LAB REPORT

Submitted by

P. KETSI DEBORAH
[RA2011003010372]

Under the Guidance of

Dr. J. Rene Beulah

Assistant Professor, Department of Computing Technologies

In partial satisfaction of the requirements for the degree of

BACHELOR OF TECHNOLOGY
in
COMPUTER SCIENCE ENGINEERING

SCHOOL OF COMPUTING

COLLEGE OF ENGINEERING AND TECHNOLOGY

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

KATTANKULATHUR - 603203

JUNE 2022

SRM INSTITUTION OF SCIENCE AND TECHNOLOGY
KATTANKULATHUR-603203

BONAFIDE CERTIFICATE

Certified that this lab report titled **FRAUD DETECTION IN CREDIT CARD SYSTEM** is the bonafide work done by **P. KETSI DEBORAH (RA2011003010372)** who carried out the lab exercises under my supervision. Certified further, that to the best of my knowledge the work reported herein does not form part of any other work.

SIGNATURE

Dr. J. Rene Beulah

SEPM – Course Faculty

Assistant Professor, Department of Computing Technologies

ABSTRACT

Credit card frauds are easy and friendly targets. E-commerce and many other online sites have increased the online payment modes, increasing the risk for online frauds. Increase in fraud rates, researchers started using different machine learning methods to detect and analyze frauds in online transactions. The main aim of the paper is to design and develop a novel fraud detection method for Streaming Transaction Data, with an objective, to analyze the past transaction details of the customers and extract the behavioural patterns. Where cardholders are clustered into different groups based on their transaction amount. Then using sliding window strategy, to aggregate the transaction made by the cardholders from different groups so that the behavioural pattern of the groups can be extracted respectively. Later different classifiers are trained over the groups separately. And then the classifier with better rating score can be chosen to be one of the best methods to predict frauds. Thus, followed by a feedback mechanism to solve the problem of concept drift . In this paper, we worked with European credit card fraud dataset.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

UML- Unified Modeling Language

HR-Human Resource

HOD-Head of Department

WBS-Work Breakdown Structure

HTML-Hypertext Markup Language

CSS-Cascading Style Sheet

XML-Extensible Markup Language

MTTR-Mean Time to Repair

CAPTCHA-Completely Automated Public Turing test to tell Computers and Humans Apart

ID-Identity Document

UI-User Interface

IDE-Integrated Development Environment

UX-User Experience

API-Application Programming Interface

DB-Database

QR code-Quick Response code

DDOS-Distributed Denial of Service

SWOT-Strength Weakness Opportunities Threats

ER-Entity Relationship

DFD-Data Flow Diagram

QA-Quality Assurance

Department of Computing Technologies

SRM IST, Kattankulathur – 603 203

Course Code: 18CSC206J

Course Name: Software Engineering and Project Management

Experiment No	1
Title of Experiment	To identify the Software Project, Create Business Case, Arrive at a Problem Statement
Name of the candidate	P. KETSI DEBORAH
Team Members	I) KASHISH KEDIA II) SAKTHI MAHALAKSHMI S
Register Number	RA2011003010372
Date of Experiment	15-03-2022

Mark Split Up

S.No	Description	Maximum Mark	Mark Obtained
1	Exercise	5	
2	Viva	5	
	Total	**10**	

Staff Signature with date

Aim

To Frame a project team, analyze and identify a Software project. To create a business case and Arrive at a Problem Statement for the project: Fraud detection in credit card system.

Team Members:

S. No	Register No	Name	Role
1	RA2011003010363	SAKTHI MAHALAKSHMI S	Lead/Rep
2	RA2011003010372	P KETSI DEBORAH	Member
3	RA2011003010355	KASHISH KEDIA	Member

Project Title: FRAUD DETECTION IN CREDIT CARD SYSTEM

Project Description:

- ✓ credit card for personal reasons while the owner and the card issuing authorities are unaware of the fact that the card is being used.
- ✓ Due to rise and acceleration of E- Commerce, there has been a tremendous use of credit cards for online shopping which led to High amount of frauds related to credit cards.
- ✓ In the era of digitalization, the need to identify credit card frauds is necessary. Fraud detection involves monitoring and analysing the behaviour of various users in order to estimate detect or avoid undesirable behaviour.

ONE PAGE BUSINESS CASE TEMPLATE

DATE	15-03-2022
SUBMITTED BY	SAKTHI MAHALAKSHMI S P KETSI DEBORAH KASHISH KEDIA
TITLE / ROLE	FRAUD DETECTION IN CREDIT CARD SYSTEM

THE PROJECT

In bullet points, describe the problem this project aims to solve or the opportunity it aims to develop.

- ✓ The use of credit cards is prevalent in modern day society. But it is obvious that the number of credit card fraud cases is constantly increasing in spite of the chip cards worldwide integration and existing protection systems. This is why the problem of fraud detection is very important now to be in the saferhand.

THE HISTORY

In bullet points, describe the current situation.

- ➢ In 2018, unauthorized financial fraud losses across payment cards and remote banking totalled £844.8million in the United Kingdom. Whereas banks and card companies prevented £1.66 billion in unauthorized fraud in 2018. That is the equivalent to £2 in every £3 of attempted fraud being stopped.
- ➢ In 2012, about 40 million sets of payment card information were compromised by a hack of Adobe Systems.[33] The information compromised included customer names, encrypted payment card numbers, expiration dates, and information relating to orders, Chief Security Officer Brad Arkin said.

LIMITATIONS

List what could prevent the success of the project, such as the need for expensive equipment, bad weather, lack of special training, etc.

- ✓ Dataset available for studying the pattern of the customer is limited.
- ✓ False positives - customers falsely rejected for fraudulent activity.
- ✓ Due to lack of sophistication and logistics, some intricate points couldn't be covered.
- ✓ A layman may find it hard to use the software.

APPROACH

List what is needed to complete the project.

- ➢ The credit card fraud detection features uses user behavior and location scanning to check for unusualpatterns. These patterns include user characteristics such as user spending patterns as well as usual user geographic locations to verify his identity. If any unusual pattern is detected, the system requires revivification.
- ➢ The system analyses user credit card data for various characteristics. These characteristics include usercountry, usual spending procedures. Based upon previous data of that user the system recognizes unusual patterns in the payment procedure. So now the system may require the user to login again or even block the user for more than 3 invalid attempts.
- ➢ CC Avenue verifies and validates credit cards of buyers for over a thousand e-commerce Web sites. Itconducts checks like IP mapping, zip code mapping and reverse lookup of telephone numbers.

BENEFITS

In bullet points, list the benefits that this project will bring to the organization.

- ✓ Due to Behavior and location analysis approach, there is a drastic reduction in the number of FalsePositives transactions identified as malicious by an FDS although they are actually genuine.
- ✓ The system stores previous transaction patterns for each user.
- ✓ Based upon previous data of that user the system recognizes unusual patterns in the paymentprocedure.
- ✓ The System will block the user for more than 3 invalid attempts.
- ✓ User can safely use his credit card for online transaction.
- ✓ Added layer of security

Result:

Thus, the project team formed, the project is described, the business case was preparedand the problem statement was arrived.

