International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

|| Volume 10, Issue 7, July 2021 ||

|DOI:10.15680/IJIRSET.2021.1006001|

Credit Card Fraud Detection with Facial Recognition

Shuar Alshi [1], Asim Upendra Chitre [2], Rutvij Mayank Dave [3], Heet Shaliesh Chheda [4], Mr. Dipak Kulkarni [5]

U.G. Student, Department of Electronics and Telecommunication Engineering, KJ Somaiya College of Engineering, Vidyavihar, Maharashtra, India[1,2,3,4]

Associate Professor, Department of Electronics and Telecommunication Engineering, KJ Somaiya College of Engineering, Vidyavihar, Maharashtra, India [5]

ABSTRACT: *Credit card fraud occurs when someone, a fraudster or a thief, uses your stolen credit card or the data from that card to make fraudulent transactions on your behalf or use your account to take out cash advances. It is important to be able to recognize fraudulent credit card purchases by credit card companies so that consumers are not charged for products they have not purchased. This can be successfully implemented using Machine Learning and Neural Networks. This project aims to demonstrate the modelling of a data set of Credit Card Fraud Detection using machine learning. The problem of detecting credit card fraud involves modelling prior credit card purchases with the data of those that turned out to be fraud. This model is then used to understand whether or not a new transaction is fraudulent. We used ANN as the basis of our model and we then applied it to payment gateway and e-commerce website which was made using web development technologies*

KEYWORDS: *Credit Card fraud, Machine Learning, Neural Networks, XGBoost, Decision Tree, Random Forest, ANN, Web development.*

I.INTRODUCTION

A credit card is a convenient financial service that can be used to buy petrol, food, and other goods and services on a regular basis. It can also be a great resource for purchasing big-ticket items such as TVs, travel packages, and jewellery because the funds for these items are not always immediately at our disposal.

When used responsibly, credit cards can be valuable tools for earning rewards, traveling, handling emergencies or unplanned expenses, and building credit.

A rewards credit card does exactly what its name implies: rewards the cardholder for making purchases. Rewards can vary by issuer and card type. Some rewards come in the form of cash back, discounts on gas station purchases, and even travel miles. For those who use their cards regularly, earning rewards is one of the primary advantages of credit cards, as cardholders can redeem them for things they were going to purchase already as well as the occasional treat.

Credit cards can also be beneficial when traveling. This is because some major car rental companies and hotels require a hold on a credit or debit card to reserve a vehicle or book a room. This procedure can take several days or longer. During this time, the amount of the hold on either a credit or debit card is not available to use. Because you may not have the necessary funds in your bank account, credit cards increase your purchasing power, providing you with the required funds at the time they are needed.

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

|| Volume 10, Issue 7, July 2021 ||

|DOI:10.15680/IJIRSET.2021.1006001|

II. RELATED WORK

Credit ‹ when someone uses your credit card or credit account to make a purchase you didn't authorize. Credit card fraud can be authorized, where the genuine customer themselves processes a payment to another account which is controlled by a criminal, or unauthorized, where the account holder does not provide authorization for the payment to proceed and the transaction is carried out by a third party.

Several literatures on anomaly or fraud detection in this area have already been published and are available for public use. Maniraj (2019) [1] deployed multiple anomaly detection algorithms such as the Local Outlier Factor and Isolation Forest algorithm on the dataset. The project was divided into two sections: one where only 10% data was used and the other where the whole dataset was used. The precision was about 28% when a tenth of data was taken into consideration whereas the precision increased to 33% when the entire dataset was fed into the algorithm. This was because of the imbalanced data between the number of fraud and genuine transactions.

Dornadula [2] concluded that there are two ways of dealing with the imbalanced dataset: Consider Matthew Coefficient Correlation of the classifier on the original dataset or make use of one-class classifiers. Different classifiers like local outline factor, Isolation Forest, Logistic regression, Decision tree and Random Forest algorithms were used and observed that Decision Tree and Random Forest algorithms gave better results.

