Industrial Engineering Journal

ISSN: 0970-2555

Volume : 53, Special Issue 4, April : 2024

LEVARAGING HMM FOR UPI FRAUD DETECTION USING DEEP LEARNING MODEL

Dr D Manendra Sai1,B Jyotsna2,V Supritha Anjali 3, B Sankeerthana4, SK Ayesha Siddiqua5
1Professor, Department of Computer Science and Engineering, Vignan's Institute of Engineering for Women, isakhapatnam, Andhra Pradesh, India
2,3,4,5students Department of Computer Science and Engineering, Vignan's Institute of Engineering for Women, Visakhapatnam, Andhra Pradesh, India

ABSTRACT

Now a day the usage of UPI has dramatically increased. As UPI becomes the most popular mode of payment for both online as well as regular purchase, cases of fraud associated with it are also rising. In this project, we model the sequence of operations in UPI transaction processing using a Convolutional Neural Network (CNN) and show how it can be used for the detection of frauds. An CNN is initially trained with the normal behavior of a cardholder. If an incoming UPI transaction is not accepted by the trained CNN with sufficiently high probability, it is considered to be fraudulent. At the same time, we try to ensure that genuine transactions are not rejected. We present detailed experimental results to show the effectiveness of our approach and compare it with other techniques available in the literature.

Keywords :- UPI Fraud, Hidden Markov Models (HMM), Convolutional Neural Networks (CNN), Artificial Neural Networks (ANN), Fraud Mitigation, Digital Transactions

INTRODUCTION

The popularity of online shopping is growing day by day. According to an ACNielsen study conducted in 2005, one-tenth of the world's population is shopping online. Germany and Great Britain have the largest number of online shoppers, and UPI is the most popular mode of payment (59 percent). About 350 million transactions per year were reportedly carried out by Barclaycard, the largest UPI company in the United Kingdom, toward the end of the last century. Retailers like Wal-Mart typically handle much larger number of UPI transactions including online and regular purchases. As the number of UPI users rises world-wide, the opportunities for attackers to steal UPI details and, subsequently, commit fraud are also increasing. The total UPI fraud in the United States itself is reported to be $2.7 billion in 2005 and estimated to be $3.0 billion in 2006, out of which $1.6 billion and $1.7 billion, respectively, are the estimates of online fraud.

Credit-card-based purchases can be categorized into two types:

1) Physical card and

2) Virtual card.

In a physical-card-based purchase, the cardholder presents his card physically to a merchant for making a payment. To carry out fraudulent transactions in this kind of purchase, an attacker has to steal the UPI. If the cardholder does not realize the loss of card, it can lead to a substantial financial loss to the UPI company. In these cond. kind of purchase, only some important information about a card (card number, expiration date, secure code) is required to make the payment. Such purchases are normally done on the

http://doi.org/10.36893/IEJ.2024.V53I4.155-163

Internet or over the telephone. To commit fraud in these types of purchases, a fraudster simply needs to know the card details. Most of the time, the genuine cardholder is not aware that someone else has seen or stolen his card information. The only way to detect this kind of fraud is to analyze the spending patterns on every card and to figure out any inconsistency with respect to the “usual” spending patterns. Fraud detection based on the analysis of existing purchase data of cardholder is a promising way to reduce the rate of successful UPI frauds. Since humans tend to exhibit specific behaviouristic profiles, every cardholder can be represented by a set of patterns containing information about the typical purchase category, the time since the last purchase, the amount of money spent, etc. Deviation from such patterns is a potential threat to the system. Several techniques for the detection of UPI fraud have been proposed in the last few years.

LITERATURE SURVEY

[1] The article titled "BLAST-SSAHA Hybridization for UPI Fraud Detection," authored by Amlan Kundu, Suvasini Panigrahi, Shamik Sural, and Arun K. Majumdar, was published in the IEEE Transactions on Dependable and Secure Computing in the October-December 2009 issue.

