INTERNATIONAL JOURNAL ENGINEERING RESEARCH

International Journal for Innovative Engineering and Management Research

PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL

www.ijiemr.org

IJIEMR Transactions, online available on 31th Feb 2023. Link

https://ijiemr.org/downloads.php?vol=Volume-12&issue= Issue02

DOI:10.48047/IJIEMR/V12/ ISSUE02/95

Title: " DEEP LEARNING INNOVATIONS: ADVANCED NEURAL NETWORK ARCHITECTURES FOR EFFECTIVE FRAUD DETECTION IN INTERNET LOAN APPLICATIONS"

Volume 12, ISSUE 02, Pages: 652- 668

Paper Authors

Dr. Persis Urbana Ivy B, Venkatesh Maheshwaram, Sri Priya Nagula Malyala

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

DEEP LEARNING INNOVATIONS: ADVANCED NEURAL NETWORK ARCHITECTURES FOR EFFECTIVE FRAUD DETECTION IN INTERNET LOAN APPLICATIONS

Dr. Persis Urbana Ivy B, Venkatesh Maheshwaram, Sri Priya Nagula Malyala
Department of Computer Science Engineering, Sree Dattha Group of Institutions, Sheriguda, Hyderabad, Telangana

ABSTRACT

The proliferation of digital technology and online transactions has resulted in a surge in diverse fraud forms, particularly within the financial sector. Internet loans, although providing simple access to rapid financial aid, have also grown susceptible to fraudulent activity. Conventional fraud detection systems often depend on rule-based approaches and statistical models. Rule-based systems employ established criteria to identify transactions that correspond to particular patterns linked to fraud. Statistical methods, such logistic regression, examine previous transaction data to detect abnormalities. Although these algorithms have proven beneficial, they frequently encounter difficulties in identifying intricate, non-linear patterns typical of fraud in online loan applications. Consequently, it is imperative to address fraudulent actions with efficacy and efficiency. Identifying fraud in online loan applications is essential for financial organizations to uphold confidence, mitigate financial losses, and adhere to regulatory mandates. Deep learning, a branch of artificial intelligence (AI), has demonstrated significant potential in improving fraud detection due to its capacity to examine extensive datasets and recognize intricate patterns. These models employ advanced methodologies to analyze extensive datasets, facilitating the detection of nuanced and intricate fraud patterns that may elude conventional approaches. This research formulates a deep learning anti-fraud model for online loan applications, focusing on augmenting model accuracy via sophisticated neural network architectures, enhancing real-time processing capabilities, incorporating explainable AI techniques for improved transparency, and utilizing unsupervised learning methods to identify previously unrecognized fraud patterns. Furthermore, the future depends on the coordinated endeavors of data scientists, cybersecurity specialists, and financial institutions to outpace fraudsters and establish a safe digital lending landscape.

Keywords: Logistic Regression, Deep Learning, Fraud Detection, Neural Network

1. INTRODUCTION

1.1 Overview

The project aims to develop an advanced neural network architecture specifically tailored for detecting fraud in internet loan applications. With the increasing prevalence of online lending platforms, the risk of fraudulent activities has become a significant concern for financial institutions. Traditional methods of fraud detection often fall short in accurately identifying fraudulent applications due to the evolving nature of fraudulent tactics. Therefore, the proposed neural network architecture seeks to leverage the power of deep learning and advanced algorithms to enhance the accuracy and efficiency of fraud detection in this domain.

At its core, the neural network architecture will employ a combination of deep learning techniques and possibly attention mechanisms to effectively analyze various aspects of loan applications. These aspects may include applicant information, financial data, transaction history, and behavioral patterns. By processing large volumes of data, the neural network will learn intricate patterns indicative of fraudulent behavior, enabling it to differentiate between genuine and suspicious applications with high precision.

One of the key challenges in fraud detection is dealing with imbalanced datasets where legitimate loan applications significantly outnumber fraudulent ones. To address this issue, the neural network architecture will incorporate techniques such as oversampling, undersampling, or the use of specialized loss functions to ensure robust performance even in imbalanced scenarios. Additionally, the model will be designed to continuously adapt and learn from new data, allowing it to stay ahead of emerging fraud schemes.

Moreover, interpretability and explainability are crucial considerations in the context of fraud detection, especially in highly regulated industries like finance. Therefore, efforts will be made to incorporate transparency mechanisms into the neural network architecture, enabling stakeholders to understand how decisions are made and providing insights into the reasoning behind fraud predictions.

1.2 Problem Statement

The existing method for detecting fraud in internet loan applications faces several challenges and limitations. Primarily, the current system relies heavily on traditional data analysis techniques, which often struggle to keep pace with evolving fraud tactics in the digital age. While the system incorporates basic data validation checks such as verifying applicant information against known databases, it lacks robust mechanisms to detect sophisticated fraudulent patterns.

Moreover, the reliance on static rule-based algorithms makes the system prone to false positives and negatives, leading to inefficient resource allocation and missed fraudulent activities. The absence of real-time monitoring further exacerbates this issue, as fraudulent activities may go undetected until after the damage is done.

Additionally, the current method struggles to effectively analyze large volumes of data, resulting in delays in fraud detection and response times. This lag in processing can allow fraudulent actors to exploit vulnerabilities within the system, resulting in significant financial losses for both lenders and borrowers.

Furthermore, the system's inability to adapt to emerging fraud trends and techniques leaves it vulnerable to exploitation by increasingly sophisticated fraudsters. Without continuous updates and enhancements, the existing method fails to keep pace with the dynamic nature of fraudulent activities in the digital lending landscape.

1.3 Research Motivation

The motivation for pursuing an advanced neural network architecture for detecting fraud in internet loan applications stems from the urgent need to address the escalating threat posed by sophisticated fraudulent activities. Traditional methods of fraud detection are proving to be increasingly inadequate in the face of evolving tactics employed by fraudsters in the realm of internet-based loan applications.