Department of Computing Technologies

SRM IST, Kattankulathur – 603 203

Course Code: 18CSC206J

Course Name: Software Engineering and Project Management

Experiment No	2
Title of Experiment	Identification of Process Methodology and Stakeholder Description
Name of the candidate	P. KETSI DEBORAH
Team Members	1) KASHISH KEDIA 2) SAKTHI MAHALAKSHMI
Register Number	RA2011003010372
Date of Experiment	22-03-2022

Mark Split Up

S.No	Description	Maximum Mark	Mark Obtained
1	Exercise	5	
2	Viva	5	
	Total	**10**	

Staff Signature with date

Aim

To identify the appropriate Process Model for the project and prepare Stakeholder and User Description.

Team Members:

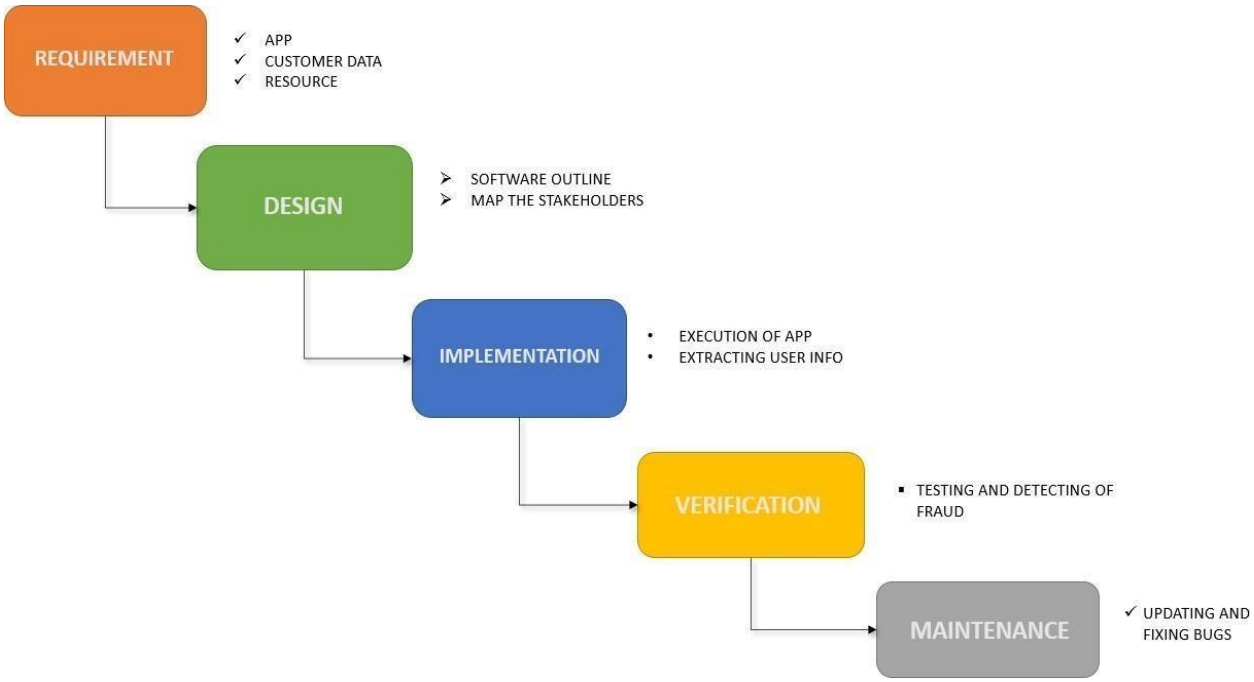
Sl No	Register No	Name	Role
1	RA2011003010363	SAKTHI MAHALAKSHMI S	Rep/Member
2	RA2011003010372	P KETSI DEBORAH	Member
3	RA2011003010355	KASHISH KEDIA	Member

Project Title: FRAUD DETECTION IN CREDIT CARD SYSTEM

Stakeholder Name	Activity/ Area/phase	Interest	Influence	Priority (High/ Medium/ Low)
Credit Card holders	User of the App: Provides correct information	High	High	1
Owners	Bank-Owns And Analayze The Application	High	High	3
Team Members	Development , Creating and Implementing Retain and upgrade sills	High	Low	2
Project Manager	Leads the team in every aspect and accountable for entire project scope ,team , success,failure	High	Medium	2
Investors	**KSK HOLMES** (COMPANY): promotes and provide necessary financial resources	High	Low	4
Resource Manager	Member of KSK HOLMES- will allocate and plan resources	High	Low	6
Suppliers	Ensuring feasible and realistic in every aspect and manage divergence from budgeted cost	High	Low	5
End Users	Client-benefited from the company	Low	High	7

WATERFALL METHODOLOGY:

The waterfall methodology is a project management approach that emphasizes a linear progression from beginning to end of a project. This methodology, often used by engineers, is front-loaded to rely on careful planning, detailed documentation, and consecutive execution



RESULT:

Thus the Project Methodology was identified and the stakeholders were described.

SRM INSTITUTE OF SCIENCE & TECHNOLOGY
(Deemed to be University u/s 3 of UGC Act, 1956)

Department Of Computing Technologies

SRM IST, Kattankulathur – 603 203

Course Code: 18CSC206J

Course Name: Software Engineering and Project Management

Experiment No	3
Title of Experiment	System, Functional and Non-Functional Requirements of the Project
Name of the candidate	P. KETSI DEBORAH
Team Members	1) KASHISH KEDIA 2) SAKTHI MAHALAKSHMI S
Register Number	RA2011003010372
Date of Experiment	28-03-2022

Mark Split Up

S.No	Description	Maximum Mark	Mark Obtained
1	Exercise	5	
2	Viva	5	
	Total	**10**	

Staff Signature with date

Aim

To identify the system, functional and non-functional requirements for the project.

Team Members:

S No	Register No	Name	Role
1	RA2011003010363	SAKTHI MAHALAKSHMI S	**Rep/Member**
2	RA2011003010372	P KETSI DEBORAH	**Member**
3	RA2011003010355	KASHISH KEDIA	**Member**

Project Title: < : FRAUD DETECTION IN CREDIT CARD SYSTEM >

System Requirements

- ✓ Security and protection
- ✓ Modern updates
- ✓ System Reliability
- ✓ Detection Accuracy
- ✓ User friendly

Functional Requirements

- ➢ Detection of unusual activity
- ➢ Support multiple users
- ➢ Provide immediate indication to the user
- ➢ Security of user details

Non-Functional Requirements

- Performance
- Scalability
- Capacity
- Availability
- Reliability
- Recoverability
- Maintainability
- Serviceability

Result

Thus the requirements were identified and accordingly described.