The main focus of this article is on the development and application of a hybrid approach combining BLAST (Basic Local Alignment Search Tool) and SSAHA (Sequence Search and Alignment by Hashing Algorithm) for the detection of fraud in UPI (Unified Payment Interface) transactions. Here's a breakdown of the key components and findings of the article:

1. Title and Objective: The title clearly indicates the objective of the research, which is to detect fraud in UPI transactions using a hybridization of BLAST and SSAHA.
2. Introduction: The introduction likely provides background information on UPI transactions, emphasizing their importance in modern financial systems, as well as the growing concern over fraudulent activities within these systems. It may also introduce the concept of sequence alignment algorithms like BLAST and SSAHA and their potential application in fraud detection.
3. Methodology: The authors describe the hybrid approach they developed, which combines the strengths of BLAST and SSAHA. BLAST is known for its speed and sensitivity in identifying similar sequences, while SSAHA utilizes hashing algorithms for efficient sequence alignment. By combining these two methods, the authors aim to enhance the accuracy and efficiency of fraud detection in UPI transactions.
4. Experimentation and Results: The article likely outlines the experimental setup used to evaluate the performance of the hybrid approach. This would include details such as the datasets used, the metrics for evaluating fraud detection performance (e.g., accuracy, precision, recall), and comparisons with other existing methods. The results section would present the findings of these experiments, demonstrating the effectiveness of the proposed hybrid approach in detecting fraudulent UPI transactions.

[2] The article titled "Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning" likely explores a novel method for detecting credit card fraud by integrating two distinct techniques: Dempster–Shafer theory and Bayesian learning.

1. Dempster–Shafer theory: This theory, also known as evidence theory, provides a mathematical framework for reasoning with uncertainty and incomplete information. It allows for the combination of evidence from different sources to make decisions or draw conclusions. In the context of fraud detection, Dempster–Shafer theory could be used to integrate various pieces of evidence related to credit card transactions to assess the likelihood of fraud.
2. Bayesian learning: Bayesian learning is a statistical approach based on Bayes' theorem, which updates probabilities as new evidence becomes available. In the context of fraud detection, Bayesian learning could be used to model the probability of a transaction being fraudulent based on historical data and other relevant factors.

The fusion approach described in the article likely involves combining these two methodologies to enhance the accuracy and reliability of credit card fraud detection. By leveraging the strengths of both Dempster– Shafer theory and Bayesian learning, the method may be capable of handling uncertainty, noise, and incomplete information more effectively than using either approach alone.

The article discuss the theoretical underpinnings of the fusion approach, provide details on the implementation and methodology used, and present experimental results demonstrating the effectiveness of the proposed method compared to existing approaches.

Overall, this fusion approach represents an innovative strategy for tackling the challenge of credit card fraud detection, leveraging principles from both evidence theory and statistical learning to improve detection accuracy and reduce false positives.

Next, the paper would delve into the methodology employed, focusing on Random Forest as the chosen algorithm. Random Forest is a popular ensemble learning method known for its effectiveness in classification tasks. It works by constructing multiple decision trees during the training phase and outputting the mode of the classes as the prediction. This ensemble approach helps in reducing overfitting and improving accuracy.

The author likely describes the dataset used for training and testing the Random Forest model. This dataset would contain features extracted from credit card transactions, such as transaction amount, location, time, type of transaction, etc., along with labels indicating whether each transaction is fraudulent or legitimate.

The paper would then explain the preprocessing steps applied to the data, including handling missing values, scaling features, and possibly feature engineering techniques to enhance model performance.The training process of the Random Forest model would be detailed, including hyperparameter tuning to optimize its performance. The author may discuss techniques for preventing overfitting, such as limiting tree depth or adjusting the number of trees in the forest.

Evaluation metrics used to assess the model's performance, such as accuracy, precision, recall, F1-score, and ROC- AUC, would be described. The paper would likely compare the Random Forest model's performance with other traditional methods or alternative machine learning algorithms commonly used for fraud detection.

Results and findings obtained from the experiments would be presented and analyzed, showcasing the efficacy of the Random Forest approach in accurately identifying fraudulent transactions while minimizing false positives.The paper might also discuss the practical implications of deploying such a model in real-world scenarios, including considerations for implementation, scalability, and computational resources required.

[4] The article titled “Fuzzy Darwinian Detection of UPI Fraud,” authored by Peter J. Bentley, Jungwon Kim, Gil-Ho Jung and Jong-Uk Choi was published In the 14th Annual Fall Symposium of the Korean Information Processing Society, 14th October 2000.

