Credit Card Fraud Prediction Using Machine Learning

Dr.S.BELWIN JOEL

[1]Professor,[23]Student

Changu Kana Thakur Arts, Commerce and Science College, New Panvel(Autonomous), Maharashtra-410206

Abstract:-

Nowadays, number of peoples uses online mode of transaction by using credit card on a very large scale. There are lots of transactions are done by the peoples every day. Number of times it may be secure or clear transaction but sometimes there once some issues happened regarding credit card transactions with some peoples or we say credit card users.

In today's world, due to cashless transaction, every human being use ATM or credit card for transaction,so frauds can also be increased. Billions of dollars of loss are caused every year by in a way that involves the action of deceiving someone, especially criminal deception, by credit card transaction. The design of efficient fraud detection algorithms is a key for reducing such losses, and more and more algorithms depends one with full trust and confidence on advanced machine learning techniques to assist or we can say to help by doing a share of the work fraud investigators.

In this research paper, we are going to predict whether a credit card user is fraud or using machine learning algorithm. The key objective of credit card fraud detection system is to identify any event or any activity that seems unusual or out of place i.e. we say suspicious activities and report them to an analyst which letting normaltransactions be automatically done. Since e-commerce websites are becoming more popular, credit card fraud is also becoming most common issue. When a credit card is stolen, it is being used for many dishonest reasons and it is called as card theft. In this paper, we used machine learning algorithm to findfraud transactions. So it is going to be helpful for society to identify the fraud credit card transactions.

Keywords:-
Credit card, Electronic commerce, Fraud detection, Genetic algorithms.

1. Introduction:-

Criminal use of credit card or its information without the knowledge of the owner is referred to as credit card fraud. Number of times it may be confident transaction but sometimes there once some issues happened regarding credit card transactions with some peoples or we say credit card users.

The online shopping is growing day to day. Credit cards are used for purchasing goods and services with the help of virtual card for online transaction and physical card for offline transaction. In a physical-card based purchase, the card holder hand over his/her card physically to a supplier for making a payment. If the cardholder does not realize the loss of card, it can lead to a considerable financial loss to the credit card company. In online payment mode, attackers need only little information for doing untrustworthy transaction (OTP, card number, expiration date etc.)

The Credit card fraud detection features uses user behavior and location scanning to check for uncommon patterns. These patterns include user spending patterns as well as usual user geographic locations to verify his identity. If any uncommon pattern is detected, the system requires renovation. The system analyses useof credit card data for various characteristics. These characteristics include user country, usual spending procedures.

1.1 Objectives:-

- To identify credit card fraud accurately and the number of wrong classifications should be minimum.

- To detect the fraud while it is in transition state.

- To predict whether it is the cardholders or the criminals using credit card.

- To classify whether credit card transaction are genuine or fraudulent using algorithms such as logistic regression etc.

- To identify considerable transactions from regular credit card transaction to detect credit card fraud.

2. Literature Review:-

KuldeepRandhawa et al. [1] proposed a technique using machine learning to detect credit card fraud detection. Initially, standard models were used after that hybrid models came into picture which made use of AdaBoost and majority voting methods. Publically available data set had been used to evaluate the model efficiency and another data set used from the financial institution and analyzed the fraud. Then the noise was added to the data sample through which the robustness of the algorithms could be measured. The experiments were conducted on the basis of the theoretical results which show that the majority of voting methods achieve good accuracy rates in order to detect the fraud in the credit cards. For further evaluation of the hybrid models noise of about 10% and 30% has been added to the sample data. Several voting methods have achieved a good score of 0.942 for 30% added noise. Thus, it was concluded that the voting method showed much stable performance in the presence of noise.

Abhimanyu Roy et al. [2] proposed deep learning topologies for the detection of fraud in online money transaction. This approach is derived from the artificial neural network with in-built time and memory components like long term short term memory and several other parameters. According to the efficiency of these components in fraud detection, almost 80 million online transactions through credit card have been pre-labeled as fraudulent and legal. They have used high performance distributed cloud computing environment. The study proposed by the researchers provides an effective guide to the sensitivity analysis of the proposed parameters as per the performance of the fraud detection. The researchers also proposed a framework for the parameter tuning of Deep Learning topologies for the detection of fraud. This enables the financial institution to decrease the losses by avoiding fraudulent activities.