Department of Computing Technologies

SRM IST, Kattankulathur – 603 203

Course Code: 18CSC206J

Course Name: Software Engineering and Project Management

Experiment No	4
Title of Experiment	Prepare Project Plan based on scope, Calculate Project effort based on resources and Job roles and responsibilities
Name of the candidate	P. KETSI DEBORAH
Team Members	1. KASHISH KEDIA 2. SAKTHI MAHALAKSHMI S
Register Number	RA2011003010372
Date of Experiment	04-04-2022

Mark Split Up

S.No	Description	Maximum Mark	Mark Obtained
1	Exercise	5	
2	Viva	5	
	Total	**10**	

Staff Signature with date

Aim

To Prepare Project Plan based on scope, Calculate Project effort based on resources, Find Job roles and responsibilities

Team Members:

Sl No	Register No	Name	Role
1	RA2011003010363	SAKTHI MAHALAKSHMI S	**Lead**
2	RA2011003010355	KASHISH KEDIA	**Member**
3	RA2011003010372	P. KETSI DEBORAH	**Member**

REQUIREMENTS:

1. **SUMMARY:**
   The use of credit cards is prevalent in modern day society. But it is obvious that the number of credit card fraud cases is constantly increasing in spite of the chip cards worldwide integration and existing protection systems. This is why the problem of fraud detection is very important now to be in the safer hand.
2. **PROJECT MANAGEMENT PLAN**:

Focus Area	Details
Integration Management	The model used for this project is **Waterfall Model**. The project will be divided into 3 parts: first member will develop the front end and backend portions of the application. The second member will be designing and detection in the app and the third member will be doing the documentation.
Scope Management	**Stakeholders: KSK Holmes ,** consumer , bank **Project Objectives:** The objective is to take immediate action against credit card fraud . **Schedule Objectives:** Project is scheduled to be completed in 5 months **Constraints:** Lack of experience, Budget
Schedule Management	**Conception & Initiation:** 1 week (approx.) **Planning:** 2 weeks (approx.) **Execution:** 4-5 weeks (approx.) **Testing:** 2-3 weeks (approx.) **Deployment:** 2-4 weeks (approx.)
Cost Management	**Total Budget:** Rs.1.6 L(Rough estimate) Most of the technical tasks will be done on free platforms and resources. **Effort:** 14 hours/week distributed among all the team members.
Quality Management	Automation level, real time operation tracking and reporting , compliance with security standards

Resource Management	**People:** Project team will undertake the task of planning, building and documentation of the project. The developer team will build the front end and backend portions of the project **Physical:** A good database with necessary facilities will be required to store data.
Stakeholder	The stakeholders are the developers KSK Holmes, customer and bank.
Communication Management	In the application, customers as well as the detectors can communicate using either the inbuilt chat or directly make a call.
Risk Management	Potential technical risks will be discussed in every meeting and managed accordingly and will provide security to prevent unauthorised access.
Procurement Management	Adhering to organization procurement process. Most of the procurement process will be online.

3. ESTIMATION:

3.1 EFFORT AND COST MANAGEMENT:

Activity description	Sub-Task	Sub-Task distribution	Effort (in hrs)	Cost in INR
Frontend	E1R1A1T1	website creation	2	1,000
	E1R1A1T2	layout of login page	4	2,000
	E1R1A1T3	Designing and development	40	20,000
	E1R1A1T4	Navigation Menu	4	2,000
Backend	E1R1A2T1	Database creation	20	10,000
	E1R1A2T2	Credit card Details	10	5,000
	E1R1A2T3	Servers	20	10,000
	E1R1A2T4	Registration and login verification	4	2,000
	E1R1A2T4	Fraud Detection	5	30,000
	E1R1A2T5	Sensors for detecting	3	50,000
Integration	E1R1A3T1	Integration of frontend and backend	10	5,000
Testing and quality assurance	E1R1A4T1	Check for glitches and optimize if required	6	3,000

Note: We used COCOMO 1 Model to calculate the cost and efforts.

Effort (hr)	Cost (INR)
1	500

3.2 INFRASTRUCTURE/RESOURCE COST

Infrastructure Requirement	Qty	Cost per qty	Cost per item
laptops	3	75000	2,25,000
Sensors (tracking and detecting)	2	1500	30,000

3.3. MAINTENANCE AND SUPPORT

Category	Details	Qty	Cost per qty per annum	Cost per item
People	Developers, Content Curators, Marketing,	3	20,00,000	8,00,000
License	Operating System Database, IDE domain	10	10000	1,00,000
Infrastructures	Server, Storage and Network Office Workspace sensors	20	20000	4,00,000

4) Project Team Formation

Name	Role	Responsibilities
SAKTHI MAHALAKSHMI	AI/ML	DESIGN THE APPLICATION
	PROGRAM MANAGER	MANAGE THE PROJECT
	SPONSORSHIP MAINTAINANCE	GETTING FUNDS, REQUIREMENTS
	CLIENT DEALING	DETAILING ABOUT THE APPLICATION
KASHISH KEDIA	FRONTEND DEVELOPER	DESIGN, DEVELOP THE WEB PAGE
	CONTENT CURATOR	SEARCH AND CREATE EFFECTIVE COURSE ROADMAPS
	AI/ML	DESIGN THE APPLICATION
KETSI DEBORAH	Frontend Developer/AI/ML	DEVELOP USER INTERFACE
	TESTER	Define Test Cases and Perform Testing
	Key Business User (Product Owner)	Provide clear business and user requirements

5) RESPONSIBILITY ASSIGNMENT M ATRIX:

RACI Matrix	Team Members		
Activity	SAKTHI (DEVELOPER)	KASHISH (PROJECT MANAGER)	KETSI (KEY BUSINESS USER)
User Requirement Documentation	C/I	I	R
Coding	A &R	A	I
Integration	R	A	R
Deployment	C	C	A
Security	R	I	I

A	Accountable
R	Responsible
C	Consult
I	Inform

REFERENCE:

1. https://www.researchgate.net/publication/337591236_Credit_Card_Fraud_Detection_Using_Machine_Learning_With_Python
2. https://www.altexsoft.com/blog/credit-card-fraud-detection/

Result:

Thus, the Project Plan was documented successfully.

SRM INSTITUTE OF SCIENCE & TECHNOLOGY
(Deemed to be University u/s 3 of UGC Act, 1956)

Department of Computing Technologies

SRM IST, Kattankulathur – 603 203

Course Code: 18CSC206J

Course Name: Software Engineering and Project Management

Experiment No	5
Title of Experiment	Prepare Work breakdown structure, Timeline chart, Risk identification table
Name of the candidate	P. KETSI DEBORAH
Team Members	1) KASHISH KEDIA 2) SAKTHI MAHALAKSHMI S
Register Number	RA2011003010372
Date of Experiment	11-04-2022

Mark Split Up

S.No	Description	Maximum Mark	Mark Obtained
1	Exercise	5	
2	Viva	5	
	Total	**10**	