1. Fuzzy Logic: Fuzzy logic is a form of multi-valued logic that deals with approximate reasoning. In the context of fraud detection, fuzzy logic can be used to handle imprecise or uncertain information.
2. Darwinian Detection: This likely refers to a detection approach inspired by principles of evolution, such as natural selection and survival of the fittest. In the context of fraud detection, it could involve evolving detection algorithms or strategies over time to adapt to changing fraud patterns.
3. Credit Card Fraud: This is the specific type of fraud being addressed in the article. Credit card fraud involves unauthorized or fraudulent use of credit card information for financial gain.

The article discuss how these concepts are integrated to create a novel approach to credit card fraud detection. It might cover aspects such as:

- The development of fuzzy logic-based rules or models for detecting fraudulent transactions.
- Evolutionary algorithms used to optimize or evolve these detection models over time.
- Performance evaluation of the proposed method compared to traditional fraud detection techniques.
- Real-world applications and potential benefits of the fuzzy Darwinian approach in combating credit card fraud.

EXISTING SYSTEM

All the existing method to detect the UPI was on the mode like the detection occurs only after the complaint of the card holder about fraud done. It is not a convenient way to avoid the loss happens to the cad holder. After getting the complaint they detected the fraud on the basis of the IP address. For this they need the help of the Cyber crime to detect the fraud and make action on it. It takes so much man power.

□ To detect counterfeit transactions, three machine-learning algorithms were presented and implemented.

□ There are many measures used to evaluate the performance of classifiers or predictors, such as the Gradient Boost Classifier, Vector Machine, Random Forest, and Decision Tree.

□ These metrics are either prevalence dependent or prevalence independent.

□ Furthermore, these techniques are used in UPI fraud detection mechanisms and the results of these algorithms have been compared.

Disadvantages

- The main disadvantage of the existing system is the detection occurs only after gets a written complaint.
- In the existing system there is physical inconvenience exists.
- The period occurs to detect the fraud will cause so many losses to the card holder.
- There is no particular security system in the existing so a hacker can easily access others card.

PROPOSED SCHEME

This project introduces a UPI fraud detection system using a Convolutional Neural Network (CNN) based on the spending profile of cardholders. The system, integrated into the bank's Fraud Detection System (FDS), monitors user spending, automatically blocking transactions deemed unusual and alerting the bank without manual intervention. Various modern techniques including artificial neural networks and machine learning algorithms such as Auto Encoder, Local Outlier Factor, and K means Clustering are compared and utilized. The primary goal is to detect fraudulent transactions and develop a method for generating test data. This algorithm serves as a heuristic approach for solving complex computational problems. Implementing an efficient fraud detection system is crucial for minimizing losses for UPI issuing companies and their clients.

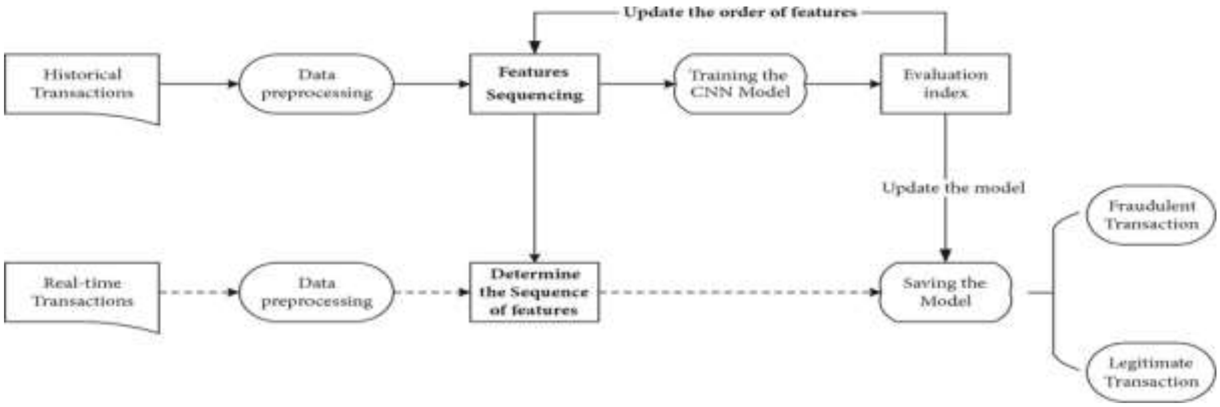


Fig: Sequence Model Of CNN

Advantages

- The main advantage is that the detection occurs much faster than any other method.
- In all the existing systems the real card holder should checked for the Fraud detection. But in our method there is no need of the physical inconveniences of the card holder. All the checking and the detection occur automatically.
- This project needs no man power for the detection.
- This project provides most accurate method in UPI fraud detection.