Varun Kumar VS [3] used Artificial Neural Network (ANN) along with Logistic Regression, Naive Bayes and Decision Tree algorithm and concluded that ANN model has better accuracy, precision and recall.

Yu [4] used distance-based outlier and outlier mining algorithm using distance sum to detect frauds in the given dataset using data standardization concept. The experiment showed that outlier mining can detect credit card fraud better than anomaly detection and also stated that If this algorithm is applied into bank credit card fraud detection system, the probability of fraud transactions can be predicted soon after credit card transactions by the banks.

Asha RB [5] showed the accuracy performance of three algorithms, namely Support Vector Machines (SVM), K-nearest neighbors (KNN) and Artificial Neural Networks (ANN) using a dataset which consisted of credit card frauds and concluded that ANN gives accuracy close to 100%. In this project, they also carry out processes like normalization and under-sampling to tackle the problems faced by an imbalanced dataset.

AA Taha [6] used F1- score as a metric to classify which algorithms suit a given dataset. F1- score consists of precision and recall which tells us how robust and effective our classifier is to a given dataset.

Dhawle [7] gave an ideal way of detecting and recognizing face using deep learning and face-recognition library, which comes under deep learning module of data science. Further, they also showed places where this module can be used like smartphone and several software applications.

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

|| Volume 10, Issue 7, July 2021 ||

|DOI:10.15680/IJIRSET.2021.1006001|

III.METHODOLOGY

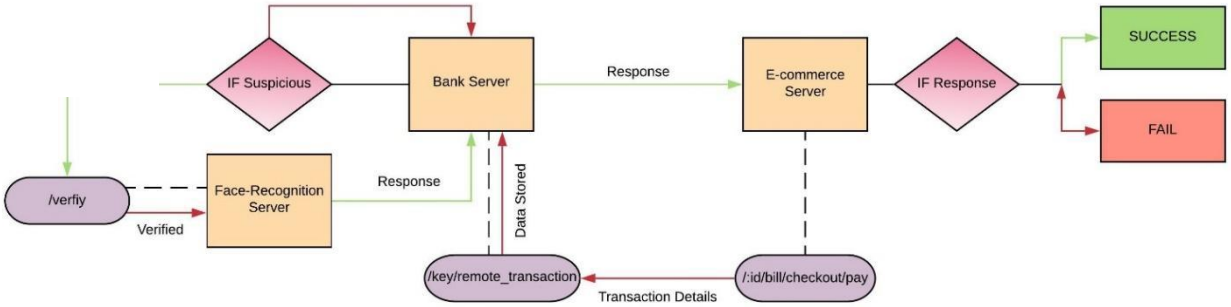

Fig.1 Flowchart of Project's working

- From the e-commerce server, payment will be made via the /id/bill/checkout/pay route to /key/remote transaction api of the bank server.
- The bank will verify developer's account. After that, it will verify merchant's bank account.
- The customer's credit card number, CVV, expiry date of the card and credit card status is also checked by the bank.
- The bank also compares the transaction amount with the credit card limit.
- The details of the transaction such as amount, date and time of transaction is fed into our deep learning model in order to check whether the transaction is fraudulent.
- It also retrieves the average amount along with the average fraud percentage of the customer's previous transactions.
- If the percentage of fraud in the current transaction is greater than 45% or if the amount is greater than the average expenditure or if the average fraud level is greater than 50%, then the system will request for the facial recognition.
- After the model classifies the transaction to be suspicious, if the facial recognition output is true, that is, if the current picture rightly corresponds to the picture which was taken at the time of making an account, then the transaction will be successful.
- On the contrary, if the model predicts the transaction to be not suspicious, then the transaction will be successful and an email will be sent to the account holder.
- If the transaction fails for the third time, then the credit card will be automatically blocked and an e-mail will be sent to the respective account holder.