Staff Signature with date

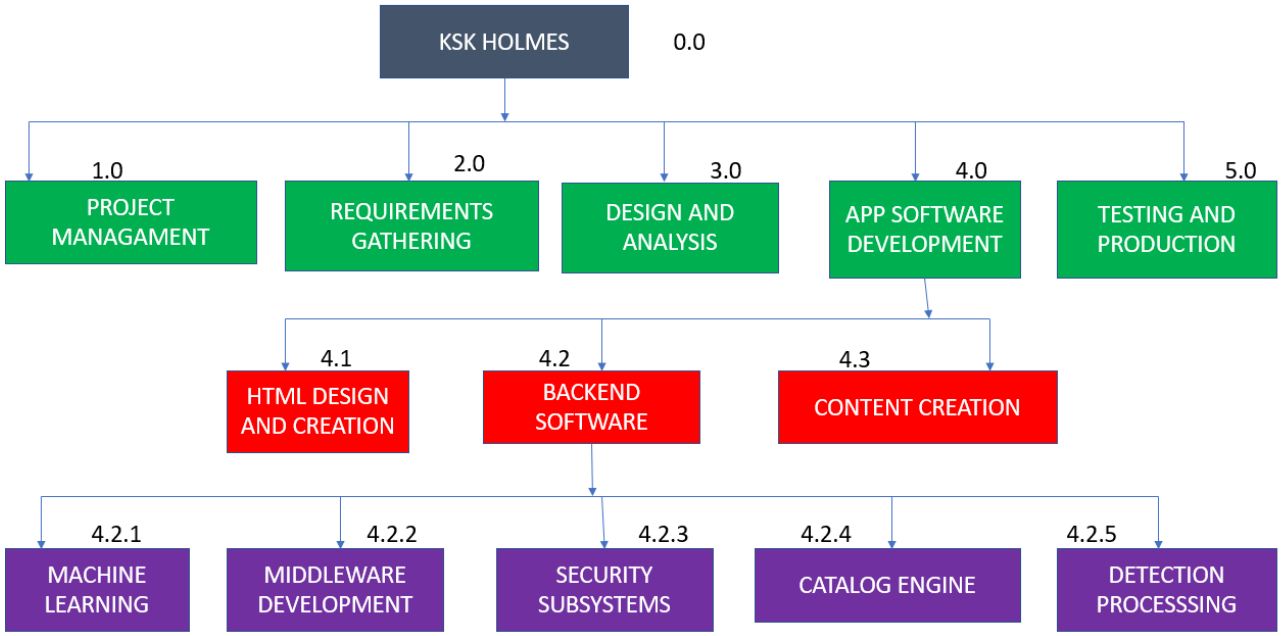
Aim

To Prepare Work breakdown structure, Timeline chart and Risk identification table

Team Members:

Sl No	Register No	Name	Role
1	RA2011003010363	SAKTHI MAHALAKSHMI S	**Rep**
2	RA2011003010355	KASHISH KEDIA	**Member**
3	RA2011003010372	KETSI DEBORAH P	**Member**

WBS



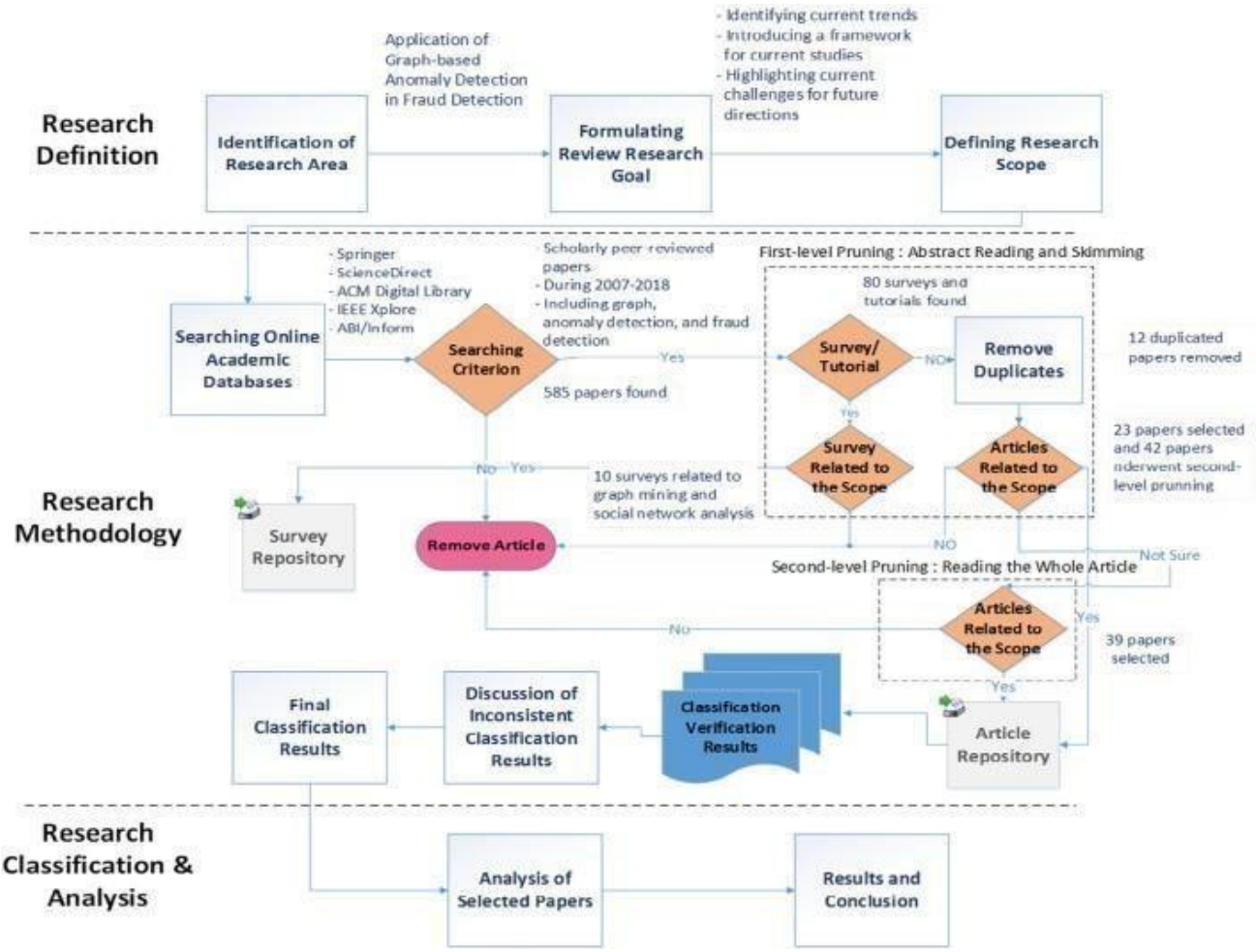
0.0 KSK HOLMES
1.0 Project Management
2.0 Requirements Gathering
3.0 Analysis & Design
4.0 APP Software Development
- 4.1 HTML Design and Creation
- 4.2 Backend Software
    - 4.2.1 Machine learning
    - 4.2.2 Middleware Developments
    - 4.2.3 Security Subsystems
    - 4.2.4 Catalog Design
    - 4.2.5 Detection Processing
- 4.3Content Creation

