A

Major Project

On

CREDIT CARD FRAUD DETECTION USING STATE OF THE ART USING MACHINE LEARNING AND DEEP LEARNING ALGORITHM

(Submitted in partial fulfillment of the requirements for the award of Degree)

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

By

M RISHITHA REDDY (207R1A0595)

GUGLAVATH SRISHANTH (207R1A0580)

MARIPEDDA PRAVEEN (217R5A0508)

Under the Guidance of

Ms. Saba Sultana

(Assistant Professor)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CMR TECHNICAL CAMPUS

UGC AUTONOMOUS

(Accredited by NAAC, NBA, Permanently Affiliated to JNTUH, Approved by AICTE, New Delhi) Recognized Under Section 2(f) & 12(B) of the UGCAct.1956, Kandlakoya (V), Medchal Road, Hyderabad-501401.

2020-2024

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the project entitled **"CREDIT CARD FRAUD DETECTION USING STATE OF THE ART USING MACHINE LEARNING AND DEEP LEARNING ALGORITHM"** being submitted by **M. RISHITHA REDDY (207R1A0595), GUGLAVATH SRISHANTH (207R1A0580) & MARIPEDDA PRAVEEN (217R5A0508)** in partial fulfillment of the requirements for the award of the degree of B.Tech in Computer Science and Engineering to the Jawaharlal Nehru Technological University Hyderabad, is a record of bonafide work carried out by them under our guidance and supervision during the year 2023-24.

The results embodied in this project have not been submitted to any other University or Institute for the award of any degree or diploma.

Ms. Saba Sultana
(Assistant Professor)
INTERNAL GUIDE

Dr. A. Raji Reddy
DIRECTOR

Dr. K. Srujan Raju
HOD

EXTERNAL EXAMINER

Submitted for viva voice Examination held on _____

ACKNOWLEDGEMENT

Apart from the efforts of us, the success of any project depends largely on the encouragement and guidelines of many others. We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project.

We take this opportunity to express my profound gratitude and deep regard tomy guide **Ms. Saba Sultana,** Assistant Professor for her exemplary guidance, monitoring and constant encouragement throughout the project work. The blessing, help and guidance given by her shall carry us a long way in the journey of life on which we are about to embark.

We also take this opportunity to express a deep sense of gratitude to the Project Review Committee (PRC) **G.Vinesh Shanker, Dr. J. Narasimharao , Ms. Shilpa,& Dr. K. Maheswari** for their cordial support, valuable information and guidance, which helped us in completing this task through various stages.

We are also thankful to **Dr. K. Srujan Raju,** Head, Department of Computer Science and Engineering for providing encouragement and support for completing this project successfully.

We are obliged to **Dr. A. Raji Reddy,** Director for being cooperative throughout the course of this project. We also express our sincere gratitude to Sri. **Ch. Gopal Reddy,** Chairman for providing excellent infrastructure and a nice atmosphere throughout the course of this project.

The guidance and support received from all the members of **CMR Technical Campus** who contributed to the completion of the project. We are grateful for their constant support and help.

Finally, we would like to take this opportunity to thank our family for their constant encouragement, without which this assignment would not be completed. We sincerely acknowledge and thank all those who gave support directly and indirectly in the completion of this project.

M. RISHITHA REDDY (207R1A0595)
GUGLAVATH SRISHANTH (207R1A0580)
MARIPEDDA PRAVEEN (217R5A0508)

ABSTRACT

Credit cards have become an integral part of modern financial transactions, providing users with a convenient and widely accepted method of payment. The widespread use of credit cards, both online and offline, has, however, given rise to the escalating threat of credit card fraud. Fraudulent activities such as unauthorized transactions, identity theft, and skimming pose substantial risks to financial institutions and cardholders. With the increasing prevalence of online transactions, the risk of credit card fraud has become a major concern for financial institutions and users alike. This study proposes an advanced Credit Card Fraud Detection system leveraging state-of-the-art machine learning and deep learning algorithms to enhance the accuracy and efficiency of fraud detection.

