A PROJECT REPORT ON

Credit Card Fraud Detection Using Machine Learning

SUBMITTED TO THE SAVITRIBAI PHULE PUNE UNIVERSITY,
PUNE IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE

BACHELOR OF ENGINEERING (Computer Engineering)

BY

Mr. Aditya Raj Singh (Exam Seat No: 71725964B)

Mr. Hrishikesh Morade (Exam Seat No: 71726130B)

Mr. Sanket Sable (Exam Seat No: 71726081L)

Mr. Sunil Hule (Exam Seat No: 71726030F)

Under The Guidance of

Prof B.R. Patle

DEPARTMENT OF COMPUTER ENGINEERING

SKN SINHGAD INSTITUTE OF TECHNOLOGY AND SCIENCE
KUSGAON (BK), LONAVALA 410401

SAVITRIBAI PHULE PUNE UNIVERSITY, PUNE
2020-2021

CERTIFICATE

This is to certify that the Project Entitled

Credit Card Fraud Detection Using Machine Learning Submitted by

Mr. Aditya Raj Singh (Exam Seat No: 71725964B)
Mr. Hrishikesh Morade (Exam Seat No: 71726130B)
Mr. Sanket Sable (Exam Seat No: 71726081L)
Mr. Sunil Hule (Exam Seat No: 71726030F)

is a bonafide work carried out by Students under the supervision of Prof Mr.Avinash Bagul and it is submitted towards the partial fulfillment of the requirement of Bach elor of Engineering (Computer Engineering) Project.

Prof Mr.B.R. Patle Prof.G.M.Kadam Internal Guide H.O.D
Dept. of Computer Engg. Dept. of Computer Engg.

Dr.M.S.Rohokale
Principal

Signature of Internal Examiner Signature of External Examiner

PROJECT APPROVAL SHEET

A Project Report Titled as

Credit Card Fraud Detection Using Machine Learning

Is verified for its originality in documentation, problem statement,proposed work and implementation successfully completed by

Mr. Aditya Raj Singh (Exam Seat No: 71725964B)
Mr. Hrishikesh Morade (Exam Seat No: 71726130B)
Mr. Sanket Sable (Exam Seat No: 71726081L)
Mr. Sunil Hule (Exam Seat No: 71726030F)

at

DEPARTMENT OF COMPUTER ENGINEERING

SKN SINHGAD INSTITUTE OF TECHNOLOGY AND SCIENCE SAVITRIBAI PHULE

PUNE UNIVERSITY,PUNE

ACADEMIC YEAR 2020-2021

Prof B.R.Patle Prof.G.M.Kadam Internal Guide H.O.D.
Dept. of Computer Engg. Dept. of Computer Engg SKNSIT, Dept. of Computer

Engineering 2020-21 I

Abstract

We aim to avoid the robberies and wrong person misuse the so that we can make them to lead their life safely and securely. The proposed system is designed based on the intelligence system to ensure the usage without any hesitation and make the world to be a part of digitization. Once customer inserts the card into the , a session is initiated, the system starts face detection using the camera located near the and builds a temporary identity database for the customer and user face verification is performed on the .Valid user would continue the normal process but the Invalid user cannot be access the card so they give the secondary password to the system automatically the unauthorized person would continue the transaction.

Acknowledgments

It gives us great pleasure in presenting the preliminary project report on Credit card Fraud Detection Using Machine Learning.

We would like to take this opportunity to thank our internal guide Prof B.R.Patle for giving us all the help and guidance we needed. We are really grateful to them for their kind support. Their valuable suggestions were very helpful.

We are also grateful to Prof.G.M.Kadam, Head of Computer Engineering Depart ment, SKN Sinhgad Institute of Technology And Science for indispensable support and suggestions.

Mr. Aditya Raj Singh (Exam Seat No: 71725964B)
Mr. Hrishikesh Morade (Exam Seat No: 71726130B)
Mr. Sanket Sable (Exam Seat No: 71726081L)
Mr. Sunil Hule (Exam Seat No: 71726030F)

INDEX

List of Figures

List of Tables

SYNOPSIS

0.1 PROJECT TITLE

“ Credit Card Fraud Detection Using Machine Learning”

0.2 PROJECT OPTION

Final Year Project

0.3 INTERNAL GUIDE

Prof Mr.Avinash Bagul

0.4 TECHNICAL KEYWORDS (AS PER ACM KEYWORDS)

0.4.1 Categories and Subject Descriptors:

K.4.4 [Electronic Commerce]: Security.
K.6.5 [Security and Protection]: Unauthorized access.

2) General Terms:
Design, Human factors, Security.

3) Keywords:
, security, fraud, face recognition, LRR, OTP

0.5 PROBLEM STATEMENT

Presently, systems use no more than an access card which usually has a mag netic stripe (magstripe) and a fixed Personal Identification Number (PIN) for iden tity verification. Some other cases utilize a chip and a PIN which sometimes has a magstripe in case the chip fails as a backup for identification purposes. This method is not very secure and prone to increase in criminal activities. The need for a novel,

simple as well as secure method of access is thus imperative. In the present work, a PIN is generated by the user and this PIN is made available

to the system by the means of a Subscriber Identity Module (SIM) in the user's Mobile Phone. This PIN can be trapped by any user and can be lead to fraud.

0.6 ABSTRACT

We aim to avoid the robberies and wrong person misuse the so that we can make them to lead their life safely and securely. The proposed system is designed based on the intelligence system to ensure the usage without any hesitation and make the world to be a part of digitization. Once customer inserts the card into the , a session is initiated, the system starts face detection using the camera located near the and builds a temporary identity database for the customer and user face verification is performed on the .Valid user would continue the normal process but the Invalid user cannot be access the card so they give the secondary password to the system automatically the unauthorized person would continue the transaction.

0.7 GOALS AND OBJECTIVES

An Automatic Teller Machine ( ) is a computerized machine that uses to with draw the cash from customer's respective bank account. As financial user prefer for cash withdrawals, cash deposits many other transaction, the banks are fo cusing a lot over the security of s. should be protected properly from the criminal activities or from any unwanted things. To overcome these issues we have proposed the system where the chances of fraud will be minimizing to the great extends.

- Carry out a survey of various authentication models for Automated Teller Ma chines.
- Design and develop an authentication model that uses two-factors i.e. biomet rics and password authentication.
- Simulate the two-factor model and evaluate its performance.

• Demonstrate the practicality of the facial recognition component.

0.8 PLAN OF PROJECT EXECUTION

Sr. No.	Month Sheduled	Phase	Work Done
1	June-August	Topic Seraching	Topic Searched
2	August-September	Topic Selection	Topic Selected
3	August-September	Project Confirmation	Project Confirmed
4	August-September	Literature Survey	Literature Survey Done
5	September-Octobe r	Requirement Analysis	Requirement Analysis Done
6	September-Octobe r	Requirement Gathering	Requirements Gathered
7	November-Decem ber	Designing	Architecture Design
8	November-Decem ber	Designing Test	GUI Tested
9	November-Decem ber	Database Creation	Database Tested
10	January-February	Coding	Coded Different modules
11	January-February	Database And Module Connectivity	Connectivity Done
12	March	Testing of Project	Project Tested
13	April	Result Analysis	Result Analysis

CHAPTER 1

TECHNICAL KEYWORDS

.

1.1 TECHNICAL KEYWORDS (AS PER ACM KEYWORDS)

K.4.4 [Electronic Commerce]: Security.

K.6.5 [Security and Protection]: Unauthorized access.

2) General Terms:

Design, Human factors, Security.

3) Keywords:

, security, fraud, face recognition, LRR, OTP

CHAPTER 2

INTRODUCTION

2.1 OVERVIEW

Abstract::

t's vital that mastercard companies are ready to identify fraudulent credit card transactions so that customers are not charged for items that they didn't purchase. Such problems are often tackled with Data Science and its importance, along with Machine Learning, cannot be overstated. This project intends for instance the modelling of a knowledge set using machine learning with mastercard Fraud Detection. The Credit Card Fraud Detection Problem includes modelling past credit card transactions with the info of those that clothed to be fraud. This model is then wont to recognize whether a replacement transaction is fraudulent or not. Our objective here is to detect 100% of the fraudulent transactions while minimizing the incorrect fraud classifications. Credit Card Fraud Detection is a typical sample of classification. In this process, we have focused on analysing and pre-processing data sets also because the deployment of multiple anomaly detection algorithms such as Local Outlier Factor and Isolation Forest algorithm on the Principle Component Analysis transformed Credit Card Transaction data.

Proposed System:

A mechanism is developed to determine whether the given transaction id is fraud or not. • • A mechanism uses simple Machine Learning Model (Logistic Regression) to detect fraud transactions. • This Machine Learning Model works on basis of spending habit of user. • When we have large number of features in dataset then feature selection is very important part in our Machine Learning. • As we use feature selection it gives us most important feature and this feature selection gives us more accuracy.

2.3 LITERATURE SURVEY:

Fraud act because the unlawful or criminal deception intended

toresult in financial or personal benefit. It is a deliberate act thatis against the rule or policy with an aim to achieveunauthorized financial benefit. Numerous literatures pertaining to anomaly or fraud detection in this domain are published already and are available for public usage. A comprehensive survey conducted by Clifton Phua and his associates have disclosed that techniques employed in this domain include data mining applications, automated fraud detection, adversarial detection. In another paper, Suman, Research Scholar, GJUS&T at Hisar HCE presented techniques like Supervised and Unsupervised Learning for credit card fraud detection. Even though these methods and algorithms fetched an unexpected success in some areas, they didn't provide a permanent and consistent solution to fraud detection. A alike research domain was presented by Wen-Fang YU and Na Wang where they used Outlier detection, Outlier detection mining and Distance sum algorithms to accurately predict

fraudulent transaction in an emulation experiment of credit card transaction data set of 1 certain commercial bank. Outlier mining may be a field of knowledge mining which is basically used in monetary and internet fields. It deals with detecting objects that are detached from the most system i.e. the transactions that aren't genuine. They have taken attributes of customer's behaviour and supported the worth of these attributes they've calculated that distance between the observed value of that attribute and its predetermined value. Unconventional techniques such as hybrid data mining/complex network classification algorithm is able to recognize illegal instances in an actual card transaction data set, based on network reconstruction algorithm that allows creating representations of the deviation of 1 instance from a reference group have showed efficient typically on medium sized online transaction. There have also been efforts to progress from a totally new aspect. Attempts have been made to improve the alert feedback interaction in case of fraudulent transaction. In case of fraud transaction, the authorised system would be alerted and a feedback would be sent to deny the continued transaction. Artificial Genetic Algorithm, one among the approaches that shed new light during this domain, countered fraud from a special direction. It proved accurate find out the fraudulent transactions and minimizing the number of false alerts. Even though, it was accompanied by classification problem with variable misclassification costs

CHAPTER 3

PROBLEM DEFINITION AND SCOPE

3.1 PROBLEM DEFINITION

With the growth of e-commerce websites, people and financial companies

rely on online services to carry out their transactions that have led to an exponential increase in the credit card frauds [1]. Fraudulent credit card transactions lead to a loss of huge amount of money. The design of an effective fraud detection system is necessary in order to reduce the losses incurred by the customers and financial companies [2]. Research has been done on many models and methods to prevent and detect credit card frauds. Some credit card fraud transaction datasets contain the problem of imbalance in datasets. A good fraud detection system should be able to identify the fraud transaction accurately and should make the detection possible in real-time transactions. Fraud detection can be divided into two groups: anomaly detection and misuse detection. Anomaly detection systems bring normal transaction to be trained and use techniques to determine novel frauds. Conversely, a misuse fraud detection system uses the labeled transaction as normal or fraud transaction to be trained in the database history. So, this misuse detection system entails a system of supervised learning and anomaly detection system a system of unsupervised learning [2]. Fraudsters masquerade the normal behavior of customers and the fraud patterns are changing rapidly so the fraud detection system needs to constantly learn and update. Credit card frauds can be broadly classified into three categories, that is, traditional card related frauds (application, stolen, account takeover, fake and counterfeit), merchant related frauds (merchant collusion and triangulation) and Internet frauds (site cloning, credit card generators and false merchant sites) [1].

3.3 METHODOLOGY USED FOR PROBLEM SOLVING

The approach that this paper proposes, uses the newest machine learning algorithms to detect anomalous activities, called outliers.

Firstly, we obtained our dataset from Kaggle, a data analysis website which provides datasets. Inside this dataset, there are 31 columns in which 28 are named as v1-v28 to protect sensitive data. The other columns represent Time, Amount and sophistication . Time shows the time gap between the primary transaction and therefore the following one. Amount is the amount

of money transacted. Class 0 represents a legitimate transaction and 1 represents a fraudulent one. We plot different graphs to see for inconsistencies within the dataset and to visually comprehend it
After checking this dataset, we plot a histogram for each column. this is often done to urge a graphical representation of the dataset which may be wont to verify that there are not any missing any values in the dataset. This is done to ensure that we don't require any missing value imputation and the machine learning algorithms can process the dataset smoothly. After this analysis, we plot a heatmap to urge a coloured representation.

3.4 APPLICATIONS

Application will be used by Systems.

3.5 HARDWARE RESOURCES REQUIRED

Sr. No.	Parameter	Minimum Requirement
1	Processor	Core I7
2	RAM	3 GB.

