CREDIT CARD FRAUD DETECTION USING MACHINE LEARNING

**Abstract –** *In the recent years, credit card is the most preferred payment mode by the customers. With the advancement of modernized technology, the number of credit card fraud cases are increasing rapidly, causing severe damage to financial institutions and individuals. Therefore, it is very crucial for the financial institutions to detect and prevent these fraudulent activities. Credit card fraud generally happens when the card is either stolen/lost or a fraudster accesses the credit card information of the user. The main aim of this project is to check whether the transactions made are genuine or fraud by implementing the supervised machine learning algorithms such as Naïve Bayes, Decision tree, K-Nearest Neighbors on the highly imbalanced dataset. Finally, we will compare the accuracy, precision, recall, F1-score results and choose the best algorithm to identify an optimal solution.*

**Keywords** credit card fraud, supervised machine learning, Naïve Bayes, Decision tree, K-Nearest Neighbors

**1.INTRODUCTION**

As the technology is advancing rapidly, the usage of credit cards gets increased as they ease our daily transactions in many ways. This leads to the raise in number of fraudulent transactions. Credit card fraud means the act of committing fraud using another person’s credit card to make purchases or request cash advances without the cardholder’s knowledge or consent. To commit credit card fraud, the fraudsters try to steal the card or steal sensitive information such as bank account, credit card number, social security number using phishing techniques and the users may not know whether their credit card information was leaked. The fraudsters try to make every fraudulent transaction legitimate which makes the fraud detection a challenging problem [4].

According to Nilson Report, the US accounted for 23.02% of global card volume in 2021 but 36.83% of worldwide losses to card fraud. Higher fraud losses in the US in 2021 were attributable to a 25.0% increase in purchase volume on credit cards after an 8.8% drop in 2020 as well as continued growth in card-not-present (CNP) transactions, which are more vulnerable to fraud [9].

Fraud can be avoided in two main ways: prevention and detection. Prevention avoids the attacks happening from the fraudsters and if the prevention has already failed then detection happens. Detection helps in identifying and alerting as soon as a fraudulent transaction takes place [1].

Credit card fraud can be classified into several categories. But in a set of transactions the two types of frauds we can observe mainly are Card-present fraud and card-not-present fraud [1]. In Card-present fraud the fraudulent party uses the stolen or counterfeit credit card. Credit-not-present fraud is a scam where the fraudster tries to make a fraudulent transaction without possessing the physical card, which is done by stealing credit card information. [16] These frauds can be further divided into types such as Application fraud, Account takeover, social engineering fraud, skimming, phishing, Information sharing etc. Our goal is to address these frauds and propose a method to detect those frauds in real time.

Researchers have been finding ways to tackle these issues, as the banking data such as credit card fraud and default payments are quite of a challenge [2]. The credit card dataset is highly imbalanced because it contains more legitimate transactions than the fraudulent one [4]. Data mining and machine learning algorithms are the solutions that can work on large datasets which is not easy for human beings, and it is the best way to find whether a transaction is fraudulent or not.

In this paper, we first perform data preprocessing on the dataset by cleaning the data, converting categorical variables into numeric, standardising and then use Spearman’s Correlation and we implement three supervised machine learning algorithms on the preprocessed dataset, compare their results and produce an efficient solution for the fraud detection.

**2. LITERATURE REVIEW:**

In earlier studies, many approaches have been proposed to find solutions to detect credit card fraud using supervised and unsupervised machine learning algorithms and data mining techniques.

In paper [1], we take use of predictive analytics done by the implemented machine learning models and an API module to decide if a particular transaction is genuine or fraudulent. We also assess a novel strategy that effectively addresses the skewed distribution of data.

Algorithms used: Logistic Regression, Support Vector Machine, Naïve Bayes.

In paper [2], we used a Multiple Classifiers System (MCS) on these two data sets: (i) credit card frauds (CCF), and (ii) credit card default payments (CCDP). The MCS employs a sequential decision combination strategy to produce accurate anomaly detection.

