IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

Credit Card Fraud Detection Using Machine Learning Algorithms

Manisha Mali **Abhay Pratap Singh**

Prathmesh Dasare **Aneesh Bhambure**

Abstract :

Credit card fraud remains prevalent raising the need to address the problem as a major threat for the monetary as well as the e-commerce industries their countermeasures. As for this challenge, neural networks algorithm has a number of equipment’s that is indispensable. This abstract is a brief description of program developed for identification of credit card scams using neural networks algorithms. This theory starts by embracing the fact that it is dynamic for credit card scam, which mainly renews itself in ways that elude conventional rule-based kind of deteciton systems. Machine learning, therefore, presents a path of analyzing data in order to identify fraudulent exchange by interpreting designs, and other abnormalities for information. Several preliminary structures and models of the neural network using logistic regression, decision trees, random forests, support vector machines, neural networks, and ensemble learning techniques are discussed. Available models are divergent, it helps the system to be more flexible and less sensitive to false positive results but increases its effectiveness of fraud detection.

Keywords: Credit card fault detection, Ada boost, Machine Learning, Fraudulent Transactions, Anomaly Detection, Classification Algorithms, Feature Engineering, Data Preprocessing Imbalanced Data, Majority voting.

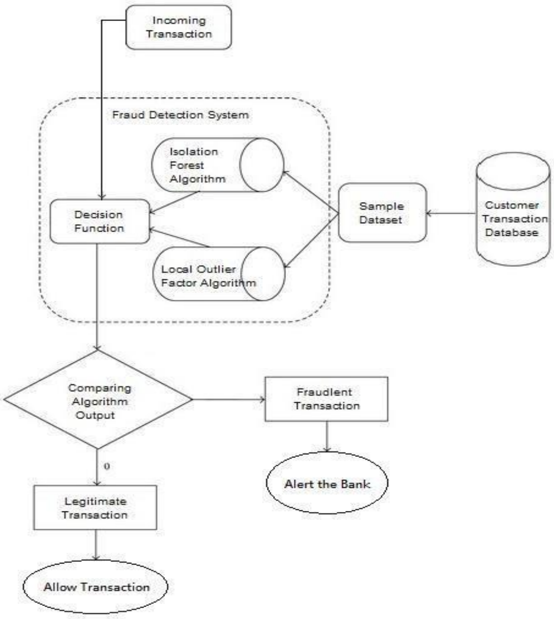
I. Introduction

Introduction of credit and debit cards which has altered the general methods of undertaking financial transactions in the business world. Nevertheless, with the use of electronic payments, which are very convenient, there is constant danger to credit card frauds. Highly complex and innovative are the methods used by fraudsters since the advent of the digital era which adversely affects financial organizations and clientele to the tune of billions of U S dollars per year. To motivate this danger, the utilisation of neural network algorithms is developed as complex and effective tool especially in the sphere of the credit card scam discovery.

As the volume and complexity for monetary transactions grows, rule-based systems are gradually less effective for the job of fraud detection. Therefore, machine learning which provides the flexibility of learning from data is considered more appropriate. As there is a tremendous volume of transaction data produced daily, the application of neural network designs is capable of identifying and improving less extensive design fluctuations or problems ineffective by other models. This work aims to discuss one of the broad categories for different approaches to the neural network for the essential problem for credit card scam detection. The aim of this research is in understanding its methodologies, algorithms such as the appropriate approach in credit card scam detection utilizing neural networks. Thus, the main purpose is to contribute to further attempts to achieve successful security as well as reliability for electronic payment systems.

In particular, we will investigate the following key aspects: Data cleaning, classification algorithms, problem of imbalance data, methods of model assessment, some limitations and direction for further research. Therefore, the objective of this research paper is to present a comprehensive literature review of credit card scam identification using the technological perspective of neural network algorithms.

While exploring intricacies for data, procedures along with evaluation techniques, we hope for assisting the ongoing progress for more strong as well as efficient scam identification models, ultimately safeguarding the financial interests of both consumers and institutions in an increasingly digital world.



