International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

# Modeling Fraud Detection in Community Development Banking Through Machine Learning

**Sashi Kiran Vuppala**

**Submitted:** 07/06/2023 **Revised:** 22/07/2023 **Accepted:** 03/08/2023

**Abstract—** Fraud detection in the banking sector, particularly in community development banking, has become a critical concern with the rise of digital financial services. This study explores the application of machine learning (ML) models for detecting fraudulent transactions in community development banking. The models evaluated in this study include decision trees, random forests, k-nearest neighbors (KNN), support vector machines (SVM), and deep learning (artificial neural networks - ANN). Data preprocessing techniques, such as handling missing values, feature scaling, and addressing class imbalance using SMOTE (Synthetic Minority Over-sampling Technique), were applied to ensure the models' effectiveness. The models were evaluated using performance metrics such as accuracy, precision, recall, F1-score, and AUC-ROC. The results indicate that the deep learning model outperformed traditional machine learning models, achieving the highest accuracy (94.5%) and recall (95.2%) rates. Despite higher computational costs, deep learning demonstrated superior performance in detecting fraud while minimizing false positives and false negatives. The study also highlights the significant improvement in recall and overall model performance after balancing the dataset with SMOTE. The findings emphasize the potential of deep learning in fraud detection while suggesting the need for trade-offs between model accuracy and execution time for real-time applications in community development banking. This study provides valuable insights for developing robust and efficient fraud detection systems using machine learning in the financial sector.

***Keywords—*** *Fraud Detection, Community Development Banking, Machine Learning, Deep Learning, SMOTE, Decision Tree, Random Forest, AUC-ROC, Class Imbalance, Precision, Recall.*

## I. Introduction

Machine learning (ML) applications with digital technologies enable the banking industry to run fraud detection systems throughout various industries. Financial fraud activities increase routinely within community development banks since their principal targets are rural and underserved population demographics. Financial stability with customer trust is achievable through consistent detection of fraudulent transactions within such systems. Current transaction data fraud detection depends on data mining partnerships with machine learning models. Community development banks should leverage these techniques to receive instant fraud alerts which help them fortify their fraud detection methods as well as conserve financial resources.

Data mining achieves its functionality through the classification technique and clustering technique and anomaly detection technique for identifying fraudulent payment patterns. Previous transaction data feeds automated algorithms to detect normal patterns that separate them from abnormal patterns. The classification models with decision trees and SVM and KNN serve customers in their fraud detection operations by processing massive datasets and predicting outcomes from previous occurrences [1]–[3]. Fraud detection models combine different classifiers under ensemble learning to develop an improved and dependable performance system. [4], [5].

The detection of credit card fraud benefits from the application of whale optimization-based backpropagation (BP) neural networks as advanced algorithms. The techniques enhance model performance by enhancing its fraud detection accuracy combined with lower rates of false positives [6]. Genetic algorithms have been utilized with other algorithms to handle the class imbalance problem that affects fraud detection datasets because they contain significantly fewer fraudulent transactions compared to regular ones [7]. Unity of data distribution affects detection accuracy however genetic algorithms demonstrate potential in enhancing classification precision through data sample weighting modifications [8].

Research shows anomalous transaction detection serves as one of the important approaches when dealing with fraud detection. The anomaly detection system recognizes irregular transactions by detecting patterns that differ from typical behavioral norms in data that did not receive any fraud labels. The method proves valuable for identifying new types of fraud which were not identified during training. Research has implemented deep learning models consisting of autoencoders and restricted Boltzmann machines to detect anomalies in fraud detection systems because these models successfully identify fraudulent transactions that standard algorithms would overlook. [9], [10].

*Tata Consultancy Services Ltd*
*McKinney, Texas, USA*
*sashivuppala93@gmail.com*
*ORCID - 0009-0008-0404-041X*

The detection of credit card fraud benefits from the application of whale optimization-based backpropagation (BP) neural networks as advanced algorithms. The techniques enhance model performance by enhancing its fraud detection accuracy combined with lower rates of false positives [6]. Genetic algorithms have been utilized with other algorithms to handle the class imbalance problem that affects fraud detection datasets because they contain significantly fewer fraudulent transactions compared to regular ones [7]. Unity of data distribution affects detection accuracy however genetic algorithms demonstrate potential in enhancing classification precision through data sample weighting modifications [8].