The financial ecosystem's reliance on credit cards necessitates the development of robust and sophisticated fraud detection systems to safeguard the integrity of transactions and protect consumers from financial losses. Traditional methods of fraud detection, such as rule-based systems and signature verification, have become increasingly inadequate in the face of evolving and sophisticated fraudulent techniques. As a response to the growing challenges, the integration of advanced technologies, particularly machine learning and deep learning, has emerged as a promising avenue for improving the accuracy and efficiency of credit card fraud detection. These technologies leverage complex algorithms and data patterns to identify abnormal behavior, detect anomalies, and adapt to the dynamic nature of fraudulent activities. This project aims to explore and implement state-of-the-art machine learning and deep learning algorithms for credit card fraud detection. By harnessing the power of these advanced techniques, the research seeks to enhance the capability of financial institutions to detect and prevent fraudulent transactions in real-time, providing a more secure environment for credit card users.

LIST OF FIGURES/TABLES

TABLE OF CONTENTS

1. INTRODUCTION

1. INTRODUCTION

1.1 PROJECT SCOPE

The project, "Credit Card Fraud Detection using State of the Art using Machine Learning and Deep Learning Algorithm," involves implementing cutting-edge machine learning and deep learning algorithms for Credit Card Fraud Detection. The study encompasses the exploration and utilization of diverse datasets containing both legitimate and fraudulent transactions. Various state-of-the-art algorithms, including machine learning models such as logistic regression, decision trees, random forests, support vector machines, and deep learning models such as neural networks and convolutional neural networks, will be employed. The project will assess the performance of each algorithm and explore ensemble methods to enhance the overall fraud detection system. Performance evaluation metrics, including precision, recall, F1 score, and the area under the Receiver Operating Characteristic curve, will be used to measure the effectiveness of the proposed models. The ultimate goal is to develop a robust and adaptive system that can accurately detect credit card fraud in real-time, contributing to the ongoing efforts in securing financial transactions against evolving fraudulent techniques.

1.2 PROJECT PURPOSE

The purpose of this project is to address the escalating threat of credit card fraud through the implementation of cutting-edge machine learning and deep learning algorithms. As the reliance on credit cards for financial transactions grows, so does the risk of fraudulent activities. The project aims to develop a sophisticated fraud detection system capable of identifying and preventing unauthorized transactions in real-time. By leveraging state-of-the-art techniques, including machine learning algorithms such as logistic regression, decision trees, and deep learning models like neural networks, the project seeks to enhance the accuracy and efficiency of fraud detection.

The ultimate goal is to provide financial institutions with a robust and adaptive tool to safeguard users from the evolving nature of credit card fraud in the digital era. This project seeks to leverage the capabilities of cutting-edge machine learning and deep learning algorithms to enhance the accuracy, sensitivity, and efficiency of fraud detection systems. By harnessing the power of these state-of-the-art technologies, the project aims to contribute to the development of a robust and proactive solution that can quickly and accurately identify fraudulent activities, thereby safeguarding financial institutions and protecting the interests of credit card users in the rapidly evolving digital ecosystem.

1.3 PROJECT FEATURES

This project encompasses a range of features that leverage cutting-edge machine learning and deep learning algorithms to create an advanced and robust system. Firstly, the project incorporates a comprehensive dataset, comprising both legitimate and fraudulent transactions, to facilitate effective model training and evaluation. The utilization of state-of-the-art machine learning algorithms, such as logistic regression, decision trees, random forests, and support vector machines, forms the foundation for establishing a baseline performance in fraud detection. The integration of deep learning models, including neural networks and convolutional neural networks (CNNs), takes the project to a higher level by capturing intricate patterns and non-linear relationships within the data. The adoption of these advanced algorithms aims to enhance the system's accuracy and efficiency, particularly in identifying subtle and evolving fraudulent patterns that may escape traditional detection methods.