Here, we modified a pre-existing dataset, which included real-life transaction data. The feature 'amt' is the transaction Amount, this feature can be used for example-dependent cost-sensitive learning. Feature 'is_fraud' is the response variable and it takes value 1 in case of fraud and 0 otherwise.

Firstly, we will see which are the features that are more important than the rest by using the random forest machine learning model.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

|| Volume 10, Issue 7, July 2021 ||

|DOI:10.15680/IJIRSET.2021.1006001|

```
from sklearn.model_selection import train_test_split
import random
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import GridSearchCV , train_test_split
from sklearn.metrics import mean_absolute_error

X = df[['cc_num','amt','zip','lat','long','city_pop',
        'unix_time','merch_lat','merch_long','day','month','year','hour','minute']]
y = df['is_fraud']

X_train, X_test, y_train, y_test = train_test_split(X, y,
                                                    test_size=0.3, random_state=42)

rf_model = RandomForestRegressor(n_estimators = 100,random_state = 42)
rf_model.fit(X_train,y_train)
```

We will individually see how much the dataset is dependent on a certain feature

The graph below shows the same:

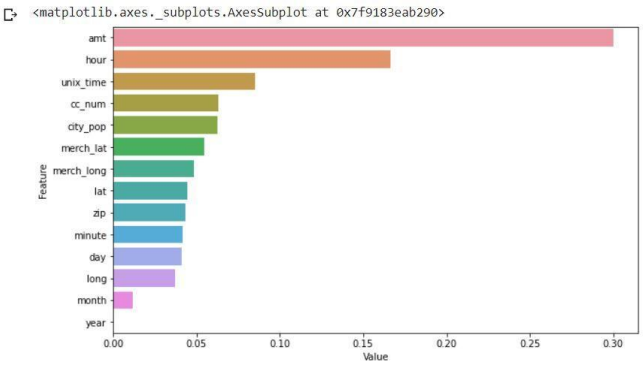

Fig 2. Bar graph of value vs features

From the graph, we can infer that “amount” has more impact than any other feature in our dataset, followed by “hour.”

1. Amount

Using amount as our main criteria, we will be plotting a graph between the mean amount spent on a certain category and the categories.

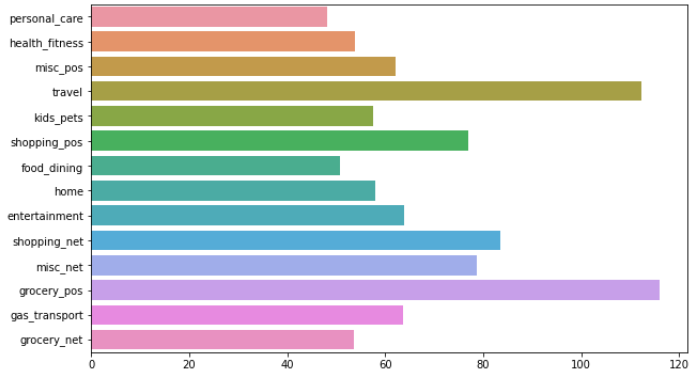

Fig 3. Bar graph for Value vs Category

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

|| Volume 10, Issue 7, July 2021 ||

|DOI:10.15680/IJIRSET.2021.1006001|

2. Nun s in certain category

Next, w number of frauds that a certain category face and we conclude that most of the frauds happen while s nternet.