5.0 Testing and Production

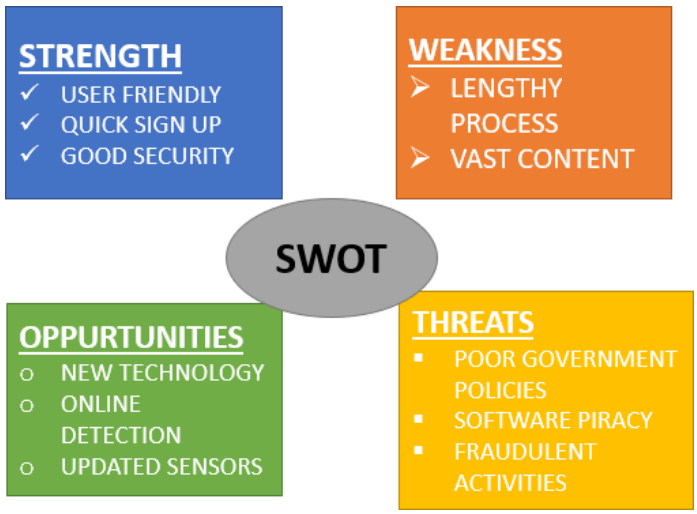
TIMELINE – GANTT CHART

Monitoring Metrics for Behavior-based Fraud Detection Solutions

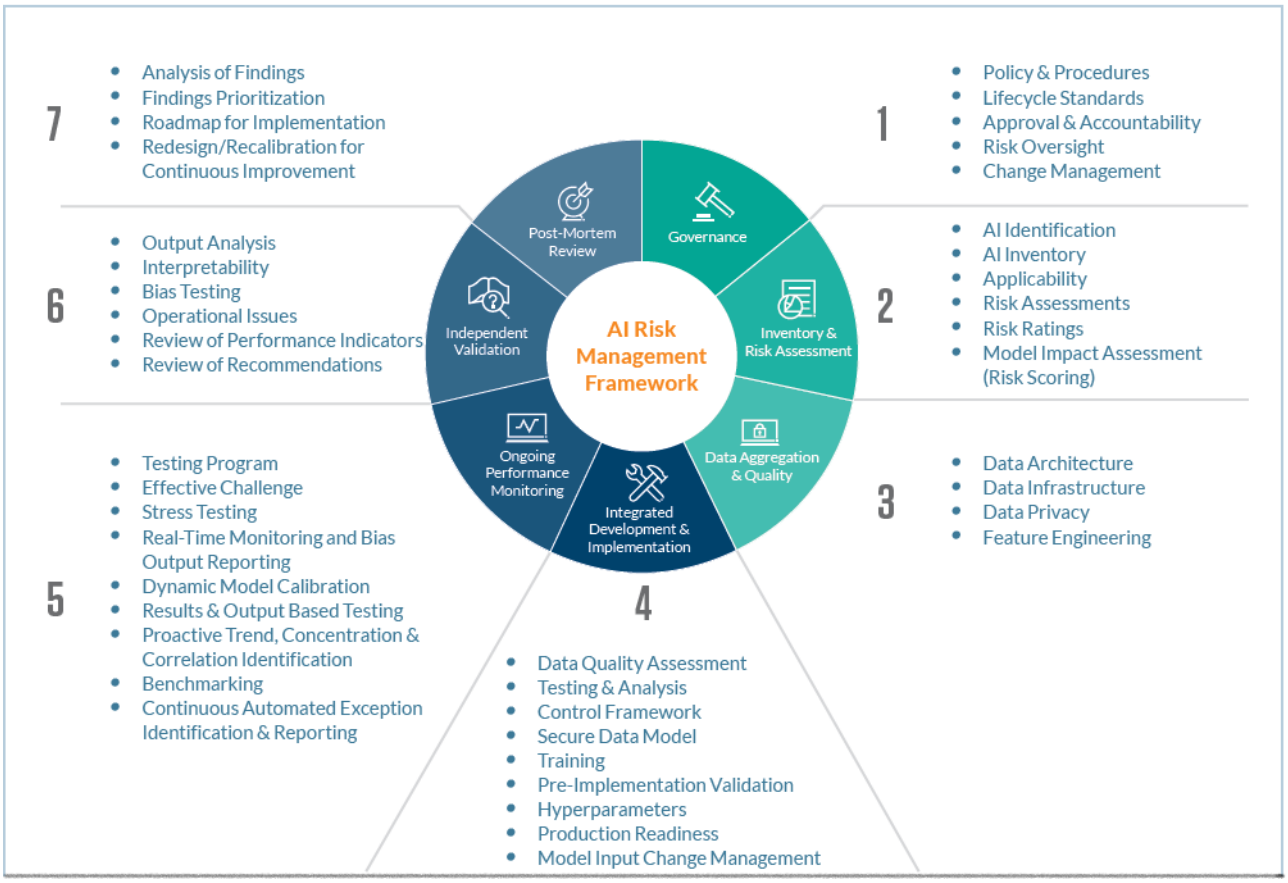
Login	Non- Transactional Activities		Transaction
• Challenges • Device • Cookie • IP Address • Time of day • Network	• View balance • View history • Updated address • Update email • Update password	• Add new user • Change limits • Set up batch • Set up template • Add payees	• ACH • Wire • Bill Pay • Loan Draw



RISK ANALYSIS – SWOT & RMMM



AI Risk Management Framework



With the continued investment in AI, the use of AI in business processes and practises is only growing larger in scope and deeper in granularity.

To stay ahead and provide effective and efficient monitoring of risk, organisations will not only utilise AI as their most comprehensive and valued tool but will need agile risk and compliance management. Competitive advantages will come not only from how organisations use AI but also from how they are able to avoid mistakes, ensure smooth customer experiences, prevent violations of law and explain what AI is intended to do to customers and regulators.

An AI tool will never be fully clear of risk, but an efficient and effective AI risk management framework will keep risk manageable and enable organisations to respond to fluctuations in the outputs and decisions generated by AI.

The key for all organisations using AI currently is to build and maintain AI in a responsible and transparent way, which, in turn, will help reduce operational cost and, more important, maintain the confidence of customers

Result:

Thus, the work breakdown structure with timeline chart and risk table were formulated successfully.

Department of Computing Technologies

SRM IST, Kattankulathur – 603 203

Course Code: 18CSC206J

Course Name: Software Engineering and Project Management

Experiment No	6
Title of Experiment	Design a System Architecture, Use Case and Class Diagram
Name of the candidate	P. KETSI DEBORAH
Team Members	1) KASHISH KEDIA 2) SAKTHI MAHALAKSHMI S
Register Number	RA2011003010372
Date of Experiment	20-04-2022

Mark Split Up

S.No	Description	Maximum Mark	Mark Obtained
1	Exercise	5	
2	Viva	5	
	Total	**10**	