There are three combination strategies when employing MCS: (i) sequential combination, (ii) parallel combination, and (iii) hybrid combination.

Algorithms used: Naïve Bayes, C4.5, K-Nearest Neighbour, Artificial Neural Network, and Support Vector Machine.

By using the proposed MCS on CCF data, they achieved the highest TPR of 0.872 for the minority class and the proposed approach also gave a good accuracy of 0.999 and a TNR of 1.000. The MCS was also tested on CCDP data set and obtained the highest TPR for the minority class, which is 0.840. The proposed approach also achieved an accuracy of 0.930 and a TNR of 0.955.

In paper [3], we will compare the performance of eight machine learning methods applied to credit card fraud detection and identify their weaknesses. More precisely, we will compare the imbalanced classification approaches and study how effective they are in the case of extreme imbalance.

Algorithms used: C5.0, Support Vector Machine, Artificial Neural Network, Naïve Bayes, Bayesian Belief Network, Logistic Regression, K-Nearest Neighbour, Negative Selection Algorithm. We found that LR, C5.0 decision tree algorithm, SVM and ANN are the best methods according to the 3 considered performance measures (Accuracy, Sensitivity and AUPRC).

In paper [4], we apply many supervised machine learning algorithms to detect credit card fraudulent transactions using a real-world dataset. Furthermore, we employ these algorithms to implement a super classifier using ensemble learning methods. We identify the most important variables that may lead to higher accuracy in credit card fraudulent transaction detection. Additionally, we compare and discuss the performance of various supervised machine learning algorithms that exist in literature against the super classifier that we implemented.

Algorithms used: Random Forest, Support Vector Machine, K-Nearest Neighbour, XG boost.

Paper [5], proposes different machine learning algorithms such as Decision Tree classifier, Random Forest, Logistic Regression, Naïve Bayes. Results shows that Random Forest classifier performs best among all with 96.7741% accuracy, 100% precision, 91.1111% recall.

**3.EXPERIMENTAL METHODOLOGY:**

In this section, we will discuss the methodology adopted in this project to classify the fraudulent transactions from the non-fraudulent transactions. Figure 1 shows the workflow and steps of our implementation. However, before we discuss the different steps of the methodology used in this work, we will first discuss about the dataset.

**3.1 DATASET**

The Dataset [8] used for our project consists of 307511 records in total, of which 24825 of them are fraudulent cases i.e., 8.0728% of fraud cases resulting in high class imbalance. The dataset consists of attributes such as Income\_Total, AMTAPPLICATION, AMT\_CREDIT and around 122 columns, of which 16 of them are containing categorical values. The target values are given in as 0 or 1 (class 0 means genuine transaction and class 1 means fraud transaction.)

**3.2 DATA PREPROCESSING**

Data preprocessing is a data mining technique used to turn raw data into clean information. It is necessary before training the machine learning models as raw data can be inconsistent or incomplete in its formatting. Effectively preprocessing raw data can increase its accuracy, which can increase the quality of projects and improve its reliability.

**3.2.1 Encoding Categorical values**

Another essential step in the data preprocessing is, converting all the categorical variables into numeric form because most of the machine learning algorithms allow features only in the numerical form. Categorical features are most often strings and therefore we are using LabelEncoder to encode them into its integer correspondents. The dataset consists of 16 columns containing categorical values. For a feature with two categories, the categories are assigned a numeric value of 1 or 0.

**3.2.2 Data cleaning**

Data cleaning is a crucial process in Data Mining. Filling the missing values and null values are the tasks we perform during the data cleaning process. There are 9152465 null values or missing values in the dataset containing 307511 transactions in total. Here, we are using mean method to fill the missing values in our dataset.