Table 3.1: Hardware Requirements

3.6 SOFTWAREWARE RESOURCES REQUIRED

Sr. No.	Parameter	Minimum Requirement
1	OPERATING SYSTEM	Windows 7/8./10
2	CODING LANGUAGE	Python

3	IDE	Spyder

Table 3.2: Software Requirements

CHAPTER 4

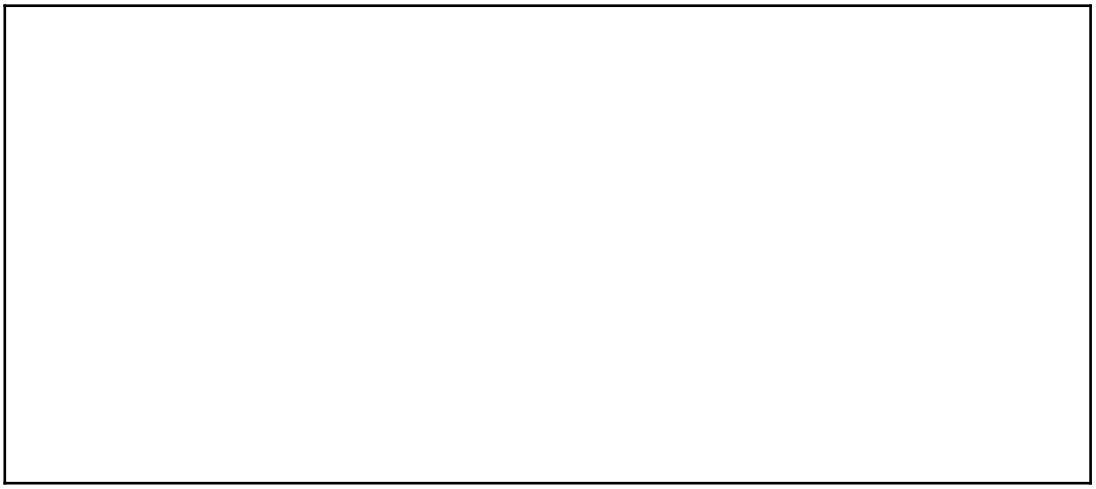
PROJECT PLAN

4.1 PROJECT ESTIMATES

4.1.1 Reconciled Estimates

4.2 RISK MANAGEMENT

4.2.1 Overview of Risk Mitigation, Monitoring, Management



4.2.1.1 Time Estimates

Approximately 11 months

4.2.2 Project Resources

Windows , eclipse , 3 GB RAM, High speed internet connection.

4.3 RISK MANAGEMENT W.R.T. NP HARD ANALYSIS This

section discusses Project risks and the approach to managing

them.

4.3.1 Risk Identification

For risks identification, review of scope document, requirements specifications and schedule is done. Answers to questionnaire revealed some risks. Each risk is catego rized as per the categories mentioned in [?]. Please refer table for all the risks. You can refereed following risk identification questionnaire.

1. Have top software and customer managers formally committed to support the project?

Ans-Not apllicaable.

2. Are end-users enthusiastically committed to the project and the system/product to be built?

Ans-Not known at this time.

3. Are requirements fully understood by the software engineering team and its cus tomers?

Ans-Yes

4. Have customers been involved fully in the definition of requirements? Ans-Not applicable

5. Do end-users have realistic expectations?

Ans-Not applicable

6. Does the software engineering team have the right mix of skills? Ans-yes

7. Are project requirements stable?

Ans-Yes

8. Is the number of people on the project team adequate to do the job? Ans-Not applicble

9. Do all customer/user constituencies agree on the importance of the project and on the requirements for the system/product to be built?

Ans-Not applicable

4.4 TASK NETWORK

Project planning is part of project management, which relates to the use of sched ules such as Gantt charts to plan and subsequently report progress within the project environment. Initially, the project scope is defined and the appropriate methods for completing the project are determined. Following this step, the durations for the various tasks necessary to complete the work are listed and grouped into a work breakdown structure. Project planning is often used to organize different areas of a project, including project plans, work loads and the management of teams and individuals. The logical dependencies between tasks are defined using an activity network diagram that enables identification of the critical path. Project

planning is

inherently uncertain as it must be done before the project is actually started. There fore the duration of the tasks is often estimated through a weighted average of op timistic, normal, and pessimistic cases. The critical chain method adds ”buffers” in the planning to anticipate potential delays in project execution. Float or slack time in the schedule can be calculated using project management software. Then the necessary resources can be estimated and costs for each activity can be allocated to each resource, giving the total project cost. At this stage, the project schedule may be optimized to achieve the appropriate balance between resource usage and project duration to comply with the project objectives. Once established and agreed, the project schedule becomes what is known as the baseline schedule. Progress will be measured against the baseline schedule throughout the life of the project. Analyzing progress compared to the baseline schedule is known as earned value management. The inputs of the project planning phase 2 include the project charter and the concept proposal. The outputs of the project planning phase include the project requirements, the project schedule, and the project management plan. The Project Planning can be done manually. However, when managing several projects, it is usually easier and faster to use project management software.

4.5 TIMELINE CHART

A Gantt chart is constructed with a horizontal axis representing the total time span of the project, broken down into increments (for example, days, weeks, or months) and a vertical axis representing the tasks that make up the project (for example, if the project is outfitting your computer with new software, the major tasks involved might be: conduct research, choose software, install software). Horizontal bars of varying lengths represent the sequences, timing, and time span for each task. Using the same example, you would put ”conduct research” at the top of the vertical axis and draw a

bar on the graph that represents the amount of time you expect to spend on the research, and then enter the other tasks below the first one and representative bars at the points in time when you expect to undertake them. The bar spans may overlap, as, for example, you may conduct research and choose software during the same time span. As the project progresses, secondary bars, arrowheads, or darkened

bars may be added to indicate completed tasks, or the portions of tasks that have been completed. A vertical line is used to represent the report date. Gantt charts give a clear illustration of project status, but one problem with them is that they don't indicate task dependencies - you cannot tell how one task falling behind schedule affects other tasks.

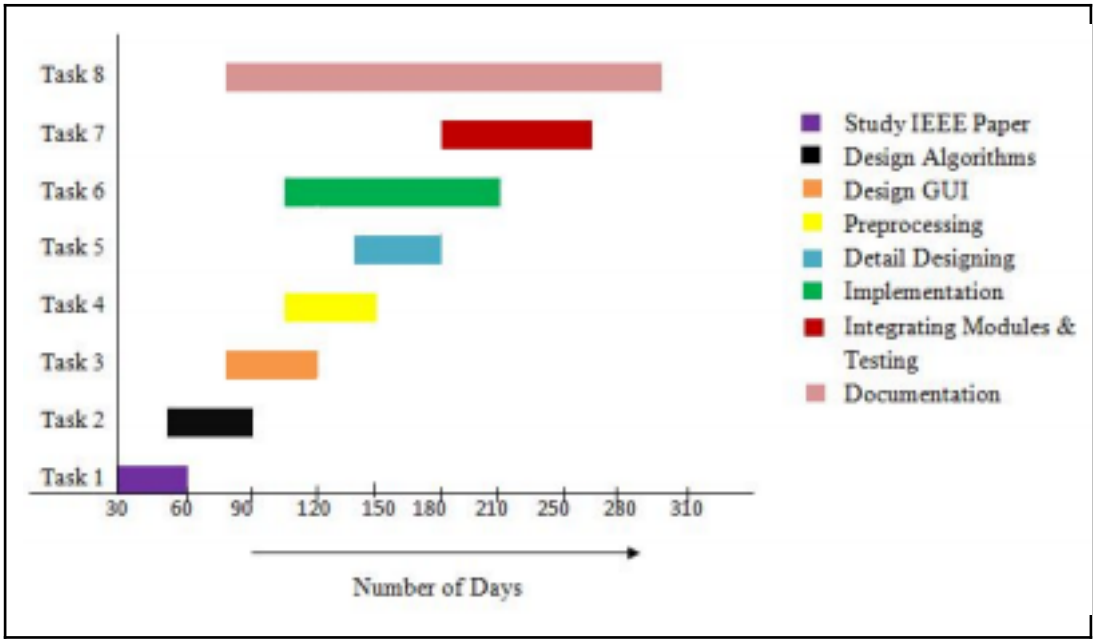


Figure 4.1: Timeline Chart

4.6 TEAM ORGANIZATION

4.6.1 Team structure

4.7 TEAM ORGANIZATION

Team consists of only Four members and proper planning mechanisms are used and role are defined.

4.7.1 Team structure

The team structure for the project is identified. Roles are defined. SKNSIT, Dept. of

4.7.2 Management reporting and communication

Well planning mechanisms are used for progress reporting and inter/intra team com munication are identified as per requirements of the project.

CHAPTER 5

SOFTWARE REQUIREMENT SPECIFICATION (SRS IS TO BE PREPARED USING RELEVANT MATHEMATICS DERIVED AND SOFTWARE ENGG.INDICATORS IN ANNEX A AND B)

5.1 INTRODUCTION

5.1.1 Purpose and Scope of Document

In order to provide reliable security solution to the people, the concept of se curity system based on face detection is emerged. The Area of work is basically fo cused on Design and Implementation of Face Detection based Security System using LRR algorithm. Limitations of existing system are overcome in our proposed system. In order to make any transaction, system will provide to option to process. First Self user, where in system will ask for “Detect Face” and allow to process trans action if it matches with Image store in banks database otherwise system will decline the transaction after couple of warnings. Second Guest User, where in system will ask for “OTP” and allow to process transaction if Guest User enter the correct OTP which has been sent to authorised User.

5.1.2 Overview of responsibilities of Developer

1. To have understanding of the problem statement.
2. To know what are the hardware and software requirements of Proposed system. 3. To have understanding of proposed system.
4. To do planning various activities with the help of planner.
5. Designing, programming, testing etc.

5.2 USAGE SCENARIO

This section provides various usage scenarios for the system to be developed.

5.2.1 Use-cases

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.

5.2.2 Use Case View

Use Case Diagram. Example is given below

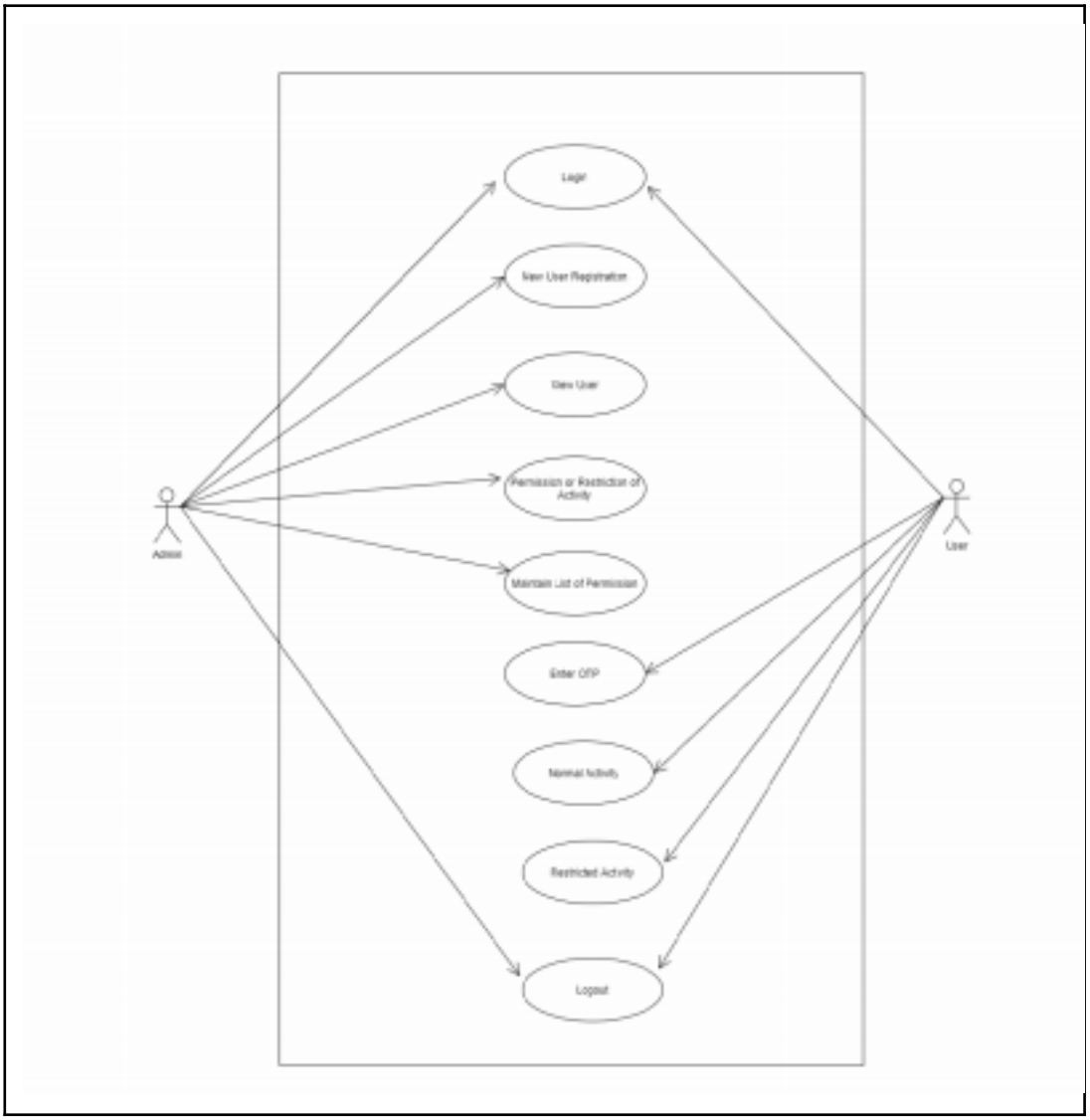


Figure 5.1: Use case diagram

5.3 DATA MODEL AND DESCRIPTION

5.3.1 Data Description

Describing and documenting data is essential in ensuring that the researcher, and others who may need to use the data, can make sense of the data and understand the processes that have been followed in the collection, processing, and analysis of

the data. Research data are any physical and/or digital materials that are collected, observed, or created in research activity for purposes of analysis to produce original research results or creative works.

5.3.2 Data objects and Relationships

A data object is a part of the repository whose content can be addressed and inter preted by the program. All data objects must be declared in the ABAP program and are not persistent, meaning that they only exist while the program is being executed. Before you can process persistent data (such as data from a database table or from a sequential file), you must read it into data objects first. Conversely, if you want to retain the contents of a data object beyond the end of the program, you must save it in a persistent form.

5.4 FUNCTIONAL MODE LAND DESCRIPTION

5.4.1 Safety Requirements

The application is designed in modules where errors can be detected and fixed easily. This makes it easier to install and update new functionality if required.

5.4.2 Security Requirement:

All data will be encrypted using strong encryption algorithm and according to loca tion encryption is done.