ShiyangXuan et al. [3] used two types of random forests which train the behavior features of normal and abnormal transactions. The researcher compares these two random forests which are differentiated on the basis of their classifiers, performance on the detection of credit card fraud. The data used is of an e-commerce company of China which is utilized to analyze the performance of these two types of random forests model. In this paper, the author has used B2C dataset for the identification and detection of fraud from the credit cards. Therefore, the researcher concluded from the result that the proposed random forests provide good results on

small dataset but there are still some problems like imbalanced data which makes it less effective than any other dataset.

Zahra Kazemi et al. [4] proposed Deep autoencoder which is used to extract the best characteristics of the information from the credit card transaction. This will further add softmax software to resolve the class labels issues. An overcompleteautoencoder is used to map the data into a high dimensional space and a sparse model was used in a descriptive manner which provides benefits for the classification of a type of fraud. Deep learning is one of the most motivated and powerful techniques being employed for the detection of fraud in the credit card. These types of networks have a complex distribution of data which is very difficult to recognize. Deep autoencoder has been used in some stages to extract the best features of the data and for the classification purposes. Also, higher accuracy and low variance are achieved within these networks.

John O. Awoyemi et al. [5] proposed an investigation through which the performances of several algorithms were evaluated when they were applied on credit card fraud data that is highly skewed. The European cardholders' 284,807 transactions were used as a source to generate the dataset of credit card transactions. On the skewed data, a hybrid approach of under-sampling and oversampling is performed. On raw and preprocessed data, there are three different techniques applied in Python. Based on certain parameters like precision, sensitivity, accuracy, balanced classification rate and so on, the performances of these techniques are evaluated. It is seen through the achieved results that in comparison to naïve Bayes and logistic regression approaches, the performance of k-NN is better.

SharmisthaDutta et al. [6] presented a study on the commonly found crime within the credit card applications. There are certain issues faced when the existing non-data mining approaches are applied to avoid identity theft. A novel data mining layer of defense is proposed for solving these issues. For detecting the frauds within various applications, two algorithms named Communal Detection and Spike Detection which generate novel layer. There is a large moving window, higher numbers of attributes and numbers of link types available which can be searched by CD and SD algorithms. Thus, results can be generated by the system by consuming a huge amount of time. Since the attackers do not get time to modify their behaviors with respect to the algorithms being deployed in real time, there is no true evaluation achieved even after a regular update of the algorithms. Therefore, it is not possible to properly demonstrate the concept of

adaptability. These issues can be resolved by making certain enhancements in the proposed algorithm in future work.

Krishna Modi et al. [7] investigated several techniques that were used for detecting the fraudulent transactions and provided a comparative study amongst them. The fraudulent transactions can be detected by utilizing either one of these or integrating any of these methods. The model can possibly be trained in a more accurate manner by adding new features. Several data mining techniques are being used by bank and credit card companies for detecting fraud behaviors. The normal usage pattern of clients depending upon their past activities can be identified by applying any of these methods. Therefore, a comparative analysis is made here by studying different fraud detection techniques proposed over the years.

3. Methodology:-

3.1   Sample Design:

In this research, we have conducted a survey online with the help of Google forms for the people who have/haven't credit cards and collected information through well-structured questionnaires. We had created a set of questionnaire. This form was shared with people. From which we got 83 numbers of responses from people.

3.2   Data Collection:

Following are the set of questionnaires from which we collected the data:

1. Do you have credit card?

2. Is credit card safe to use?

3. It is more helpful than debit card?

4. While doing payment OTP on your mobile is beneficial or not?

5. Are you satisfied with your domestic credit card?

6. Have you ever transferred a balance to your credit card?

7. How did you apply for your credit card?

8. How many times do you use your domestic credit card per month?

9. What are the benefits of your credit card in your opinion?

10. Would you prefer to use credit card in your day to day life?

3.3 Algorithm :-

In this paper, Random Forest Classifier Algorithm is used to count an accuracy. Random Forest is a popular machine learning algorithm which belongs to the supervised learning technique. It can be used in Machine Learning for both Classification and Regression problems. It is based on the concept of group learning, which is a process of combining multiple classifiers to solve a composite i.e. complex problem and for better performance of the model.