**3.2.3 Feature scaling**

Standardisation or Z-score normalisation is a scaling technique whereby the values in a column are rescaled so that they demonstrate the properties of a standard Gaussian distribution, that is mean = 0 and variance = 1. Here we use StandardScalar which is the Scikit-learn function for standardisation [10].

The standard score of a sample x is calculated as:

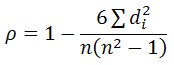
z = (x - u) / s

where **u** is the mean of the training samples or zero if with\_mean=False, and **s** is the standard deviation of the training samples or one if with\_std=False [11].

**3.2.4 Data Reduction**

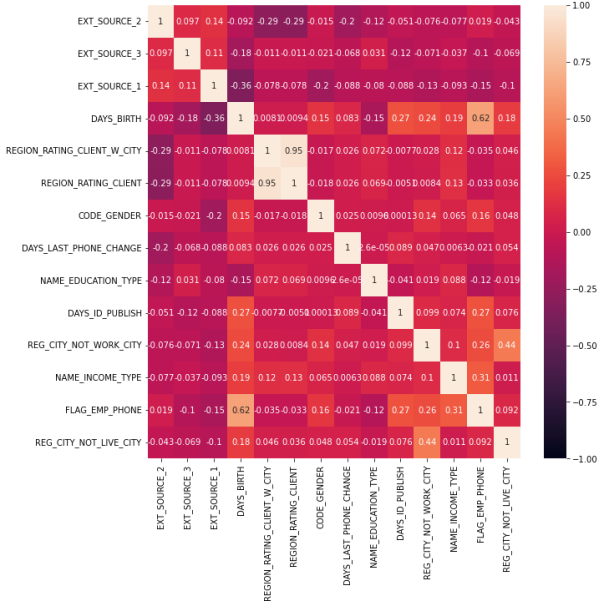
All the features we obtain in the dataset might not be beneficial in building a machine learning model to make necessary prediction. Using some of the features might improve the prediction accuracy. So, feature correlation performs a tremendous purpose in creating a better machine learning model. Features with high correlation are more likely to be linearly dependent and have almost the same impact on the dependent variable.  
Therefore, when two features produce a high correlation, we can drop one of the two features [7]. In our project we have used Spearman’s correlation.

Spearman’s correlation coefficient is a non-parametric measure of the strength and direction of association that exists between two variables measured on at least an ordinal scale.



**Spearman’s correlation coefficient formula**

where n = total number of observations, di = (xi-yi) where xi and yi are the observations [12].



**Figure 1: Correlation heatmap of attributes selected after attribute reduction**

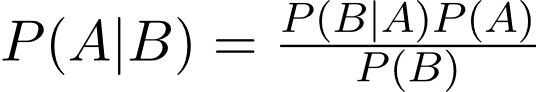
**3.3 MACHINE LEARNING MODELS**

In this project, we have used supervised machine learning algorithms to classify the fraudulent transactions. Before implementing the machine learning algorithms, we have divided our dataset into train data (70%) and test data(30%).

**Naïve Bayes**

Naive Bayes classifiers are a collection of classification algorithms based on **Bayes’ Theorem**. They are a collection of classification algorithms or a family of algorithms where all of them share a common principle, i.e., every pair of features being classified is independent of each other.

**Bayes’ Theorem:** Bayes’ Theorem finds the probability of an event occurring given the probability of another event that has already occurred. Bayes’ theorem is stated mathematically as the following equation:



where A and B are events and P(B) ≠ 0.

We used Gaussian Naïve Bayes for our project: In Gaussian Naive Bayes, continuous values associated with each feature are assumed to be distributed according to a **Gaussian distribution**. A Gaussian distribution is also called Normal Distribution. When plotted, it gives a bell-shaped curve which is symmetric about the mean of the feature values. [13]

**Decision Tree**

[14] Decision Tree is a **Supervised learning technique** that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree-structured classifier, where **internal nodes represent the features of a dataset, branches represent the decision rules** and **each leaf node represents the outcome.**

We have used decision tree for our project because:

* Decision Trees usually mimic human thinking ability while deciding, so it is easy to understand.
* The logic behind the decision tree can be easily understood because it shows a tree-like structure.