5.4.3 Software Quality Attributes

Our software has many quality attribute that are given below:-

- Adaptability: This software is adaptable by all users.
- Availability: This software is freely available to all users. The availability of the software is easy for everyone.
- Maintainability: After the deployment of the project if any error occurs

then it can be easily maintained by the software developer.

- Reliability: The performance of the software is better which will increase the reliability of the Software.
- User Friendliness: Since, the software is a GUI application; the output gener ated is much user friendly in its behavior.
- Integrity: Integrity refers to the extent to which access to software or data by unauthorized persons can be controlled.
- Security: Users are authenticated using many security phases so reliable secu rity is provided.
- Testability: The software will be tested considering all the aspects.

5.4.4 Activity Diagram:

- Activity diagrams are graphical representations of workflows of stepwise ac tivities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control.

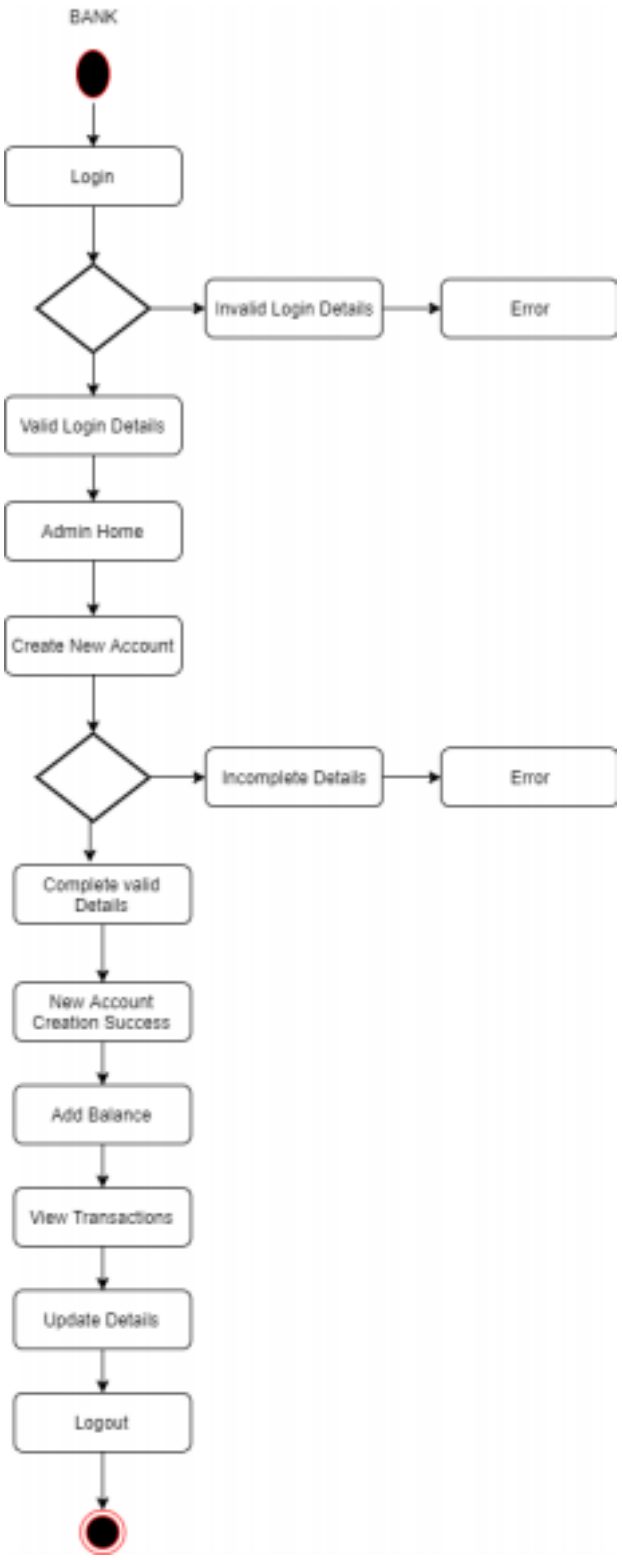


Figure 5.2: Activity Diagram - Bank Admin

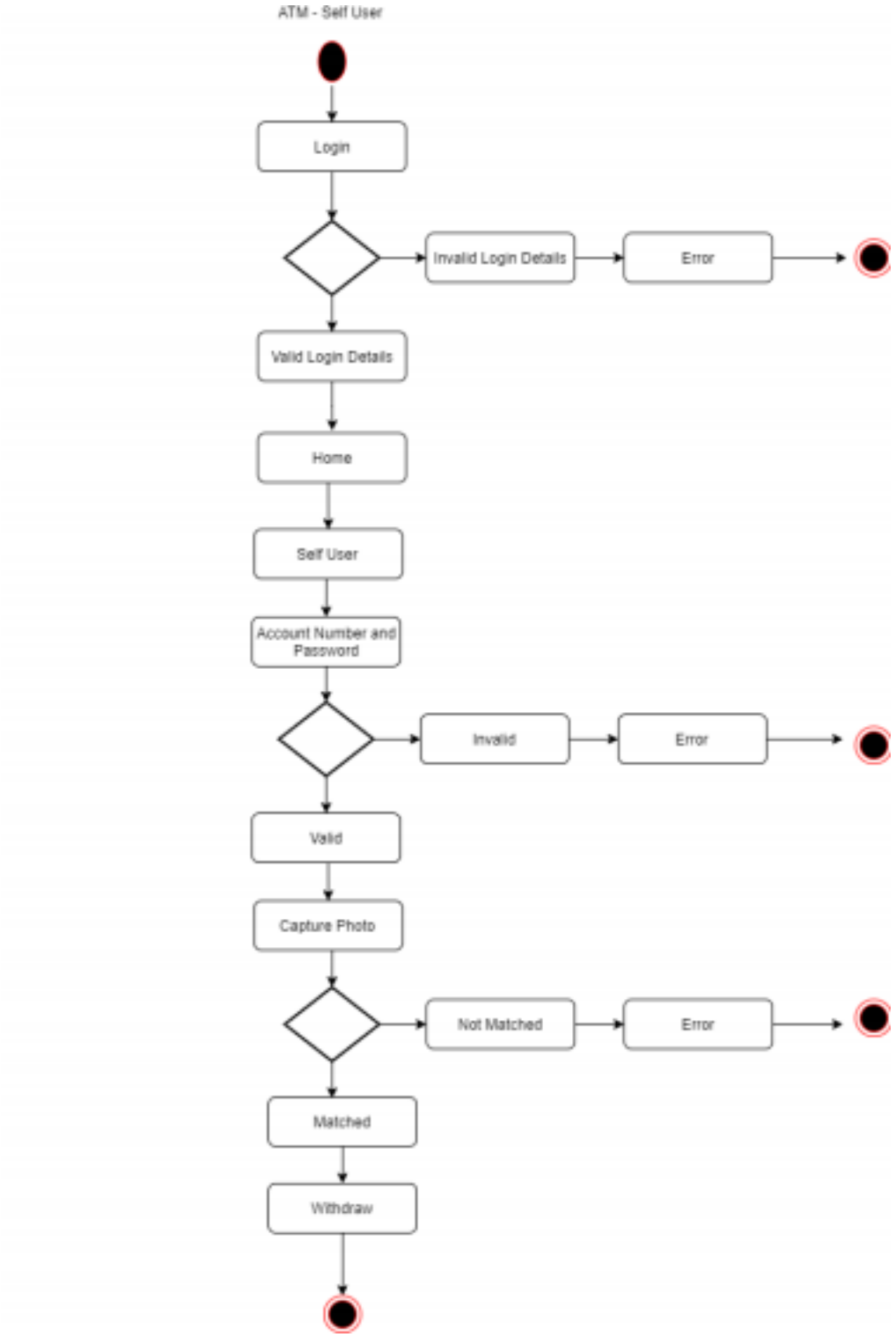


Figure 5.3: Activity Diagram - Self User

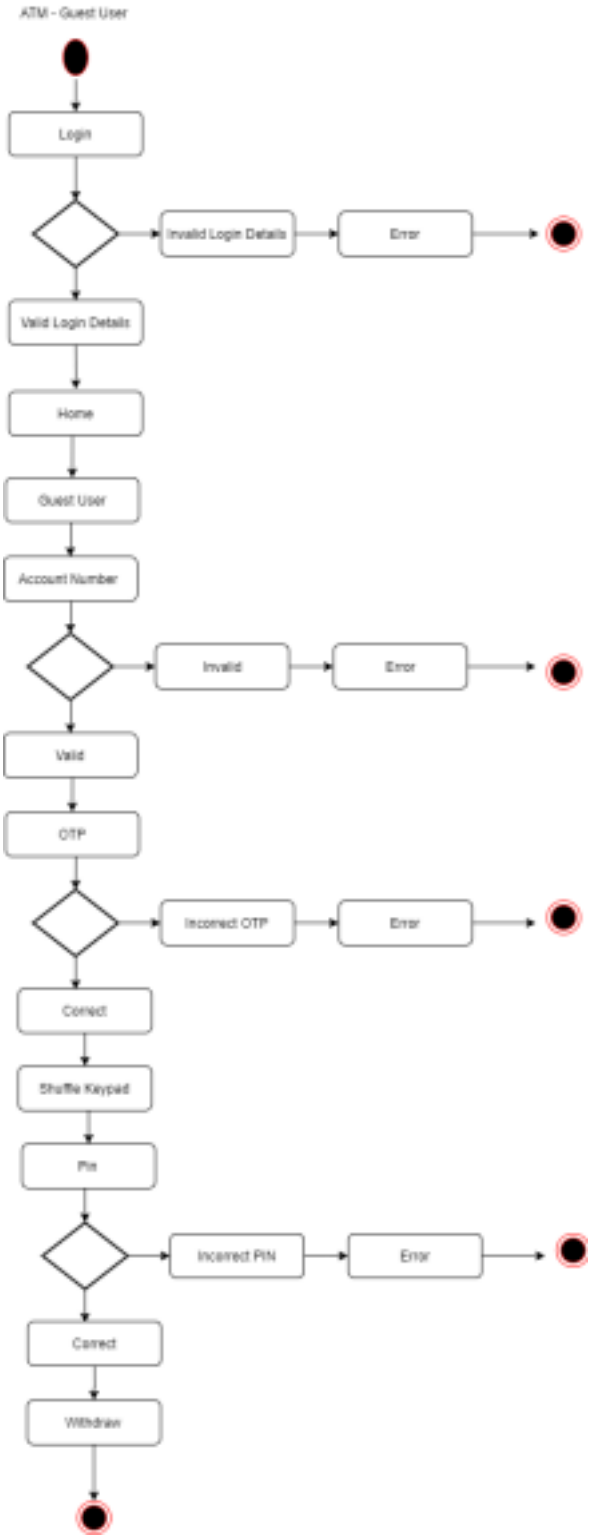


Figure 5.4: Activity Diagram - Guest User

5.4.5 Sequence Diagram:

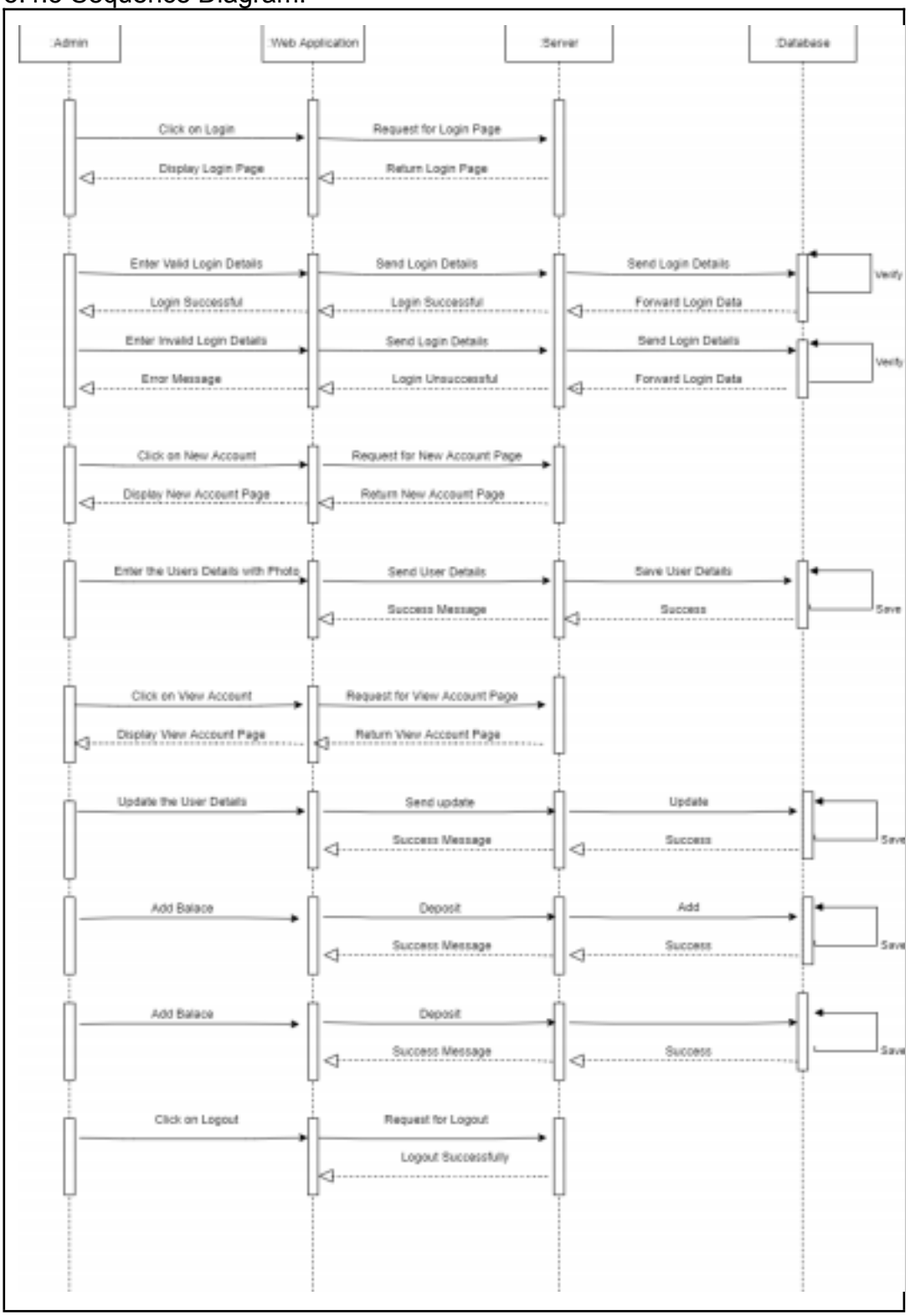


Figure 5.5: Sequence Diagram - Bank Admin

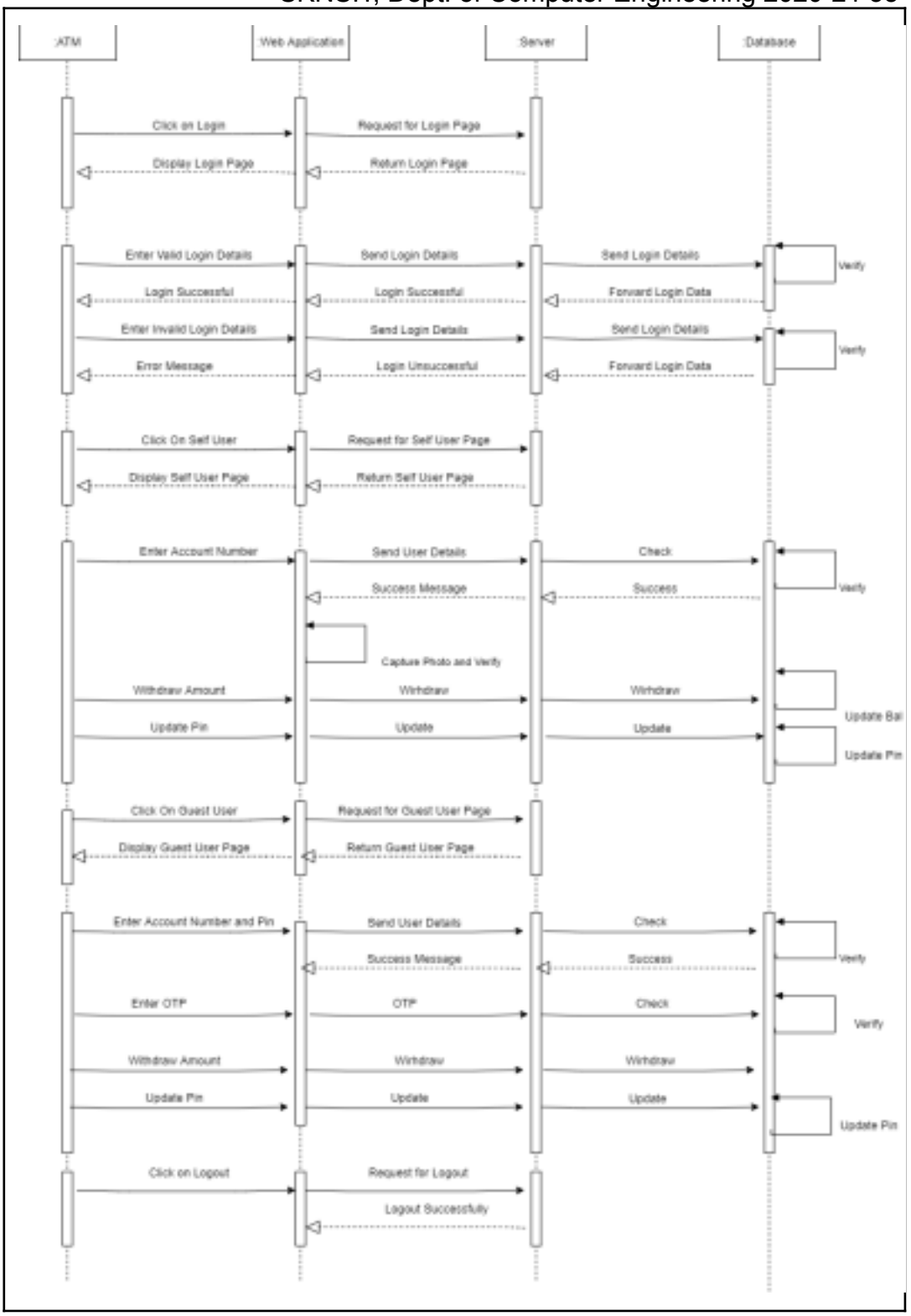


Figure 5.6: Sequence Diagram - User

5.4.6 Non Functional Requirements:

5.4.6.1 Interface Requirements

- High Speed Internet
- Router

5.4.6.2 Performance Requirements

- Laptops with latest configuration

5.4.7 Design Constraints

1. Apache Tomcat webserver.
2. SQLYog community/XAMPP Server.

5.4.8 Software Interface Description

The software interface(s)to the outside world is(are) described. The requirements for interfaces to other devices/systems/networks/human are stated.

CHAPTER 6

DETAILED DESIGN DOCUMENT USING APPENDIX A AND B

6.1 INTRODUCTION