Fraud detection methods are continuously developed to defend criminals in adapting to their fraudulent strategies. These frauds are classified as:

- Credit Card Frauds: Online and Offline
- Card Theft
- Account Bankruptcy
- Device Intrusion
- Application Fraud
- Counterfeit Card
- Telecommunication Fraud

Some of the currently used approaches to detection of such fraud are:

- Artificial Neural Network
- Fuzzy Logic
- Genetic Algorithm
- Logistic Regression
- Decision tree
- Support Vector Machines
- Bayesian Networks
- Hidden Markov Model
- K-Nearest Neighbour

II. LITERATURE REVIEW:

With the credit card scam becoming increasingly common in the new world, people are focusing more on the improved development of efficient credit card anti-scam methods. This has proved critically important because neural network, with its adaptability and capability to identify intricate patterns, appears to offer a feasible solution to this highly sensitive problem. This paper gives a brief explanation of the notable research works that are related to credit card sca, with emphasis on its use in concerns to several neural network models. In the past, credit card scam identification mainly involved rule-based system where there was a set of default rules that were used in the identification of the credit card scammers. Despite being somewhat effective, such systems failed to be on par with such changes in paces along with the strategies of the scammers. A study was done to compare the rule based system input with the machine learning based system input. In their studies, authors established that the accuracy of machine learning was significantly higher than rule-based systems particularly when the fraud patterns are intricate and ever-changing.

III. ASSOSCIATED RESEARCH

In this paper, we have studied every single and hybrid neural network procedures for economic programs are assessed. Many economic programs starting with credit card scam to monetary statement fraud have been analyzed.

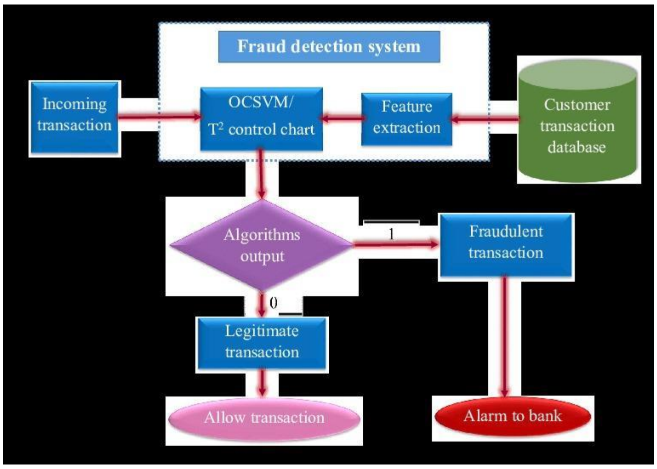
A. SINGULAR FRAMEWORKS

One of the neural network frameworks widely employed for credit card scam identification is Random Forest Classifier. Random Forest is a collective study technique which is nothing but a combination of various decision trees for making more accurate forecasts. It is famous for its resilience, capacity for managing large datasets, and capability to capture complex fraud patterns. Here, we have provided a synopsis of the Random Forest model, its key characteristics, as well as its uses in the context of credit card scam identification.

Model Description:

Ensemble Method: Random Forest is a collective technique which leverages wisdom for crowd. It focuses on making a forest of decision trees, each well studied on a random subset of the information as well as features. Assumptions from multiple trees are then combined to make the final prediction. Decision Trees: The base component of a Random Forest is the decision tree. Decision trees are built by iteratively dividing available information predominantly on communicative applications, leading to a tree-like structure. In Random Forest every tree contributes to the overall decision. Bagging: Random Forest hires a method known as bagging (Bootstrap Aggregating), which includes bootstrapping the dataset (sampling with replacement) tocreate multiple subsets of the data. Some portion are utilized for instructing an entity in the decision tree. Application Significance: Random Forest provides a survey for characteristic significance, that indicates participation for every characteristic of model's assumptions. The available data is beneficial for characteristic evaluation while understanding which variables play a significant role in fraud detection. Parallelization: Random Forest is parallelizable, making it suitable for efficient computation on modern hardware and distributed systems.

Figure :



B. HYBRID MODELS

In this world of credit card scam identification, hybrid models has emerged into a powerful approach for improving its reliability along with robustness for scam identification applications. These models leverage strengths for multiple neural network algorithms, combining them in a synergistic way to improve fraud detection performance. This section introduces the concept of hybrid models andexplores their potential in credit card fraud detection.

Hybrid Model Overview: Hybrid models, for some circumstances of credit card scam identification, refer to a combination for different machine learning algorithms and methodologies to create a unified, more effective fraud detection system. These models are organized in various sections depending on the design along with utilization for multiple algorithms, some are mentioned below:

- **Sequential Hybrid Models**: In sequential hybrid models, different algorithms are applied sequentially in a predetermined order. Each algorithm in the sequence processes the data, and its output serves as input to the next algorithm. This sequential approach can effective in improving the overall detection accuracy.

