e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Conference:01/December-2022 Impact Factor- 6.752 www.irjmets.com
1st National Conference on Applications of soft Computing Techniques in Engineering NCASCTE-2022
Organized by Department of Electrical & Electronics Engineering, ACE Engineering College, Hyderabad

CREDIT CARD FRAUD DETECTION USING MACHINE LEARNING

Dr. Khaleel Ur Rahman Khan*1, A. Jahnavi*2, A. Nikitha*3, C. Ravishankar*4

*1Profesor, Computer Science Engineering, Ace Engineering College, Ghatkesar, Telangana, India.

*2,3,4Student, Computer Science Engineering, Ace Engineering College, Ghatkesar, Telangana, India.

DOI: https://www.doi.org/10.56726/IRJMETS-NCASCTE202220

ABSTRACT

The detection of credit card theft is undoubtedly the most common problem in the modern day. This is a result of growing internet shopping and electronic commerce platforms. Since online payment methods save time and are more practical for transportation as digitalization grows more pervasive in today's world, more people are choosing them. The use of credit cards for e-commerce has skyrocketed, which has greatly boosted credit card theft. Fraudsters attempt to profit from the card and the transparency of online transactions. Credit card fraud typically happens when a card is misplaced or stolen, or when the card's details are known. Thus, it is essential to put a halt to fraudsters' actions. Today's population faces a wide range of credit card issues. use K-Nearest Neighbor (KNN) and A few approaches and machine learning algorithms can be utilized to address this issue including the can be utilized to address this issue include Knn classifier. Protecting electronic payments is the fundamental goal in order for people to use e-banking conveniently and safely.

Keywords: Credit Card, Machine Learning, Detection, K-Nearest Neighbor, KNN Classifier.

I. INTRODUCTION

Credit card fraud is one of the major threats now facing businesses. But in order to effectively combat the fraud, it's imperative to first understand how it's executed. When committing fraud, credit card thieves employ a wide range of techniques. The use of another person's credit card for personal expenses without the cardholder's or card issuer's knowledge is referred to as credit card fraud. Furthermore, the cardholder or issuer is not related to the person using the card, and the person using the card has no intention of contacting them or returning the purchases they made. Credit card fraud is done using the following techniques: unauthorized or illegal use of a credit card account. giving false account information to obtain goods or services Using an unauthorized account or personal information to commit a crime of fraud (intentional misrepresentation). Since attackers can obtain a sizable sum in a short period of time with little risk and the fraud is found after a few days, credit cards are regarded as "excellent targets of fraud". CNP fraud, also known as card, not present fraud, is when a con artist tries to trick the network by pretending to be someone else. Both legitimate mail orders and web shops are impacted by the use of the internet and the postal service as essential conduits. Through the skimming method, they are able to extract private information from a credit card that belonged to someone else and was used in a typical purchase.

II. METHODOLOGY

The K-nearest neighbors (KNN) algorithm forecasts the values of fresh purpose information based on "feature similarity," meaning the new knowledge point will be given a price supported regardless of how closely it resembles the points in the coaching set. Several anomaly detection methods have employed the K-nearest neighbors analysis paradigm.

Existing System

In the current system, fraud is discovered after a robbery has occurred, fraud is discovered following a cardholder complaint, and because all transactions are recorded in a log, we are required to keep a large amount of data. Since most transactions these days take place online, we only record the IP address for verification. Therefore, cybercrime must contribute to the investigation of the scam. We provide a system to identify fraud in the best and simplest way possible to avoid all of the aforementioned disadvantages. They have used the Support Vector Algorithm. However, they are having some trouble detecting credit card fraud.

Proposed System (k-nearest neighbors Algorithm)

The K-Nearest Neighbor method, a supervised learning technique in which the outcome of a new instance query is classified based on the majority of the K-Nearest Neighbors category, is one of the finest classifier algorithms that have been utilized in credit card fraud detection.

Three primary elements affect how well the KNN algorithm performs:

1. The measurement of distance used to identify the closest neighbors.

2. The distance formula used to the k-nearest neighbor's categorization.

3. The number of neighbors categorized the fresh sample

In the process of KNN, we classify any incoming transaction by calculating the nearest point to a new one. Then if the nearest neighbors were fraudulent, then the transaction indicates fraud. The value of K is used as, a small and odd to break the ties (typically 1, 3, or 5). Larger K values can help to reduce the effect of the noisy dataset. In this algorithm, the distance between two data instances can be calculated in different ways.