- **Rising Incidence of Fraud:** The financial industry is experiencing a surge in internet loan application fraud, with perpetrators continually adapting their techniques to exploit vulnerabilities in existing detection systems. This necessitates a proactive approach to develop more robust and adaptive fraud detection mechanisms.
- **Financial Losses and Reputation Damage:** Fraudulent activities result in substantial financial losses for both financial institutions and individuals. Moreover, the aftermath of successful fraud can severely damage the reputation of lending institutions, eroding trust among clients. An advanced neural network architecture offers the potential to significantly mitigate these financial and reputational risks.
- **Complexity of Fraud Patterns:** Fraudulent activities in internet loan applications are becoming increasingly sophisticated, often involving intricate patterns and subtle anomalies that are challenging to detect using conventional methods. Advanced neural networks can leverage deep learning techniques to unravel complex patterns and identify subtle indicators of fraudulent behavior.
- **Data Volume and Dynamics:** The sheer volume of data generated by internet loan applications requires advanced techniques to extract meaningful insights. Conventional methods struggle to keep pace with the dynamic nature of fraud, necessitating the development of neural network architectures that can handle large datasets and adapt to evolving fraud patterns in real- time.
- **Regulatory Compliance:** Regulatory bodies are placing greater emphasis on robust fraud prevention measures within the financial sector. The implementation of advanced neural network architectures aligns with regulatory expectations, demonstrating a commitment to safeguarding financial transactions and ensuring compliance with evolving industry standards.
- **Advancements in Neural Networks:** Recent advancements in neural network architectures, including deep learning algorithms and architectures such as recurrent neural networks and convolutional neural networks, provide an opportunity to significantly enhance the accuracy and efficiency of fraud detection systems. Leveraging these advancements is crucial to staying ahead of sophisticated fraudsters.

2. LITERATURE SURVEY

Xu, et al. [1] Proposed experimental results showed that the fraud prediction model based on the GTWE algorithm achieved outstanding classification effect and stability with satisfactory interpretability. Meanwhile, the fraud probability of customers detected by the fraud prediction model was as high as 84.19%, indicating that App behaviors had a considerable impact on predicting fraud in online loan applications.

Mytnyk, et al. [2] Proposed model was based on an artificial neural network, effectively improved the accuracy of fraudulent transaction detection. The results of the different algorithms were visualized, and the logistic regression algorithm performed the best, with an output AUC value of approximately 0.946. The stacked generalization showed a better AUC of 0.954. The recognition of banking fraud using artificial intelligence algorithms was a topical issue in our digital society.

Lakshmi, et al. [3] Proposed, the survey of current strategies utilized in credit card fraud detection was depicted. This study employed Principal Component Analysis (PCA) to perform feature selection and speed up the learning process. The comparison outcomes demonstrated that Random Forest (RF) outperformed Decision Tree.

Sharma, et al. [4] Proposed work analyzed the performance of unsupervised learning techniques such as k-means clustering on a credit card fraud detection dataset. A Particle swarm optimization and k-means clustering hybrid model were proposed for the same, aiming to further research in this area. The model based on the proposed approach improved the performance of the k-means clustering approach. Our Hybrid approach showed better accuracy, precision, and recall than the k-means clustering approach.

Yedukondalu, et al. [5] Proposed research work made use of random forest and XGBOOST algorithms. A big public loan dataset, such as that from Lending Club, was used to detect fraud. A random forest was used to fill in the missing values initially. The most discriminating features were then chosen using the XGBoost algorithm. Such a basic and successful model could have improved the use of machine learning for detecting frauds in Internet loan.

Zhan, et al. [6] Proposed a new way to extract features automatically from a borrower's phone network graph using neural networks to detect fraudulent loans, which not only overcame the above issue but also captured features that were hard to fake. This method yielded strong results in reality.

Nwade, et al. [7] Proposed the evaluation of results, that was done by comparing its performance with the classifier using accuracy metrics. The model implementation was done using the Python programming language. The data was passed into MLP with an algorithm classifier and the results were obtained with an accuracy of 93% and 99% respectively.

Reddy, et al. [8] Proposed method solved the problem by first cleaning and normalizing the data, then using Kernel principal component analysis to extract features. Finally, it utilized these features to train a model with CNN-BiLS TM, a neural network architecture that combined the best parts of the Bidirectional Long Short-Term Memory (BiLS TM) network and the Convolution Neural Network (CNN).

Mizher, et al. [9] Proposed model was evaluated and compared when dealing with large amounts of data using a highly imbalanced real-world credit card fraud detection dataset. Python programming languages were used to preprocess the data and test the model's measurements and performance. As observed in the results, an accuracy of 99.7% using the Random Forest classifier was obtained.

Achary, et al. [10] Proposed algorithm was utilized to analyze the resampled dataset, minimizing the high class imbalance. Numerous intelligent algorithms were analyzed on a public dataset to determine the correlation of certain factors with fraudulence. Data was analyzed using the proposed algorithm for enhanced accuracy.

Bajracharya, et al. [11] Proposed some key potential directions to inspire intelligent solutions for defending and mitigating against cyberattacks. Analyzed the current scenario of cybersecurity risks and provided a comprehensive overview of the recent approaches in evolving cybersecurity and fraud detection practices at scale. Reviewed new challenges in effective cybersecurity measures and financial fraud detection.

Fanai, et al. [12] Proposed approach was found to improve the performance of the employed deep learning-based classifiers in the experimental evaluations. Specifically, the utilized deep learning classifiers trained on the transformed dataset by the deep Autoencoder significantly outperformed their baseline classifiers trained on the original data in terms of all performance measures.

Singh, et al. [13] Proposed, the chapter presented real-world examples to illustrate the effectiveness of these techniques in detecting fraud and reducing financial losses. Overall, the chapter provided a comprehensive overview of the application of classification and regression techniques in bank fraud detection, which was beneficial for researchers interested in the field of ML and fraud detection.

Zhang, et al. [14] Proposed comprehensive experiments was conducted that compare TBCCA with state-of-the-art fraudster detection approaches. Experimental results showed that TBCCA effectively identified fraudsters in real review networks, achieving a 6%–10% performance improvement over other baselines.

Aburbeian, et al. [15] Proposed the comprehensive summary of several peer-reviewed research papers on GAN synthetic generation techniques for fraud detection in the financial sector. Additionally, various solutions proposed by different researchers were included to balance imbalanced classes. The work concluded by pointing out the limitations of the most recent research articles.

3. PROPOSED METHOD

Overview

The proposed system for the project, "Advanced Neural Network Architecture for Detecting Fraud in Internet Loan Applications," aims to revolutionize the detection of fraudulent activities in online loan applications using cutting-edge neural network technology. This system integrates a multi-layered neural network architecture trained on vast datasets of historical loan application data, including both legitimate and fraudulent cases. By leveraging sophisticated algorithms and advanced machine learning techniques, the model can effectively identify patterns, anomalies, and subtle indicators associated with fraudulent behavior.