Staff Signature with date

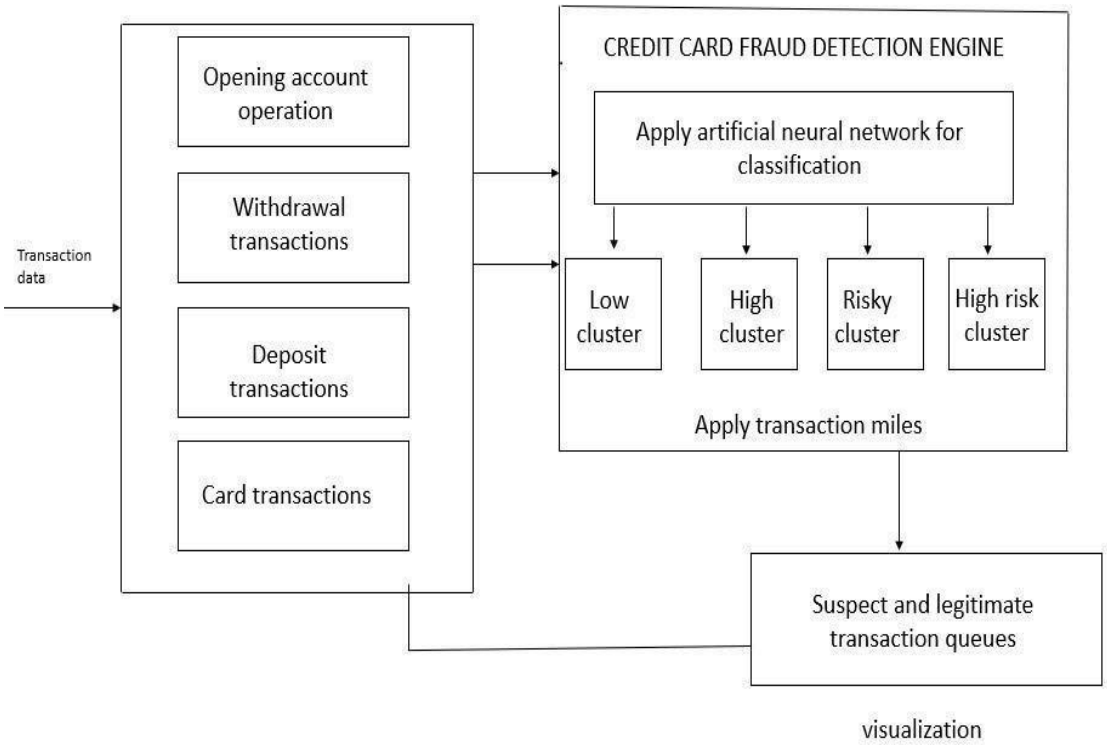
Aim

To Design a System Architecture, Use case and Class Diagram

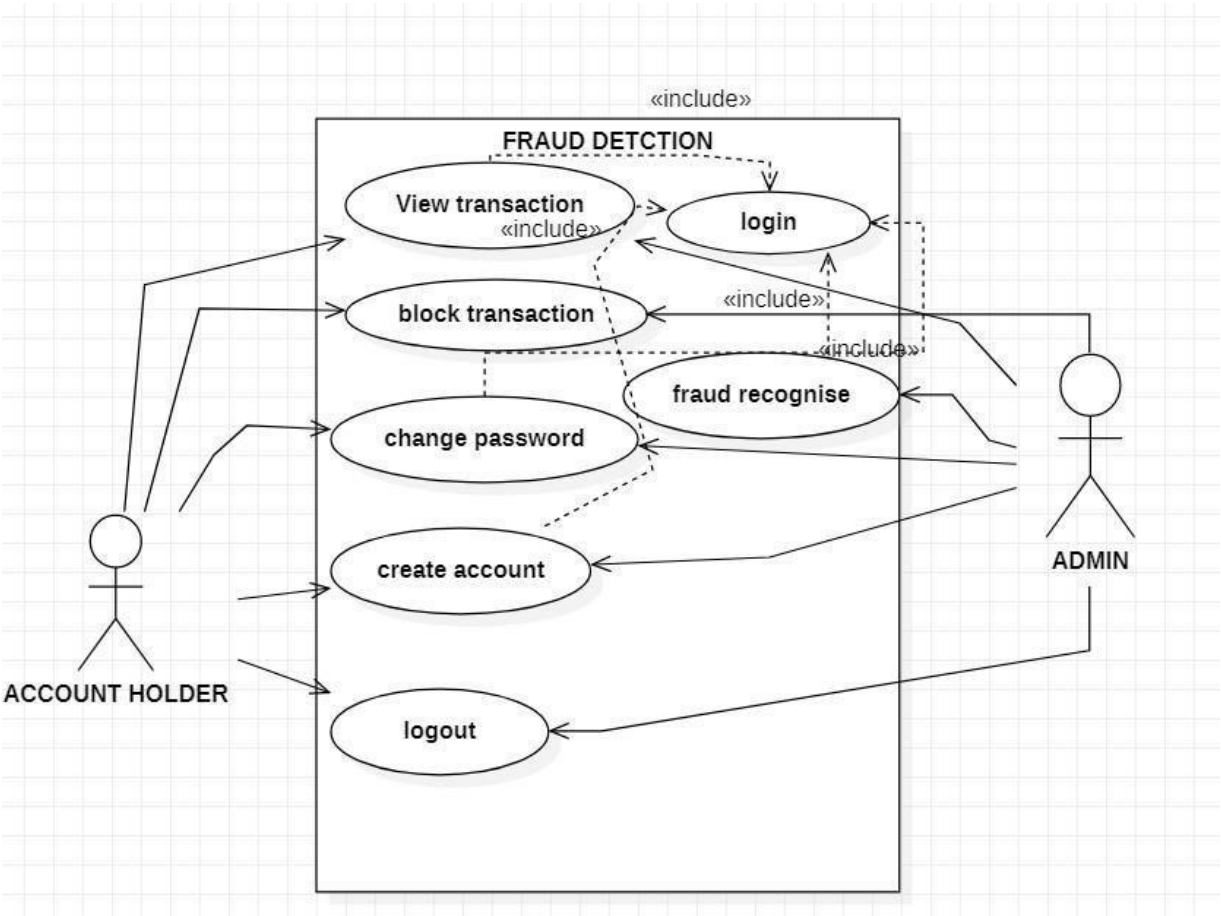
Team Members:

Sl No	Register No	Name	Role
1	RA2011003010363	SAKTHI MAHALAKSHMI S	**Rep**
2	RA2011003010355	KASHISH KEDIA	**Member**
3	RA2011003010372	KETSI DEBORAH P	**Member**