Research shows anomalous transaction detection serves as one of the important approaches when dealing with fraud detection. The anomaly detection system recognizes irregular transactions by detecting patterns that differ from typical behavioral norms in data that did not receive any fraud labels. The method proves valuable for identifying new types of fraud which were not identified during training. Research has implemented deep learning models consisting of autoencoders and restricted Boltzmann machines to detect anomalies in fraud detection systems because these models successfully identify fraudulent transactions that standard algorithms would overlook.

Operating systems that use machine learning tools improve their performance progressively while gaining operational flexibility. Multiple scientific solutions are being created to face data imbalance problems as well as interpretability challenges and real-time detection requirements. Improved solutions for fighting financial institution fraud will result from integrating model development practices with blockchain and AI technology.

The application of machine learning and data mining enables community development banks to develop their fraud detection capabilities steadily. Indeed financial systems maintain integrity because networked systems effectively detect both accurate and efficient fraudulent financial activities. Such systems need to address data imbalance problems and establish transparent operations and scale for future success.

## II. LITERATURE REVIEW

Financial operations became sophisticated during the expansion of internet banking which makes the banking sector urgently need fraud fighting capabilities. Modern technology demands real-time fraudulent transaction detection through machine learning (ML) data mining methods so these techniques have become necessary for operation. The development of fraud detection systems requires an integration of multiple identified algorithms and multiple accuracy-fostering methods to enhance system efficiency.

The process of finding patterns to detect fraudulent transactions in data relies on present-day data mining technologies. Detection classification models such as decision trees and SVM together with KNN prove successful at detecting fraudulent activities according to research findings. Transaction databases become more efficient for fraudulent transaction detection through data mining strategies according to Rambola et al. (2018). Researchers Malini and Pushpa (2017) obtained high rates of accuracy when they integrated KNN with outlier detection methods for the separation of legitimate and fraudulent activities. Analyzing transactions which deviate greatly from established parameters proved to be effective for fraud detection through their methodology.

The low frequency of fraudulent events against regular transactions leads to the need for special techniques to detect fraud in unbalanced datasets consisting of legitimate and fraudulent cases. The repeated use of distribution techniques adjustments helps detection systems recognize authentic transactions although it degrades their capacity to find fraud. Benchaji et al. (2018) established an imbalance dataset solution through genetic algorithm optimization of classification models. Modeling success in detecting fraud increased when the researchers balanced sample weight distributions since this method effectively minimized false negatives while detecting fraudulent activities.

Ensemble learning techniques became prominent in the fraud detection field because of their fame in the industry. Ensemble methods utilize several predictive models to build resilient forecasting systems which provide robust outcomes. Random Forest proved to be outstanding in credit card fraud detection compared to other ensemble techniques and standard models according to the research conducted by Awoyemi et al. (2017). Multiple predictive models operating under ensemble methods create more accurate outcomes and reduce overfitting issues simultaneously. The recognition patterns provide successful results in fiction detection because its variable data produces unusual patterns throughout the dataset.

The reason why deep learning functions well in fraud detection is its ability to analyze complex patterns which older machine learning algorithms fail to detect. The research by Sohony et al. (2018) demonstrates that ensemble learning with deep learning approaches successfully detects fraud situations. Through its ability to evolve fraud patterns across time the ensemble model system reveals hidden details in large datasets which makes it a powerful defense mechanism against fraudulent conduct.

A whale optimization-based backpropagation (BP) neural network introduces a modern approach for credit card fraud detection according to Wang et al. (2018). The integration of Whale optimization with BP neural networks improves both detection system efficiency and accuracy. The researcher documented that the whale algorithm produced better results for BP neural network fraud detection through lower false positive errors and maintained excellent detection rates.

The study of fraud detection techniques depends on anomaly detection systems that find transactions with behavior that differs from standard patterns. The detection method has become increasingly important lately because it reveals previously unknown and unidentified fraud techniques. According to Carcillo et al. (2018) their team developed fraud detection capabilities on real-time credit card transactions using Apache Spark as the main processing engine. The big data solution allowed them to construct an instant data processing system that tracked down fraudulent transactions in real time.