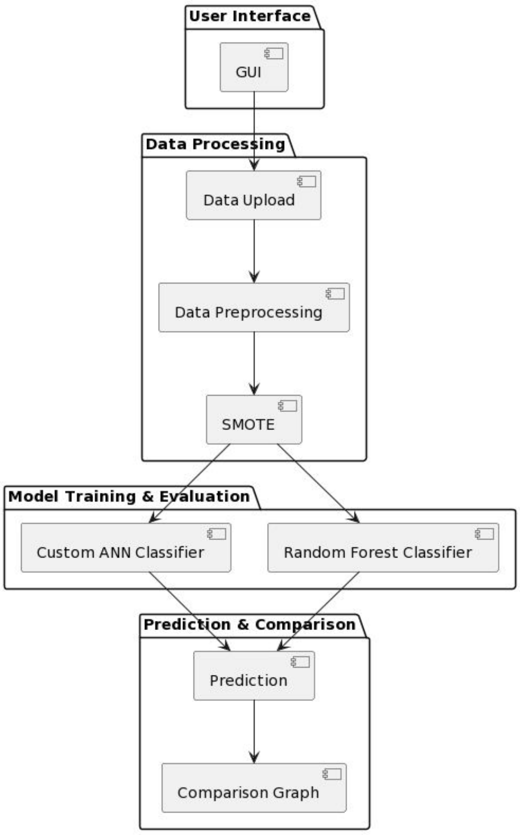


Fig 1: Block Diagram of Proposed System.

ANN Classifier

Although today the Perceptron is widely recognized as an algorithm, it was initially intended as an image recognition machine. It gets its name from performing the human-like function of perception, seeing, and recognizing images. Interest has been centered on the idea of a machine which would be capable of conceptualizing inputs impinging directly from the physical environment of light, sound, temperature, etc. — the "phenomenal world" with which we are all familiar — rather than requiring the intervention of a human agent to digest and code the necessary information. Rosenblatt's perceptron machine relied on a basic unit of computation, the neuron. Just like in previous models, each neuron has a cell that receives a series of pairs of inputs and weights. The major difference in Rosenblatt's model is that inputs are combined in a weighted sum and, if the weighted sum exceeds a predefined threshold, the neuron fires and produces an output.

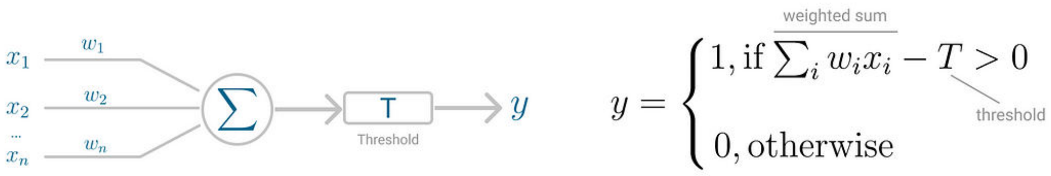

x_1
w_1
x_2
w_2
...
x_n
w_n
$\sum$
T
Threshold
y
weighted sum
$y = \begin{cases} 1, \text{if } \sum_i w_i x_i - T > 0 \\ 0, \text{otherwise} \end{cases}$
threshold

Fig. 2: Perceptron neuron model (left) and threshold logic (right).

Threshold T represents the activation function. If the weighted sum of the inputs is greater than zero the neuron outputs the value 1, otherwise the output value is zero.

Perceptron for Binary Classification

With this discrete output, controlled by the activation function, the perceptron can be used as a binary classification model, defining a linear decision boundary. It finds the separating hyperplane that minimizes the distance between misclassified points and the decision boundary.

The perceptron loss function is defined as below:

$$\underbrace{D(w,c)}_{\text{distance}} = -\sum_{\underset{\text{misclassified observations}}{i\in \mathrm{M}}} \overset{\text{output}}{y_i}\,(x_i w_i + c)$$

To minimize this distance, perceptron uses stochastic gradient descent (SGD) as the optimization function. If the data is linearly separable, it is guaranteed that SGD will converge in a finite number of steps. The last piece that Perceptron needs is the activation function, the function that determines if the neuron will fire or not. Initial Perceptron models used sigmoid function, and just by looking at its shape, it makes a lot of sense! The sigmoid function maps any real input to a value that is either 0 or 1 and encodes a non-linear function. The neuron can receive negative numbers as input, and it will still be able to produce an output that is either 0 or 1.

But, if you look at Deep Learning papers and algorithms from the last decade, you'll see the most of them use the Rectified Linear Unit (ReLU) as the neuron's activation function. The reason why ReLU became more adopted is that it allows better optimization using SGD, more efficient computation and is scale-invariant, meaning, its characteristics are not affected by the scale of the input. The neuron receives inputs and picks an initial set of weights random. These are combined in weighted sum and then ReLU, the activation function, determines the value of the output.

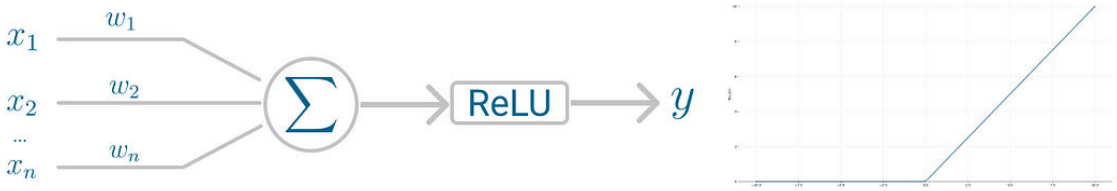


Fig. 3: Perceptron neuron model (left) and activation function (right).

Perceptron uses SGD to find, or you might say learn, the set of weight that minimizes the distance between the misclassified points and the decision boundary. Once SGD converges, the dataset is separated into two regions by a linear hyperplane. Although it was said the Perceptron could represent any circuit and logic, the biggest criticism was that it couldn't represent the XOR gate, exclusive OR, where the gate only returns 1 if the inputs are different. This was proved almost a decade later and highlights the fact that Perceptron, with only one neuron, can't be applied to non-linear data.

ANN

The ANN was developed to tackle this limitation. It is a neural network where the mapping between inputs and output is non-linear. A ANN has input and output layers, and one or more hidden layers with many neurons stacked together. And while in the Perceptron the neuron must have an activation function that imposes a threshold, like ReLU or sigmoid, neurons in a ANN can use any arbitrary activation function. ANN falls under the category of feedforward algorithms because inputs are combined with the initial weights in a weighted sum and subjected to the activation function, just like in the Perceptron. But the difference is that each linear combination is propagated to the next layer. Each layer is feeding the next one with the result of their computation, their internal representation of the data. This goes all the way through the hidden layers to the output layer. If the algorithm only computed the weighted sums in each neuron, propagated results to the output layer, and stopped there, it wouldn't be able to learn the weights that minimize the cost function. If the algorithm only computed one iteration, there would be no actual learning. This is where Backpropagation comes into play.