2. SYSTEM ANALYSIS

2. SYSTEM ANALYSIS

2. SYSTEM ANALYSIS

System Analysis for Credit Card Fraud Detection using state-of-the-art machine learning and deep learning algorithms involves a comprehensive examination of the various components and processes within the proposed framework. At the core of this analysis is the exploration of diverse machine learning and deep learning models, each designed to uncover intricate patterns and anomalous behavior indicative of credit card fraud. The initial phase entails the acquisition and preprocessing of a comprehensive dataset, comprising both legitimate and fraudulent transactions, to facilitate effective model training.

2.1 PROBLEM DEFINITION

The problem addressed is the increasing prevalence of credit card fraud poses a significant challenge for financial institutions and users alike. Traditional methods of fraud detection are proving insufficient in addressing the sophisticated techniques employed by fraudsters. This study aims to address this problem by leveraging state-of-the-art machine learning and deep learning algorithms for credit card fraud detection. The primary objective is to enhance the accuracy and efficiency of fraud detection systems, providing financial institutions with a more robust tool to identify and prevent fraudulent transactions in real-time. The research aims to contribute to the ongoing efforts in developing advanced solutions to safeguard financial transactions and protect users from the evolving landscape of credit card fraud.

2.2 EXISTING SYSTEM

ML has many branches, and each branch can deal with different learning tasks. However, ML learning has different framework types. The ML approach provides a solution for CCF, such as random forest (RF). The ensemble of the decision tree is the random forest. Most researchers use the RF approach. To combine the model, we can use (RF) along with network analysis. This method is called APATE. Researchers can use different ML techniques, such as supervised learning and unsupervised techniques. ML algorithms, such as LR, ANN, DT, SVM and NB, are commonly used for CCF detection. The researcher can combine these techniques with ensemble techniques to construct solid detection classifiers. The linking of multiple neurons and nodes is known as an artificial neural network. A feed-forward perceptron multilayer is built up of numerous layers: an input layer, an output layer and one or more hidden layers. For the representation of the exploratory variables, the first layer contains the input nodes. With a precise weight, these input layers are multiplied, and each of the hidden layer nodes is transferred with a certain bias, and they are added together.

An activation function is then applied to create the output of each neuron for this summation, which is then transferred to the next layer. Finally, the algorithm's reply is provided by the output layer. The first set randomly used weights and formerly used the training set to minimise the error. All these weights were adjusted by detailed algorithms such as back propagation. The graphic model for contingency relationships between a set of variables is called the Bayesian belief network. The independence assumption in naïve Bayes is that it was developed to relax and allow for dependencies among variables.

2.2.1 DISADVANTAGES OF EXISTING SYSTEM

Following are the disadvantages of existing system:

- The system is not implemented Classification on Imbalanced Data.
- The system is not implemented CONVOLUTIONAL NEURAL NETWORK (CNN) for test and train the datasets.

2.3 PROPOSED SYSTEM

In this project, Feature selection algorithms are used to rank the top features from the CCF transaction dataset, which help in class label predictions. The deep learning model is proposed by adding a number of additional layers that are then used to extract the features and classification from the credit card farad detection dataset. To analyse the performance CNN model, apply different architecture of CNN layers. To perform a comparative analysis between ML with DL algorithms and proposed CNN with baseline model, the results prove that the proposed approach outperforms existing approaches. To assess the accuracy of the classifiers, performance evaluation measures, accuracy, precision, and recall are used. Experiments are performed on the latest credit cards dataset. It integrates cutting-edge machine learning and deep learning algorithms to enhance the accuracy and efficiency of fraud detection in credit card transactions.

2.3.1 ADVANTAGES OF THE PROPOSED SYSTEM

- The proposed system uses SUPERVISED MACHINE LEARNING APPROACHES which are effective for testing and training datasets.
- The proposed system implemented CNN is to minimise processing without losing key features by reducing the image to make predictions

2.4 FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and a business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. Three key considerations involved in the feasibility analysis:

- Economic Feasibility
- Technical Feasibility
- Social Feasibility

2.4.1 ECONOMIC FEASIBILITY

The developing system must be justified by cost and benefit. Criteria to ensure that effort is concentrated on a project, which will give best, return at the earliest. One of the factors, which affect the development of a new system, is the cost it would require.

