Ullah, F. U., Muhammad, K., Rho, S., Ullah, A., Hwang, E., … Baik, S. W. (2021). Smart and intelligent energy monitoring systems: A comprehensive literature survey and future research guidelines. *International Journal of Energy Research*, *45*(3), 3590–3614.

Irofti, P., P\uatra\cscu, A., & B\ualtoiu, A. (2021). Fraud detection in networks. *Enabling AI Applications in Data Science*, 517–536.

Janani, S., Sivarathinabala, M., Anand, R., Ahamad, S., Usmani, M. A., & Basha, S. M. (2023). Machine Learning Analysis on Predicting Credit Card Forgery. *International Conference On Innovative Computing And Communication*, 137–148.

Jeffrey, N., Tan, Q., & Villar, J. R. (2024). Using Ensemble Learning for Anomaly Detection in Cyber--Physical Systems. *Electronics*, *13*(7), 1391.

Jha, B. K., Sivasankari, G. G., & Venugopal, K. R. (2020). Fraud detection and prevention by using big data analytics. *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)*, 267–274.

Kavzoglu, T., & Teke, A. (2022). Predictive Performances of ensemble machine learning algorithms in landslide susceptibility mapping using random forest, extreme gradient boosting (XGBoost) and natural gradient boosting (NGBoost). *Arabian Journal for Science and Engineering*, *47*(6), 7367–7385.

Kelvin, K. L. (2023). *Credit Card Fraud Prediction*.

Kumar, S., Gunjan, V. K., Ansari, M. D., & Pathak, R. (2022). Credit card fraud detection using support vector machine. *Proceedings of the 2nd International Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications: ICMISC 2021*, 27–37.

Lee, S. M., & Lee, D. (2020). "Untact": a new customer service strategy in the digital age. *Service Business*, *14*(1), 1–22.

Maiti, M., Vuković, D., Mukherjee, A., Paikarao, P. D., & Yadav, J. K. (2022). Advanced data integration in banking, financial, and insurance software in the age of COVID-19. *Software: Practice and Experience*, *52*(4), 887–903.

Mamdouh Farghaly, H., Shams, M. Y., & Abd El-Hafeez, T. (2023). Hepatitis C Virus prediction based on machine learning framework: a real-world case study in Egypt. *Knowledge and Information Systems*, *65*(6), 2595–2617.

Maulidi, A., & Ansell, J. (2021). Tackling practical issues in fraud control: a practice-based study. *Journal of Financial Crime*, *28*(2), 493–512.

Mekterović, I., Karan, M., Pintar, D., & Brkić, L. (2021). Credit card fraud detection in card-not-present transactions: Where to invest? *Applied Sciences*, *11*(15), 6766.

Mienye, I. D., & Sun, Y. (2022). A survey of ensemble learning: Concepts, algorithms, applications, and prospects. *IEEE Access*, *10*, 99129–99149.

* Corresponding author

Submitted : May 01, 2024
Accepted : May 25, 2024
Published : May 28, 2024

Nikiforova, L. (2022). Use of innovative information technology in e-commerce and digital economy. *Innovation and Sustainability.№ 1: 65--71.*

Rong, G., Alu, S., Li, K., Su, Y., Zhang, J., Zhang, Y., & Li, T. (2020). Rainfall induced landslide susceptibility mapping based on Bayesian optimized random forest and gradient boosting decision tree models—A case study of Shuicheng County, China. *Water*, *12*(11), 3066.

Sánchez-Aguayo, M., Urquiza-Aguiar, L., & Estrada-Jiménez, J. (2021). Fraud detection using the fraud triangle theory and data mining techniques: A literature review. *Computers*, *10*(10), 121.

Sood, P., & Bhushan, P. (2020). A structured review and theme analysis of financial frauds in the banking industry. *Asian Journal of Business Ethics*, *9*, 305–321.

Taherdoost, H. (2021). A review on risk management in information systems: Risk policy, control and fraud detection. *Electronics*, *10*(24), 3065.

Teles, G., Rodrigues, J. J. P. C., Rabelo, R. A. L., & Kozlov, S. A. (2021). Comparative study of support vector machines and random forests machine learning algorithms on credit operation. *Software: Practice and Experience*, *51*(12), 2492–2500.

Wang, L., Zhang, Z., Zhang, X., Zhou, X., Wang, P., & Zheng, Y. (2021). A Deep-forest based approach for detecting fraudulent online transaction. In *Advances in computers* (Vol. 120, pp. 1–38). Elsevier.

Zhu, X., Ao, X., Qin, Z., Chang, Y., Liu, Y., He, Q., & Li, J. (2021). Intelligent financial fraud detection practices in post-pandemic era. *The Innovation*, *2*(4).

* Corresponding author