System Block Diagram

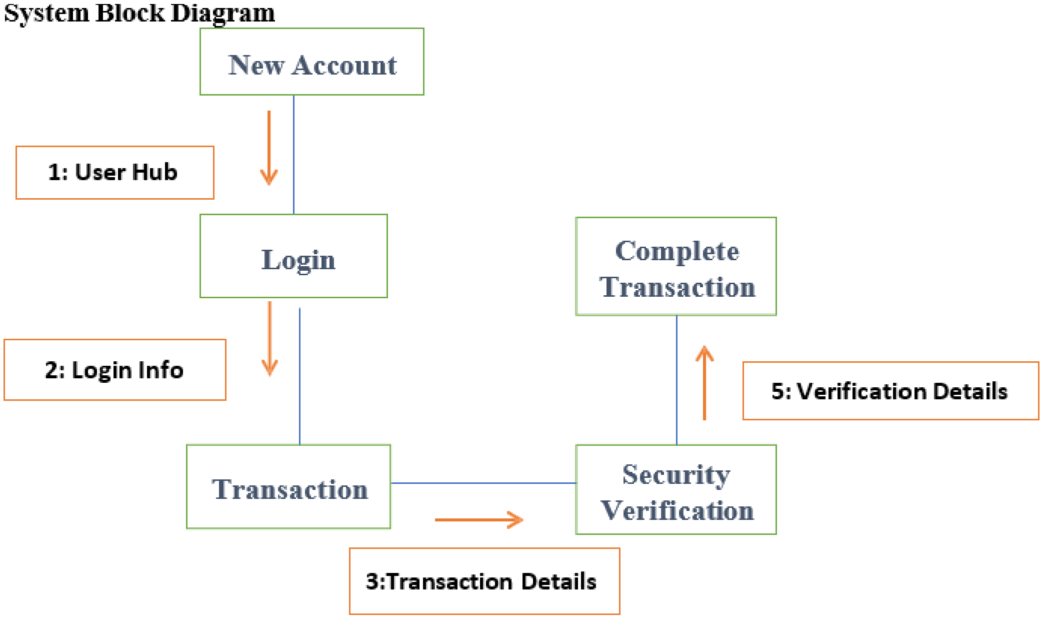


Fig1: System Architecture

OUTPUT SCREENS

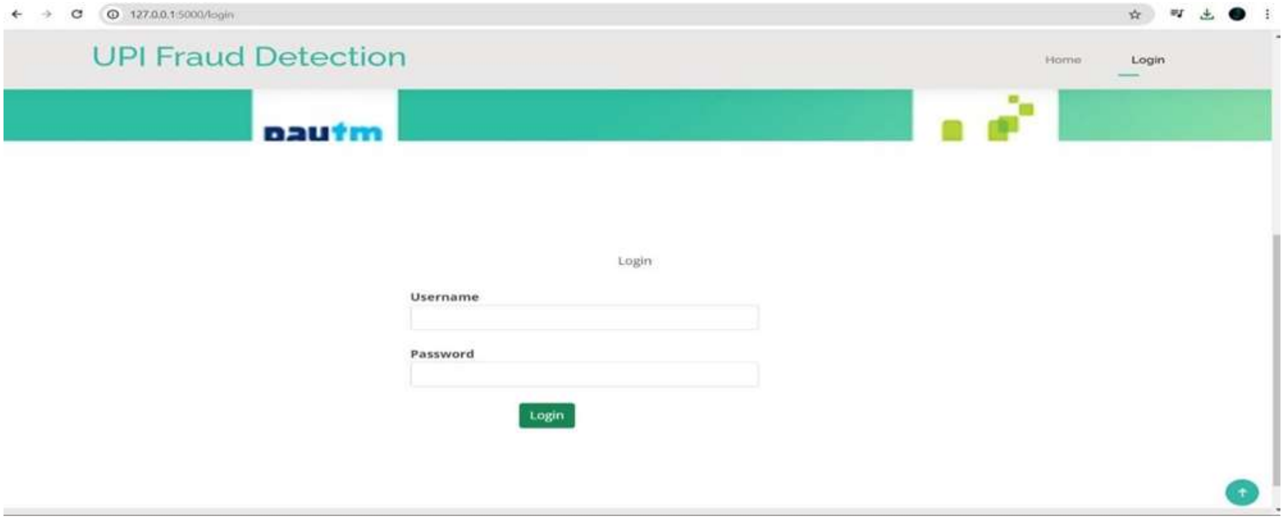
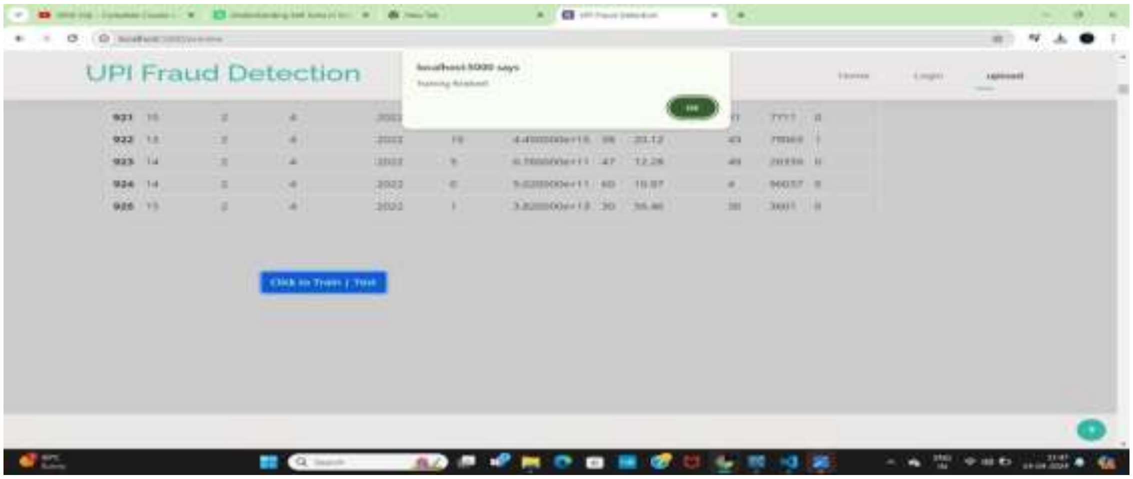


Fig: Login Page



921	15	2	4	2022	[illegible]	[illegible]	[illegible]	[illegible]	11	7717	0
922	13	2	4	2022	19	4.490000e+15	39	20.12	43	78663	1
923	14	2	4	2022	9	6.780000e+11	47	12.26	49	26938	0
924	14	2	4	2022	6	5.020000e+11	60	10.07	4	96037	0
925	15	2	4	2022	1	3.820000e+12	30	35.46	30	3001	0