The following are some of the important financial questions asked during preliminaryinvestigation:

- The costs conduct a full system investigation.
- The cost of the hardware and software.
- The benefits in the form of reduced costs or fewer costly errors.

Since the system is developed as part of project work, there is no manual cost to spend for the proposed system. Also all the resources are already available, it give an indication that the system is economically possible for development.

2.4.2 TECHNICAL FEASIBILITY

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

2.4.3 BEHAVIORAL FEASIBILITY

This includes the following questions:

- Is there sufficient support for the users?
- Will the proposed system cause harm?

The project would be beneficial because it satisfies the objectives when developed and installed. All behavioral aspects are considered carefully and conclude that the project is behaviorally feasible

2.5 HARDWARE & SOFTWARE REQUIREMENTS

2.5.1 HARDWARE REQUIREMENTS:

Hardware interfaces specify the logical characteristics of each interface between the software product and the hardware components of the system. The following are some hardware requirements.

- Processor : Pentium-IV
- Hard disk : 20 GB
- RAM : 4 GB (min)
- Key Board : Standard Windows Keyboard
- Mouse : Two or Three Button Mouse
- Monitor : SVGA

2.5.2 SOFTWARE REQUIREMENTS:

Software Requirements specifies the logical characteristics of each interface and software components of the system. The following are some software requirements,

- Operating system : Windows 7 Ultimate
- Coding Language : Python
- Front-End : Python
- Back-End : Django-ORM
- Designing : HTML, CSS, Javascript
- Data Base : MySQL (WAMP Server)

3. ARCHITECTURE

3. ARCHITECTURE

3.1 PROJECT ARCHITECTURE

This project architecture shows the procedure followed for classification, starting from input to final prediction.

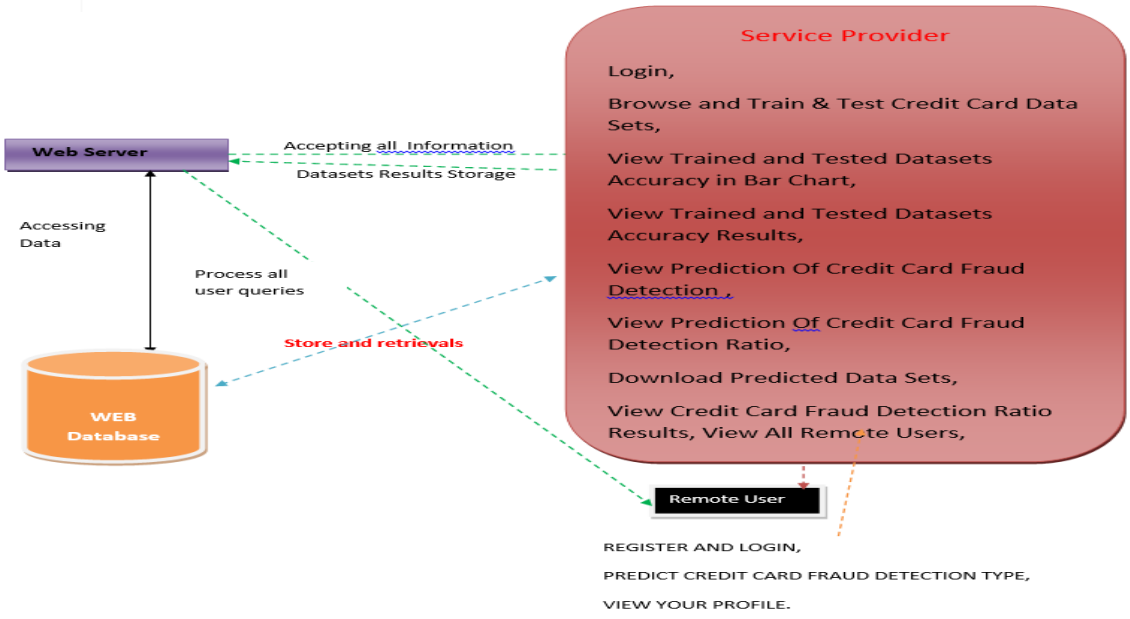

Figure 3.1: Project Architecture for Credit Card Fraud Detection using State of Art Machine Learning and Deep Learning Algorithms