- **Parallel Hybrid Models:** The parallel hybrid models are those in which more than one algorithm runs concurrently on the same data set. Each algorithm gives its conclusion, and the final decision is mostly taken from the combination of all separate conclusions. Some of the approaches of processing output of parallel models include: voting, weighted scoring, or stacking.
- **Feature-Level Hybrids**: The feature-level hybrid approaches utilize more than one feature selection/extraction method to improve the quality of the input data needed for the individual models. These models are centered on enhancing the data as opposed to compounding the algorithms.
- **Data-Level Hybrids**: Data-level hybrids refer to use of different forms of data, either taken from the same source or from different sources to form a bigger data set. This approach is important while working with multiple source data, or data from various payment gateways.

Challenges: As with any type of model, there are advantages and disadvantages when it comes to hybrid models: One advantage is that there is rarely the need to tune the models greatly, which may be a disadvantage if it is required, Another disadvantage is that the models may be computationally intensive and a downside to this is that the models may be difficult to interpret.

IV. MACHINE LEARNING ALGORITHMS

Credit card scam identification is a crucial system of neural network that helps monetary institutions and businesses protect their customers from fraudulent transactions. Several machine learning algorithms can be employed for this purpose, and often a combination of these algorithms is used to improve detection accuracy. Here are some commonly used algorithms incase of credit card scam identification.

A. ALGORITHMS

Naive Bayes a probabilistic algorithm that can be utilized in case of credit card scam identification. It assumes independence between features, which might not hold in all cases, but it can still perform well in some situations.

A decision stump is a simple machine learning model that consists of a single decision node, used for binary classification tasks. It is essentially a decision tree with a depth of 1,meaning it makes decisions based on a single feature and a threshold value. Decision stumps are also referred to as "one-level decision trees" because they have only one node.

Decision stumps are simple and easy to interpret, but they have limited expressive power compared to more complex decision tree models. However, they can be useful in ensemble methods like boosting, where multiple decision stumps are combined to create astronger model. In boosting, decision stumps are often used as weak learners, and their predictions are weighted and combined to improve classification accuracy. Decision stumps are particularly helpful in problems where you need to create a diverse set of base models for boosting algorithms or when you want to quickly establish a baseline for classification tasks.

Here are some key characteristics along with components of a decision tree: Root Node: The top node of the tree is known as root node. This depicts a whole dataset or a subset of it. Internal Nodes: These are non-leaf nodes in the tree, which represent a feature and a threshold for splitting the data. Internal nodes have branches that lead to child nodes. Leaf Nodes (Terminal Nodes): These nodes depict last output along with decision. In this case of characterization, every leaf node consists to class label, amidst degeneration, it represents mathematical evaluation. Edges (Branches): Edges or divisions attach junctions in the tree, indicating outcome for feature comparison. Based on the feature's value for a given data point, the tree follows the appropriate branch to reach a leaf node. Credit card scam identification is very crucial and sensitive application, and it's essential to handle imbalanced datasets, where fraudulent transactions are typically a small fraction of the data . Techniques such as oversampling, undersampling, and other measures as well as different measures of performance may be required to solve this problem.

Logistic regression is basic powerful tool in terms of bipartite characterization assignments such as credit card scam detection. This eventually is easy for interpreting and summarizes a prediction by providing a probability score from 0-1. Should you realise that logistic regression no longer answers your performance expectations, then there are other techniques including decision trees, random forest, support vector machines or even other machine learning samples that can in turn be used in improving precision of a ‘’scam’’ label. Further, it is also worthy to address the issue of imbalance data using over or under sampling, or even employing different measures for evaluating computed results in case of categories’ disproportion.

SVMs are a very powerful and commonly utilized neural network algorithm for the credit card scam detecting. SVMs are well reputed for maximum margin separation between the classes which is beneficial for generalization and for the FFT credit card fraud detection. They are particularly useful when working with unbalanced datasets, which is almost always true for fraud detection. In any case, SVMs might need a little fine tuning in terms of parameters in order to best suit the dataset in question.