As the name suggests, "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the prophetical i.e. predictive accuracy of that dataset." Instead of depending on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.The greater number of trees in the forest leads to more accuracy and prevent the problem of over fitting.

Random Forest Classifier Algorithm is used to count an accuracy. Accuracy in this algorithm is depends on the correlation between any two trees in the forest.

4. Data Analysis:
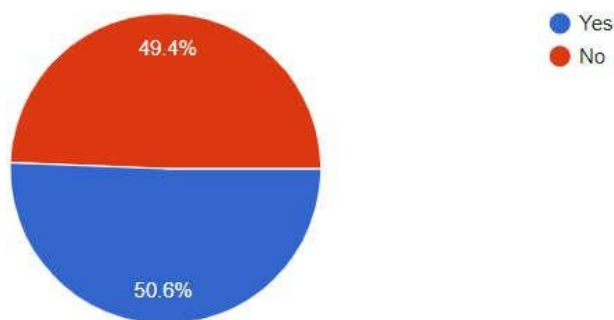
1. Do you have credit card?

fig.1

In fig.1, according to a survey of peoples through google form, among 83 collected responses, 50.6% peoples have credit card, and 49.4% peoples haven't credit card.
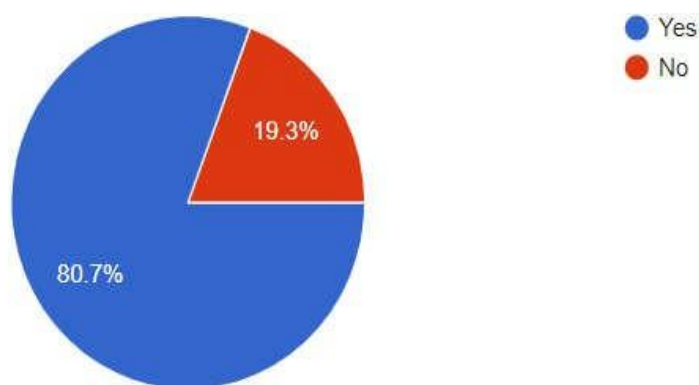
2) Is credit card safe to use?

fig.2

In fig.2, around 80.7% peoples says credit card is safe to use for them, and 19.3% peoples says credit card is not safe for them.
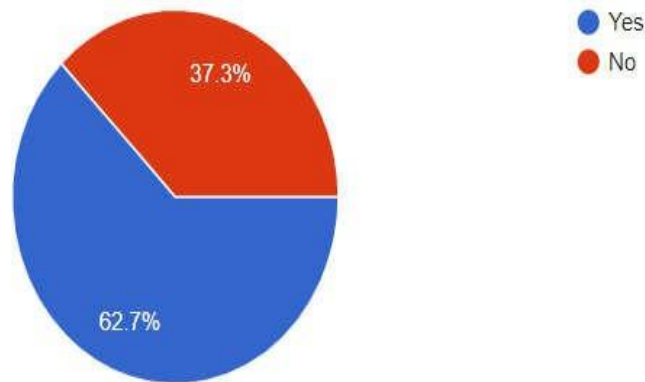
3) Is it more helpful then debit card?

fig.3

In fig.3, according to a survey, 62.7% peoples say credit card is more helpful than debit card, and 37.3% say credit card is not helpful than debit card.

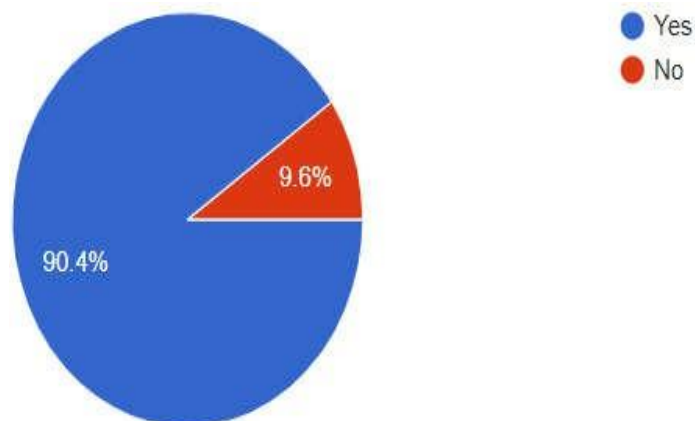4) While doing payment OTP on your mobile is beneficial or not?