3.2 DESCRIPTION

This project involves the integration of feature selection algorithms, machine learning (ML), and deep learning (DL) techniques for fraud detection in credit card transactions. Feature selection algorithms are employed to rank the top features from the CCF (Credit Card Fraud) transaction dataset. Different architectures of Convolutional Neural Networks (CNN) layers are applied to analyze the performance of the CNN model. Experiments are conducted on the latest credit card dataset, suggesting that the project incorporates up-to-date data for training and evaluation.

3.3 USE CASE DIAGRAM

This use case diagram represents how a system can be used to detect credit card fraud. This use case diagram provides a high-level overview of the functionalities and interactions involved in the credit card fraud detection project, emphasizing the user's role in initiating and overseeing various tasks within the system.

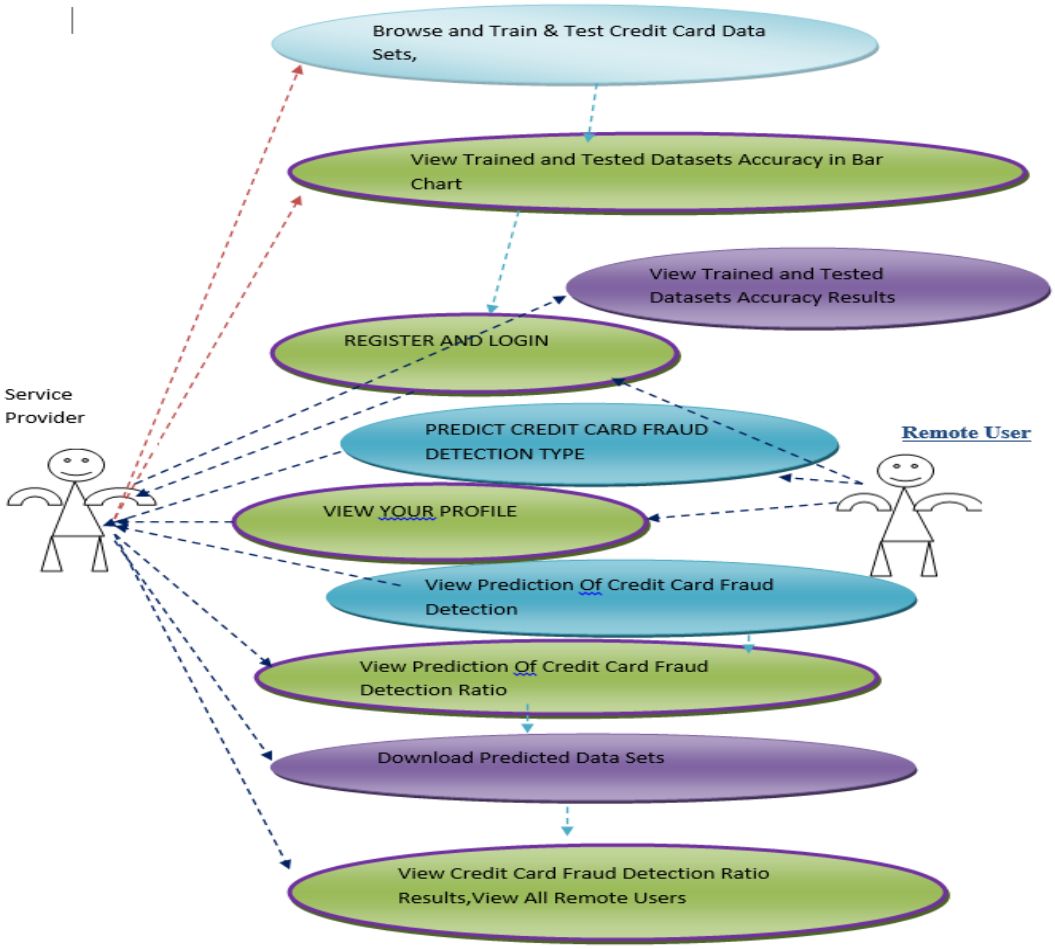

Figure 3.2: Use Case Diagram for Credit Card Fraud Detection using State of Art Machine Learning and Deep Learning Algorithms