B. MAJORITY VOTING

$$V_k(x \in C_i) = \{1,\ \square\square\ p_k(x) = i, i \in \Lambda\ 0,\ \square\square h\square\square\square\square\square\square$$

The majority voting, also known as collective understanding is a technique used for improving the accuracy as well as the robustness for credit card fraud detect&model involving prediction for multiple deep learning algorithms. The thought process behind this concept is to take advantage of the particular proficiency of each model that in turn masks for the deficiencies that those may hold. weaknesses. Here's how you can implement majority voting for credit card fraud detection:

Select Multiple Base Models: Choose options for deep learning algorithms to serve you your base models. Some common choices for credit card scam identification consists of logistic regression, decision trees, random forests, support vector machines, k-nearest neighbors, along with deep learning. By selecting diverse algorithms, you increase the chances of capturing different aspects of the data and improving overall performance.

Train Base Models: Train each of the selected base models on your training dataset. Ensure that you use the same data splits (training, validation, and test sets) for all base models. Predict Probabilities or Labels: For each transaction in your test set,

obtain the predictions from each base model. This can be done by predicting either class labels (fraudulent or non-fraudulent) or, ideally, the class probabilities (i.e., the probability of being fraudulent or non-fraudulent).

Majority Voting: Implement the majority voting mechanism to make the final decision. There are two common approaches:
Hard Voting: Each base model makes a binary prediction (0 or 1), and category with majority of ballots is taken into consideration for final prediction.
Soft Voting: Each base model provides a probability estimate. The probabilities are averaged for each class (fraudulent and non-fraudulent), and a category with highest average likelihood is taken into consideration for final prediction.

Evaluate Ensemble Technique: Analyze its execution based on ensemble of your test group while harnessing appropriate evaluation parameters like reliability, finesse, reminisce, F1 score, along with ROCAUC. Adjust Thresholds (if needed): Depending on the performance characteristics of your ensemble, you may need to adjust the threshold for making the final decision. For example, you can choose a threshold for the averaged probability in the case of soft voting for getting the wanted deal among finesse along with reminisce .

C. ADABOOST

$$F_T(x) = \sum_{t=1}^{T} \quad f_t(x)(2)E_t = \sum_{i} \quad E[F_{t-1}(x_i) + a_i h(x_i)]$$

AdaBoost (Adaptive Boosting) is a very popular collective understanding technique which can be used efficiently in credit card fraud detection. It functions through building a collection of several weak classifiers – typically decision trees or another simple classifier. Here's how to implement AdaBoost for credit card fraud detection:

1. Data Preprocessing: Process your data where you have features ofcredit card transactions as well as the labels if they are fraudulent or not. Standardized or scale them so that all the features can be put to a similar scale.

2. Model Selection: Select an easy very poor performing classifier as the base model. AdaBoost employs weak classifiers where typical decision trees are sprawling decision trees with shallow depth, commonly referred to as decision stumps. The weak classifier is actually the best choice for AdaBoost to be successful.

3. Training Weak Classifiers: They are the juice of thousands of weak classifiers trained on your training dataset. Begin with a small count for the weak classifiers and increase it, as and when the output is evaluated. The weak classifiers on the other hand are supposed to be just slightly better than constructing a classifier that would guess at random.

4. Weighting Data Points: In every pass of the AdaBoost, the information set consists of patterns which are misclassified by previous weak classifiers that gets increased learner's mass while the points correctly classified gets lower learner's mass. This enhances the capacity of the subsequent weak classifiers to pay more attention to the wrongly classified data.

5. Calculating Weak Classifier Weight: Assign a weight for each trivial classifier proportional to its performance working on the training set. Highly accurate classifiers are the ones given higher Wt.

6. Ensemble Creation: This makes the weak classifiers to be blended to form an intense strong ensemble model. The average of weighted of their predictions is then used to come up with the final decision making process.

7. Thresholding: As in many other ensemble learning models when the weak classifiers are trained and combined, the weighted sum of the predictions can be further passed through a threshold for the purpose of transaction classification. Choose a defining number that marks the required level between precise results and a broad search targeting.

8. Analyzing: AdaBoost parameters left for analysis after using validation set and right evaluation parameters are reliability, finesse, reminisce, F1 score, and ROC AUC.

9. Final Evaluation: After you have decided on the best tuning method and set the threshold you can analyze the model on the test set to see how well it improves on identifying fraudulent transactions .