fig.4

In fig.4, round 90.4% peoples say OTP is beneficial while doing any payment, and 9.6% peoples say OTP is not beneficial while doing any payment.
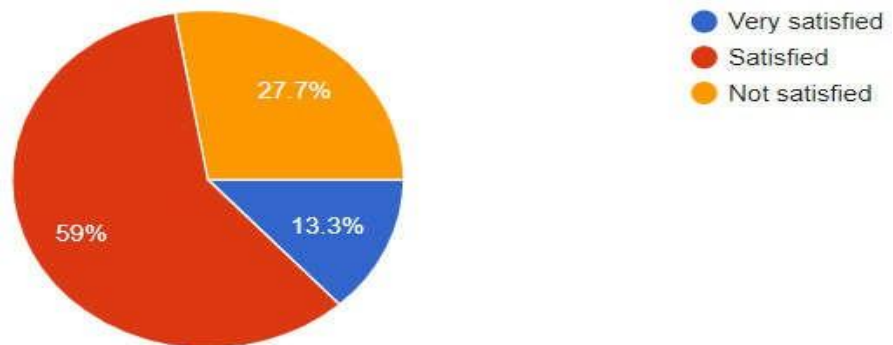
5) Are you satisfied with your domestic credit card?

fig.5

In fig.5, around 59% peoples are satisfied with their domestic credit card, 13.3% peoples are very satisfied with their domestic credit card, and 27.7% peoples are not satisfied with their domestic credit card according to a survey.

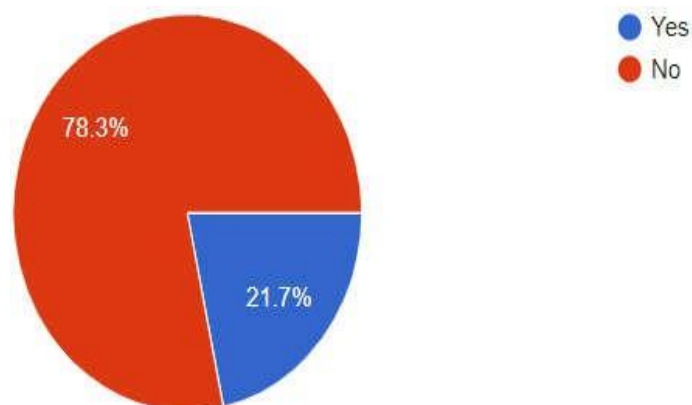6) Have you ever transferred a balance to your credit card?

fig.6

In fig.6, from the survey, among 83 responses, 78.3% peoples have transferred a balance to their credit card, and 21.7% peoples haven't transferred a balance to their credit card.

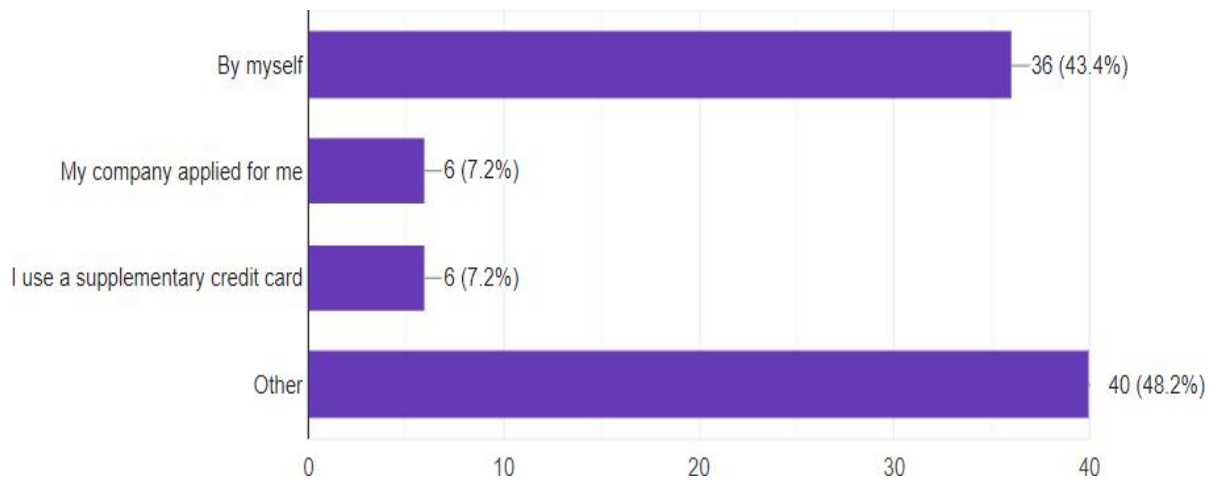7) How did you apply for your credit card?