In order to provide reliable security solution to the people, the concept of se curity system based on face detection is emerged. The Area of work is basically fo cused on Design and Implementation of Face Detection based Security System using LRR algorithm. Limitations of existing system are overcome in our proposed system. In order to make any transaction, system will provide to option to process. First Self user, where in system will ask for “Detect Face” and allow to process trans action if it matches with Image store in banks database otherwise system will decline the transaction after couple of warnings. Second Guest User, where in system

will ask for “OTP” and allow to process transaction if Guest User enter the correct OTP which has been sent to authorised User.

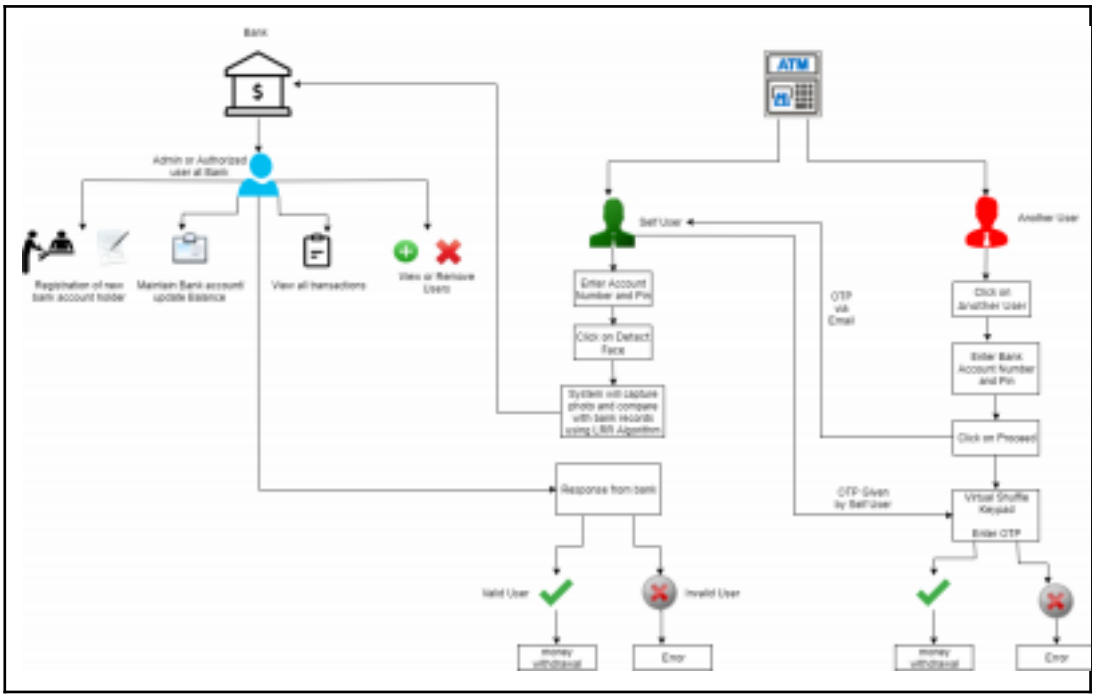


Figure 6.1: System Architecture

6.2 STEPS

Bank Admin:

1. Bank Admin can create New Bank Account of User with Profile Photo. 2. View all Users.
3. Deposit the Amount
4. View all Transactions.
5. Delete the User.

6. Update the Details of Existing User – Pin, Email, and

Photo. :

User will login into the System.

A. Self User:

1. Self User will enter Account Number and Password.
2. After Verification System will capture photo of User.
3. Execution of LRR Algorithm.
4. If User is Valid – Withdraw Process.
5. User can change pin as well if he wants.

B. Guest User:

1. Guest User will enter Account Number.
2. System will send OTP to Bank Account Holder.
3. Guest User will enter the OTP.
4. If OTP is correct then Shuffle keypad will get Enable.
5. Guest User will enter Pin.
6. Withdraw Process.

6.3 DATA DESIGN (USING APPENDICES A AND B)

A description of all data structures including internal, global, and temporary data structures, database design (tables), file formats.

6.3.1 Internal software data structure

Protects the data confidentiality and integrity .

6.3.2 Global data structure

No global data structure used

6.3.3 Database description

Database(s) / Files created/used as part of the application is(are) described. SKNSIT,

6.4 DATA FLOW DIAGRAMS

6.4.1 Level 1 Data Flow Diagram

DFD 1 Diagram shows our project at High Level, As Mentioned in the below Dia gram We can get Idea of Our Flow of the Project.

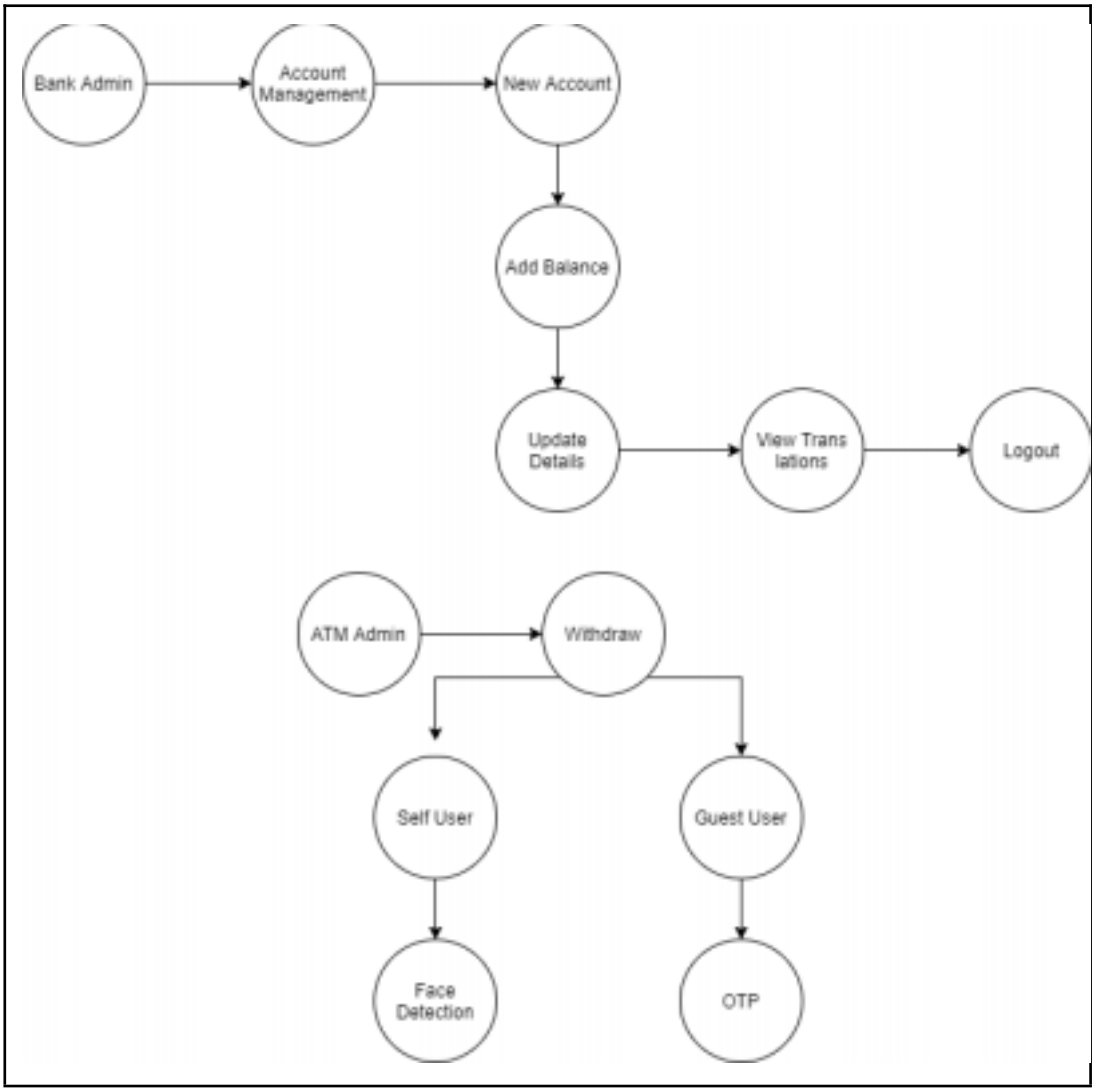


Figure 6.2: DFD 1 level

6.4.2 Level 2 Data Flow Diagram

DFD 2 Diagram shows our project at Detail Level, As Mentioned in the below Dia gram We can get Flow of Our the Project in Detail.

