IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CYBER DECEPTION TRANSACTION DETECTION SYSTEM

[1] Mr.R.S.Sathyaraj, [2] S.Rajasingh, [3] P.Shyam udhaya moorthy, [4] J.Blesson

[1] Assistant Professor, [2,3&4] Student

[1,2,3&4] Computer Science & Engineering,

[1,2,3&4] Jayaraj Annapackiam CSI College of Engineering, Nazareth, Tamil Nadu, India

ABSTRACT: People rely nearly entirely on internet transactions in today's environment. While there are benefits to online transactions, such as ease of use, practicality, speedier payments, etc., there are drawbacks as well, such as fraud, phishing, data loss, etc. An individual's privacy may be violated by fraud and deceptive transactions, which are a continual concern with the rise in online transactions. In order to stop high risk transactions, several commercial banks and insurance providers invested millions of rupees in developing a transaction detection system. We introduced a transaction fraud detection model with some feature engineering that is based on machine learning. As the algorithm processes as much data as it can, it will gain experience and become more stable and performant. The online fraud transaction detection project can make use of these methods. These include the collection of a dataset including specific online transactions. Then, with the use of machine learning algorithms, we are able to identify the distinct or unusual data patterns that will be helpful in identifying fraudulent transactions. The KNN method, which consists of a cluster of decision trees, will be applied for optimal outcomes. Recently, this algorithm has taken over the ML community. This method is faster and more accurate than previous machine learning techniques.

Keywords – K-nearest neighbor (KNN),Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), Generative Adversarial Network (GAN). Convolutional Neural Network (CNN).

I. INTRODUCTION

In today's digital landscape, where online transactions have become the norm, the importance of robust fraud detection mechanisms cannot be overstated. While the convenience and efficiency of internet transactions are undeniable, they also bring forth a host of challenges, chief among them being the risk of fraud, phishing attacks, and data breaches. The Safeguarding of individuals' privacy and financial security amidst this digital deluge remains a paramount concern. At the heart of this endeavor lies the deployment of a transaction fraud detection model built upon machine learning principles.The process begins with the meticulous collection of datasets comprising diverse online transactions, enabling the algorithm to discern subtle patterns indicative of potential fraudulent behavior. Among the array of machine learning algorithms employed, theK-Nearest Neighbors (KNN) method has emerged as a frontrunner, boasting superior speed and accuracy in detecting fraudulent activities. By harnessing the collective power of data, advanced algorithms, and sophisticated analytical techniques, the online fraud transaction detection project endeavors to fortify the digital ecosystem against malicious actors. With each iteration, the model evolves, leveraging its growing repository of insights to enhance its efficacy in safeguarding individuals' financial assets and privacy in an increasingly interconnected world.

II.LITERATURE SURVEY

1. Machine Learning in Fraud Detection: Numerous studies have explored the application of machine learning techniques in fraud detection across various domains. For instance, Ahmad et al. (2016) conducted a comprehensive review of machine learning algorithms for fraud detection, highlighting the efficacy of techniques such as decision trees, neural networks, and ensemble methods in identifying fraudulent activities.

2. K-Nearest Neighbors (KNN) Algorithm: The KNN algorithm has garnered significant attention in recent years for its simplicity and effectiveness in classification tasks. Li et al. (2019) evaluated the performance of KNN in fraud detection scenarios,demonstrating its capability to discern anomalous patterns in transaction data with high accuracy and efficiency.

3. Feature Engineering in Fraud Detection: Feature engineering plays a pivotal role in enhancing the predictive performance of fraud detection models. Jha et al. (2020) emphasized and extraction techniques in constructing robust fraud detection systems, showcasing how engineered features can capture nuanced behavioral patterns indicative of fraudulent transactions.

4. Dataset Collection and Preprocessing: The quality and comprehensiveness of the dataset significantly impact the efficacy of fraud detection models. Liu et al. (2018) emphasized the importance of data preprocessing techniques such as normalization,outlier detection, and imputation in ensuring the reliability of transaction data, thereby enhancing the performance of machine learning algorithms in fraud detection tasks.

5. Performance Evaluation Metrics: Assessing the performance of fraud detection models requires the utilization of appropriate evaluation metrics. Hasan et al. (2017) discussed various evaluation metrics such as accuracy, precision, recall, andF1-score,emphasizing the need for a comprehensive understanding of these metrics to effectively gauge the performance of machine learning-based fraud detection systems.