KNN is a lazy algorithm that simply stores the complete data set and executes the task when new data is entered. It does not develop a hypothesis based on the training data. The K-nn Algorithm diagram is displayed. By calculating the distance between the new data point and the existing data points, we may determine which category the new data point belongs to.

1.Euclidean = square root((x2-x1)^2 – (y2-y1)^2)

2.Manhattan = |(x2-x1)|+|(y2-y1)|

Algorithm:

1. Correct the value for "k."

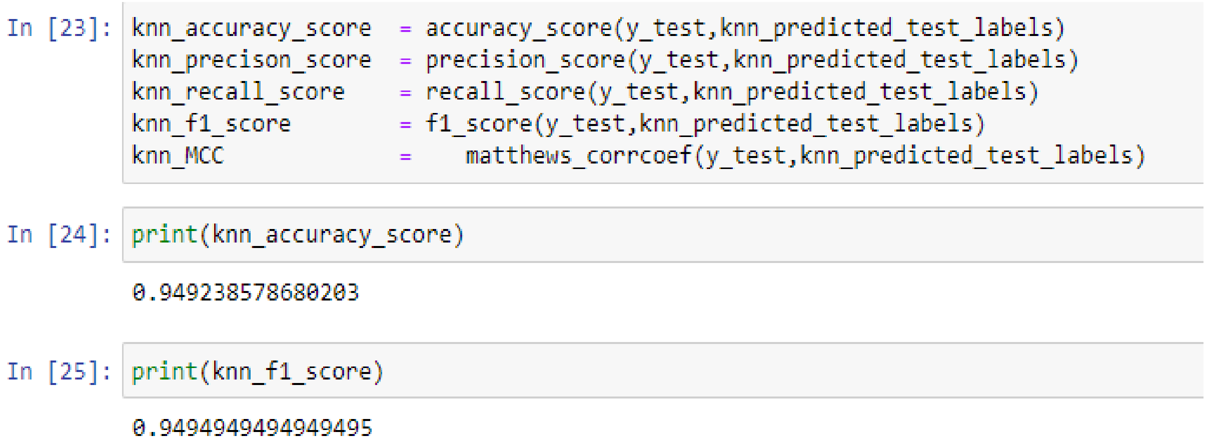
2.Utilize a distance formula to determine the separation between each point in the dataset (Euclidean, manhattan)

3. Based on the distance, locate the "k" nearest points.

4. The new point behaves similarly to the closest point to it, therefore find the distance between it and every other point in the dataset.