3.4 CLASS DIAGRAM

A class diagram is a static structure diagram that illustrates a system's structure by presenting its classes, attributes, methods (operations), and object relationships.

Creating a class diagram based on the abstract you provided involves identifying key entities, their attributes, and relationships. Class diagrams are essential for visualizing the structure of a system and understanding how different classes collaborate to achieve specific functionalities.

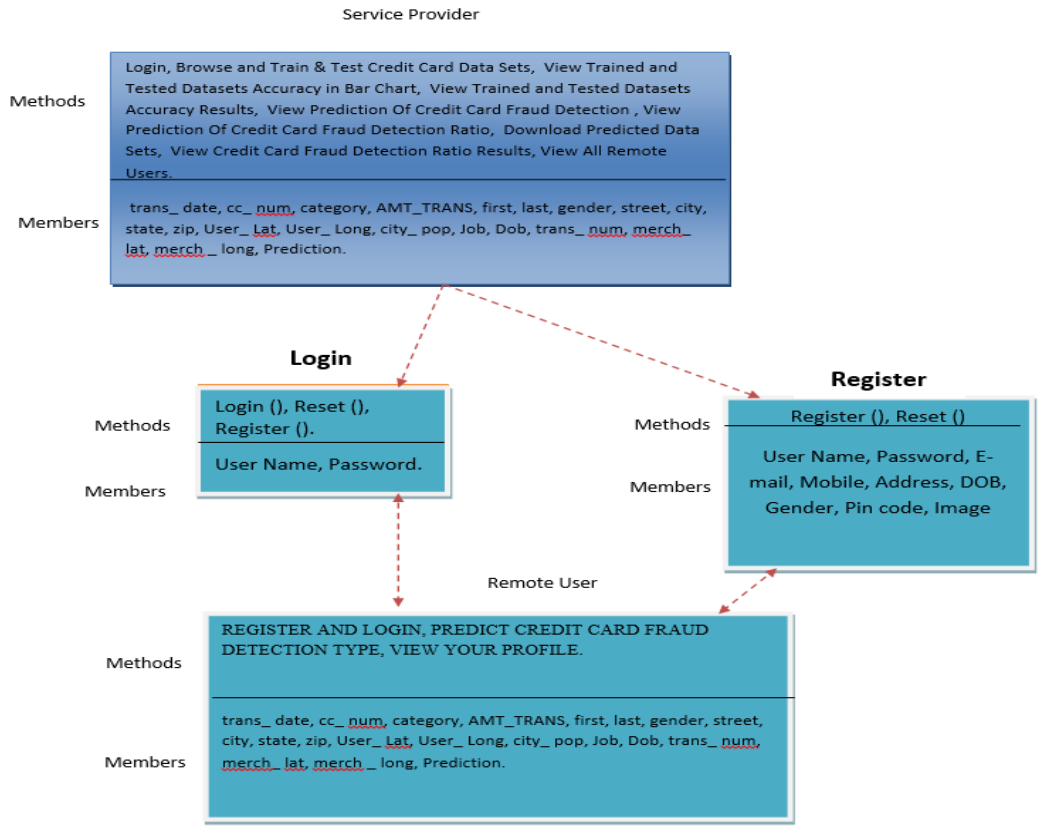

Figure 3.3: Class Diagram for Credit Card Fraud Detection using State of Art Machine Learning and Deep Learning Algorithms

3.5 SEQUENCE DIAGRAM

A sequence diagram visually depicts object interactions in chronological order, showcasing the involved objects and their message exchanges to execute a scenario. These diagrams are often linked to use case realizations in the system's logical development view. Sequence diagrams are valuable for visualizing the flow of control and collaboration among objects in a system, especially in scenarios where the order of interactions is crucial.