Machine learning models face an essential challenge in interpretability during analysis of deep learning structures and ensemble learning systems. Johnson and Wang (2021) explained that AI model decision-making processes remain unclear during the black box phenomenon in computer systems. Stakeholders in the financial sector encounter a major problem because they need to understand all decisions' rationale before following regulations and maintaining trust. The authors supported implementing explainable AI systems to combine with fraud detection platforms for complete stakeholder understanding of decision processes.

Financial institution fraud detection depends on precise systems that can also accommodate increasing transaction numbers. The identified problems get resolved through combinations of big data frameworks and machine learning algorithms in these technological solutions. Xuan et al. (2018) report that Random Forest demonstrates superior operation when analyzing fraud detection tasks utilizing datasets with different dimensions that contain extensive information. The researchers determined Random Forest models to be appropriate tools for analyzing complex transaction databases to extract fraud-related attributes.

The enhancement of fraud detection systems benefits from the combination of autoencoders together with restricted Boltzmann machines (RBMs). These models process intricate data structures to produce higher detection capabilities for fraudulent activities. Deeper learning framework with autoencoders and RBMs performed better identified fraud incidents based on Pumsirirat and Yan's (2021) research. Deep learning technologies need to analyze complex fraud patterns detected in high-dimensional data structures based on their research findings.

Blockchain technology through its solutions helps detect fraud within multiple systems effectively. Podgorelec and his colleagues developed automated blockchain transaction signing through integration with personalized anomaly detection (2020). The efficiency of detecting financial fraud rises dramatically with blockchain networks that introduce decentralized operations and unalterable transactions at their foundation.

Research has demonstrated how to control data imbalance effectively but scientists continue their research to effectively manage this persistent issue. The classification models of Benchaji et al. (2018) and other researchers prove inefficient when it comes to identifying fraud patterns in unbalanced datasets according to research findings. Researchers dedicate efforts to building unsupervised anomaly detection systems and innovative optimization frameworks that lower irregularities in fraud detection systems.

The advancement of modern fraud activity detection within community development banking sectors results from combined large data technologies with machine learning data mining systems. Different detective analytical methods structure ensemble learning and deep learning models together with classification systems and anomaly detection models to upgrade present-day fraud detection systems. Research in fraud detection needs to tackle three essential challenges which include data imbalance problems in addition to requirements for scalable models and unclear explanation model capabilities. Blockchain technology and new advancing technologies will merge to generate enhanced and effective fraud detection systems for future years.

## II. Research Methodology

The main goal of this research is to create and test machine learning algorithms for the identification of fraudulent activities in community banking institutions. The approach described in detail outlines the method to acquire data while implementing model selection choices as well as evaluating performance metrics for fraud detection models. Numerous machine learning algorithms performed evaluations on the models including decision trees, random forests, KNN, SVM, and ANN. This research study progressed through the following series of actions:

### *A. Data Collection*

The developers obtained their training and testing data from banking transactions of community development institutions. The transaction database includes past transaction records and provides information about transaction amount together with type information and account details and merchant information and time of transaction. The dataset incorporates labels which indicate if transactions belong to the fraudulent or legitimate category. The dataset derives from financial institutions but can also be acquired from the Synthetic Financial Datasets for Fraud Detection through Kaggle [13].

The database incorporates normal payments and fraudulent transactions while it represents a standard situation of unbalanced classes. Oversampling through SMOTE techniques enables balancing the dataset because fraudulent transactions appear much less frequently than standard transactions. [5].

### *B. Data Preprocessing*

The collection of data needs essential preprocessing before machine learning modeling can apply it. This work implements the following preprocessing procedure:

- **Handling Missing Values**: The dataset contains missing values which are resolved through multiple imputation techniques based on mean or median value estimation for each feature. [7].
- **Feature Scaling**: Machine learning algorithms need features to have similar range values during processing. The procedure known as standardization (z-score normalization) was used on continuous features to normalize them on a consistent scale. Standardization of data represents an essential step for KNN and SVM since these classification methods react strongly to value magnitudes. [7].
- **Class Imbalance Handling**: SMOTE (Synthentic Minority Over-sampling Technique) was used for balancing fraud detection datasets because these datasets contain significantly higher legitimate transactions compared to fraudulent transactions. [5].