fig.7

This fig shows, around 43.4% peoples applied for credit card by themselves, company applied for credit card to 7.2% peoples, then 7.2% peoples use a supplementary credit card, and 48.2% replied other.

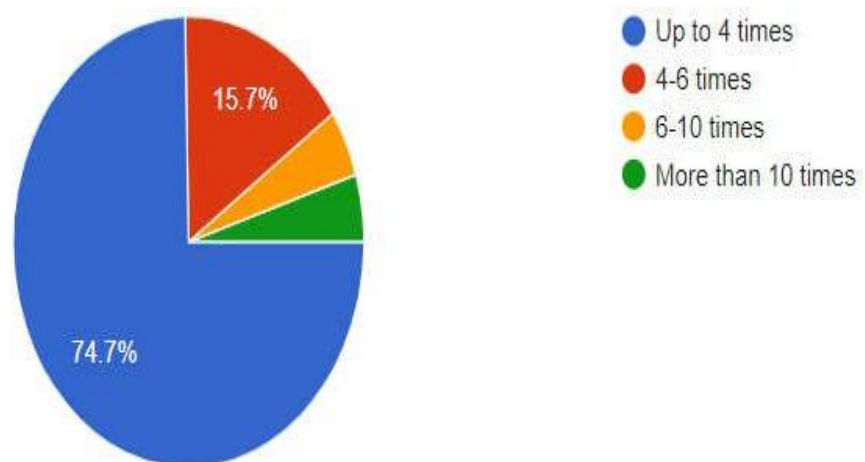8) How many times do you use your domestic credit card per month?

fig.8

In fig.8, 74.7% peoples uses credit card up to 4 times per month, 15.7% peoples uses credit card 4-6 times per month, and remaining 9.6% peoples peoples uses credit card 6-10 times and more than 10 times per month.

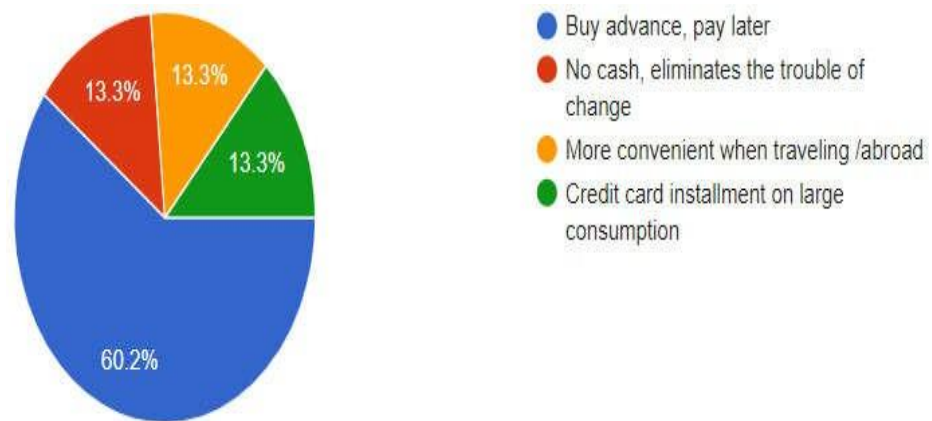9) What are the benefits of your credit card in your opinion?

fig.9

This fig. shows, beneficial level of credit card according to user's opinion. 60.2% peoples says buy advance, pay later is beneficial for them, 13.3% peoples says no cash, eliminates the trouble of change is beneficial for them, 13.3% peoples says credit card is more convenient when travelling/abroad, and 13.3% says credit card installment on large consumption is beneficial for them