AdaBoost is known collectively because of its ability for handling imbalanced datasets as well as its propensity for emphasizing misclassified data, which can be of great value in credit card fraud detection. The overfitting is also a less serious problem with this method compared to other ensemble learning techniques. However, the selection of the weak classifier is very important since it should be able to learn a decision model that best fits the data but not overly complex such that it produces high risk of over-fitting.

V. EXPERIMENTAL WORK

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Experimental work for credit card scam identification including neural network techniques involves setting up experiments for analyzing achievements for various systems, algorithms, along with techniques for practical credit card business data. Here's a step-by-step guide on how to conduct experimental work in case of credit card scam identification:

Information Collection along with Preprocessing: Obtain information set for credit card agreement, ensuring it includes both features (transaction details) and labels (fraudulent or non fraudulent indicators). Preprocess the data for managing absent standards, standardize applications, along with transforming absolute attributes if required.

Selecting Algorithms and Models: Choose different kinds of neural network algorithms along with models suitable for binary classification. Some common choices include logistic regression, decision trees, random forests, support vector machines, AdaBoost, machine learning, and more. Additionally, take into consideration collective techniques like majority voting or stacking.

Hyperparameter Tuning: Experiment with different hyperparameters for each selected model. Perform hyperparameter tuning using the validation set. Common hyperparameters include learning rates, regularization strengths, kernel types, maximum tree depth, etc**.**

Ensemble Understanding: Perhaps, the performance of the models can be improved combining them through majority voting, stacked or AdaBoost. The performance of the ensemble should also be assessed on the validation set.

Threshold Optimization: You can always adjust the threshold of your models for classification to hit the balance between precision and recall based on how stringent you want your fraud check.

Final Model Selection: Depending on the validation of results, select the optimal or the set of the most effective models for credit card fraud detection.

Final Evaluation: Result of the selected system's implementation shall be evaluated by assessment suite in order to receive the final overall effectiveness estimate. This evaluation gives an authentic idea about how much your model will be efficient in approach when at large.

Monitoring and Maintenance: After building the model, regularly use the model in production, and feed new data of fraud patterns to the model to help enhance it.

Ethical Considerations: Always note principles of, and relevant legislation on data protection and privacy, when it comes to recognizing credit card scams.
Remember that credit card scam identification is a highly important application, however, the chosen model should be always a trade-off between given fraudulent agreements and untruth positive. The model must be brought up to date as often as is necessary and its performance checked for the system to remain effective.

VI. EXISTING SYSTEM

Identification of credit card scam is one very important application observed in monetary and e commerce industries.. Various machine learning algorithms and systems have been developed and are in use to address this problem. These systems typically consist of the following components:

Data Collection and Preprocessing: Gathering transaction data, including features like transaction amount, location, time, and more. Preprocessing information for managing absent parameters, anomalies, along with encoding categorical variables.

Choosing System: Choosing proper neural network models and techniques in case of classification, like logistic retrogression, decision trees, random forests, support vector machines, machine learning, along with collective techniques.

Data Splitting: Splitting information set for educating, validation, along test parameters with the system development and evaluation.

Application Engineering: Making relevant applications or changing current ones for enhancing the system's ability to detect fraud. Feature engineering might involve dimensionality reduction, creating time-based features, and more.

System Education: Training the chosen system on dataset for evaluating the designs associated with fraudulent along with non-fraudulent transactions. Model Evaluation: Evaluating the systems' accomplishments by taking into consideration different parameters like reliability, precision, recall, F1 score, along with ROC AUC on validation set.

Hyperparameter Tuning: Optimizing hyperparameters of the models to improve their performance. Common hyperparameters include learning rates, regularization strengths, and kernel types.

Ensemble Techniques: Using classifiers sequentially, in parallel, or in cascade because creating whole new systems from multiple classifiers leads to increased accuracy in fraud detection.

Threshold Optimization: Tuning the values of the classification thresholds so as to, depending on the specifics of the business case and accepted level of risk, either maximize the recall while at the same time also maintaining relatively high precision, or vice versa maximize the precision while also keeping the recall acceptable.

Final System Decision: Choosing among outstanding evaluation system or ensemble based on its validation outcome. Test Set Evaluation: Validation which involves evaluating the actual performance of the selected model on another test set in order to arrive at the true efficiency of the model.

Real-Time Scoring: Applying the chosen model into a near real-time scoring environment where it can scan, evaluate transactions for risk of fraud.