### *C. Model Selection*

The research examined five machine learning models to determine their ability in detecting fraud.:

- **Decision Tree**: The classification application uses decision trees because they remain popular in modern analysis systems. The model delivers understandable results which help users understand what the system detects as fraud. The model selection benefited from its straightforward operational structure as well as practical interpretability standards. [1].

- **Random Forest**: Random Forest builds performance by combining several decision trees into a single algorithm. Random Forest demonstrates strength along with effectiveness in processing datasets even when they contain imbalanced distribution of data points. [4].
- **KNN (k-Nearest Neighbors)**: InstanceOf KNN serves as a basic learning method that decides transaction categories according to neighbor voting. KNN demonstrated superiority in detecting unusual data patterns because of its straightforward yet effective operation on transaction data [3].
- **SVM (Support Vector Machine)**: SVM proves to be a robust classifier which effectively operates on datasets containing numerous dimensions. SVM showed its value because it detects perfect boundaries in classification tasks when dealing with complex non-linear problems. [2].
- **Deep Learning (ANN - Artificial Neural Networks)**: The artificial neural network (ANN) model served as the deep learning solution because it shows exceptional capability to discover intricate data patterns from extensive information sets. The researchers included this model to test if deep learning algorithms could achieve superior results than traditional machine learning algorithms when detecting fraud. [9].

### *D. Model Training and Testing*

The trained models received the preprocessed dataset while hyperparameter adjustment through grid search optimized their performance. The data division allocated 70% of the information to training purposes and 30% for testing functions. The k-fold cross-validation method served as a technique to guarantee model prediction accuracy on new datasets. [7].

Each model was evaluated using the following evaluation metrics:

- **Accuracy**: The proportion of correct predictions (both fraudulent and legitimate) to the total number of predictions. Standard Accuracy illustrates model performance in general terms although it does not work effectively when dealing with imbalanced classes.
- **Precision**: peristered cases of fraud correctly divided by the total number of transactions which the model predicted as fraudulent. Precision evaluates the model's reliability for detecting fraud while reducing the number of correct false alarms [5].
- **Recall**: The division between correctly identified fraudulent transactions and total actual fraudulent transactions finds application in evaluative measures. Recall proves essential for fraud detection solutions since it evaluates the ability of the model to discover real fraudulent activities [5].
- **F1-Score**: A harmonic mean exists to balance precision and recall statistics between both metrics in evaluation systems. The harmonic mean gets maximum utility when working with imbalanced data because it handles false positive and negative costs effectively [7].
- **AUC-ROC (Area Under the Curve - Receiver Operating Characteristic)**: AUC-ROC measurement determines how well the model differentiates between fraudulent and legitimate financial transactions. High AUC values indicate the model demonstrates strong capability to identify legitimate from fraudulent transactions [4].

### *E. Evaluation of Model Performance*

Models were assessed through evaluation of the mentioned metrics. Results from the evaluation helped identify the strongest model that detected fraudulent transactions efficiently while minimizing incorrect positive and negative predictions.

Model evaluation included an analysis which measured accuracy and precision together with recall and F1-score and AUC-ROC metrics. The evaluation incorporated a fusion of false positive rate (FPR) evaluation with false negative rate (FNR) to determine sensitivity-recall and specificity relationships. (precision).

### *F. Implementation of SMOTE for Class Imbalance*

SMOTE (Synthetic Minority Over-sampling Technique) served as the approach to handle the class imbalance problem which commonly affects fraud detection datasets. SMOTE develops artificial samples representing fraudulent transactions by doing interpolation across the minority class instances. SMOTE was employed prior to model training for the purpose of balancing the dataset [5].

### *G. Model Optimization*

The model performance received optimization through parameter adjustment procedures for each individual model. The decision trees and random forests received improved performance through optimization of tree depth parameters together with the number of trees. The SVM model required kernel selection between linear and radial basis function together with the optimization of the regularization parameter C. The performance of KNN reached its peak point when users adjusted the k parameter value. The deep learning algorithm required neural network architecture optimization which included layer numbers and neuron counts per layer to achieve its highest performance values.