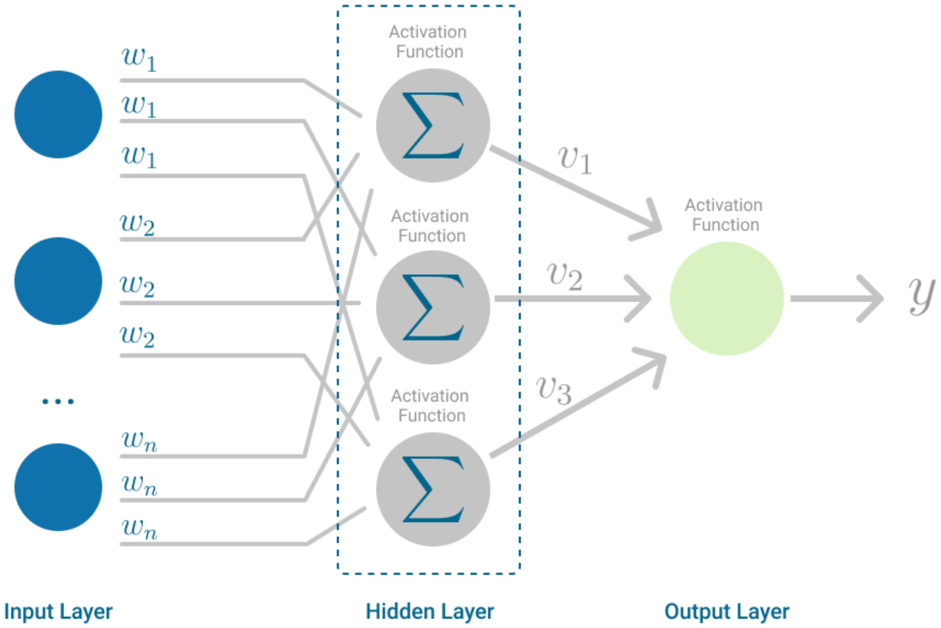


Fig. 4: Architecture of ANN.

Backpropagation: Backpropagation is the learning mechanism that allows the ANN to iteratively adjust the weights in the network, with the goal of minimizing the cost function. There is one hard requirement for backpropagation to work properly. The function that combines inputs and weights in a neuron, for instance the weighted sum, and the threshold function, for instance ReLU, must be differentiable. These functions must have a bounded derivative because Gradient Descent is typically the optimization function used in ANN. In each iteration, after the weighted sums are forwarded through all layers, the gradient of the Mean Squared Error is computed across all input and output pairs. Then, to propagate it back, the weights of the first hidden layer are updated with the value of the gradient. That's how the weights are propagated back to the starting point of the neural network. One iteration of Gradient Descent is defined as follows:

$$\underbrace{\Delta_w(t)}_{\text{Gradient Current Iteration}} = -\overset{\text{Bias}}{\varepsilon}\underbrace{\frac{\overset{\text{Error}}{dE}}{dw_{(t)}}}_{\text{Weight vector}} + \overset{\text{Learning Rate}}{\alpha}\underbrace{\Delta_{w(t-1)}}_{\text{Gradient Previous Iteration}}$$

This process keeps going until gradient for each input-output pair has converged, meaning the newly computed gradient hasn't changed more than a specified convergence threshold, compared to the previous iteration.

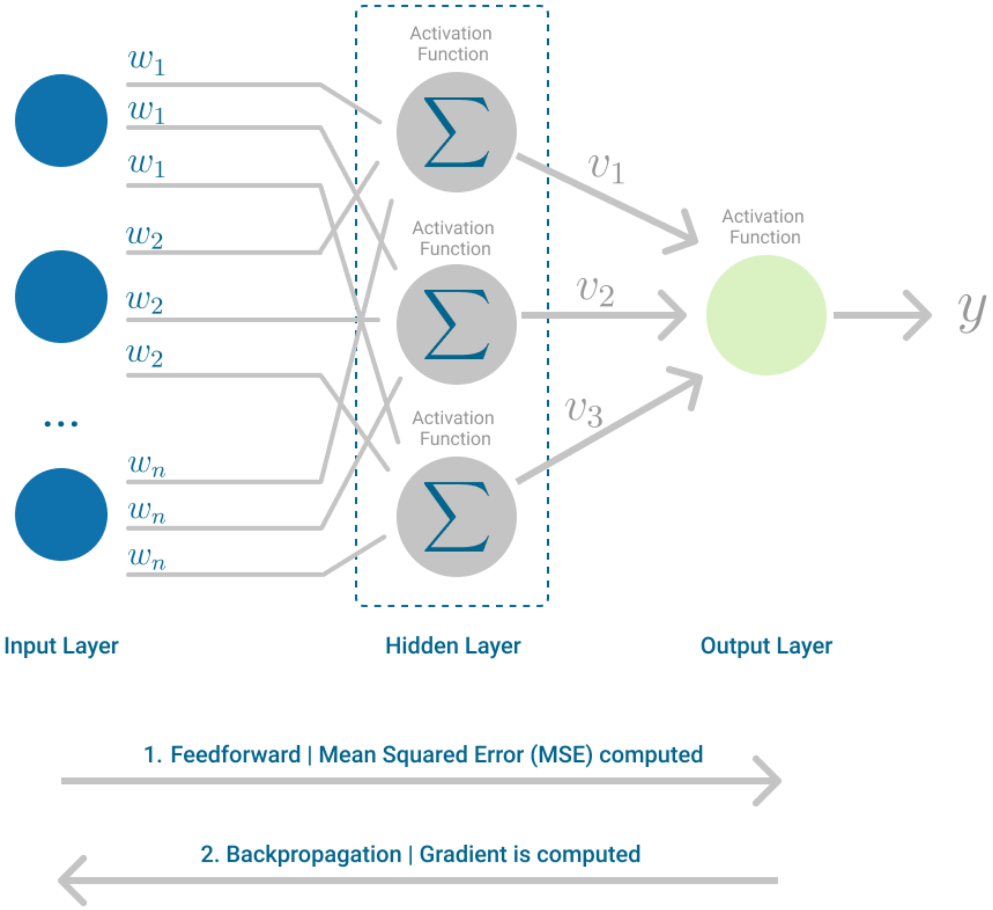


Fig. 5: ANN, highlighting the Feedforward and Backpropagation steps.