Fig: Training and testing completed

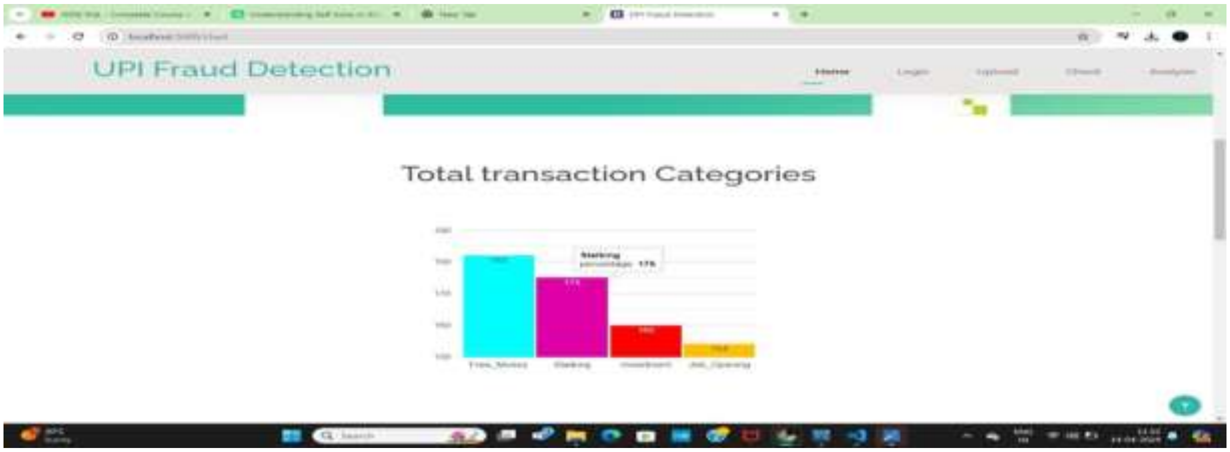


Fig: Transaction categories

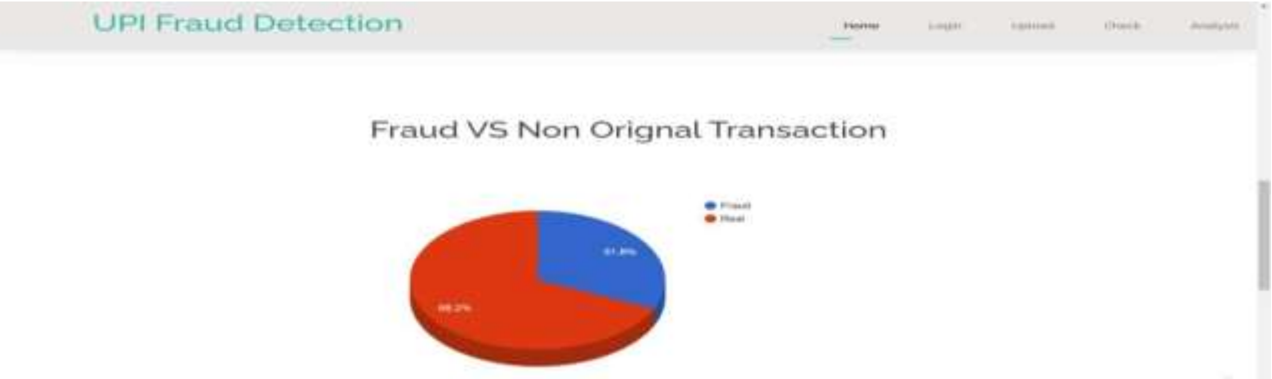


Fig: Fraud data and Real data

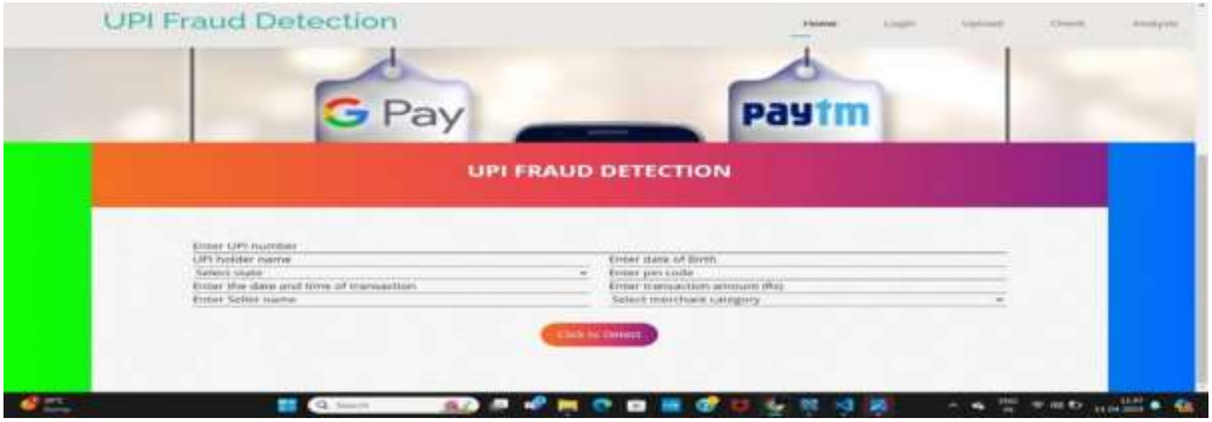


Fig: Transaction Detail

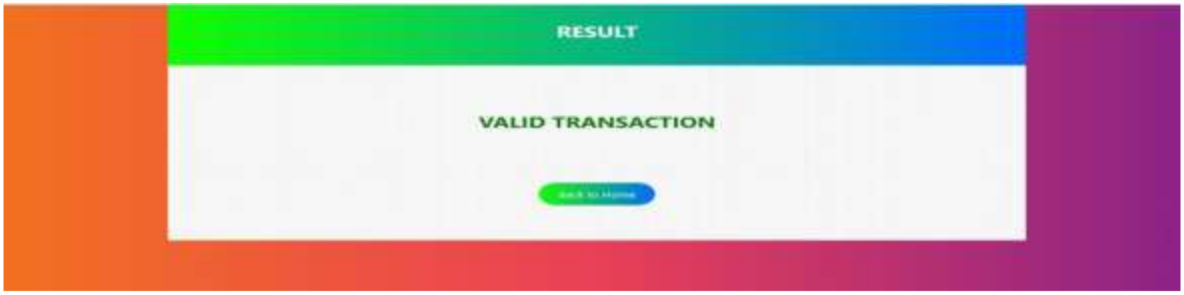


Fig:Fraud Detection Result

CONCLUSION

In this project, we have proposed an application of CNN in UPI fraud detection. The different steps in UPI transaction processing are represented as the underlying stochastic process of an CNN. We have used the ranges of transaction amount as the observation symbols, whereas the types of item have been considered to be states of the CNN. We have suggested a method for finding the spending profile of cardholders, as well as application of this knowledge in deciding the value of observation symbols and initial estimate of the model parameters. It has also been explained how the CNN can detect whether an incoming transaction is fraudulent or not. Experimental results show the performance and effectiveness of our system and demonstrate the usefulness of learning the spending profile of the cardholders. Comparative studies reveal that the Accuracy of the system is close to 80 percent over a wide variation in the input data. The system is also scalable for handling large volumes of transactions.

REFERENCES

1) Amlan Kundu, Suvasini Panigrahi, Shamik Sural and Arun K. Majumdar, “BLAST-SSAHA Hybridization for UPI Fraud Detection,” IEEE Transactions On Dependable And Secure Computing, vol. 6, Issue no. 4, pp.309-315, October-December 2009.