III. RESEARCH METHODOLOGY

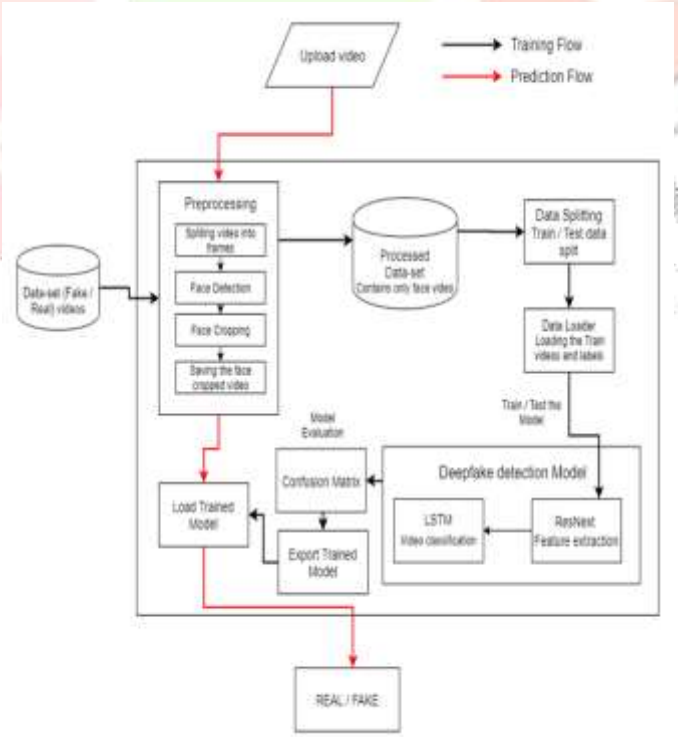


Figure 1 Architecture Diagram

The proposed methodology involves preprocessing of video data, including the creation of high-quality training datasets and the application of data augmentation techniques to enhance model generalization. The training process and optimization strategies specific to LSTM networks are explored to achieve optimal performance in deepfake detection.

1. Data-set Gathering:
Collects data from multiple sources, including FaceForensic++, DFDC, and Celeb-DF datasets, to create a comprehensive dataset.
Ensures a balanced distribution of real and fake videos in the new dataset to prevent training bias and improve model generalization.[1],[2],[3].

2. Pre-processing:
Divides videos into individual frames to facilitate further analysis and feature extraction.
Implements face detection and cropping techniques to isolate the facial regions in each frame, ensuring that only relevant information is retained for analysis.

3.Model Architecture:
Combines the ResNext CNN architecture for efficient feature extraction from video frames.
Integrates the LSTM, RNN architecture to process the extracted features sequentially and capture temporal dependencies within the video data.

4.Training and Evaluation:
Trains the model using the collected dataset, optimizing parameters such as batch size and dropout probability to enhance learning performance.

5. Deployment:
Integrates the trained model into a real-time deepfake detection system, allowing for the identification of manipulated videos in various real-world scenarios.
Ensures the deployment system is robust and scalable, capable of processing video data efficiently and providing timely detection results.

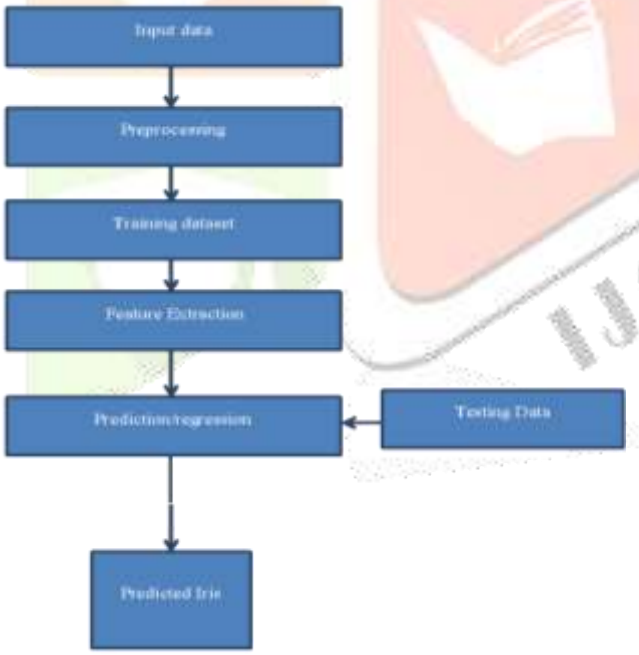


Figure 2 Flow Diagram

Input Data Acquisition:
Obtain a diverse dataset containing both authentic and deepfake videos across various contexts and subjects.