In a sequence diagram, objects are represented by lifelines, and the interactions between these objects are visualized through messages exchanged along the lifelines. Each message between objects is annotated with the specific operation or communication occurring, helping to depict the flow of control and collaboration in the system.

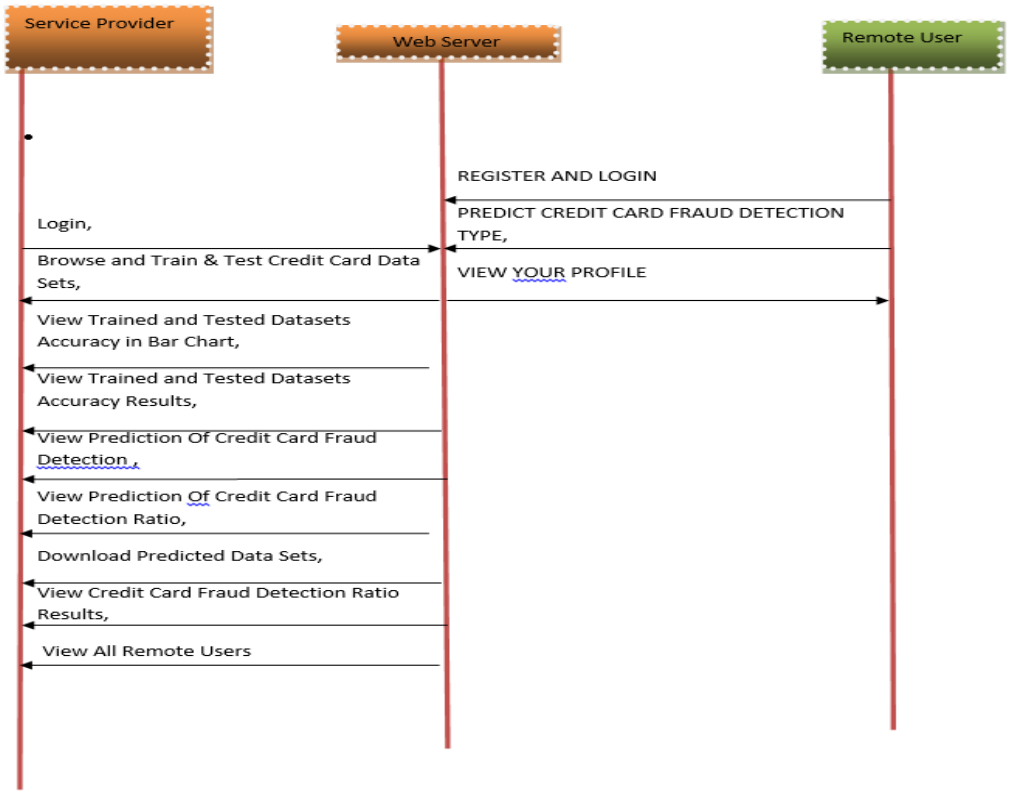

Figure 3.4: Sequence Diagram for Credit Card Fraud Detection using State of Art Machine Learning and Deep Learning Algorithms

3.6 ACTIVITY DIAGRAM

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. They can also include elements showing the flow of data between activities through one or more data stores. It provides a high-level view of the dynamic aspects of a system, emphasizing the sequence of actions and the decisions made during the execution of a particular process. In an activity diagram, nodes represent activities, which can range from simple operations to complex processes, and arrows depict the flow of control between these activities. Decision points and branches are articulated through conditional and looping constructs, enhancing the diagram's ability to model complex logic and parallel activities.