Surveillance along with Maintenance: Continually monitoring system's accomplishment while updating it with new information for enhancing it to be useful in changing fraud patterns.

Reporting and Documentation: Recording various aspects of the system which consist of the dataset, specified model, and evaluation of the model.

Ethical Considerations: It is in the interest of the system that the relevant laws in data privacy and security are met hence avoiding any scandalous leak of information.

There is always an improvement when it comes to developing credit card fraud detection models and algum are continually improving and updating, as the fraud patterns evolve. Finally, rule-based systems and or anomaly detection methods blend well with machine learning models to improve the likelihood of detection and greatly minimize cases of false alarms.

VII. PROPOSED SYSTEM

When applying a problem of credit card scam identification using neural network techniques, there are some components and features to be taken into consideration. The proposed system should have high resilience and maintain explicit improvements to respond to new trends in fraud and possessed efficiency in detecting credit card frauds. It should also ensure that the customer's interest protection and safeguarding of information security.

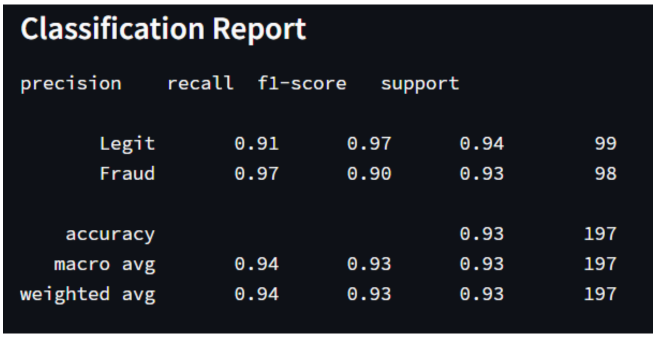
Dataset: Dataset of credit card transactions Training: Making a system for means of learning decent principles for all the parameters and the offset from labeled instances. In guided training, a neural network technique assembles a model by scrutinizing numerous instances and striving to discover a model that diminishes loss; this procedure is termed as empirical risk minimization. Characteristics extraction: The retrieval and correlation of characteristics are grounded on these evaluations. Aside from the basic trait, a more sophisticated form of characteristic is also introduced. Characteristic extraction method is applied to derive the attributes by retaining as considerable information as attainable from an extensive assortment of image data.

Categorization: Categorization is a procedure for classifying a provided collection of information into groupings; it can be carried out on both organized or disorganized data. The procedure kicks off by foretelling the grouping of provided data instances. The groupings are frequently denoted as objectives, tags, or groupings. Predictive modeling for categorization involves approximating the mapping mechanism from input factors to distinct output factors. The primary aim is to determine which group or category the fresh data will fit into. Categorization is accomplished using machine understandable techniques.

VIII. Reasult and Work

Table.1. Performance Metrics

Accuracy	Precision	Recall	F1-score
93.40	91.	97	94



Classification Report

```
precision    recall  f1-score   support

       Legit       0.91      0.97      0.94        99
       Fraud       0.97      0.90      0.93        98

    accuracy                           0.93       197
   macro avg       0.94      0.93      0.93       197
weighted avg       0.94      0.93      0.93       197
```