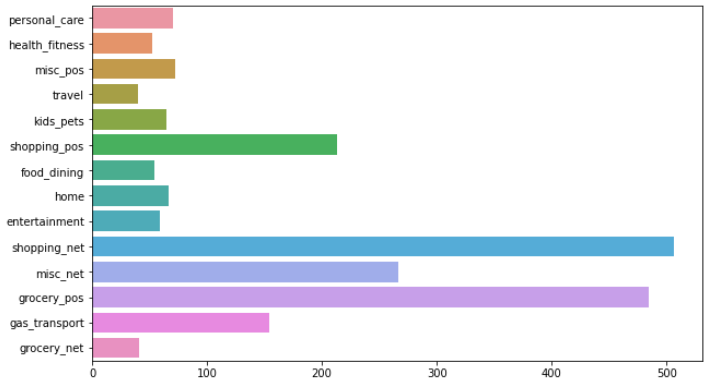

Fig 4. Bar Graph for no. of frauds vs Category

Hence, we decided to make a model which will reduce the frauds while shopping on the internet.

```
# Importing tensorflow and its components which we will use in making our model.

import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import Sequential
from tensorflow.keras.layers import Dense
from keras.callbacks import ModelCheckpoint, TensorBoard
from keras.models import Model, load_model

# Making a sequential model

classifier = Sequential()
classifier.add(Dense(units =24 , kernel_initializer = 'uniform', activation = 'relu', input_dim = X_train.shape[1]))
classifier.add(Dense(units = 24, kernel_initializer = 'uniform', activation = 'relu'))
classifier.add(Dense(units = 24, kernel_initializer = 'uniform', activation = 'relu'))
classifier.add(Dense(units = 24, kernel_initializer = 'uniform', activation = 'relu'))
classifier.add(Dense(units = 1, kernel_initializer = 'uniform', activation = 'sigmoid'))

# Compiling the ANN

classifier.compile(optimizer = 'adam', loss = 'binary_crossentropy', metrics = ['accuracy'])

checkpoint_path = "training_1/cp.ckpt"
checkpoint_dir = os.path.dirname(checkpoint_path)

# Create a callback that saves the model's weights
cp_callback = tf.keras.callbacks.ModelCheckpoint(filepath=checkpoint_path,
                                                 save_weights_only=True,
                                                 verbose=1)


history = classifier.fit(X_train, y_train, batch_size = 32, epochs = 30 , callbacks=[cp_callback], shuffle=True,verbose=1).history
```

The fitted model is then used to predict on the test dataset and hence the confusion matrix is obtained in the form of test and train result

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

|| Volume 10, Issue 7, July 2021 ||

|DOI:10.15680/IJIRSET.2021.1006001|

```
Train Result:
================================================
Accuracy Score: 93.57%

________________________________________________
Classification Report:
                     0             1  accuracy      macro avg   weighted avg
precision     0.947582      0.924445  0.935683       0.936013       0.935993
recall        0.922147      0.949170  0.935683       0.935659       0.935683
f1-score      0.934692      0.936644  0.935683       0.935668       0.935670
support    72984.000000  73244.000000  0.935683  146228.000000  146228.000000
________________________________________________
Confusion Matrix:
 [[67302  5682]
 [ 3723 69521]]

Test Result:
================================================
Accuracy Score: 93.57%

________________________________________________
Classification Report:
                     0             1  accuracy     macro avg  weighted avg
precision     0.948533      0.923458  0.935695      0.935996      0.936048
recall        0.921945      0.949559  0.935695      0.935752      0.935695
f1-score      0.935050      0.936327  0.935695      0.935689      0.935686
support    31465.000000  31205.000000  0.935695  62670.000000  62670.000000
________________________________________________
Confusion Matrix:
 [[29009  2456]
 [ 1574 29631]]
```

Fig 5. Classification report of ANN model

Then we save the model using TensorFlow function of .save () and to check the model, we first load the saved model and use the .evaluate() function to check whether the saved model accuracy is equal to the test accuracy.

```
result = new_model.evaluate(X_test, y_test)
```

```
1959/1959 [==============================] - 1s 637us/step - loss: 0.1991 - accuracy: 0.9357
```