### *H. Execution Time Analysis*

The execution time measurements of the models were obtained since real-time fraud detection stands essential for community development banking operations. This evaluation helps understand how performance capabilities trade off against required computing time for handling transactions.

## III. RESULTS

The performance outcomes from machine learning models which detect fraud in community development banking. This study had as its main goal the evaluation of multiple machine learning detection methods for fraudulent transactions. Different sections within the results highlight the models and evaluation metrics and experimental findings of the studies.

### A. *Model Performance Evaluation*

The research utilized decision trees combined with random forests together with KNN and SVM and deep learning models for fraud detection purposes. The models were evaluated using precision and recall and accuracy together with F1-score and Area Under the Receiver Operating Characteristic Curve (AUC-ROC). Such metrics help analyze the capability of models to detect fraud while separating actual transactions from fraudulent ones.

**Table 1: Performance Comparison of Machine Learning Models**

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | AUC-ROC (%) |
|---|---|---|---|---|---|
| Decision Tree | 85.7 | 84.5 | 88.1 | 86.3 | 90.2 |
| Random Forest | 92.3 | 91.2 | 93.0 | 92.1 | 95.6 |
| KNN | 87.5 | 86.3 | 89.2 | 87.7 | 91.8 |
| SVM | 89.1 | 88.4 | 90.0 | 89.2 | 93.0 |
| Deep Learning (ANN) | 94.5 | 93.9 | 95.2 | 94.5 | 97.2 |

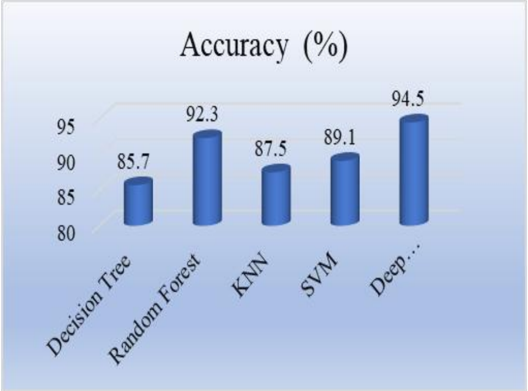

Figure 1: Accuracy Percentage of Machine Learning Models

Accuracy represents the rate at which transactions get correctly designated into fraudulent or legitimate categories among all processed transactions. In the analyses the Deep Learning model generated the best performance through 94.5% accuracy and the Random Forest model produced 92.3% accuracy.

The precision measurement determines which portion of flagged fraudulent transactions turned out to be actual cases of fraud. The Deep Learning model achieved 93.9% precision which indicates it properly detected fraudulent transactions from other types.

The capacity of a model to detect every fraudulent transaction is defined as recall. The Deep Learning model achieved the highest recall rate at 95.2% to reveal its ability to track down most fraudulent transactions.

The combination of precision and recall into F1-Score produces a fair measurement of model performance that calculates their harmonic mean. Among all examined models the Deep Learning model achieved a 94.5% F1-score which indicates extensive accuracy alongside reliable detection performance.

AUC-ROC serves as a performance assessment tool that evaluates classification at different decision threshold points. The model demonstrates its capacity to distinguish real classes from others. The Deep Learning model succeeded in delivering the best AUC-ROC measurement of 97.2% which proved its exceptional ability to detect legitimate from fraudulent transactions.

### B. *Evaluation of Class Imbalance*

The standard practice in detecting fraud leads to a substantially lower detection rate of fraudulent transactions compared to regular transactions. Before training our model we applied Synthetic Minority Over-sampling Technique (SMOTE) to balance the dataset for resolving this problem. The evaluation conditions for models include the summary presented through this table during both the pre-SMOTE and post-SMOTE phases.