Fig, 4. ROC curve

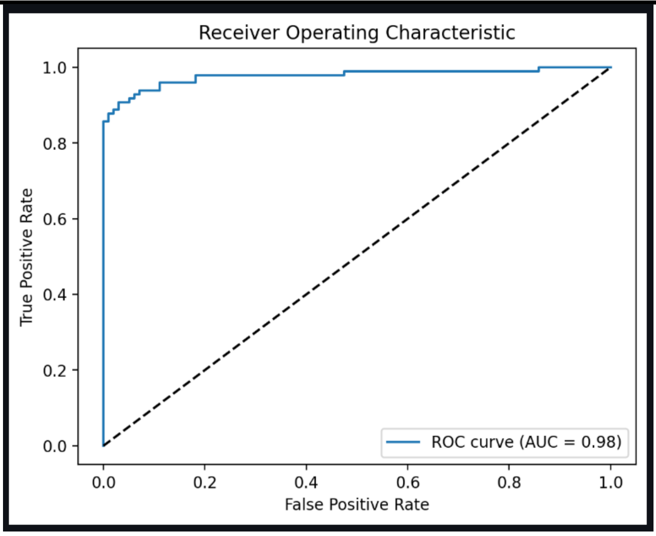
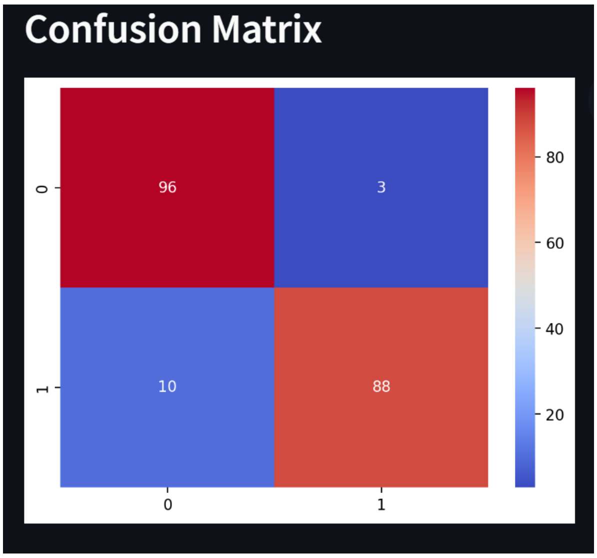
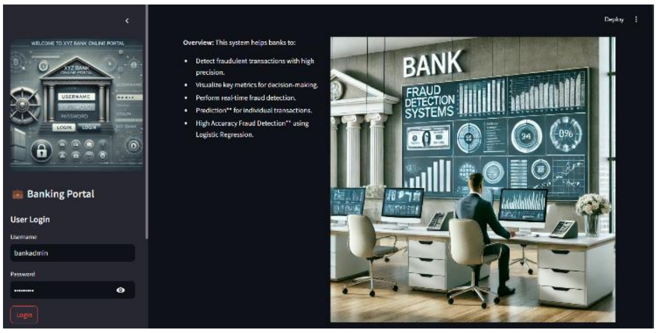


Fig.5. Confusion matrix



Interface GUI



Banking Portal

User Login

Username

bankadmin

Password

••••••••••

Login

Overview: This system helps banks to:

- Detect fraudulent transactions with high precision.
- Visualize key metrics for decision-making.
- Perform real-time fraud detection.
- Prediction** for individual transactions.
- High Accuracy Fraud Detection** using Logistic Regression.

IX. CONCLUSION

Credit card scam identification along with neural network algorithms which is a critical and evolving area in the financial and e-commerce industries. The application of deep learning models offers significant advantages for recognizing and preventing deceptive agreements, ultimately safeguarding interests for both financial institutions and customers. Credit card fraud is a pervasive problem, and its methods are constantly evolving. Traditional rule-based systems are limited in their ability to adapt to new fraud patterns. Machine learning algorithms offer a more dynamic and data-driven approach to fraud detection. Various machine learning models and algorithms can be applied to the task, like logistic retrogression, decision trees, random forests, support vector machines, machine learning, along with collective techniques. The diversity of models allows for robust fraud detection and minimizes false positives. Effective initial information work, comprising managing absent parameters, anomalies identification, along with feature engineering, significantly impacts its criteria for neural network techniques. Ensemble learning methods, like majority voting, stacking, along with AdaBoost, improve accuracy by combining the predictions of multiple models. Effective communication with customers is crucial. They need to be informed about potentially fraudulent activities and provided with resources to report suspicious transactions

X. REFERENCES

[1] Y. Sahin, S. Bulkan, E. Duman, "A cost-sensitive decision tree approach for fraud detection", Expert Syst. Appl., vol. 40, pp. 5916-5923, 2013.

[2] A. O. Adewumi, A. A. Akinyelu, "A survey of machine learning and nature inspired based credit card fraud detection techniques", Int. J. Syst. Assurance Eng. Manage.,vol. 8, no. 2, pp. 937-953, 2017.

[3] A. Srivastava, A. Kundu, S. Sural, A. Majumdar, "Credit card fraud detection using hidden Markov model", IEEE Trans. Depend. Sec. Comput., vol. 5, no. 1, pp. 37- 48,Jan. 2008.

[4] The Nilson Report, Oct. 2016, [online] Available: https://www.nilsonreport.com/upload/content_promo/The_Nilson_Report_10- 17-2016.pdf.

[5] J. T. Quah, M. Sriganesh, "Real-time credit card fraud detection using computational intelligence", Expert Syst. Appl., vol. 35, no. 4, pp. 1721-1732, 2008.