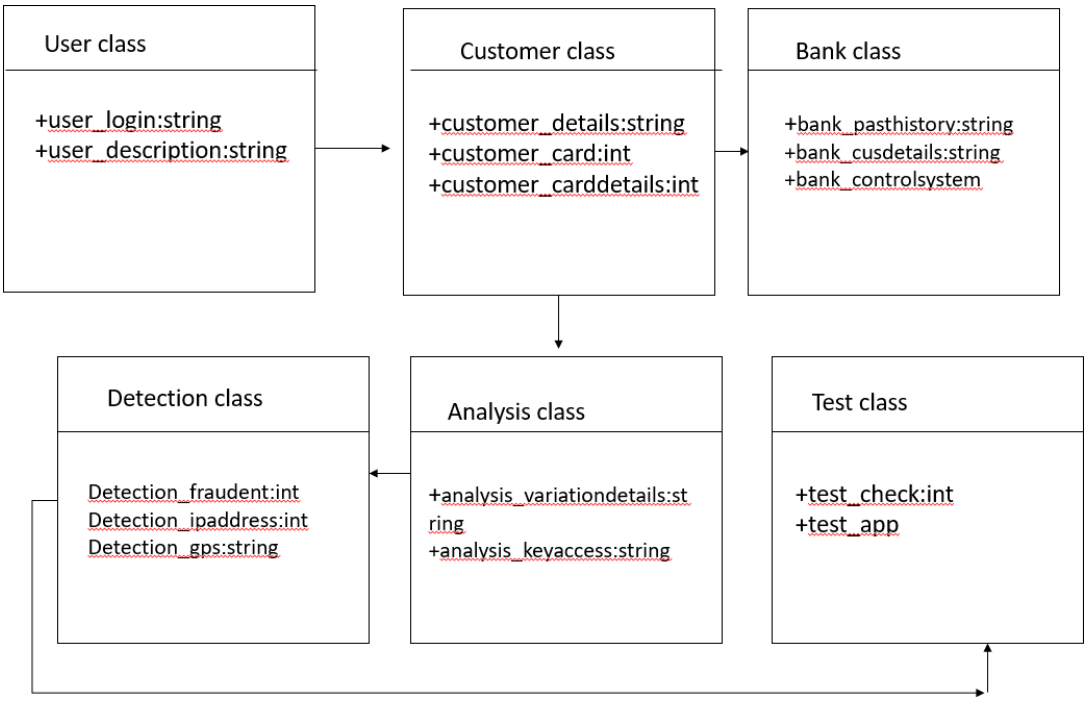
SYSTEM ARCHITECTURE



USE CASE DIAGRAM



CLASS DIAGRAM



Result:

Thus, the system architecture, use case and class diagram created successfully.

Department of Computing Technologies

SRM IST, Kattankulathur – 603 203

Course Code: 18CSC206J

Course Name: Software Engineering and Project Management

Experiment No	7
Title of Experiment	Design a Entity relationship diagram
Name of the candidate	P. KETSI DEBORAH
Team Members	1) KASHISH KEDIA 2) SAKTHI MAHALAKSHMI S
Register Number	RA2011003010372
Date of Experiment	27-04-2022

Mark Split Up

S. No	Description	Maximum Mark	Mark Obtained
1	Exercise	5	
2	Viva	5	
	Total	**10**	

Staff Signature with date

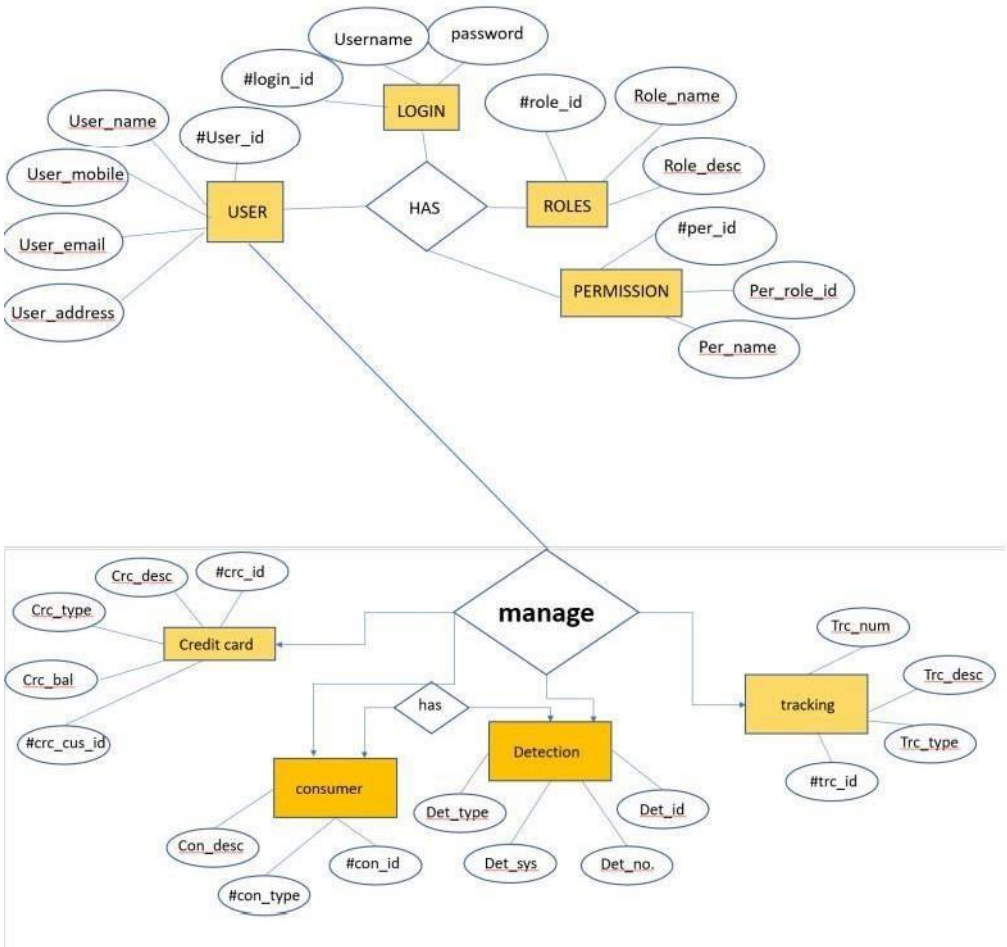
Aim

To create the Entity Relationship Diagram

Team Members:

S No	Register No	Name	Role
1	RA2011003010363	SAKTHI MAHALAKSHMI S	**Rep**
2	RA2011003010355	KASHISH KEDIA	**Member**
3	RA2011003010372	KETSI DEBORAH P	**Member**

<ER Diagram >



Result:

Thus, the entity relationship diagram was created successfully.

Department of Computing Technologies

SRM IST, Kattankulathur – 603 203

Course Code: 18CSC206J

Course Name: Software Engineering and Project Management

Experiment No	8
Title of Experiment	Develop a Data Flow Diagram (Process-Up to Level 1)
Name of the candidate	P. KETSI DEBORAH
Team Members	1) KASHISH KEDIA 2) SAKTHI MAHALAKSHMI S
Register Number	RA2011003010372
Date of Experiment	05-05-2022

Mark Split Up

S. No	Description	Maximum Mark	Mark Obtained
1	Exercise	5	
2	Viva	5	
	Total	**10**	

Staff Signature with date

Aim

To develop the data flow diagram up to level 1 for the KSK HOLMES credit card fraud detection software .

Team Members:

S No	Register No	Name	Role
1	RA2011003010363	SAKTHI MAHALAKSHMI S	Rep
2	RA2011003010355	KASHISH KEDIA	Member
3	RA2011003010372	P .KETSI DEBORAH	Member