4. RESULTS AND DISCUSSION

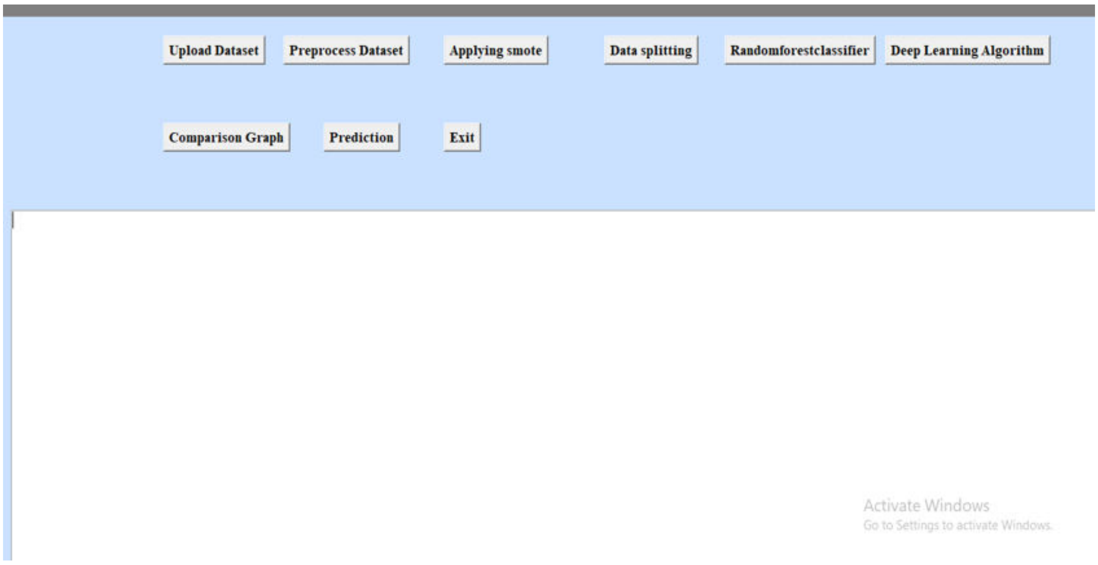
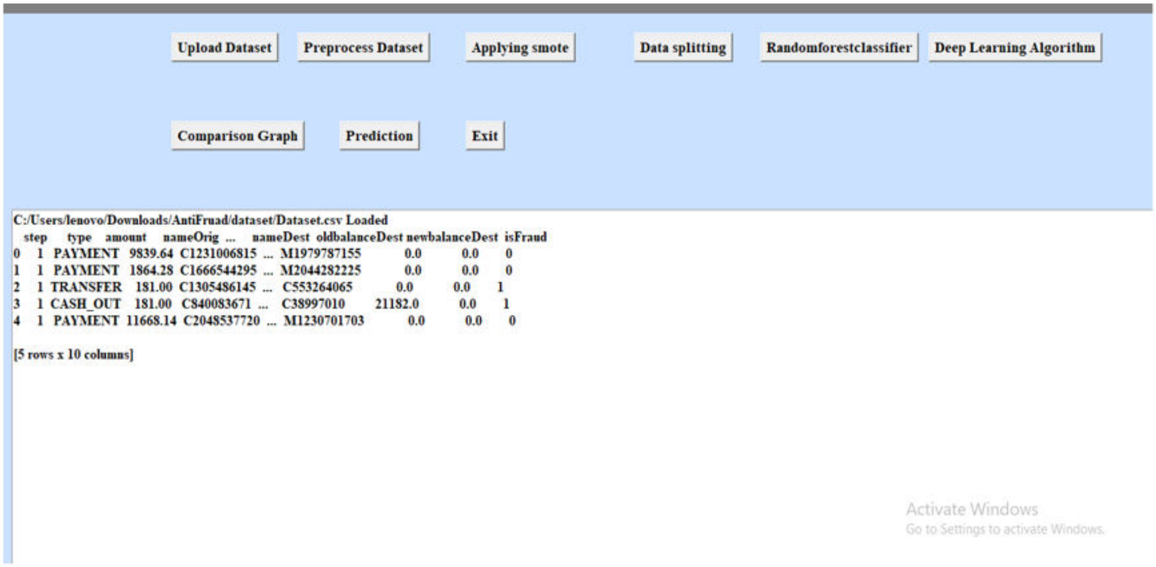


Fig 1: GUI of Fraud detection in Internet Loan Applications



```
C:/Users/lenovo/Downloads/AntiFruad/dataset/Dataset.csv Loaded
  step    type   amount    nameOrig  ...    nameDest  oldbalanceDest newbalanceDest  isFraud
0    1  PAYMENT  9839.64  C1231006815  ...  M1979787155           0.0          0.0       0
1    1  PAYMENT  1864.28  C1666544295  ...  M2044282225           0.0          0.0       0
2    1 TRANSFER   181.00  C1305486145  ...   C553264065          0.0          0.0      1
3    1 CASH_OUT   181.00   C840083671  ...    C38997010      21182.0          0.0       1
4    1  PAYMENT 11668.14  C2048537720  ...  M1230701703           0.0          0.0       0

[5 rows x 10 columns]
```

Fig 2: Upload Dataset in the GUI.

Fig 2 Uploading a dataset refers to the process of importing a dataset containing relevant information for training a deep learning model to detect fraudulent activities. This dataset usually includes features such as step, type, amount, NameOrig, NewbalanceOrig, NameDest, OldbalanceDest, NewbalanceDest, isFraud and other relevant information that can help the model learn patterns indicative of fraud. Once the dataset is uploaded,then it displays the first five rows from the dataset.

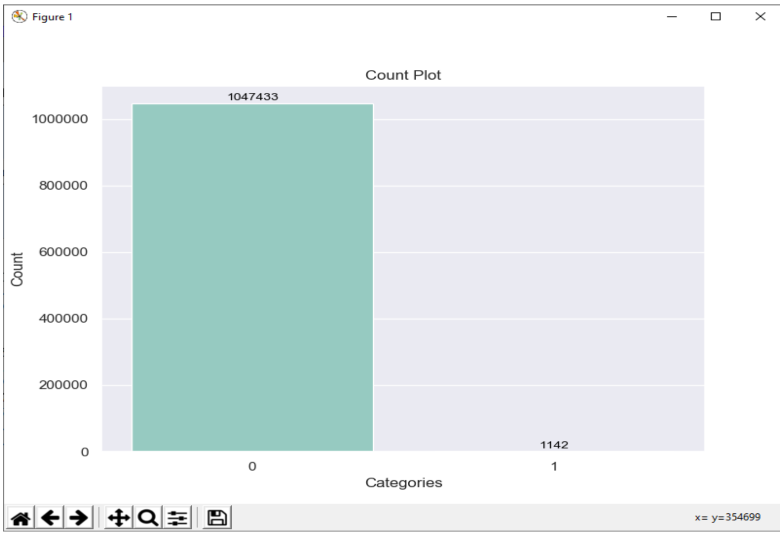


Fig 3: Count Plot of Categories Count in Dataset.

Fig 3 Preprocessing a dataset refers to the steps taken to clean, transform, and prepare the data for analysis or modeling. This can include removing duplicates, handling missing values, scaling or normalizing numerical features, encoding categorical variables among other tasks. Essentially, preprocessing ensures that the data is in a suitable format for further analysis or deep learning algorithms.