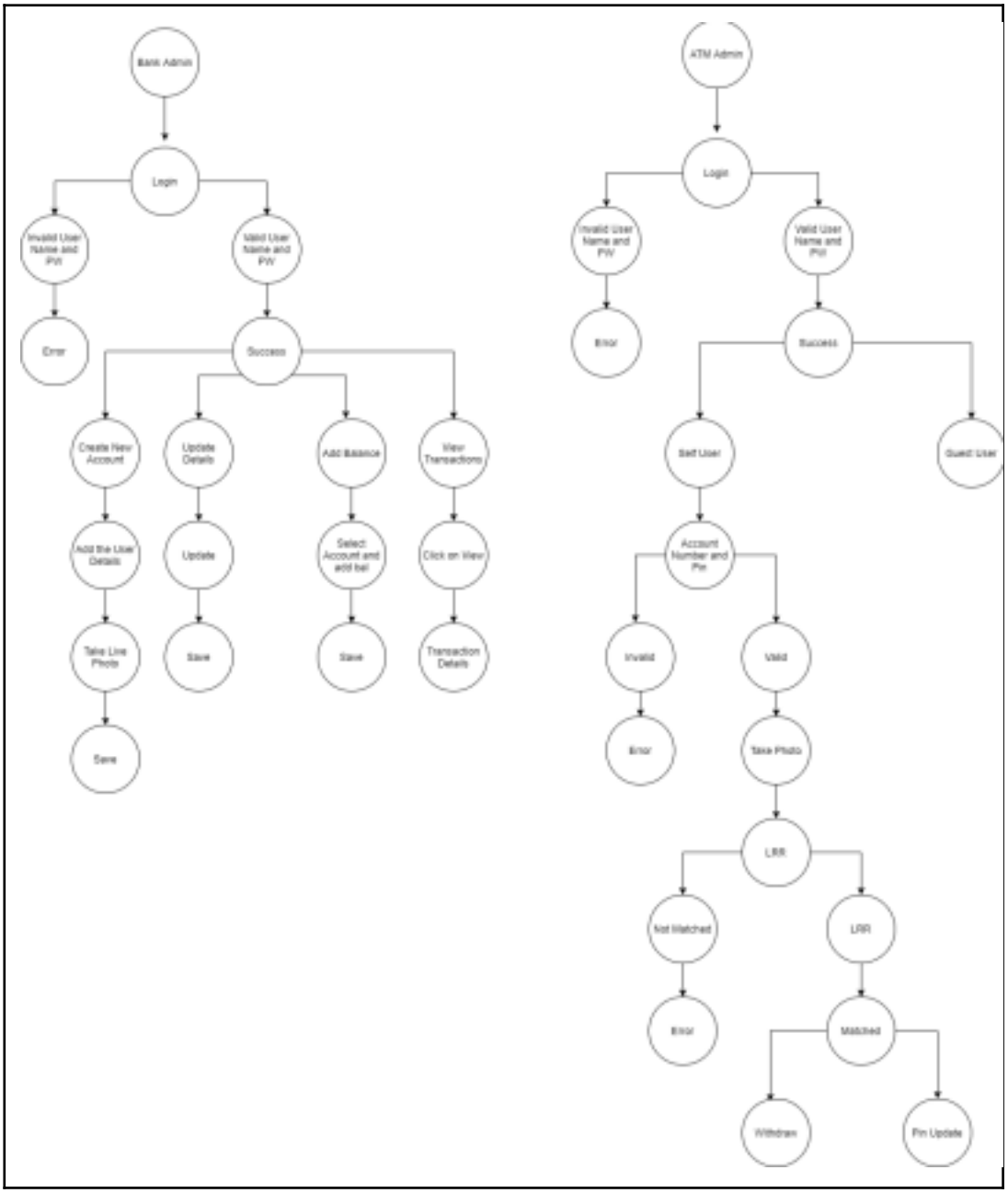


Figure 6.3: DFD 1 level

6.5 COMPOENT DESIGN

6.5.1 Class Diagram

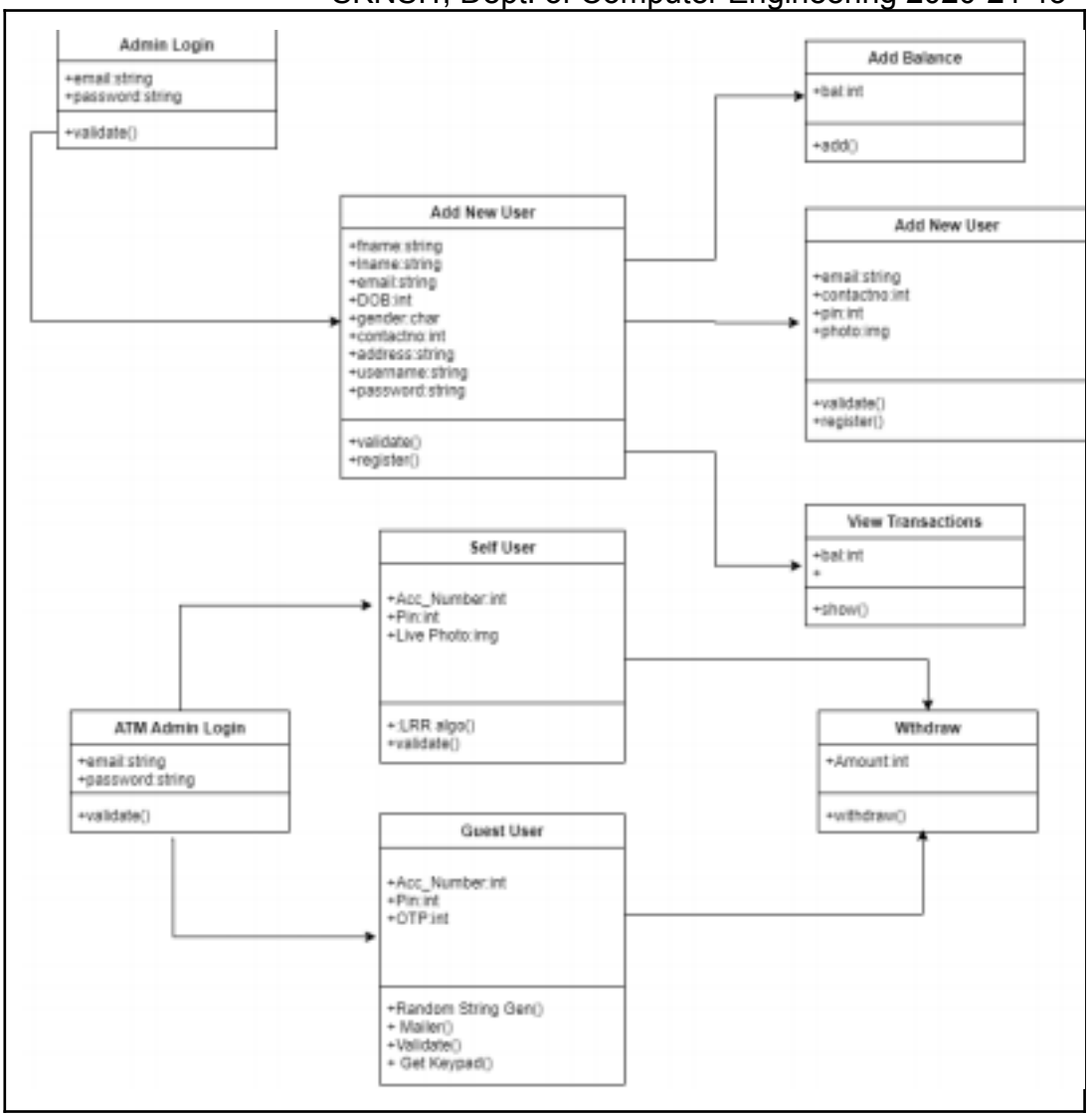


Figure 6.4: Class Diagram

CHAPTER 7

PROJECT IMPLEMENTATION

7.1 INTRODUCTION

In order to provide reliable security solution to the people, the concept of se curity system based on face detection is emerged. The Area of work is basically fo cused on Design and Implementation of Face Detection based Security System using LRR algorithm. Limitations of existing system are overcome in our proposed system. In order to make any transaction, system will provide to option to process. First Self user, where in system will ask for “Detect Face” and allow to process trans action if it matches with Image store in banks database otherwise system will decline the transaction after couple of warnings. Second Guest User, where in system will ask for “OTP” and allow to process transaction if Guest User enter the correct OTP which has been sent to authorised User.

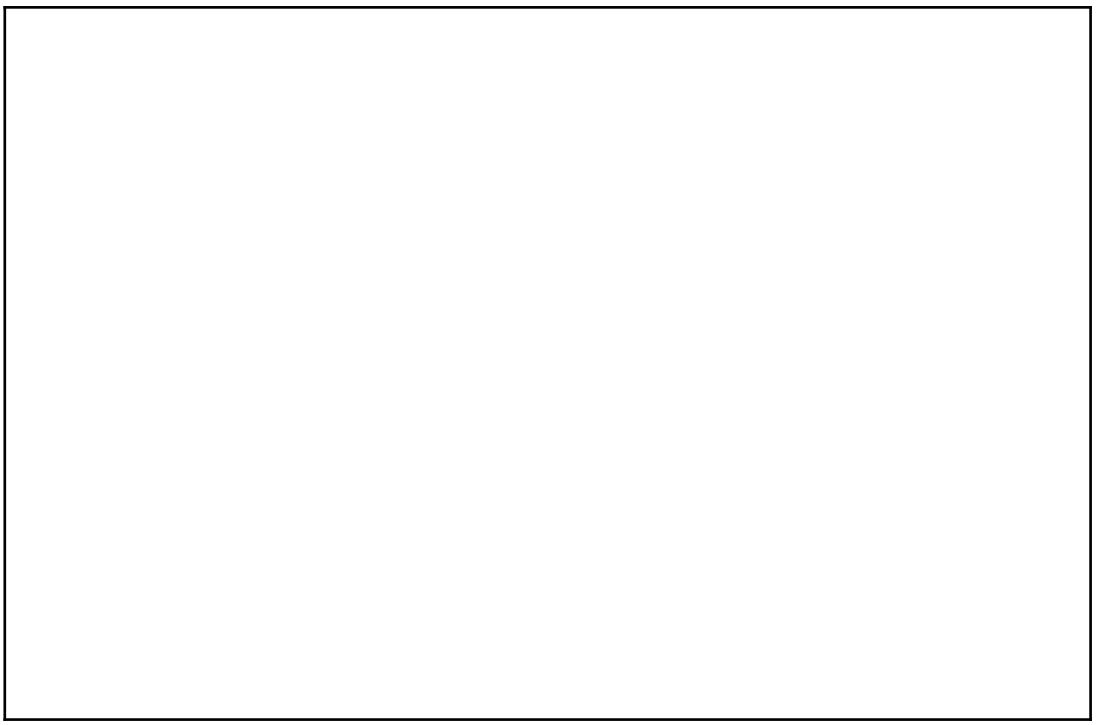
7.2 TOOLS AND TECHNOLOGIES USED

Sr. No.	Parameter	Requirement
1	OPERATING SYSTEM	Windows 7/8/10.
2	CODING LANGUAGE	JAVA/J2EE
3	IDE	Eclipse Kepler,Android SDK
4	DATABASE	SQLYog community/XAMPP Server.

5	Web Server	Apache Tomcat.

Table 7.1: Tools and Technologies Used

7.3 METHODOLOGIES



EJB modules, which are also server-side J2EE component types. These are used in conjunction with client-side components such as applets (part of the Java 2 Platform, Standard Edition specification) and application client programs. An application may consist of any number of any of these components.

CHAPTER 8

SOFTWARE TESTING

8.1 TYPE OF TESTING USED

8.1.1 Testing Strategy

Software testing methods are traditionally divided into white- and black-box testing. These two approaches are used to describe the point of view that a test engineer takes when designing test cases.

8.1.1.1 White-box testing

In white-box testing an internal perspective of the system, as well as programming skills, are used to design test cases.

8.1.1.2 Black-box testing

Black-box testing treats the software as a ”black box”, examining functionality with out any knowledge of internal implementation. The testers are only aware of what the software is supposed to do, not how it does it.

8.1.1.3 Grey-box testing

Grey-box testing involves having knowledge of internal data structures and algo rithms for purposes of designing tests, while executing those tests at the user, or black-box level. The tester is not required to have full access to the software’s source code.

8.1.2 Testing Levels

8.1.2.1 Unit testing

It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level

and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

8.1.2.2 Integration testing

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that al though the components were individually

satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

8.1.2.3 Functional test

Functional tests provide systematic demonstrations that functions tested are avail able as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be rejected. Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised. Systems/Procedures: interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to iden tify Business process flows; data fields, predefined processes, and successive pro cesses must be considered for testing. Before functional testing is complete, addi tional tests are identified and the effective value of current tests is determined.

8.1.2.4 System Test

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration oriented system

integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points

CHAPTER 9

RESULTS

9.0.1 Outcomes

Basically, there are five basic steps for the data mining process which defines the problem. 1) preparing data 2) exploring the data 3) development of the model 4) exploration and validation of the models 5) deployment and updation in the models. In this project, Neural network is used as the data mining technique and it utilized above mentioned steps for accurate and reliable result. Moreover, Neural network was used as it has the capability of adaption and generalization. Moreover, H2O [3] is also a good option for the experiment purpose. H2O flow is a notebook style open source interface for H2O. It is an interactive web-based environment that allows persons to combine text, plot, mathematics, executable code in a single document, very similar to iPython notebooks.