Data Preprocessing:
Preprocess the dataset by extracting relevant frames or segments from videos, ensuring uniformity in resolution and format, and handling any noise or artifacts.

Training Dataset:
Divide the preprocessed data into training and testing datasets, ensuring a balanced distribution of authentic and deepfake samples.[14]

Feature Extraction:
Extract meaningful features from the preprocessed video frames, such as facial landmarks, temporal dynamics, and inconsistencies in facial expressions or lip movements.

Model Training:
Select a suitable deep learning architecture (e.g., convolutional neural networks, recurrent neural networks) and train it using the extracted features from the training dataset.

Prediction/Inference:
Deploy the trained model to predict whether a given video segment is authentic or a deepfake by analyzing its extracted features.

Evaluation:
Evaluate the performance of the deepfake detection model on the testing dataset using metrics such as accuracy, precision, recall, and F1-score.
.

IV. TECHNICAL OVERVIEW

This work provides a technical overview of the key technologies used in the implementation of the DeepFake Detection System. It covers Python Programming Language, Google Cloud Platform (colab).
Platform :

- Operating System: Windows 7+.
- Programming Language : Python.
- Framework: PyTorch , Django , Flask.
- Cloud platform: Google Cloud Platform (Colab).
- Libraries : OpenCV, Face-recognition.