**K-Nearest Neighbours**

[15] KNN is a non-parametric classification method that is used to solve classification and regression problems. KNN is termed as a lazy algorithm as it does not do any generalization therefore training process is pretty much fast.

## *Why do we need a K-NN Algorithm?*

Suppose there are two categories, i.e., Category A and Category B, and we have a new data point x1, so this data point will lie in which of these categories. To solve this type of problem, we need a K-NN algorithm.



The Distance between the 2 points is calculated by Euclidean Distance:

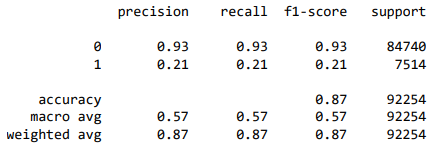
**d = √ [(x2 – x1)2 + (y2 – y1)2]**

**4. FINDINGS AND RESULTS:**

Here we have measured accuracy, precision, recall and f1 score for all three algorithms and confusion matrix for finding the best model for our project.

Now we will see the results after performing data mining techniques and compare them with other algorithms as we need to get knowledge about which algorithm works perfectly.

Firstly, we apply the dataset for the *Naïve Bayes* model and the results are as follows:



**Figure 2: output for Naïve Bayes**

The precision, recall, f1 score are same for that of the non-fraud cases and differ for that of the fraud cases.

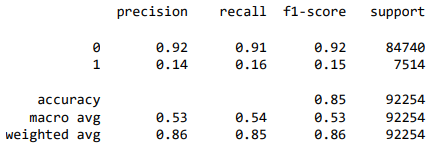
**Confusion matrix on test data**

[[78712, 6028]

[ 5902, 1612]]

The confusion matrix shows that for the test data, the true positives are 78712 and false positives are 6028, the true negatives are 5902 and the false negatives are 1612.

Now the dataset is applied for the *Decision Tree* algorithm:



**Figure 3: output for Decision Tree**

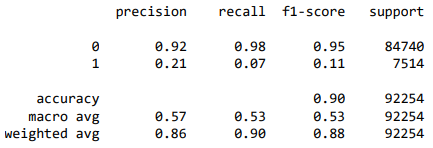
**Confusion matrix on test data**

[[77423, 7317]

[ 6319, 1195]]

The confusion matrix shows that for the test data, the true positives are 77423 and false positives are 7317, the true negatives are 6319 and the false negatives are 1195.

Finally, we apply the dataset for the *K-Nearest Neighbors* algorithm:



**Figure 4: output for K-NN**

**Confusion matrix on test data**

[[82670, 2070]

[ 6952, 562 ]]

The confusion matrix shows that for the test data, the true positives are 82670 and false positives are 2070, the true negatives are 6952 and the false negatives are 562.

**5.CONCLUSION:**

Even though there are many fraud detection techniques we cannot say that this algorithm works perfectly for the fraud detection. In, this paper we studied the algorithms like Naïve Bayes, Decision tree and K-NN. From our analysis, we can conclude that the accuracy is more for the K-Nearest Neighbors algorithm with 90% and 92% precision, 98% recall and 95% F1-score. When we consider the evaluation criteria the K-NN algorithm has the highest value than the other two models. Hence, we can conclude that the K-NN model works best to detect the credit card fraud.

**6. REFERENCES:**

[1] Thennakoon A., Bhagyani C., Premadasa S., Mihiranga S., & Kuruwitaarachchi N. (2019). “Real-time Credit Card Fraud Detection Using Machine Learning.” 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence). doi:10.1109/confluence.2019.8776942.