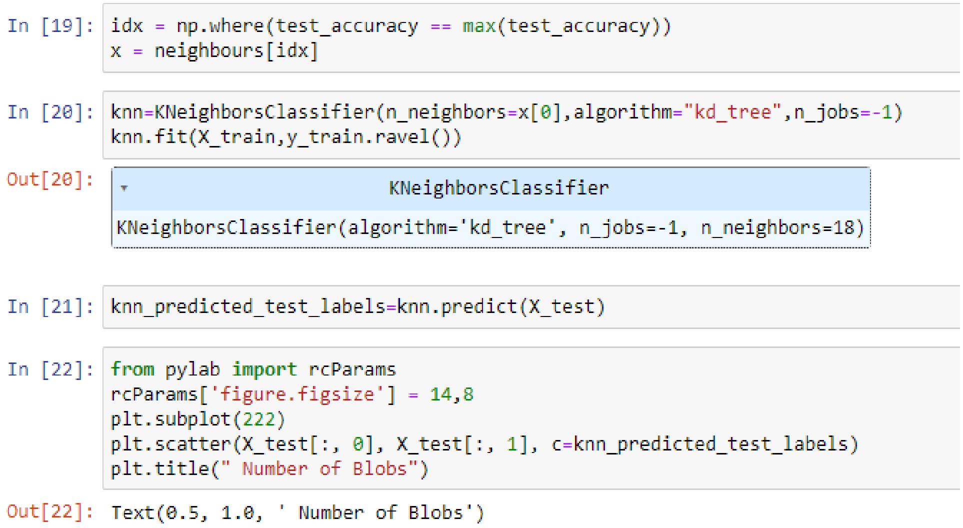
III. MODELING AND ANALYSIS

The accuracy for the K-nearest neighbor event is 0.949238578680203. Accuracy and f1-score for non-fraudule nt transactions are comparable



```
In [23]: knn_accuracy_score  = accuracy_score(y_test,knn_predicted_test_labels)
         knn_precison_score  = precision_score(y_test,knn_predicted_test_labels)
         knn_recall_score    = recall_score(y_test,knn_predicted_test_labels)
         knn_f1_score        = f1_score(y_test,knn_predicted_test_labels)
         knn_MCC             =    matthews_corrcoef(y_test,knn_predicted_test_labels)

In [24]: print(knn_accuracy_score)

         0.949238578680203

In [25]: print(knn_f1_score)

         0.9494949494949495
```

Figure 1: output of Accuracy and f1-score



```
In [19]: idx = np.where(test_accuracy == max(test_accuracy))
         x = neighbours[idx]

In [20]: knn=KNeighborsClassifier(n_neighbors=x[0],algorithm="kd_tree",n_jobs=-1)
         knn.fit(X_train,y_train.ravel())

Out[20]:                KNeighborsClassifier
         KNeighborsClassifier(algorithm='kd_tree', n_jobs=-1, n_neighbors=18)

In [21]: knn_predicted_test_labels=knn.predict(X_test)

In [22]: from pylab import rcParams
         rcParams['figure.figsize'] = 14,8
         plt.subplot(222)
         plt.scatter(X_test[:, 0], X_test[:, 1], c=knn_predicted_test_labels)
         plt.title(" Number of Blobs")

Out[22]: Text(0.5, 1.0, ' Number of Blobs')
```

Figure 2: finding of knn classifier

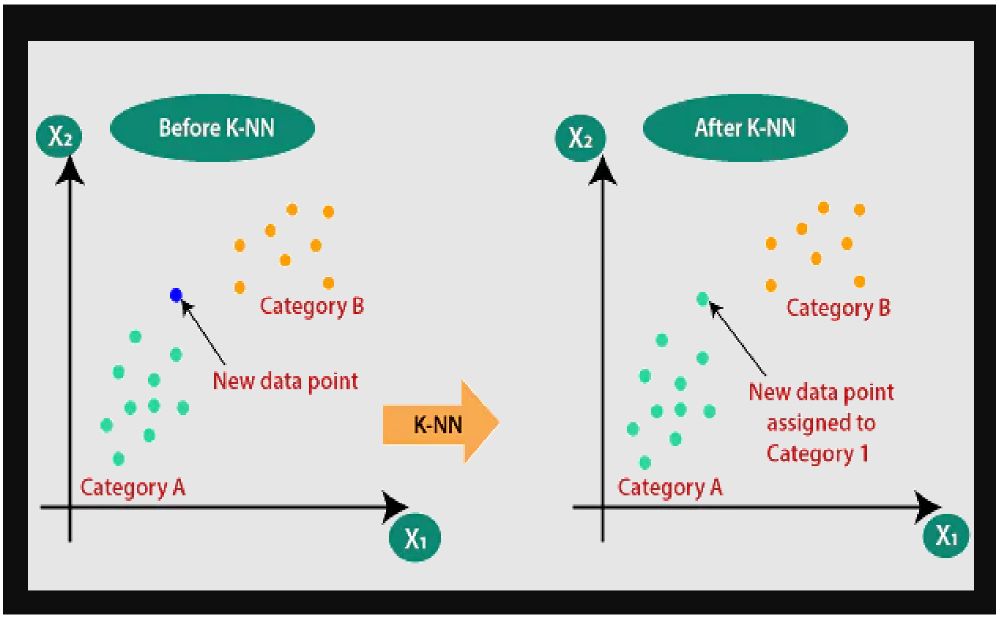


Figure 3: Knn Algorithm

IV. RESULTS AND DISCUSSION

Accuracy and f1-score are similar for non_fraudulent transactions and differ from that of fraud cases. Along with that Scatter plotted graph and the line graph is 0.94.