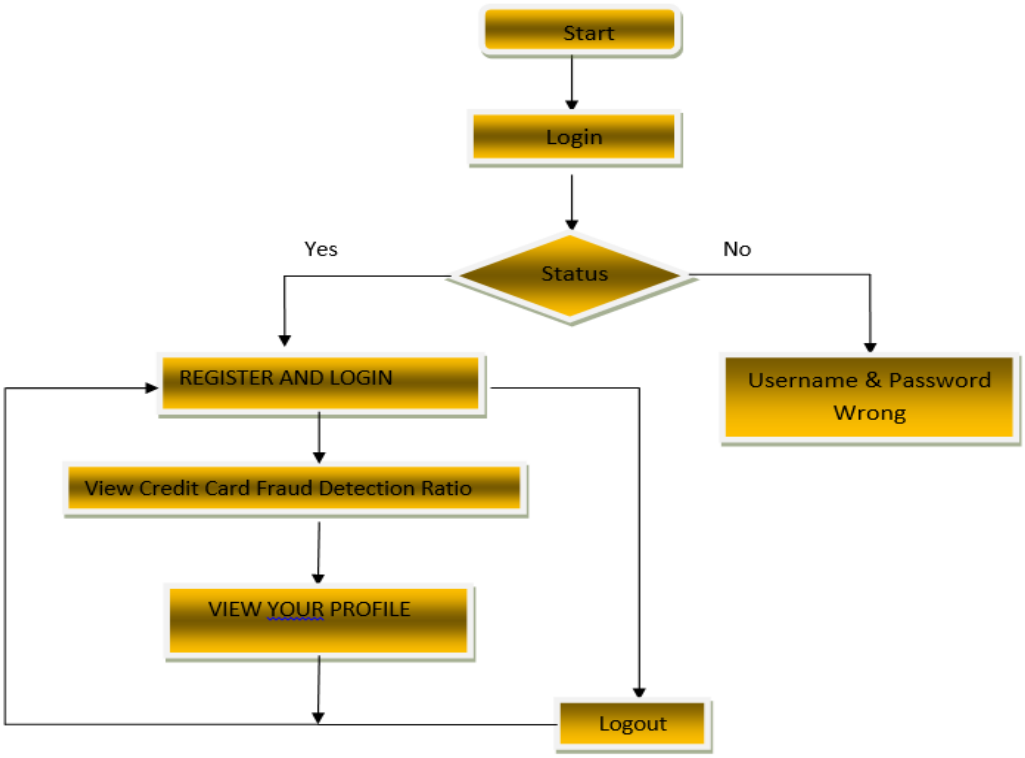

Figure 3.5: Activity Diagram for Credit Card Fraud Detection using State of Art Machine Learning and Deep Learning Algorithms

4.CONCLUSION

In conclusion, credit card fraud detection using machine learning and deep learning algorithms has proven to be a powerful and effective approach to safeguarding financial transactions. These techniques have enabled the development of robust fraud detection systems that can learn and adapt to evolving fraud patterns, ultimately reducing false positives and increasing the accuracy of identifying fraudulent transactions. The combination of feature engineering, anomaly detection, and predictive modeling has significantly enhanced the security of credit card transactions, providing both financial institutions and customers with greater peace of mind. However, it's important to remain vigilant in continuously updating and improving these algorithms to stay one step ahead of increasingly sophisticated fraudsters in the ever-evolving landscape of cybercrime.

Ensemble methods, combining multiple algorithms, can enhance overall performance. Feature engineering is crucial for extracting meaningful information, and model interpretability is essential for understanding decision-making processes. Continuous monitoring and adaptation of models are necessary due to evolving fraud patterns. Striking a balance between accuracy and computational efficiency is vital for real-time fraud detection in the dynamic landscape of financial transactions. Regular updates and collaboration with domain experts help ensure robust and adaptive fraud detection systems.