[2] S. N. Kalid, K. -H. Ng, G. -K. Tong and K. -C. Khor, "A Multiple Classifiers System for Anomaly Detection in Credit Card Data With Unbalanced and Overlapped Classes," in IEEE Access, vol. 8, pp. 28210-28221, 2020, doi: 10.1109/ACCESS.2020.2972009.

[3] S. Makki, Z. Assaghir, Y. Taher, R. Haque, M. -S. Hacid and H. Zeineddine, "An Experimental Study With Imbalanced Classification Approaches for Credit Card Fraud Detection," in IEEE Access, vol. 7, pp. 93010-93022, 2019, doi: 10.1109/ACCESS.2019.2927266.

[4] Dhankhad, S., Mohammed, E., & Far, B. (2018). “Supervised Machine Learning Algorithms for Credit Card Fraudulent Transaction Detection: A Comparative Study.” 2018 IEEE International Conference on Information Reuse and Integration (IRI). doi:10.1109/iri.2018.00025.

[5] Tanouz, D., Subramanian, R. R., Eswar, D., Reddy, G. V. P., Kumar, A. R., & Praneeth, C. V. N. M. (2021). “Credit Card Fraud Detection Using Machine Learning. 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS).” doi:10.1109/iciccs51141.2021.9432308.

[6] R. Sailusha, V. Gnaneswar, R. Ramesh and G. R. Rao, "Credit Card Fraud Detection Using Machine Learning," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), 2020, pp. 1264-1270, doi: 10.1109/ICICCS48265.2020.9121114.

[7] Ayorinde, K. (2021). A methodology for detecting credit card fraud [Master’s thesis, Minnesota State University, Mankato] Cornerstone: A Collection of Scholarly and Creative Works for Minnesota State University, Mankato.<https://cornerstone.lib.mnsu.edu/etds/1168>

[8] [Credit Card Fraud Detection | Kaggle](https://www.kaggle.com/datasets/mishra5001/credit-card)

[9] <https://nilsonreport.com/publication_the_current_issue.php>

[10] <https://towardsdatascience.com/what-is-feature-scaling-why-is-it-important-in-machine-learning-2854ae877048>

[11] <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html>

[12] <https://medium.com/analytics-vidhya/correlation-and-machine-learning-fee0ffc5faac#:~:text=Application%20in%20Machine%20Learning%20Correlation%20is%20a%20highly,model.%20Data%20having%20non-correlated%20features%20have%20many%20benefits>

[13] <https://www.geeksforgeeks.org/naive-bayes-classifiers/>

[14] <https://www.javatpoint.com/machine-learning-decision-tree-classification-algorithm>

[15] <https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning>

[16] <https://en.wikipedia.org/wiki/Credit_card_fraud>

**ACKNOWLEDGEMENT:**

We wish to express our heartfelt gratitude to the all the people who have played a crucial role in the research for this project, without their active cooperation the preparation of this project could not have been completed within the specified time limit.

We are thankful to our respected Principal **G. Chandra Mohan Reddy** and Management, MGIT for their encouragement and cooperation to complete this project with complete focus and attention.

We are also very much thankful to our project coordinator **Ms. Sreedevi** who supported us throughout this project with utmost cooperation and patience and for helping us in doing this Project. We thank all the faculty of Department of CSE, whose suggestions during reviews helped us in accomplishment of our project.

We would like to take this opportunity to express our gratitude to all our group members. The project would not have been successful without their cooperation and inputs.

**PROJECT MEMBERS:**

SAFFA SAMREEN (20261A6751)

S. RITHVIK (20261A6752)

S. ROHAN (20261A6753)

S. DHANUSH KUMAR (20261A6754)

M. SRI HARSHINI (20261A6755)

V.L.N.S. SRI PREETHI (20261A6756)

V. ANIL KUMAR (20261A6757)

V.JOSHI MANOHAR (20261A6758)

D. VISHNU HARSHINI (20261A6759)

V. ROOP SHANKAR (20261A6760)