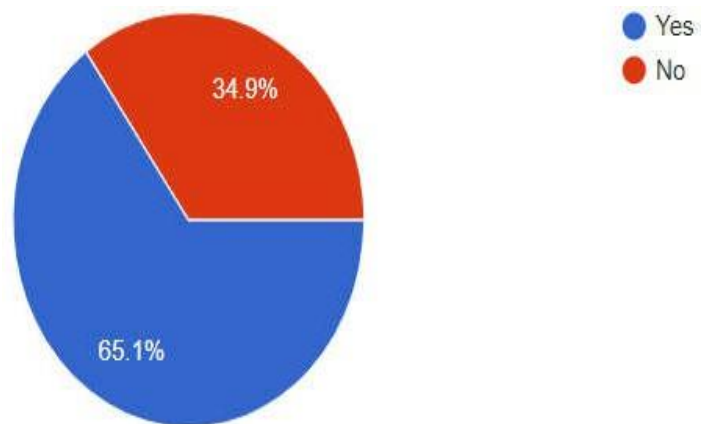10) Would you prefer to use credit card in your day to day life?

fig.10

In fig. 10, among 83 responses, 65.1% peoples preferred to use credit card in your day to day life and 34.9% peoples are not preferred to use credit card in your day to day life.

5. Result and Discussion:-

Questionnaires prepared for the survey are framed by a combination of open-ended and multiple-choice questions with demographic questions. Awareness about Credit card fraud transactions especially in rural area's people survey is conducted through google form and around 83 responses are collected through the survey.

Data validation ensures that online survey questionnaires are complete and provide consistent data.

In this research, we have done the data analysis through survey. According to the survey, we found that among 83 collected responses, 50.6% peoples have credit card, and 49.4% peoples haven't credit card.

5.1 Conclusion:-

Preventing known and unknown fraud in real time is not easy but it is feasible. The proposed architecture was originally designed to detect credit card fraud in online payments and emphasize on providing a fraud prevention mechanism to verify the transaction is fraud or genuine.

This research paper outlined the most frequent types of frauds, as well as how to detect them. It also includes an explanation of how machine learning can be used to improve fraud detection findings, as well as the algorithms. Algorithm provides accurate in deducting fraudulent transaction and minimizing the number of false alerts.

If these algorithms applied into bank credit card fraud detection system, the probability of fraud transaction can be predicted soon after credit card transaction. And a series of anti-fraud strategies can be adopted to prevent banks from great losses and reduce risks.

Fraud detection is a complex issue that requires a substantial amount of planning before throwing machine learning algorithms at it. Nonetheless, it is also an application of

data science and machine learning for the good, which makes sure that the customer's money is safe and not easily tampered with.

References:-

[1] KuldeepRandhawa, Chu Kiong Loo, ManjeevanSeera, CheePeng Lim and Asoke K. Nandi, "Credit card fraud detection using AdaBoost and majority voting," IEEE Access, vol. 6, pp. 14277-14284, 2018. [2] A. Roy and J. Sun and R. Mahoney and L. Alonzi and S. Adams and P. Beling, "Deep learning detecting fraud in credit card transactions," in Systems and Information Engineering Design Symposium (SIEDS), pp. 129-134, 2018.

[3] Guanjun Liu, Zhenchuan Li, LutaoZheng, Shuo Wang and Changjun Jiang ShiyangXuan, "Random Forest for Credit Card Fraud Detection," in IEEE 15th International Conference On Networking, Sensing and Control (ICNSC), pp.1-6, 2018.

[4] Zarrabi, H. Kazemi, "Using deep networks for fraud detection in the credit card transaction," IEEE 4th International Conference In Knowledge-Based Engineering and Innovation (KBEI), pp. 0630-0633, 2017.

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

[6] S. Dutta, A. K. Gupta and N. Narayan, "Identity Crime Detection Using Data Mining, "3rd International Conference on Computational Intelligence and Networks (CINE), Odisha, pp. 1-5, 2017.

[7] K. Modi and R. Dayma, "Review on fraud detection methods in credit card transactions, "International Conference on Intelligent Computing and Control (I2C2), Coimbatore, pp. 1-5, 2017.