[6] S. Bhattacharyya, S. Jha, K. Tharakunnel, J. C. Westland, "Data mining for credit card fraud: A comparative study", Decision Support Syst., vol. 50, no. 3, pp. 602-613, 2011.

[7] N. S. Halvaiee, M. K. Akbari, "A novel model for credit card fraud detection using artificial immune systems", Appl. Soft Comput., vol. 24, pp. 40-49, Nov. 2014.

[8] S. Panigrahi, A. Kundu, S. Sural, A. K. Majumdar, "Credit card fraud detection: A fusion approach using Dempster– Shafer theory and Bayesian learning", Inf. Fusion,vol. 10, no. 4, pp. 354-363, 2009.

[9] N. Mahmoudi, E. Duman, "Detecting credit card fraud by modified fisher discriminant analysis", Expert Syst. Appl., vol. 42, no. 5, pp. 2510-2516, 2015.

[10] D. Sánchez, M. A. Vila, L. Cerda, J. M. Serrano, "Association rules applied to credit card fraud detection", Expert Syst. Appl., vol. 36, no. 2, pp. 3630-3640, 2009.

[11] E. Duman, M. H. Ozcelik, "Detecting credit card fraud by genetic algorithm and scatter search", Expert Syst. Appl., vol. 38, no. 10, pp. 13057-13063, 2011.

[12] P. Ravisankar, V. Ravi, G. R. Rao, I. Bose, "Detection of financial statement fraud and feature selection using data mining techniques", Decision Support Syst., vol. 50, no. 2, pp.491-500, 2011.

[13] E. Kirkos, C. Spathis, Y. Manolopoulos, "Data mining techniques for the detection of fraudulent financial statements", Expert Syst. Appl., vol. 32, no. 4, pp. 995-1003, 2007.

[14] F. H. Glancy, S. B. Yadav, "A computational model for financial reporting fraud detection", Decision Support Syst., vol. 50, no. 3, pp. 595-601, 2011.

[15] D. Olszewski, "Fraud detection using self-organizing map visualizing the user profiles",Knowl.-Based Syst., vol. 70, pp. 324-334, Nov. 2014.

[16] 16. I. T. Christou, M. Bakopoulos, T. Dimitriou, E. Amolochitis, S. Tsekeridou, C. Dimitriadis, "Detecting fraud in online games of chance and lotteries", Expert Syst. Appl., vol. 38, no. 10, pp. 13158-13169, 2011.

[17] C.-F. Tsai, "Combining cluster analysis with classifier ensemblesto predict financial distress", Inf. Fusion, vol. 16, pp. 46-58, Mar. 2014.

[18] F. H. Chen, D. J. Chi, J. Y. Zhu, "Application of random forest rough set theory decision tree and neural network to detect financial statement fraud—Taking corporate governance into consideration", Proc. Int. Conf. Intell. Comput., pp. 221-234, 2014.

[19] Y. Li, C. Yan, W. Liu, M. Li, "A principle component analysis-based random forest with the potential nearest neighbor method for automobile insurance fraud identification", Appl. Soft Comput..

[20] S. Subudhi, S. Panigrahi, "Use of optimized Fuzzy C-Means clustering and supervised classifiers for automobile insurance fraud detection", J. King Saud Univ.-Comput. Inf. Sci.

[21] M. Seera, C. P. Lim, K. S.Tan, W. S. Liew, "Classification of transcranial Doppler signals using individual and ensemble recurrent neural networks", Neuro computing,vol. 249, pp. 337-344, Aug. 2017.

[22] C. Phua, K. Smith-Miles, V. Lee, R. Gayler, "Resilient identity crime detection ",IEEE Trans. Knowl. Data Eng., vol. 24, no. 3, pp. 533-546, Mar. 2012.

[23] M. W. Powers, "Evaluation: From precision recall and F measure to ROC informedness markedness and correlation", J. Mach. Learn. Technol., vol. 2, no. 1, pp.37-63, 2011.

[24] Credit Card Fraud Detection, Nov. 2017, [online] Available: https://www.kaggle.com/dalpozz/creditcardfraud.

[25] K. Randhawa, C. K. Loo, M. Seera, C. P. Lim, A. K. Nandi, ―Credit Card Fraud Detection Using AdaBoost and Majority Voting‖, IEEE Access, vol. 6, pp. 14277- 14284, 2018