Table 1. Comparison of displacement of all 4 cases

SN.	Type	Values
1	Accuracy	0.949238578680203
2	Precision	0.9494949494949495
3	Recall	0.9494949494949495
4	F1-Score	0.9494949494949495

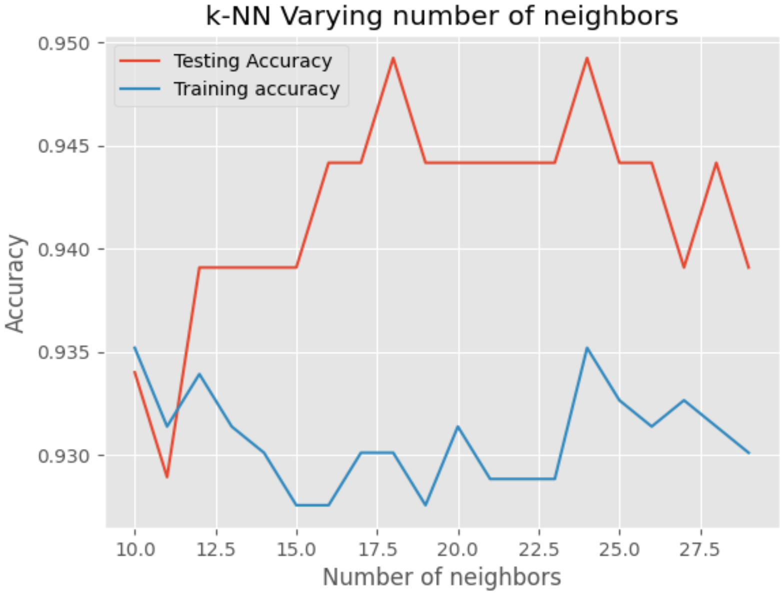


Figure 4: line graph for Accuracy

V. CONCLUSION

Since it is currently not feasible to compile reports for individual transactions, this project's goal is to identify frauds when data is collected in big quantities. The goal of this project is to quickly and accurately identify fraud in transactions on a personal level, alerting the user to the implications at that moment. In summary, because mid-sized enterprises will have a lot of data, this project is appropriate to them. With the increase in people using credit cards for transactions, the likelihood of credit card fraud is significantly increasing. The identification of credit card fraud using a range of machine learning techniques is examined in this research utilizing an online dataset. In PYTHON, the proposed machine is operated. The dataset analysis produced the highest accuracy KNN charge.

ACKNOWLEDGEMENTS

Without the supervision of our mentor Dr. Khaleel Ur Rahman Khan and project coordinator Mr. Prem Kumar, the research for the survey could not have been effectively completed. For his invaluable and consistent support during the completion of this study, Dr. M. V. VIJAYA SARADHI, Head of the Department of Computer Science and Engineering at Ace Engineering College, has our sincere gratitude.

VI. REFERENCES

[1] "Credit Card Fraud Detection" IJERT NREST - 2021 Conference Proceedings, Munira Ansari, Hashim Malik, Siddhesh Jadhav, and Zaiyyan Khan J.

[2] Neetha Natesh, Vanishree Abhay, Satish B Basapur, and Vinaya D S 2018 Systems and Information Engineering Design Symposium (SIEDS), 2019. "Deep Learning Detecting Fraud in Credit Card Transactions."

[3] Credit Card Fraud Detection, Kaggle Machine Learning Group ULB, 03 May 2021.

[4] M. Pichler, V. Boreux, A. Klein, M. Schleuning, and F. Hartig, "Machine learning approaches to infer traitmatching and predict species interactions in ecological networks," Methods in Ecology and Evolution, vol. 144, no. 1, 2019, pp. 1–13.

[5] O. Harrison, "K-Nearest Neighbors Algorithm for Machine Learning Fundamentals," 1. [Accessed: September 18, 2019].

[6] KUMAR, V., Yuvraj, S., Vengatesan, K., and Sabnis, S., 2020. Detecting credit card fraud using data analysis methods. Scientific Journal, 9(3), 1185–1196. Advances in Mathematics.

[7] S. P. Maniraj, A. Saini, S. Ahmed, and S. Sarkar, 2019. combining data science and machine learning to

detect credit card fraud. Engineering Research and International Journal, 8 (09).

[8] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, October 2017. A comparison of machine learning methods for credit card fraud detection 2017 saw the first-ever ICCNI (International Conference on Computing, Networking, and Informatics) (pp. 1-9). IEEE.

[9] May 2020, Sailusha, R., Gnaneswar, V., Ramesh, and Rao, G.R. Machine learning for the identification of credit card fraud. 2020 will see the fourth iteration of the International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 1264-1270). IEEE.