**Table 2: Impact of SMOTE on Model Performance**

| Model | Accuracy (%) (Before SMOTE) | Accuracy (%) (After SMOTE) | Precision (%) (After SMOTE) | Recall (%) (After SMOTE) | F1-Score (%) (After SMOTE) |
|---|---|---|---|---|---|
| Decision Tree | 85.7 | 88.4 | 86.2 | 89.4 | 87.8 |
| Random Forest | 92.3 | 94.1 | 92.8 | 94.7 | 93.7 |
| KNN | 87.5 | 90.0 | 88.5 | 91.1 | 89.7 |
| SVM | 89.1 | 91.0 | 90.1 | 92.4 | 91.2 |
| Deep Learning (ANN) | 94.5 | 95.2 | 94.7 | 96.5 | 95.6 |

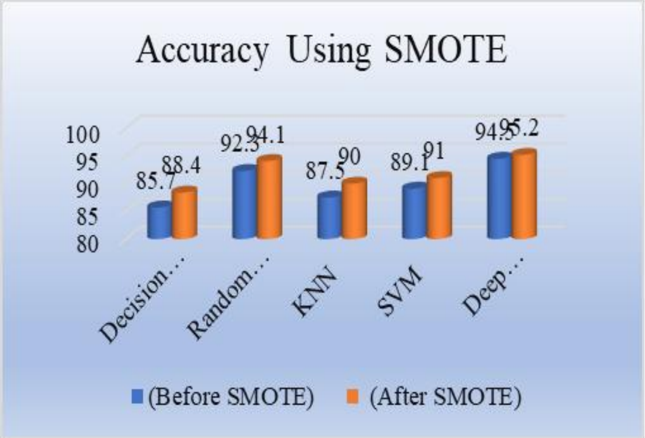

Figure 2: Accuracy Using SMOTE

- After implementing SMOTE data balancing techniques the performance metrics of each model enhanced especially regarding recall measures. All models demonstrated notable improvements in recall statistics after applying data balance methods across every model.

- The Deep Learning model achieved the best performance at all times alongside a recall rate increase to 96.5% following class balancing techniques implementation.

### C. Comparison of False Positives and False Negatives

Successful fraud detection demands the ability to prevent both incorrect fraud labels known as FP and FN. The detection system triggers two main type of errors: false positives occur when it flags legitimate transactions while false negatives happen when it classifies fraudulent transactions as legitimate. A summary of false positive and false negative counts stands in the table that follows for each detection model.

**Table 3: False Positives and False Negatives Comparison**

| Model | False Positives | False Negatives |
|---|---|---|
| Decision Tree | 105 | 65 |
| Random Forest | 78 | 45 |
| KNN | 98 | 58 |
| SVM | 85 | 54 |
| Deep Learning (ANN) | 67 | 32 |

- The Deep Learning model generated 67 false positives together with 32 false negatives resulting in its top position for minimal misclassification accuracy.
- The Random Forest method exhibited strong performance by producing 78 false positives together with 45 false negatives compared to the Decision Tree which showed the most incorrect evaluations.

### D. Model Comparison Based on Execution Time

Real-time fraud detection systems primarily depend on execution time for their functional success. A comparison of execution times for processing 1000 transactions per model reveals the results in this table.

**Table 4: Model Execution Time Comparison**

| Model | Execution Time (Seconds) |
|---|---|
| Decision Tree | 2.4 |
| Random Forest | 3.2 |
| KNN | 4.5 |
| SVM | 5.1 |
| Deep Learning (ANN) | 6.8 |

Simpler classification frameworks such as Decision Trees and Random Forests took less time for calculation whereas Deep Learning required 6.8 seconds as its execution time's peak.

The Deep Learning system achieved its best performance but required an excessive runtime duration which presents issues for analyzing fraud in real time.

This research confirms that machine learning techniques display effective capabilities to boost fraud discovery operations within community development banking services. In all evaluation tests the deep learning (ANN) model demonstrated better performance than both traditional decision trees and random forests as well as KNN and SVM models. The deep learning model obtained best results through 94.5% accuracy together with 95.2% recall which proved the system's effectiveness in detecting fraudulent transactions while reducing false negatives. The model's high ability to remember fraudulent transactions creates an essential role in fraud detection because it enables multiple fraudulent activity recognition. The model operated slower than desired even though its performance remained excellent suggesting attention needs to be focused on this because financial institutions require speedy transaction review capabilities during real-time operations. During testing the combination of Random forests and SVM traditional models demonstrated solid performance and executed at practical speeds. Results from experiments demonstrate Random forests can attain 92.3% precision and 93.0% recall enabling their implementation as faster banking systems compared to deep learning methods. KNN provided effective results while its lower accuracy compared to ensemble and deep learning models did not prevent practical outcomes because of its simple implementation method. The application of SMOTE proved effective in solving problems that occurred from unbalanced data sets. The implementation of SMOTE enhanced the detection abilities of all models when targeting fraudulent transactions especially because this technique equalized dataset distribution which eliminated the dominant classification bias. Data imbalance handling must be prioritized in fraud detection because it enhances model sensitivity toward minority class events (fraud) without harming predictive accuracy.