9.0.2 Screen Shots

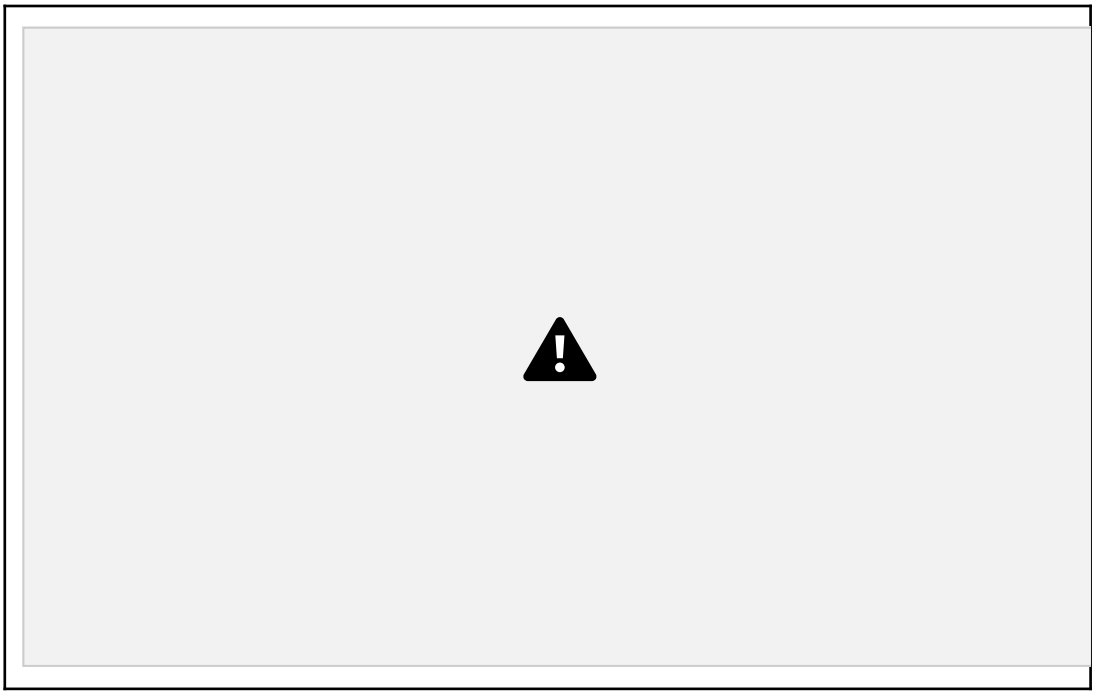


Figure 9.1: Home Page

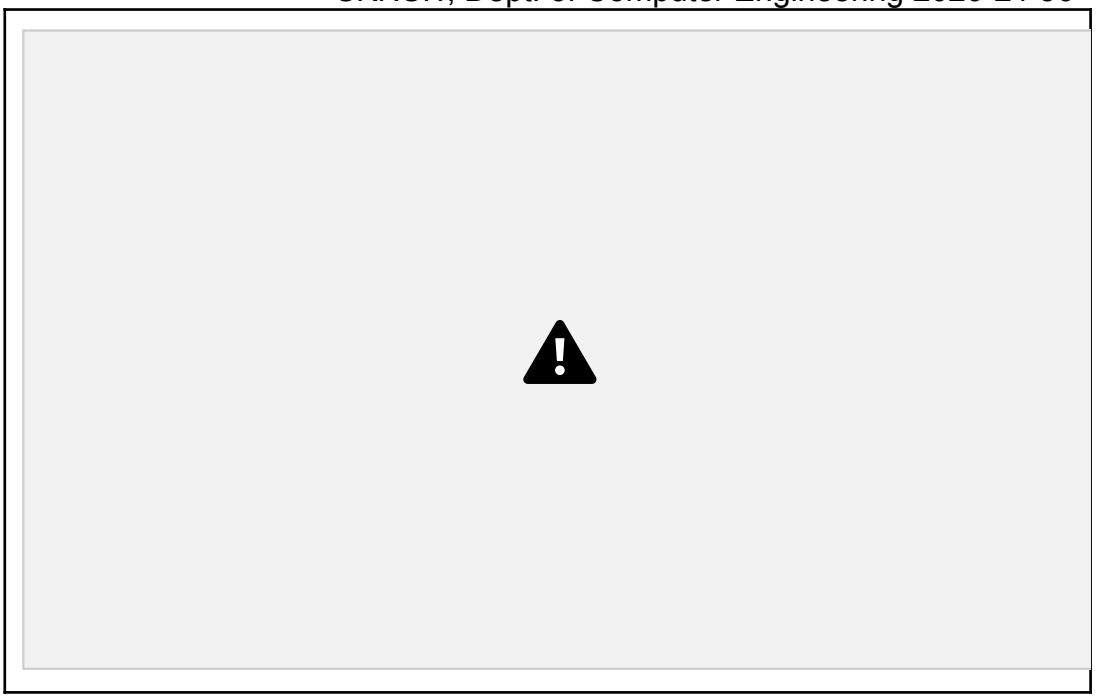


Figure 9.2: Services

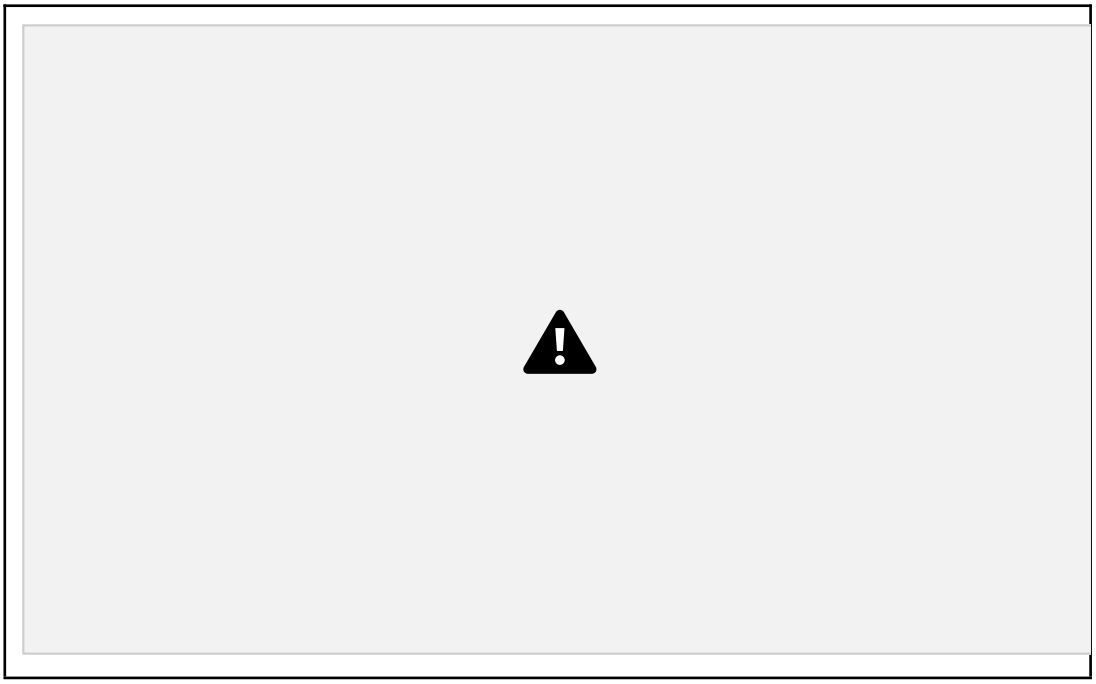


Figure 9.3: Admin Login

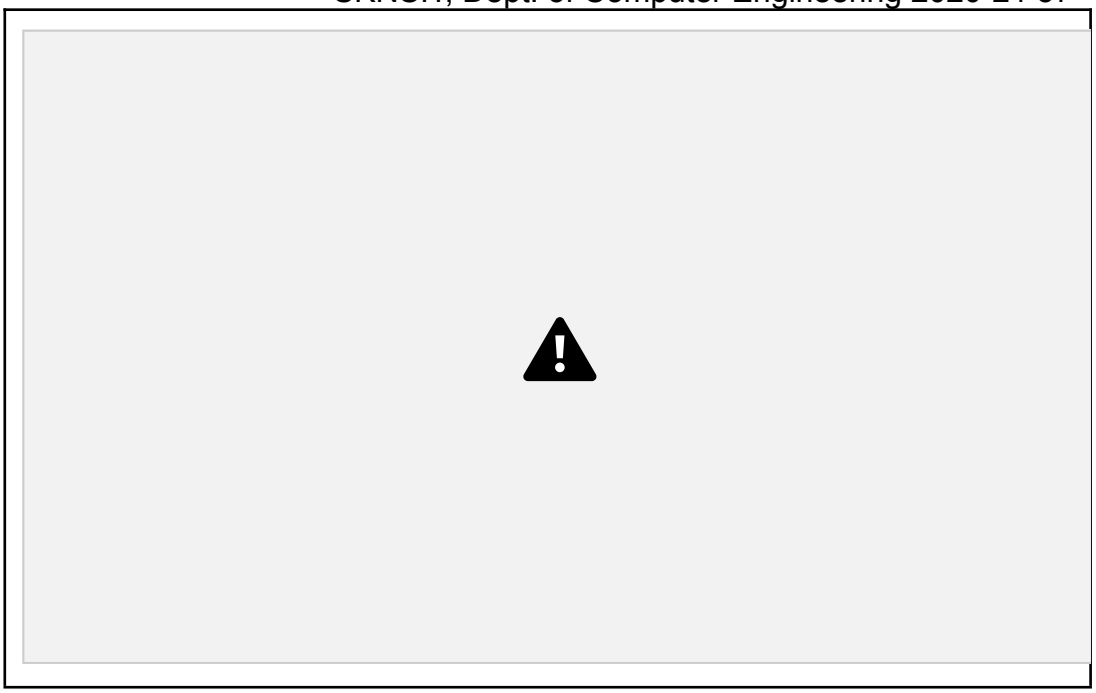


Figure 9.4: Admin Home Page

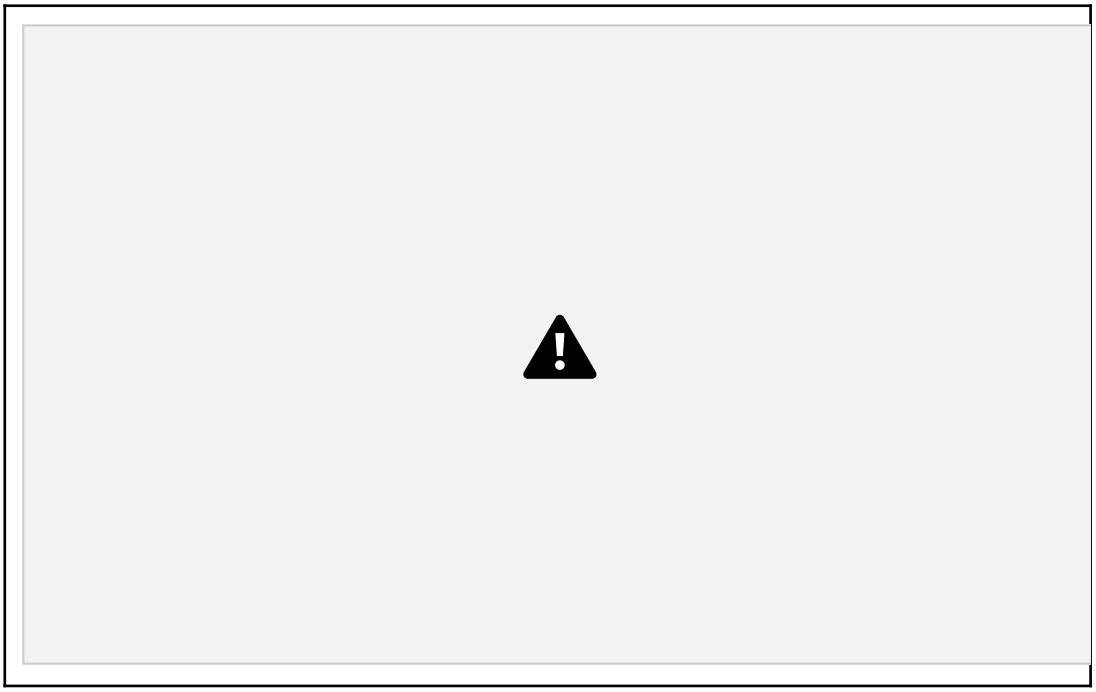


Figure 9.5: Add New Account Holder

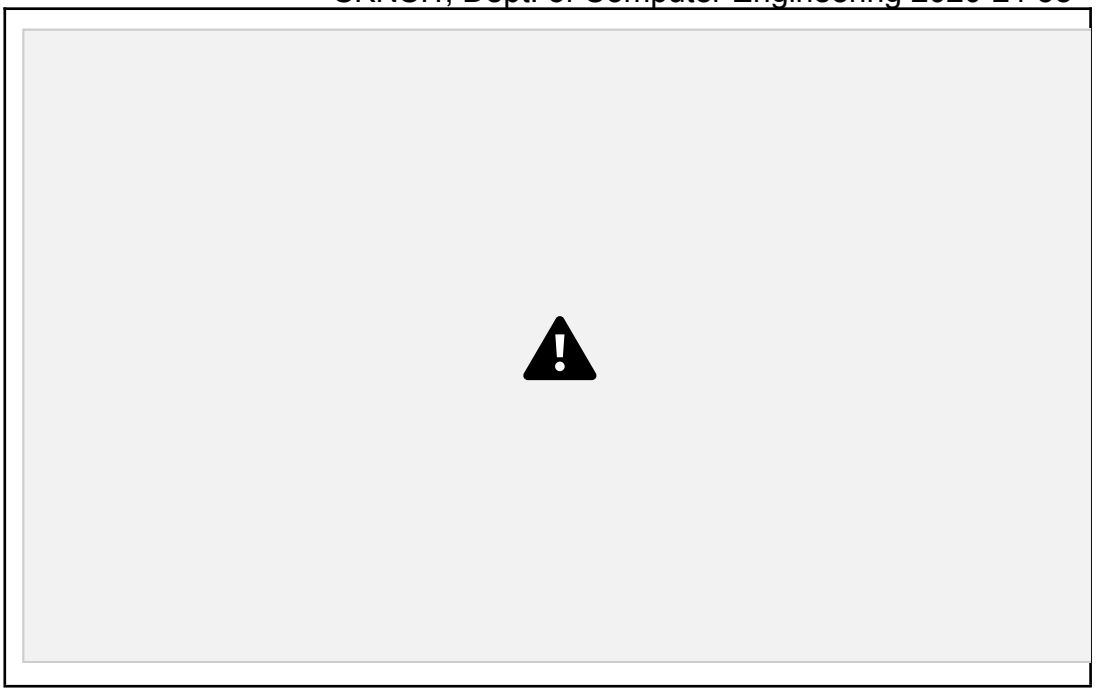


Figure 9.6: View Account Holder List

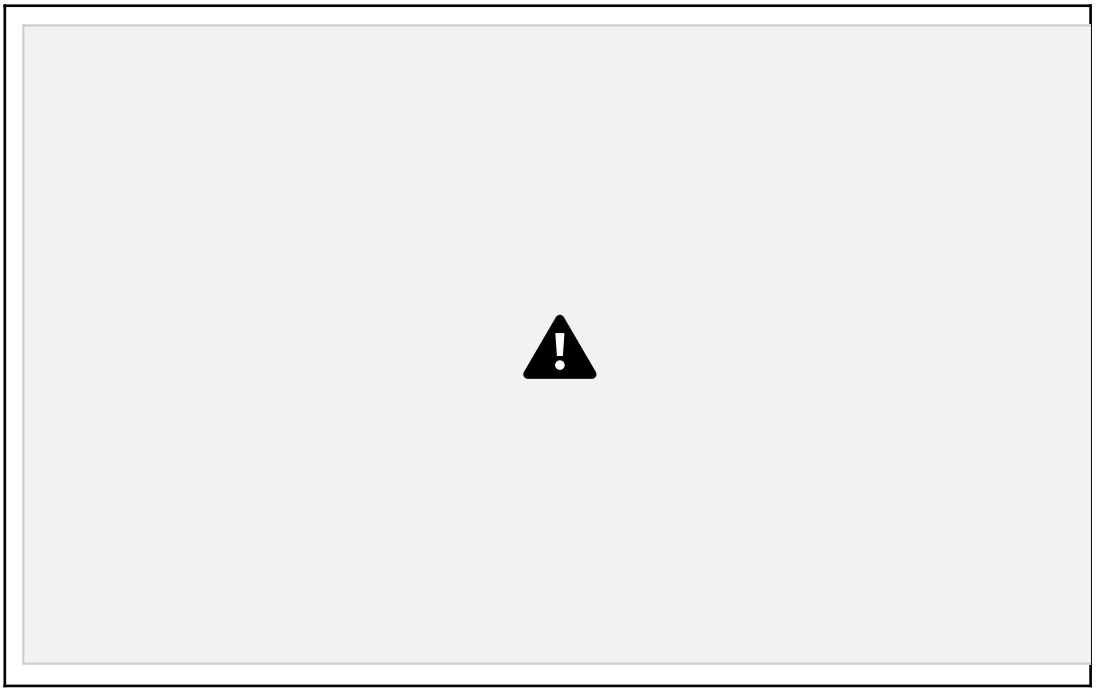


Figure 9.7: Transfer Money

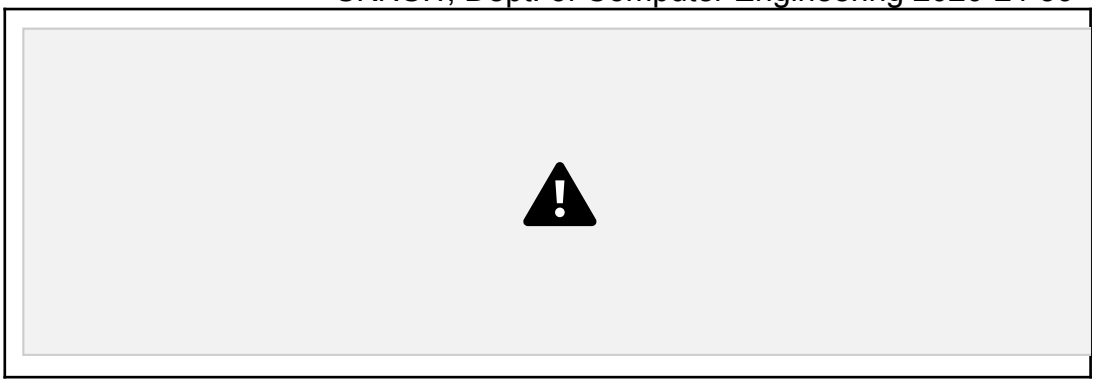


Figure 9.8: Transaction History

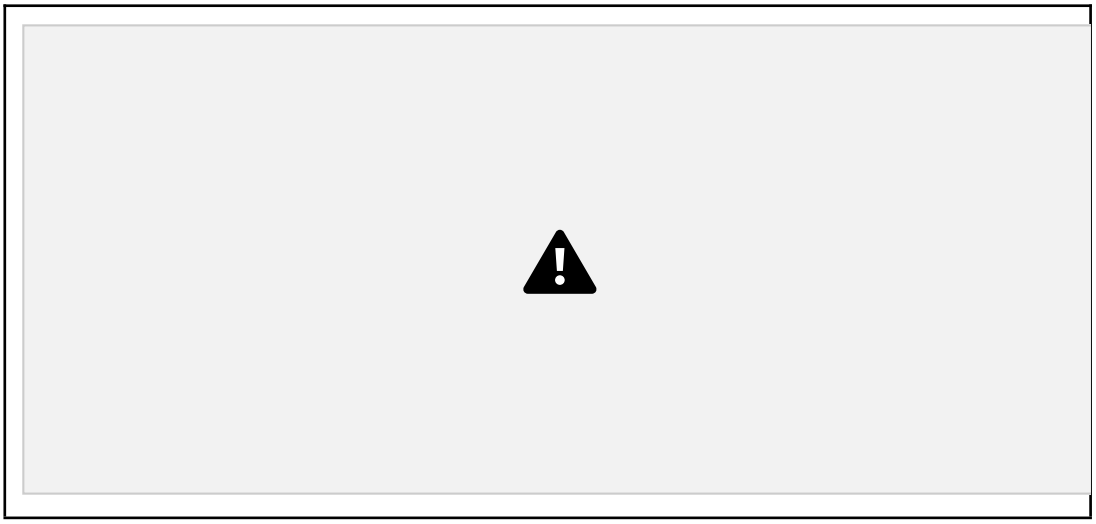


Figure 9.9: Login Options

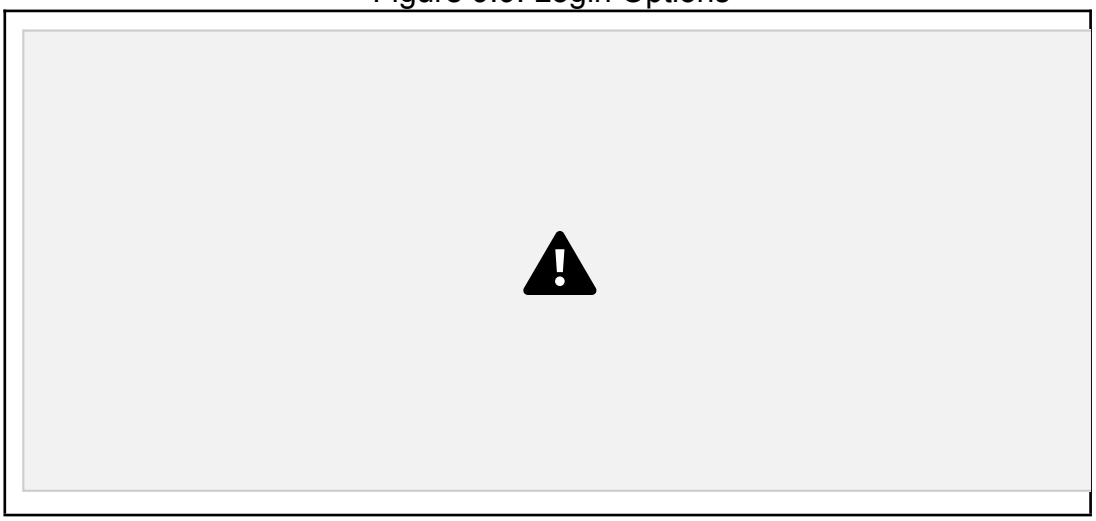


Figure 9.10: Self User

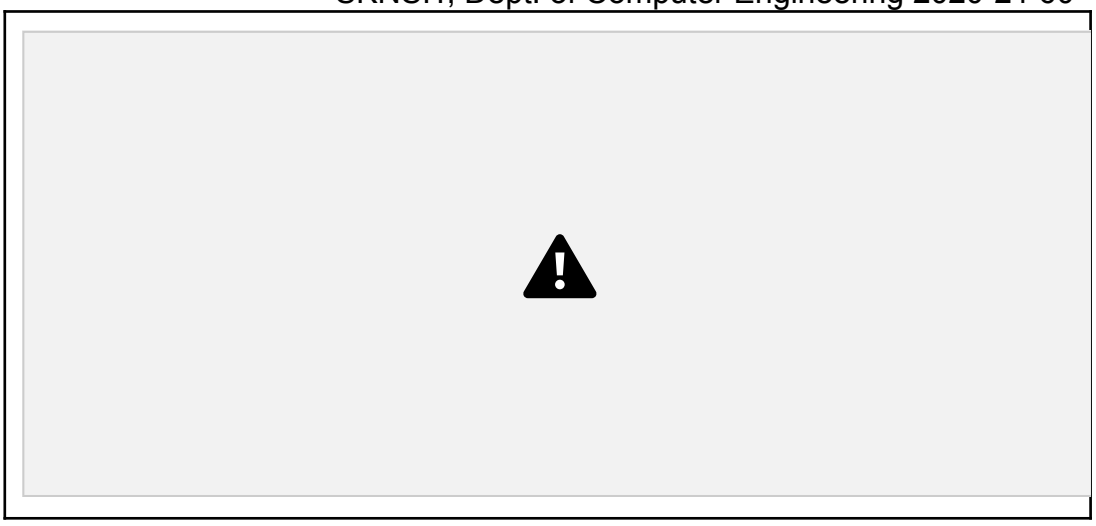


Figure 9.11: Enter Credential

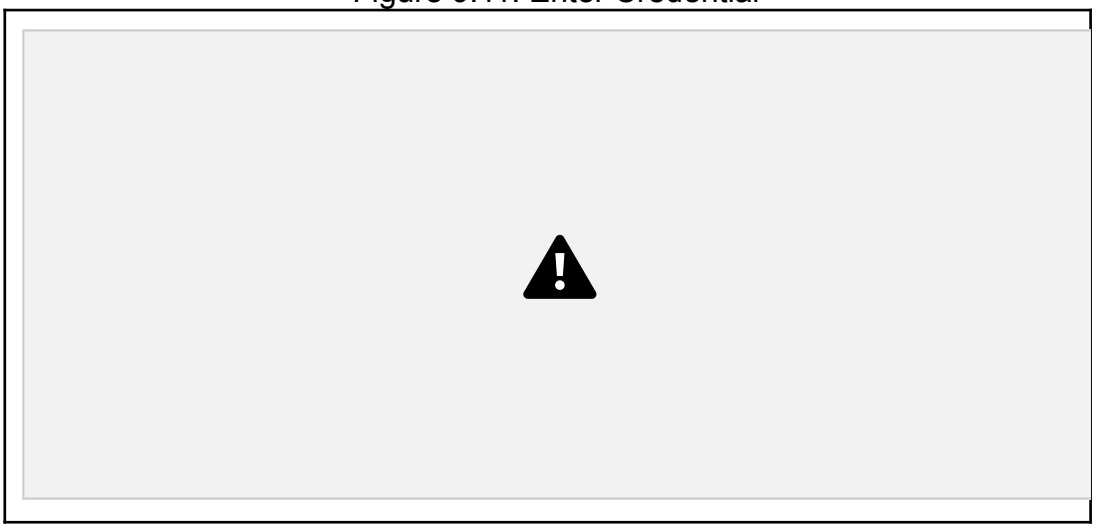


Figure 9.12: Enter Amount To Withdraw

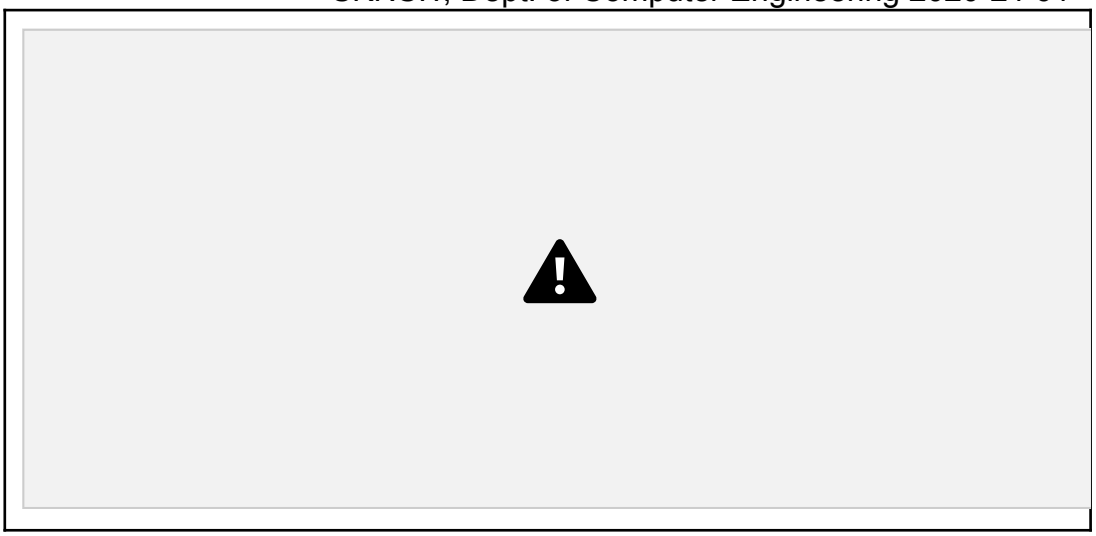


Figure 9.13: Enter Details

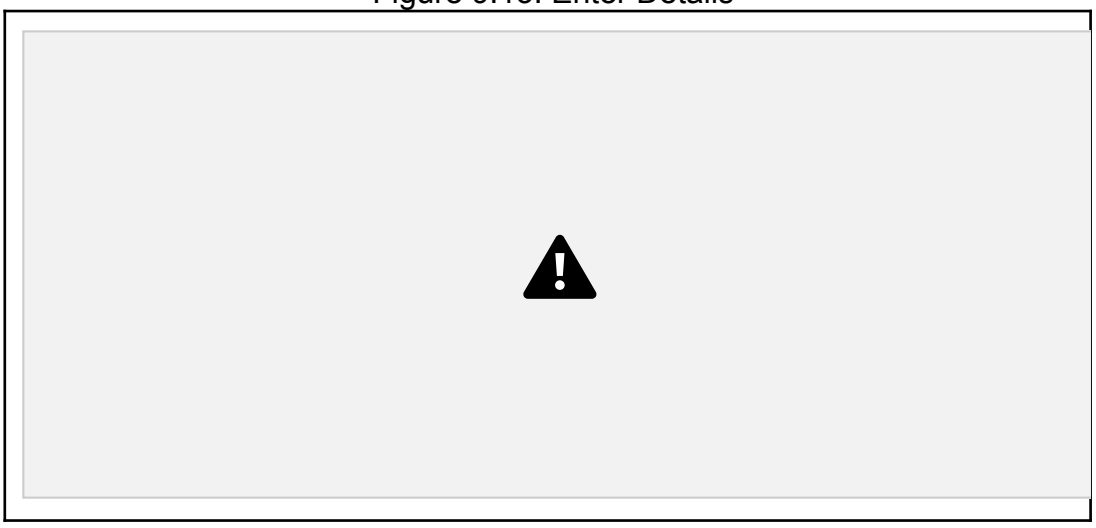


Figure 9.14: Guest User

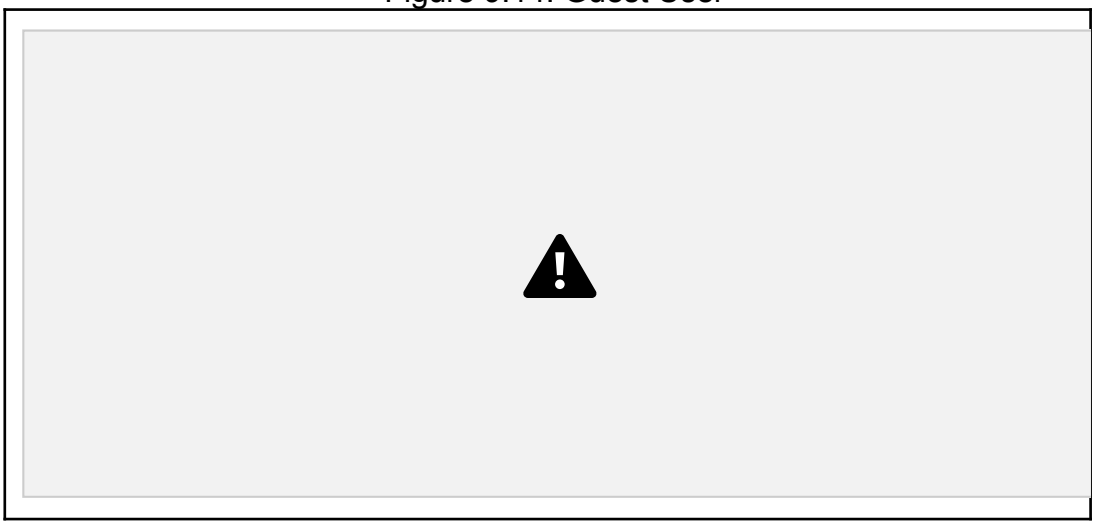


Figure 9.15: Enter Account Details

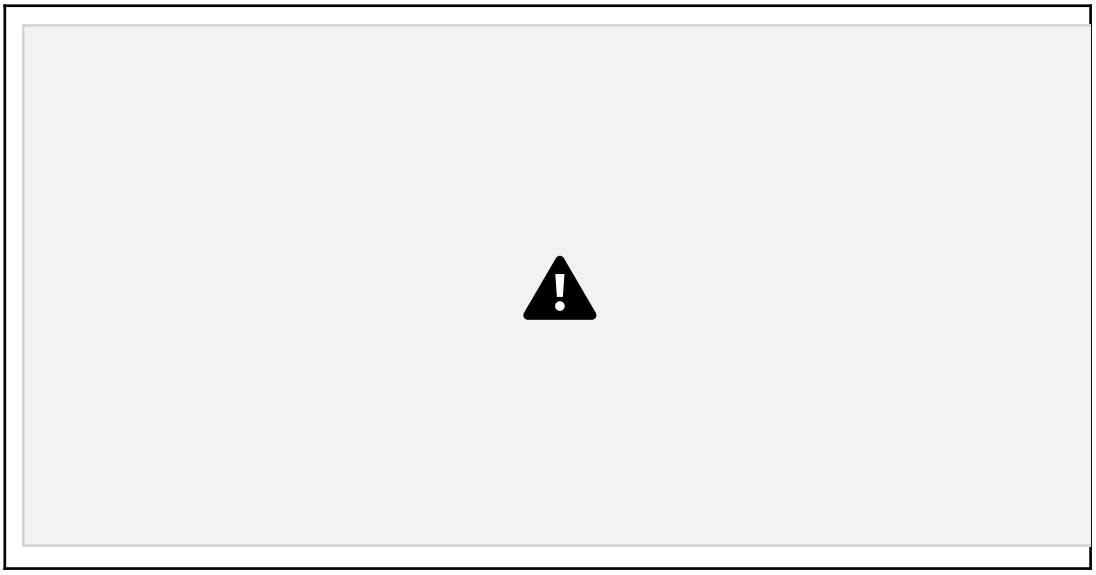


Figure 9.16: Enter Password

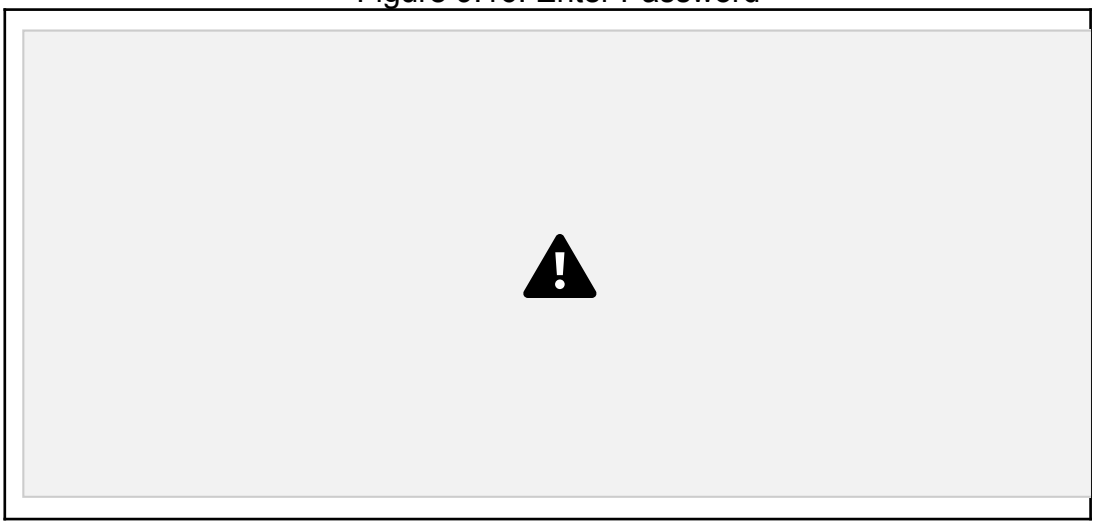


Figure 9.17: Sent OTP On Mail

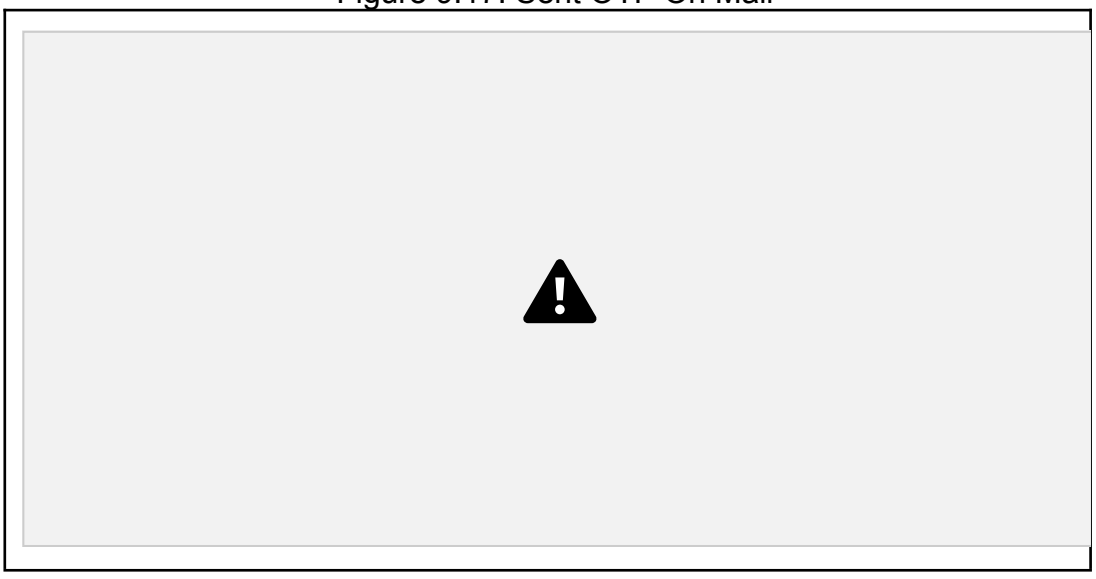


Figure 9.18: Enter OTP

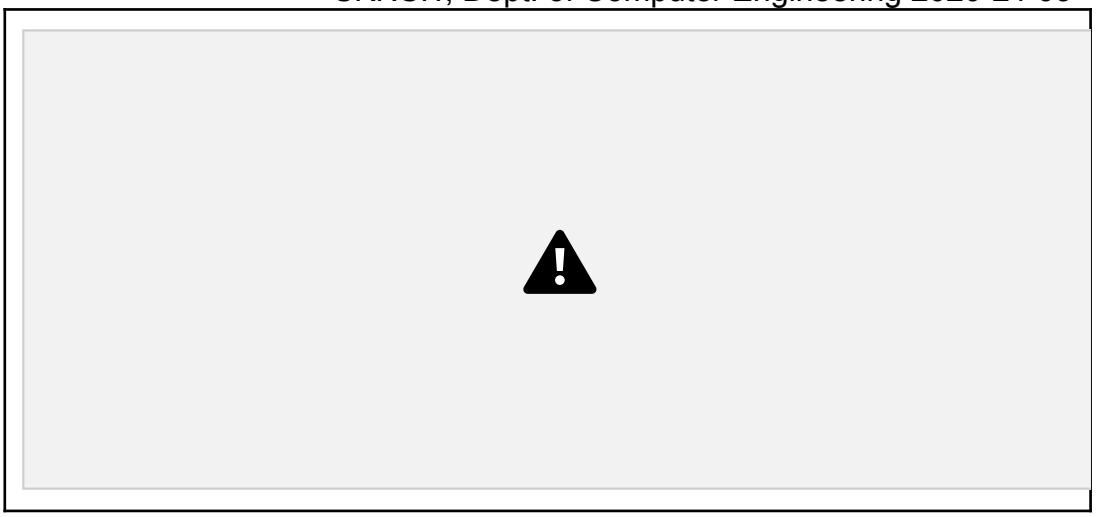


Figure 9.19: Enter Amount To Withdraw

CHAPTER 10

SUMMARY AND CONCLUSION

Credit card fraud is without a doubt a move of criminal dishonesty. this text has listed out the foremost common methods of fraud along side their detection methods and reviewed recent findings in this field. This paper has also explained intimately , how machine learning are often applied to get better leads to fraud detection along side the algorithm, pseudocode, explanation its implementation and experimentation results. While the algorithm does reach over 99.6 percentage accuracy, its precision remains only at 28% when a tenth of the info set is taken into consideration.