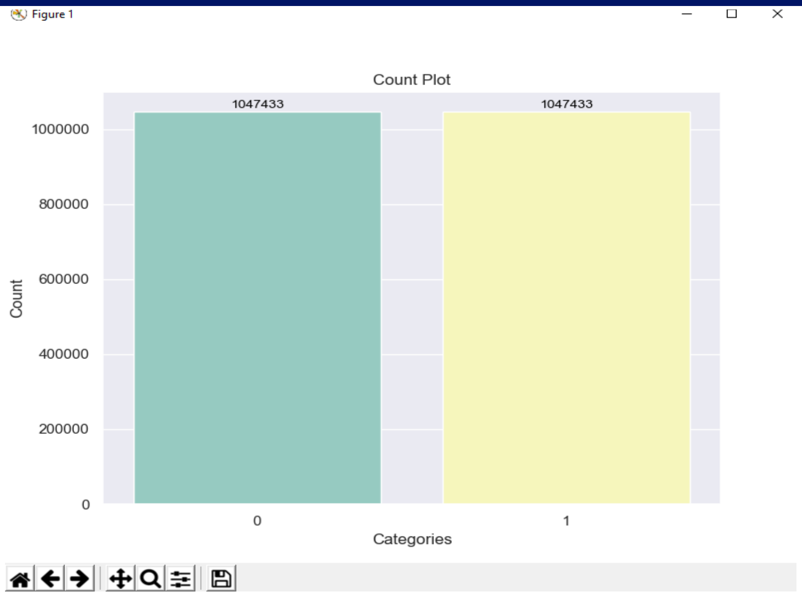
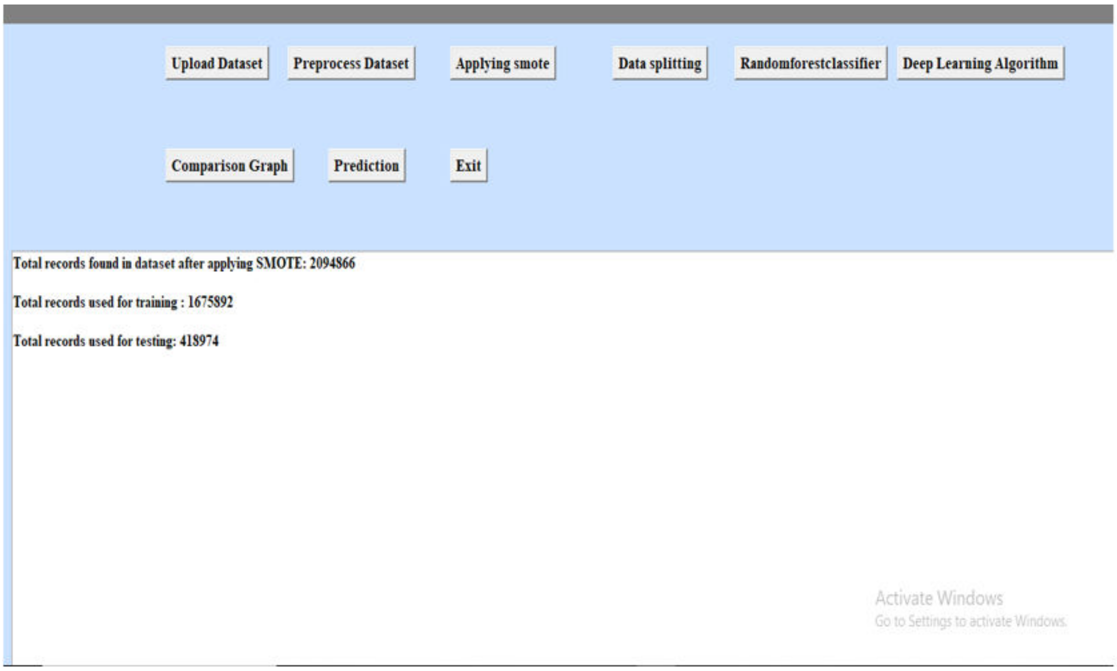


Fig 4: Applying SMOTE for Categories Label.

Fig 4 SMOTE stands for Synthetic Minority Over-sampling Technique. It is a technique to address class imbalance by generating synthetic samples for the minority class. Applying SMOTE involves creating new synthetic instances of the minority class by interpolating between existing minority class instances. This helps balance the class distribution and improve the performance of machine learning models, especially in scenarios where one class is significantly underrepresented.



Total records found in dataset after applying SMOTE: 2094866

Total records used for training : 1675892

Total records used for testing: 418974

Fig 5: Data Preprocessing and Splitting.

Data splitting refers to the process of dividing a dataset into separate subsets for different purposes, such as training, validation, and testing in machine learning. This helps assess the performance of a model on unseen data and prevent overfitting by ensuring the model's generalization ability.

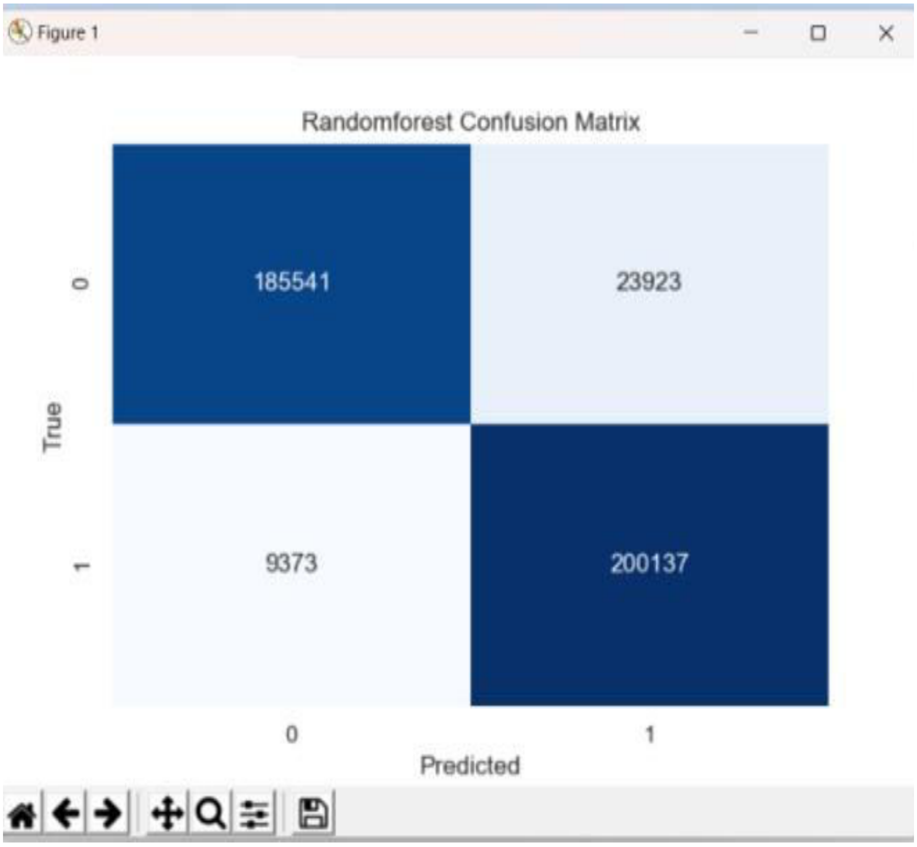


Fig 6: Confusion Matrix of Random Forest Model.