IV. Experimental results

4.1. Successful Transaction

```
2021-05-08T11:27:26.609506+00:00 app[web.1]: IP address: 10.101.190.181
2021-05-08T11:27:26.697996+00:00 app[web.1]: {'cc_no': '2661700698687952', 'cvv'
: '161', 'expiry': '2024-05-07', 'amount': ' 575', 'description': 'Harry Potter
and the Goblet of Fire', 'merchant_account_no': '30460052324114', 'clientImg': '
https://res.cloudinary.com/dxbi7mov0/image/upload/v1620472954/suw6dsjkabyorfo0k0
2h.jpg', 'visitorsIP': '::ffff:10.32.238.58', 'card_holder_name': 'Asim Chitre'}
2021-05-08T11:27:27.033618+00:00 app[web.1]: before model
2021-05-08T11:27:27.162508+00:00 app[web.1]: after model
2021-05-08T11:27:27.162609+00:00 app[web.1]: Model Prediction =  0.018924743
2021-05-08T11:27:27.246284+00:00 app[web.1]: avg. expenditure =  318.92857142857
144
2021-05-08T11:27:27.246366+00:00 app[web.1]: avg. fraud level =  0.3834751278571
4283
2021-05-08T11:27:27.328775+00:00 app[web.1]: in
2021-05-08T11:27:28.234089+00:00 app[web.1]: <Response [200]>
```

Fig 6. Results of success

- The model predicts about 1.89% fraud. The average expenditure is $318.92 and average fraud level is 0.38. According to our algorithm, if the percentage of fraud is above 45% or current expenditure is greater than

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

|| Volume 10, Issue 7, July 2021 ||

|DOI:10.15680/IJIRSET.2021.1006001|

average expenditure and fraud level of current transaction is greater than average fraud transaction, then face will be verified.

- Over here, current expenditure ($575) is greater than average expenditure ($318.92). Hence face is verified.

4.2. Un ransaction (Invalid Face verification)

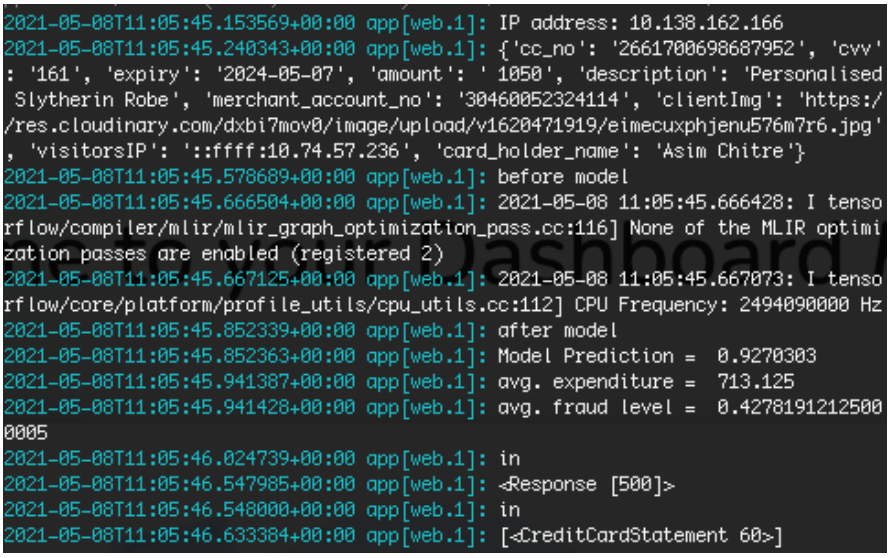

Fig 7. Results of unsuccessful transaction with invalid face verification

The model predicts the transaction to be 92.7% fraudulent and the face is not verified. Hence, transaction is unsuccessful.