However, when the entire dataset is fed into the algorithm, the precision rises to 33 percentage. This high percentage of accuracy is to be expected thanks to the large imbalance between the amount of valid and number of genuine transactions.Since the whole dataset consists of only two days' transaction records, its only a fraction of knowledge which will be made available if this project were to be used on the billboard scale. Being based on machine learning algorithms, program will only increase its efficiency over time the more data is put into it. VII.FUTURE ENHANCEMENTS While we couldn't reach out goal of 100 percentage accuracy in fraud detection, we did find yourself creating a system which will , with enough time and data, get very on the brink of that goal. As with any such project, there's some room for improvement here. The very nature of the project allows for multiple algorithms to be integrated together as a modules and their results are often combined to extend the accuracy of the ultimate result. This model can be further improved with the addition of more algorithms into it. However, the output of these algorithms needs to be within the same format because the others. Once that condition is satisfied, the modules are easy to feature as wiped out the code. This provides a great degree of modularity and versatility to the project. More room for improvement are often found within the dataset. As demonstrated before, the precision of the algorithms increases when the size of dataset is increased. Hence, more data will surely make the model more accurate in detecting the frauds and reduce the number of false positives. However, this requires official support from the banks themselves

ANNEXURE A

REFERENCES

:

[1]“Credit Card Fraud Detection supported Transaction Behaviour - published by Proc. of the 2017 IEEE

[2]CLIFTON PHUA1, VINCENT LEE1, KATE SMITH1 and ROSS GAYLER2 “ A Survey of knowledge Miningbased Fraud Detection Research” - by the School of Business, Faculty of Information Technology, Monash University, Wellington Road, Clayton, Victoria 3800, Australia

[3]“Survey Paper on the Credit Card Fraud Detection by Suman” , Scholar, GJUS&T Hisar HCE, Sonepat published by International Journal of the Advanced Research in Computer Engineering & Technology (IJARCET) Volume 3 Issue 3, March 2014

[4]“Research on Credit card Fraud Detection Model supported Distance Sum – by Wen-Fang YU and Na Wang” published on 2009 International Joint Conference on Artificial Intelligence

[5]“Credit Card Fraud Detection through the Parenclitic Network Analysis - Massimiliano Zanin, Miguel Romance, Regino Criado, and SantiagoMoral” published by HindawiComplexity Volume 2018, Article ID 5764370, 9 pages

[6]“Credit Card Fraud Detection:- a sensible Modeling and a completely unique Learning Strategy” published by IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, VOL. 29, NO. 8, AUGUST 2018

[7]“Credit Card Fraud Detection- Ishu Trivedi, Monika, Mrigya, Mridushi” published by the International Journal of Research in Computer and Communication Engineering Vol. 5, Issue 1, January 2016

[8]David J.Wetson,David J.Hand,M Adams,Whitrow and Piotr Jusczak “Plastic Card Fraud Detection using coevals Analysis” Springer, Issue 2008