2) Amlan Kundu, Suvasini Panigrahi, Shamik Sural and Arun K. Majumdar, “UPI fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning,” Special Issue on Information Fusion in Computer Security, Vol. 10, Issue no 4, pp.354- 363, October 2009.

3) Abhinav Srivastava, Amlan Kundu, Shamik Sural, Arun K. Majumdar, “UPI Fraud Detection using Hidden Markov Model,” IEEE Transactions On Dependable And Secure Computing, vol. 5, Issue no. 1, pp.37-48, January-March 2008.

4) Peter J. Bentley, Jungwon Kim, Gil-Ho Jung and Jong-Uk Choi, “Fuzzy Darwinian Detection of UPI Fraud,” In the 14th Annual Fall Symposium of the Korean Information Processing Society, 14th October 2000.

5) Sam Maes, Karl Tuyls, Bram Vanschoenwinkel, Bernard Manderick, “UPI fraud detection using Bayesian and neural networks,” Interactive image-guided neurosurgery, pp.261- 270, 1993.

6) Amlan Kundu, S. Sural, A.K. Majumdar, “Two-Stage UPI Fraud Detection Using Sequence Alignment,” Lecture Notes in Computer Science, Springer Verlag, Proceedings of the International Conference on Information Systems Security, Vol. 4332/2006, pp.260- 275, 2006.

7) Simon Haykin, “Neural Networks: A Comprehensive Foundation,” 2nd Edition, pp.842, 1999.

8) Dr .D. Manendra Sai , ET Al.,(2023) “Machine Learning Techniques Based Prediction For Crops In Agriculture”, Journal of Survey in Fisheries Sciences ,Volume 10,Issue 1s, Pages 3710-3717, February 2023 (Scopus, Web of Science).

9) Dr. D. Manendra Sai, Et Al ., (2023)“Utilizing Machine Learning Algorithms For Kidney Diseases Prognosis”, European Journal of Molecular & Clinical Medicine ,Volume 10,Issue 01, Pages 37- 50.