4.3. Successful Transaction (Without face verification)

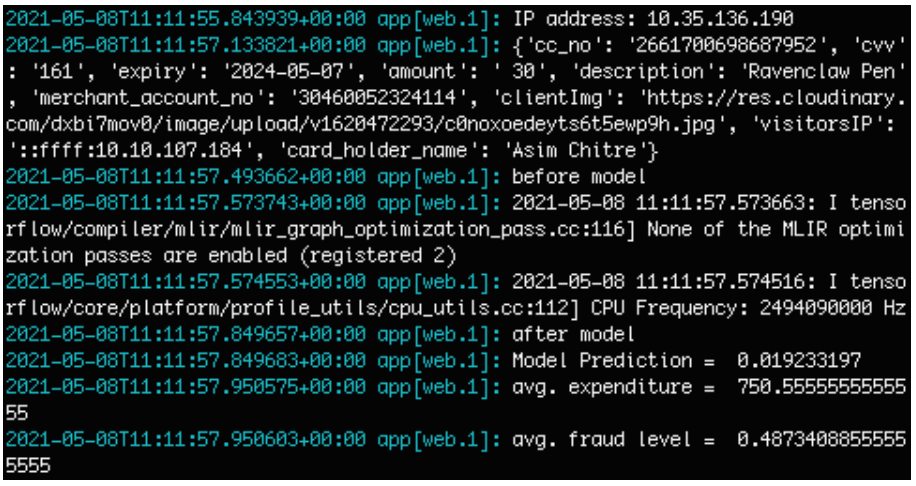

Fig 8. Results of successful transaction without face verification

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

|| Volume 10, Issue 7, July 2021 ||

|DOI:10.15680/IJIRSET.2021.1006001|

The model prediction is less that 45% (19.2%). Current expenditure ($30) is less than average expenditure ($750.5) and average fraud level (48%) is less than 50%. Hence, the transaction is classified as an non-fraudulent transaction.

V. CONCLUSION

Credit ‹ without a doubt an act of criminal dishonesty. This project has listed out the common methods of detecting a credit card fraud from the given dataset. We also developed a novel method using Artificial Neural Networks, which had substantial effect on the accuracy of the whole model and concluded that it is the best model in detecting a fraud in the credit card dataset. This project also explained in detail the different machine learning techniques and concepts required to find the accuracy, precision and recall along with the confusion matrix of a particular classifier along with their respective algorithms.

We made a payment gateway website and an e-commerce website in order to demonstrate our model. As an extra security feature, we were able to add face recognition which will ensure that minimal frauds will happen as oppose to a model which is absent of facial recognition as a tool for security.

Thus, we were successfully able to implement a model that will detect credit card frauds using neural networks as well as basic machine learning algorithms and applied it to our payment gateway and e-commerce website using present day web development technologies.

REFERENCES

[1] Maniraj, S & Saini, Aditya & Ahmed, Shadab & Sarkar, Swarna. (2019). Credit Card Fraud Detection using Machine Learning and Data Science. International Journal of Engineering Research and. 08. 10.17577/IJERT V8I 090031.

[2] Credit Card Fraud Detection using Machine Learning Algorithms,Procedia Computer Science, Volume 165, 2019, Pages 631-641, ISSN 1877-0509.

[3] Varun Kumar K S, Vijaya Kumar V G, Vijayshankar A, Pratibha K, 2020, Credit Card Fraud Detection using Machine Learning Algorithms, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 09, Issue 07 (July 2020),

[4] W. Yu and N. Wang, "Research on Credit Card Fraud Detection Model Based on Distance Sum," 2009 International Joint Conference on Artificial Intelligence, Hainan Island, 2009, pp. 353-356, doi: 10.1109/JCAI.2009.146.

[5] Asha RB, Suresh Kumar KR,Credit card fraud detection using artificial neural network, Global Transitions Proceedings, Volume 2, Issue 1, 2021, Pages 35-41,ISSN 2666-285X.

[6] A. A. Taha and S. J. Malebary, "An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine," in IEEE Access, vol. 8, pp. 25579-25587, 2020, doi: 10.1109/ACCESS.2020.2971354.

[7] Tejashree Dhawle, Urvashi Ukay, Rakshandha Choudante, Face Detection and Recognition using OpenCV and Python, Department of Computer Engineering, Dr. Babasaheb Ambedkar Technological University Raigad, India