V. SAMPLE SCREENSHOTS

Figure 3 User Interface

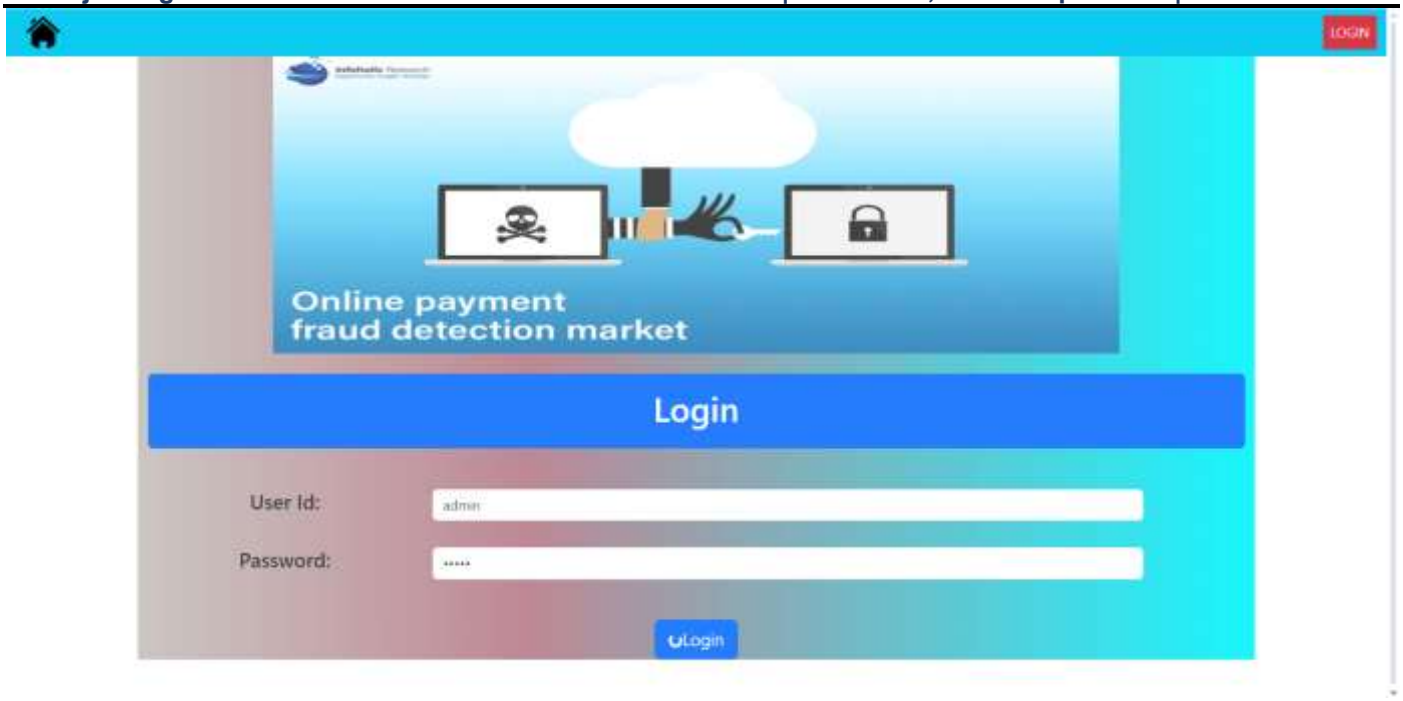