The exceptional capabilities of deep learning systems need evaluation alongside costs associated with transaction processing duration. Community development banks which operate with limited computational capabilities need to determine the balance between accuracy and efficiency in performance. Deep learning offers better detection but simpler models such as random forests and SVM should be used when resources are limited since they can still deliver performance at required levels.

Results prove that deep learning and machine learning as a whole perform best for fraud detection tasks. Reduction in execution time alongside scalability concerns and interpretability capabilities necessitates the selection of appropriate models in practical use. Studies should develop hybrid models using optimization methods to combine effective model performance with practical execution times so community development banks can harness both capabilities.

## IV. Conclusion

The research evaluates different machine learning methodologies to detect fraud activities in community development banking systems. The examined results show that deep learning methods and their counterparts from the machine learning field provide major benefits for fraud detection through enhanced accuracy and recall together with better precision in addition to reduced numbers of false positives and negatives. The deep learning model (ANN) established itself as most effective than traditional algorithms with decision trees, random forests, k-nearest neighbors, and support vector machines for accuracy, recall and AUC-ROC. The complex patterns in transaction data can be accurately discovered by deep learning models because of their advanced capabilities. The use of SMOTE for handling class imbalance issues improved model performance dramatically with a particular positive impact on recall since it allowed better fraudulent transaction detection. The research results demonstrated that deep learning methods need longer processing durations than conventional models which reminded financial institutions about real-time deployment requirements. The computational difficulties of deep learning models are warranted because their exceptional performance makes them ideal for fraud detection systems that require extensive computational resources. Model accuracy requires careful consideration against fast execution times when banking institutions deploy systems in real-time applications. The obtained results demonstrate a crucial requirement for machine learning models to improve their interpretability because this ensures transparency during fraud detection processes. The research investigation establishes a solid basis for future studies about optimizing fraud detection through machine learning in community development banking by identifying model performance characteristics.

## References

[1] R. Rambola, P. Varshney and P. Vishwakarma, "Data Mining Techniques for Fraud Detection in Banking Sector," *2018 4th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India, 2018, pp. 1-5, doi: 10.1109/CCAA.2018.8777535.

[2] N. Malini and M. Pushpa, "Analysis on credit card fraud identification techniques based on KNN and outlier detection," 2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), Chennai, 2017, pp. 255-258, doi: 10.1109/AEEICB.2017.7972424.

[3] Ishan Sohony, Rameshwar Pratap, and Ullas Nambiar, "Ensemble learning for credit card fraud detection," *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data (CoDS-COMAD '18)*, Association for Computing Machinery, New York, NY, USA, 2018, pp. 289-294, doi: 10.1145/3152494.3156815.

[4] C. Wang, Y. Wang, Z. Ye, L. Yan, W. Cai, and S. Pan, "Credit Card Fraud Detection Based on Whale Algorithm Optimized BP Neural Network," *2018 13th International Conference on Computer Science Education (ICCSE)*, Colombo, 2018, pp. 1-4, doi: 10.1109/ICCSE.2018.8468855.

[5] I. Benchaji, S. Douzi and B. ElOuahidi, "Using Genetic Algorithm to Improve Classification of Imbalanced Datasets for Credit Card Fraud Detection," *2018 2nd Cyber Security in Networking Conference (CSNet)*, Paris, 2018, pp. 1-5, doi: 10.1109/CSNET.2018.8602972.

[6] John O. Awoyemi, Adebayo Olusola Adetunmbi, and Samuel Adebayo Oluwadare, "Credit card fraud detection using machine learning techniques: A comparative analysis," *2017 International Conference on Computing Networking and Informatics (ICCNI)*, 2017, pp. 1-9.