Fig 6 Compare the predicted labels with the actual labels from the testing set to create a confusion matrix. The confusion matrix will have four components: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

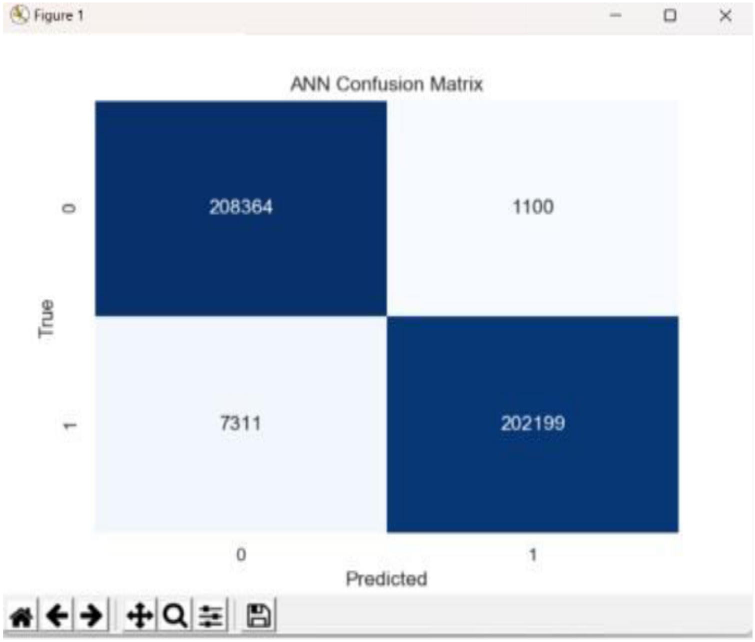


Fig 7: Confusion Matrix of ANN Model.

Fig 7 A confusion matrix is a useful tool for evaluating the performance of a classification model, including artificial neural networks (ANNs), in detecting internet loan fraud. In a confusion matrix, the true labels are compared against the predicted labels generated by the model. The matrix typically consists of four quadrants: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Here's how you can interpret each quadrant in the context of internet loan fraud detection:

- True Positives (TP): The cases where the model correctly predicts instances of internet loan fraud. These are the cases where the model predicted fraud, and it was indeed fraud.
- True Negatives (TN): The cases where the model correctly predicts non-fraudulent instances. These are the cases where the model predicted no fraud, and there was indeed no fraud.
- False Positives (FP): The cases where the model incorrectly predicts fraud. These are the cases where the model predicted fraud, but it was not fraud.
- False Negatives (FN): The cases where the model incorrectly predicts non-fraudulent instances as fraudulent. These are the cases where the model predicted no fraud, but it was fraud

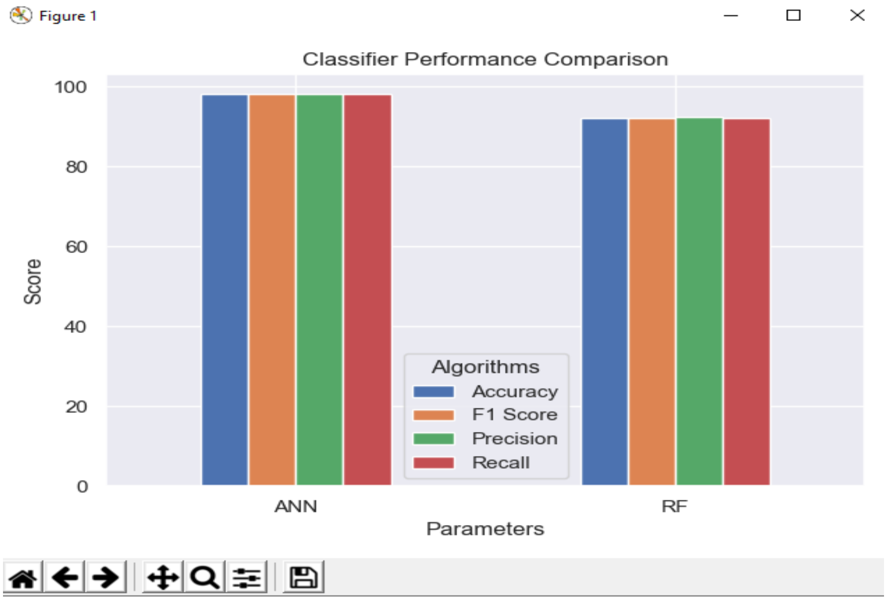
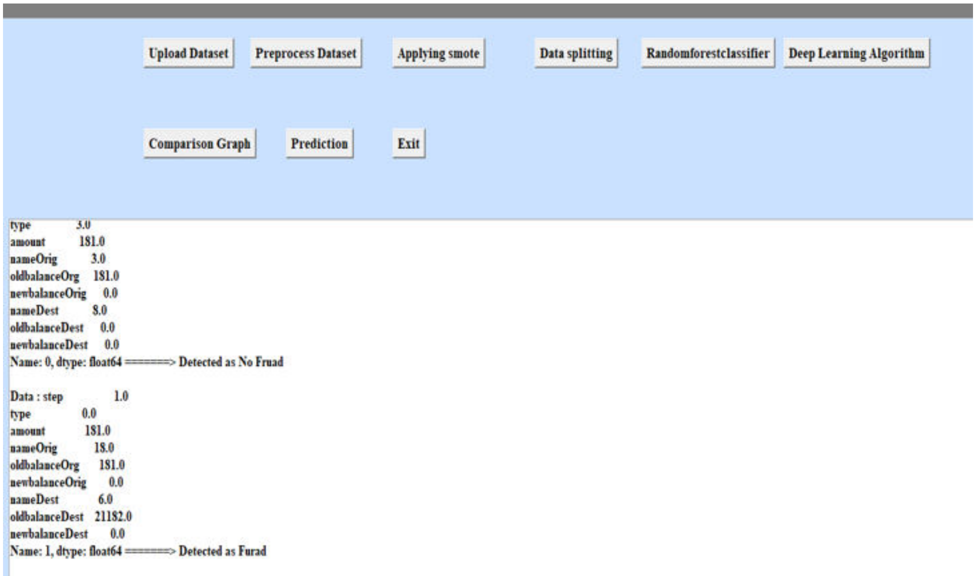


Fig 8: Performance Comparison of ANN and RFC models.