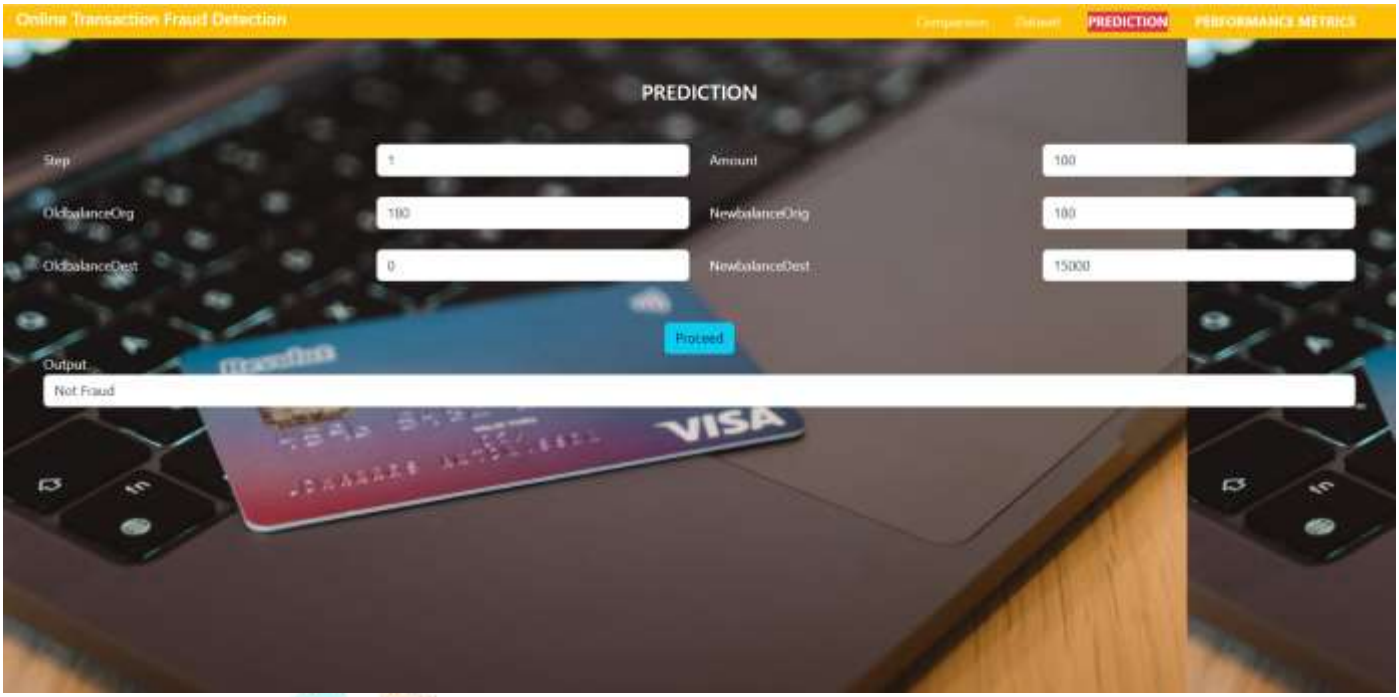
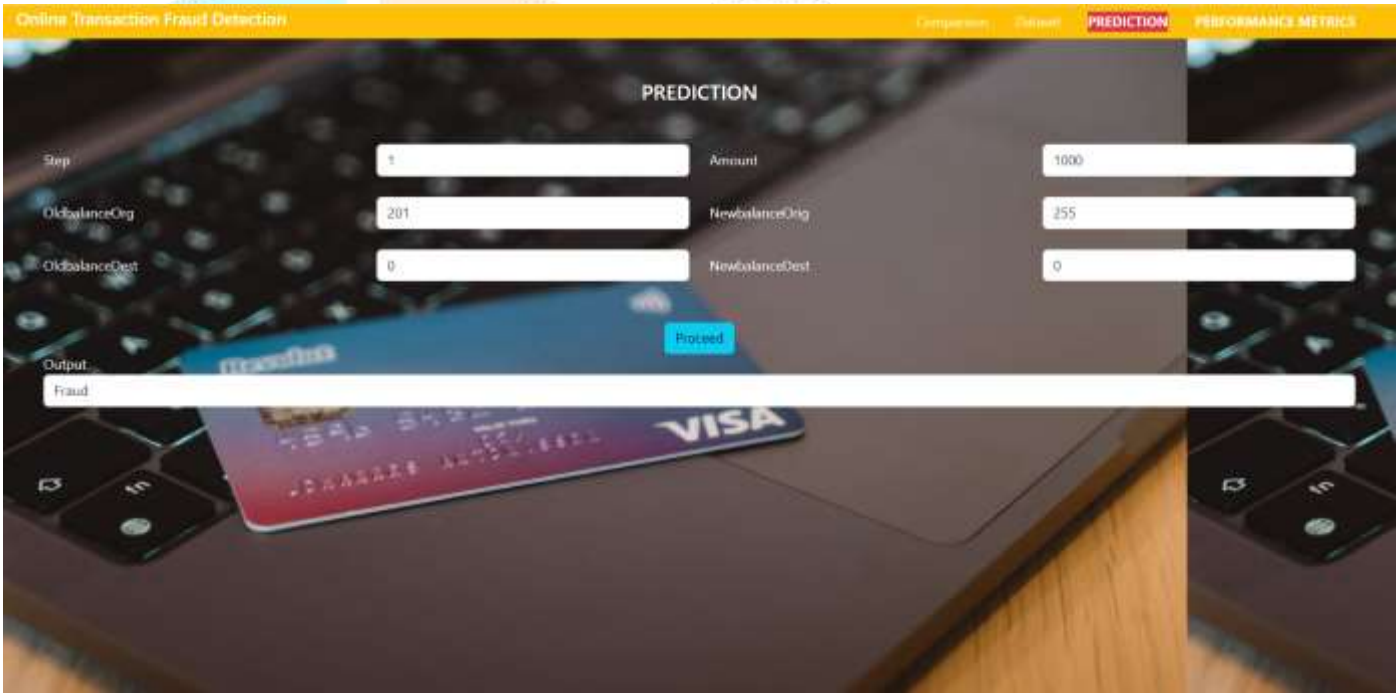