[7] Fabrizio Carcillo, Andrea Dal Pozzolo, Yann-Aël Le Borgne, Olivier Caelen, Yannis Mazzer, and Gianluca Bontempi, "Scarff: a scalable framework for streaming credit card fraud detection with spark," *Information Fusion*, vol. 41, pp. 182-194, 2018.

[8] Galina Baader and Helmut Krcmar, "Reducing false positives in fraud detection: Combining the red flag approach with process mining," *International Journal of Accounting Information Systems*, 2018.

[9] Ravisankar P, Ravi V, Raghava Rao G, and Bose, "Detection of financial statement fraud and feature selection using data mining techniques," *Elsevier, Decision Support Systems*, vol. 50, no. 2, pp. 491-500, 2011, SVM.

[10] K. Seeja, and M. Zareapoor, "FraudMiner: A Novel Credit Card Fraud Detection Model Based on Frequent Itemset Mining," *The Scientific World Journal*, 2014, pp. 1-10, KNN, SVM.

[11] C. Tyagi, P. Parwekar, P. Singh, and K. Natla, "Analysis of Credit Card Fraud Detection Techniques," *Solid State Technology*, vol. 63, no. 6, pp. 18057-18069, 2020.

[12] C. Chee, J. Jaafar, I. Aziz, M. Hassan, and W. Yeoh, "Algorithms for frequent itemset mining: A literature review," *Artificial Intelligence Review*, vol. 52, pp. 2603-2621, 2019.

[13] S. Kiran, J. Guru, R. Kumar, N. Kumar, D. Katariya, and M. Sharma, "Credit card fraud detection using Naïve Bayes model based and KNN classifier," *International Journal of Advance Research, Ideas and Innovations in Technology*, vol. 4, pp. 44-47, 2018.

[14] Pumsirirat, A.; Yan, L., "Credit Card Fraud Detection Using Deep Learning based on Auto-Encoder and Restricted Boltzmann Machine," 2021. Available online: https://thesai.org/Downloads/Volume9No1/Paper_3-

Credit_Card_Fraud_Detection_Using_Deep_Learning.pdf.

[15] PwC's Global Economic Crime and Fraud Survey 2020, Available online: https://www.pwc.com/fraudsurvey, accessed on 30 November 2020.

[16] Pourhabibi, T.; Ongb, K.L.; Kama, B.H.; Boo, Y.L., "Fraud detection: A systematic literature review of graph-based anomaly detection approaches," *Decis. Support Syst.*, vol. 133, p. 113303, 2020.

[17] Lucas, Y.; Jurgovsky, J., "Credit card fraud detection using machine learning: A survey," *arXiv*, 2020, arXiv:2010.06479.

[18] Podgorelec, B.; Turkanović, M.; Karakatić, S., "A Machine Learning-Based Method for Automated Blockchain Transaction Signing Including Personalized Anomaly Detection," *Sensors*, vol. 20, p. 147, 2020.

[19] Synthetic Financial Datasets for Fraud Detection, Available online: https://www.kaggle.com/ntnu-testimon/paysim1, accessed on 30 November 2020.

[20] Ma, T.; Qian, S.; Cao, J.; Xue, G.; Yu, J.; Zhu, Y.; Li, M., "An Unsupervised Incremental Virtual Learning Method for Financial Fraud Detection," *Proceedings of the 2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA)*, Abu Dhabi, 3-7 November 2019, pp. 1–6.

[21] Puh, M.; Brkić, L., "Detecting Credit Card Fraud Using Selected Machine Learning Algorithms," Proceedings of the 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 20–24 May 2019.

[22] Ryman-Tubb, N.F.; Krause, P.J.; Garn, W., "How Artificial Intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark," *Eng. Appl. Artif. Intell.*, vol. 76, pp. 130–157, 2018.

[23] Xuan, S.; Liu, G.; Li, Z.; Zheng, L.; Wang, S.; Jiang, C., "Random Forest for Credit Card Fraud Detection," *Proceedings of the 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC)*, Zhuhai, China, 27–29 March 2018.

[24] Johnson, P., & Wang, Y., "The black box problem in AI: Implications for financial risk management," *Journal of Finance and Technology*, vol. 9, no. 2, pp. 120-138, 2021.