Fig 8 A comparison graph for Random Forest (RF) and Artificial Neural Network (ANN) typically illustrates their performance across various metrics, such as accuracy, precision, recall, F1-score, Accuracy. These metrics help in evaluating how well each model detects fraud in loan applications. The graph would visually depict which model performs better across different evaluation criteria.



```
type          3.0
amount        181.0
nameOrig      3.0
oldbalanceOrg    181.0
newbalanceOrig     0.0
nameDest        8.0
oldbalanceDest     0.0
newbalanceDest     0.0
Name: 0, dtype: float64 =======> Detected as No Fruad

Data : step             1.0
type            0.0
amount          181.0
nameOrig        18.0
oldbalanceOrg     181.0
newbalanceOrig      0.0
nameDest          6.0
oldbalanceDest   21182.0
newbalanceDest      0.0
Name: 1, dtype: float64 =======> Detected as Furad
```

Fig 9: ANN Model Prediction on test data.

Fig 9 Prediction in fraud loan detection refers to the process of using deep learning models to forecast whether a loan application or transaction is likely to be fraudulent. This involves analyzing various features and patterns in the data to identify potentially fraudulent behavior, such as unusual transaction patterns, inconsistencies in applicant information, or suspicious activities. The goal is to accurately classify loan applications as either legitimate or fraudulent to prevent financial losses for lenders and institutions.

Title 1: Performance Comparison of Random Forest Classifier and ANN Model for Fraud Detection

Metric	**Random Forest Classifier**	**ANN Model**
Accuracy	92%	97%
Precision	92%	98%
Recall	92%	97%
F1 Score	92%	97%

Description: The table above illustrates the performance comparison between the Random Forest Classifier and the Artificial Neural Network (ANN) model for the task of fraud detection in internet loan applications. Four key performance metrics, namely accuracy, precision, recall, and F1 score, are evaluated for both models.

5. CONCLUSION

In conclusion, the implementation of advanced neural network architectures offers a promising solution for detecting fraud in internet loan applications. These models leverage intricate patterns and vast data sets to effectively identify fraudulent behaviors with high accuracy. By employing sophisticated algorithms, they enhance the detection capabilities, minimizing risks for lenders and safeguarding against fraudulent activities. The adaptability of these architectures enables continuous learning and refinement, ensuring robust performance in combating evolving fraud tactics. Overall, their deployment signifies a crucial step forward in bolstering the security and integrity of online loan processing systems.

REFERENCES

[1] Xu, Meiling, Yongqiang Fu, and Boping Tian. "An ensemble fraud detection approach for online loans based on application usage patterns." *Journal of Intelligent & Fuzzy Systems* Preprint (2023): 1-14.

[2] Mytnyk, Bohdan, Oleksandr Tkachyk, Nataliya Shakhovska, Solomiia Fedushko, and Yuriy Syerov. "Application of Artificial Intelligence for Fraudulent Banking Operations Recognition." *Big Data and Cognitive Computing* 7, no. 2 (2023): 93.

[3] Lakshmi, Y. Vijaya, Y. Sahithi Priyanka, A. Harika, N. Rajitha, and D. Bhargavi. "Machine Learning based Credit Card Fraud Detection." In *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, pp. 299-305. IEEE, 2023.

[4] Sharma, Nityanand, and Vivek Ranjan. "Credit Card Fraud Detection: A Hybrid of PSO and K-Means Clustering Unsupervised Approach." In *2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 445-450. IEEE, 2023.

[5] Yedukondalu, G., K. Thrilokya, T. Manish Reddy, and K. Sri Vasavi. "Antifraud Model For Internet Loan Using Machine Learning." In *2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 1534-1537. IEEE, 2021.

[6] Zhan, Qing, and Hang Yin. "A loan application fraud detection method based on knowledge graph and neural network." In *Proceedings of the 2nd International Conference on Innovation in Artificial Intelligence*, pp. 111-115. 2018.

[7] Nwade, I., P. Ozoh, M. Olayiwola, M. Ibrahim, M. Kolawole, O. Olubusayo, A. Adigun, and K. Ogundoyin. "DEVELOPMENT OF CREDIT CARDS FRAUD DETECTION MODEL." *LAUTECH Journal of Engineering and Technology* 17, no. 2 (2023): 1-8.

[8] Reddy, N. Madhusudhana, K. A. Sharada, Daniel Pilli, R. Nithya Paranthaman, K. Subba Reddy, and Amit Chauhan. "CNN-Bidirectional LSTM based Approach for Financial Fraud Detection and Prevention System." In *2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, pp. 541-546. IEEE, 2023.

[9] Mizher, Mohammad Ziad, and Ali Bou Nassif. "Deep CNN approach for Unbalanced Credit Card Fraud Detection Data." In *2023 Advances in Science and Engineering Technology International Conferences (ASET)*, pp. 1-7. IEEE, 2023.

[10] Achary, Rathnakar, and Chetan J. Shelke. "Fraud Detection in Banking Transactions Using Machine Learning." In *2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)*, pp. 221-226. IEEE, 2023.

[11] Bajracharya, Aakriti, Barron Harvey, and Danda B. Rawat. "Recent Advances in Cybersecurity and Fraud Detection in Financial Services: A Survey." In *2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 0368-0374. IEEE, 2023.

[12] Fanai, Hosein, and Hossein Abbasimehr. "A novel combined approach based on deep Autoencoder and deep classifiers for credit card fraud detection." *Expert Systems with Applications* 217 (2023): 119562.

[13] Singh, Nikita. "Application of Classification and Regression Techniques in Bank Fraud Detection." In *Machine Learning in Healthcare and Security*, pp. 3-24. CRC Press, 2024.

[14] Zhang, Zheng, Jun Wan, Mingyang Zhou, Zhihui Lai, Claudio J. Tessone, Guoliang Chen, and Hao Liao. "Temporal burstiness and collaborative camouflage aware fraud detection." *Information Processing & Management* 60, no. 2 (2023): 103170.

[15] Aburbeian, AlsharifHasan Mohamad, and Manuel Fernández-Veiga. "Secure Internet Financial Transactions: A Framework Integrating Multi-Factor Authentication and Machine Learning." *AI* 5, no. 1 (2024): 177-194.