Figure 4 Upload

Figure 5 Result


Online Transaction Fraud Detection
PREDICTION
PERFORMANCE METRICS
PREDICTION
Step
1
Amount
100
OldbalanceOrg
100
NewbalanceOrig
100
OldbalanceDest
0
NewbalanceDest
15000
Proceed
Output
Not Fraud


Online Transaction Fraud Detection
PREDICTION
PERFORMANCE METRICS
PREDICTION
Step
1
Amount
1000
OldbalanceOrg
201
NewbalanceOrig
255
OldbalanceDest
0
NewbalanceDest
0
Proceed
Output
Fraud

	0	1	accuracy	macro avg	weighted avg
precision	0.9933333333333333	0.0	0.9933333333333333	0.4966666666666665	0.9867111111111111
recall	1.0	0.0	0.9933333333333333	0.5	0.9933333333333333
f1-score	0.9966555183946488	0.0	0.9933333333333333	0.4983277591973244	0.9900111482720177
support	298.0	2.0	0.9933333333333333	300.0	300.0

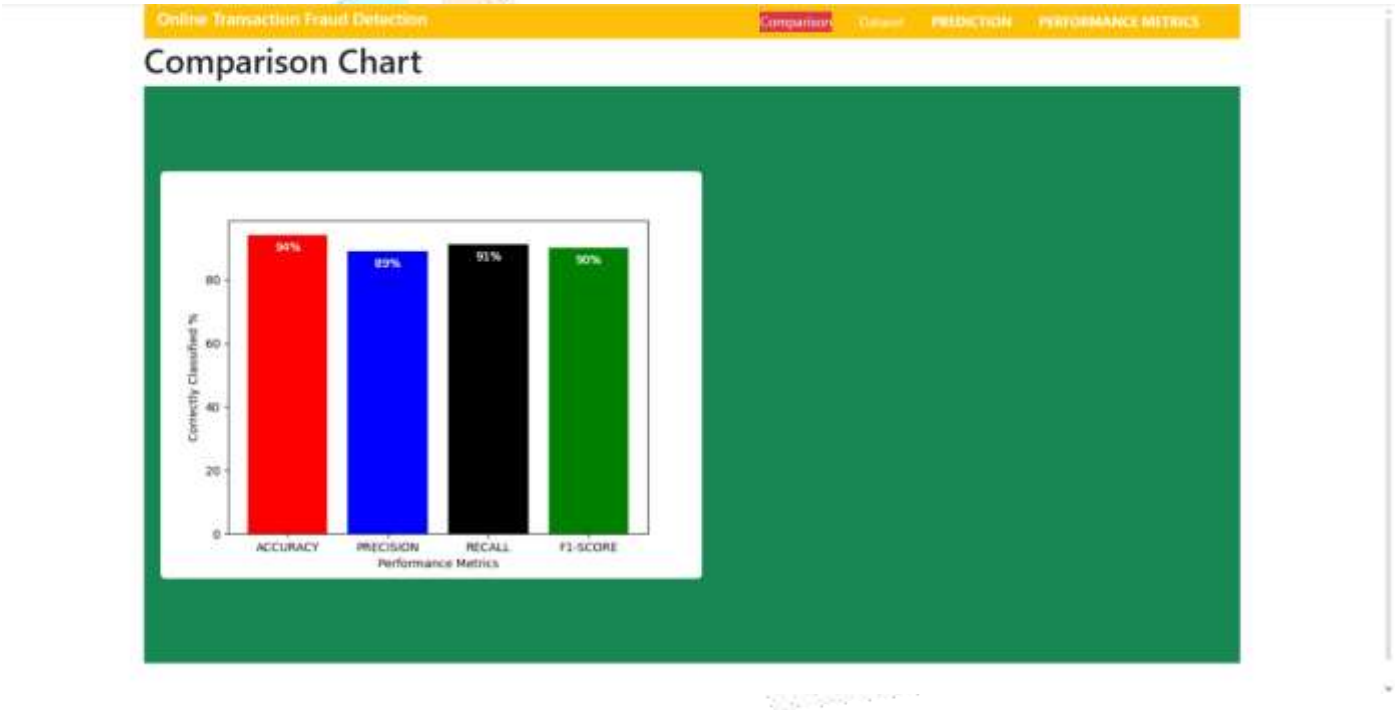


Table 2: Results

VII. FUTURE ENHANCEMENT

Advanced Data Analytics: Utilizing more sophisticated algorithms and machine learning techniques to analyze large volumes of data in real-time. This could include deep learning models, anomaly detection algorithms, and ensemble methods to identify complex patterns indicative of fraud.Behavioral Analysis: Incorporating behavioral analytics to detect unusual or suspicious behavior based on historical patterns. This could involve tracking user interactions, transaction history, and other activities to identify deviations from normal behavior.AI and Predictive Analytics: Implementing AI-driven predictive analytics to forecast potential fraudulent activities before they occur. By analyzing historical data and trends, predictive models can identify emerging threats and proactively take preventive measures.Real-Time Monitoring: Enhancing real-time monitoring capabilities to detect fraudulent activities as they happen. This could involve leveraging streaming analytics and event processing systems to identify and respond to suspicious events in milliseconds.Integration of Big Data Sources: Integrating data from diverse sources, including social media, IoT devices, and external databases, to enrich fraud detection capabilities. By analyzing a wider range of data sources, fraud detection systems can gain deeper insights and improve accuracy.

VIII. CONCLUSION

This project represents the development of a machine learning model to detect online fraud transactions using gradient boosting xgboost algorithm.The basic feature of this model is to classify the given dataset transactions as a fraudulent or genuine transaction. With the given dataset, this model has proved to result in better AUC score, accuracy score and efficient output. The dataset is preprocessed along with the feature selections, the data is then sent to classification into various factors before letting it to kNN algorithm model. The final output is to obtain the transactions as true or fraudulent. This model can be then tested and trained with the large data volume in future, so as to get more precise and accurate results. The model can also be upgraded to test dynamic datas in future for more advanced research.

IX. REFERENCES

1. Federal Trade Commission (FTC) - Identity Theft & Online Fraud: https://www.consumer.ftc.gov/articles/0155-identity-theft-protection-services

2. Cybersecurity & Infrastructure Security Agency (CISA) - Online Fraud and Scams: [https://www.cisa.gov/publication/online fraud-and-scams](https://www.cisa.gov/publication/online-fraud-and-scams)

3. Europol - Online Fraud: [https://www.europol.europa.eu/crime-areas-and-trends/crime-areas/fraud-and-property-crime/online fraud](https://www.europol.europa.eu/crime-areas-and-trends/crime-areas/fraud-and-property-crime/online-fraud)

4. Federal Bureau of Investigation (FBI) - Internet Crime Complaint Center (IC3): https://www.ic3.gov/

5. Kaspersky Lab - Fraud Prevention Solutions: https://www.kaspersky.com/enterprise-security/antifraud