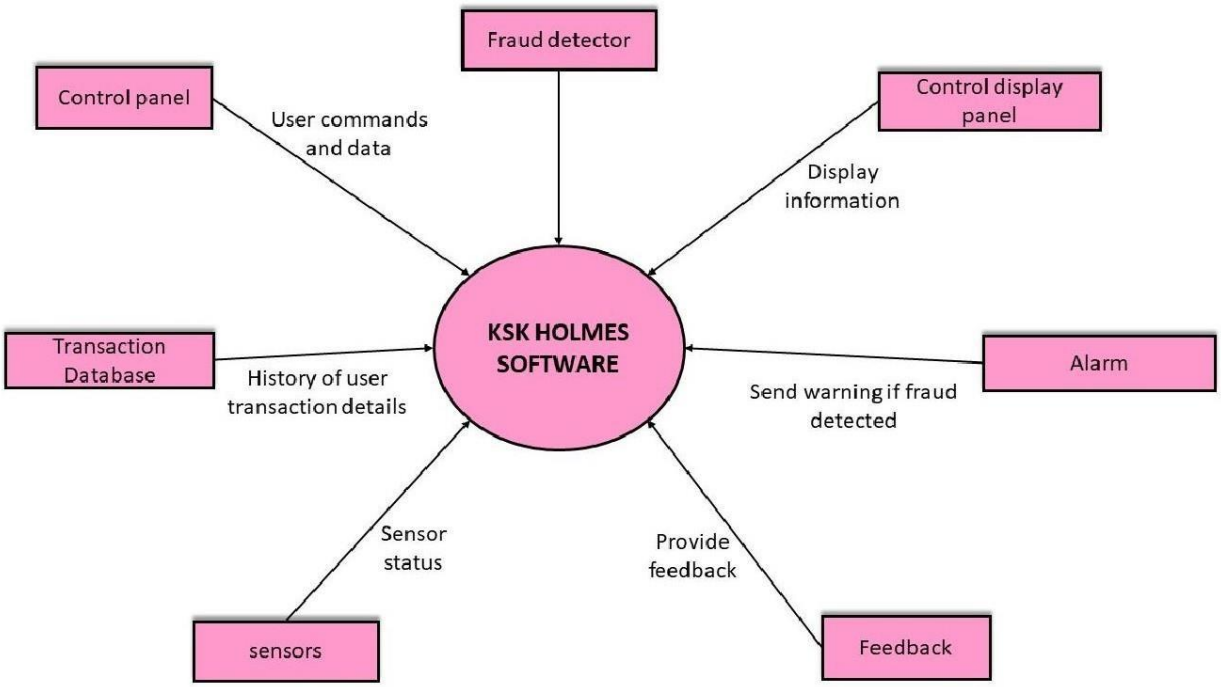
Data Flow Diagram

The DFD takes an input-process-output view of a system. That is, data objects flow into the software, are transformed by processing elements, and resultant data objects flow out of the software. Data objects are represented by labeled arrows, and transformations are represented by circles (also called bubbles). The DFD is presented in a hierarchical fashion. That is, the first data flow model (sometimes called a level 0 DFD or context diagram) represents the system as a whole. Subsequent data flow diagrams refine the context diagram, providing increasing detail with each subsequent level.

The data flow diagram enables you to develop models of the information domain andfunctional domain. As the DFD is refined into greater levels of detail, you perform animplicit functional decomposition of the system. At the same time, the DFD refinementresults in a corresponding refinement of data as it moves through the processes that embody the application.
A few simple guidelines can aid immeasurably during the derivation of a data flow diagram:
(1) Level 0 data flow diagram should depict the software/system as a single bubble;
(2) Primary input and output should be carefully noted;
(3) Refinement should begin by isolating candidate processes, data objects, and data stores to be represented at the next level;
(4) All arrows and bubbles should be labeled with meaningful names;
(5) Information flow continuity must be maintained from level to level and
(6) One bubble at a time should be refined. There is a natural tendency to overcomplicate the data flow diagram. This occurs when you attempt to show too much detail too early or represent procedural aspects of the software in lieu of information flow.

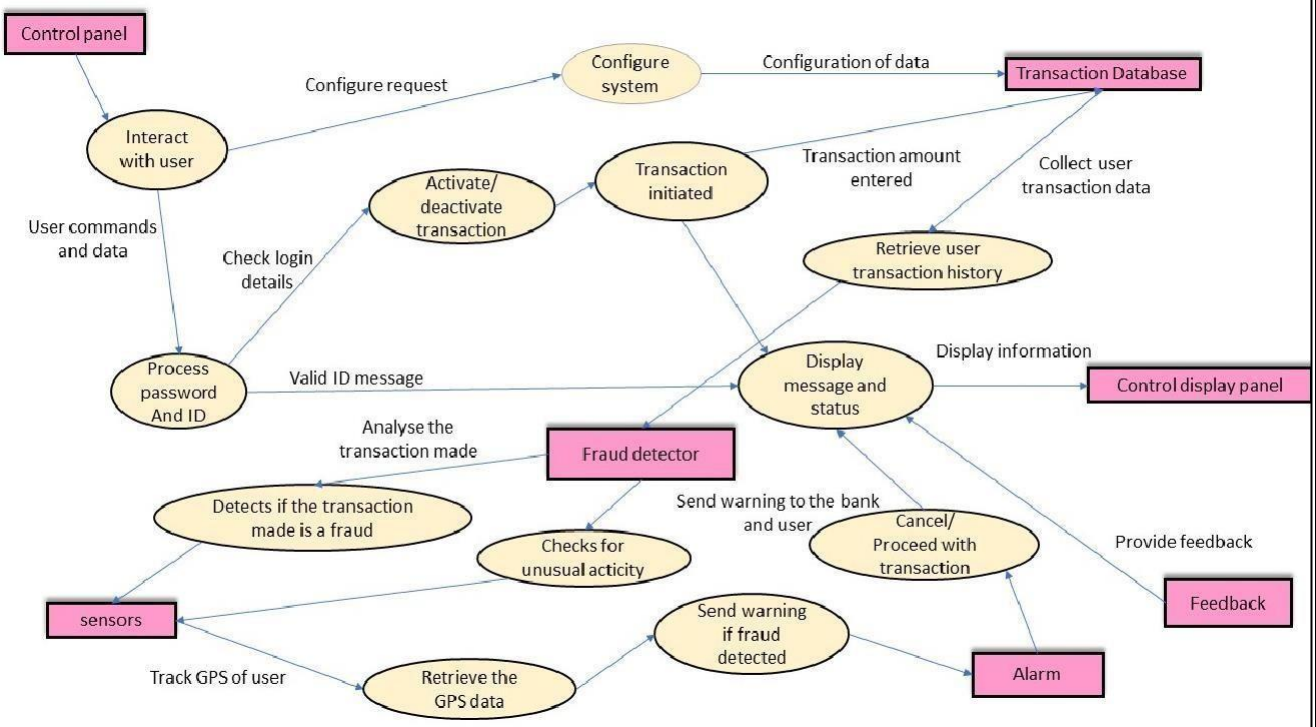
DFD LEVEL 0

Level 0 DFD for KSK HOLMES credit card fraud detection software



DFD LEVEL 1

Level 1 DFD for KSK HOLMES credit card fraud detection software



Result:

Thus, the data flow diagrams have been created for the KSK HOLMES credit card fraud detection software.

Department of Computing Technologies

SRM IST, Kattankulathur – 603 203

Course Code: 18CSC206J

Course Name: Software Engineering and Project Management

Experiment No	9
Title of Experiment	Design a Sequence and Collaboration Diagram
Name of the candidate	P. KETSI DEBORAH
Team Members	1) KASHISH KEDIA 2) SAKTHI MAHALAKSHMI S
Register Number	RA2011003010372
Date of Experiment	12-05-2022

Mark Split Up

S. No	Description	Maximum Mark	Mark Obtained
1	Exercise	5	
2	Viva	5	
	Total	**10**	

Staff Signature with date

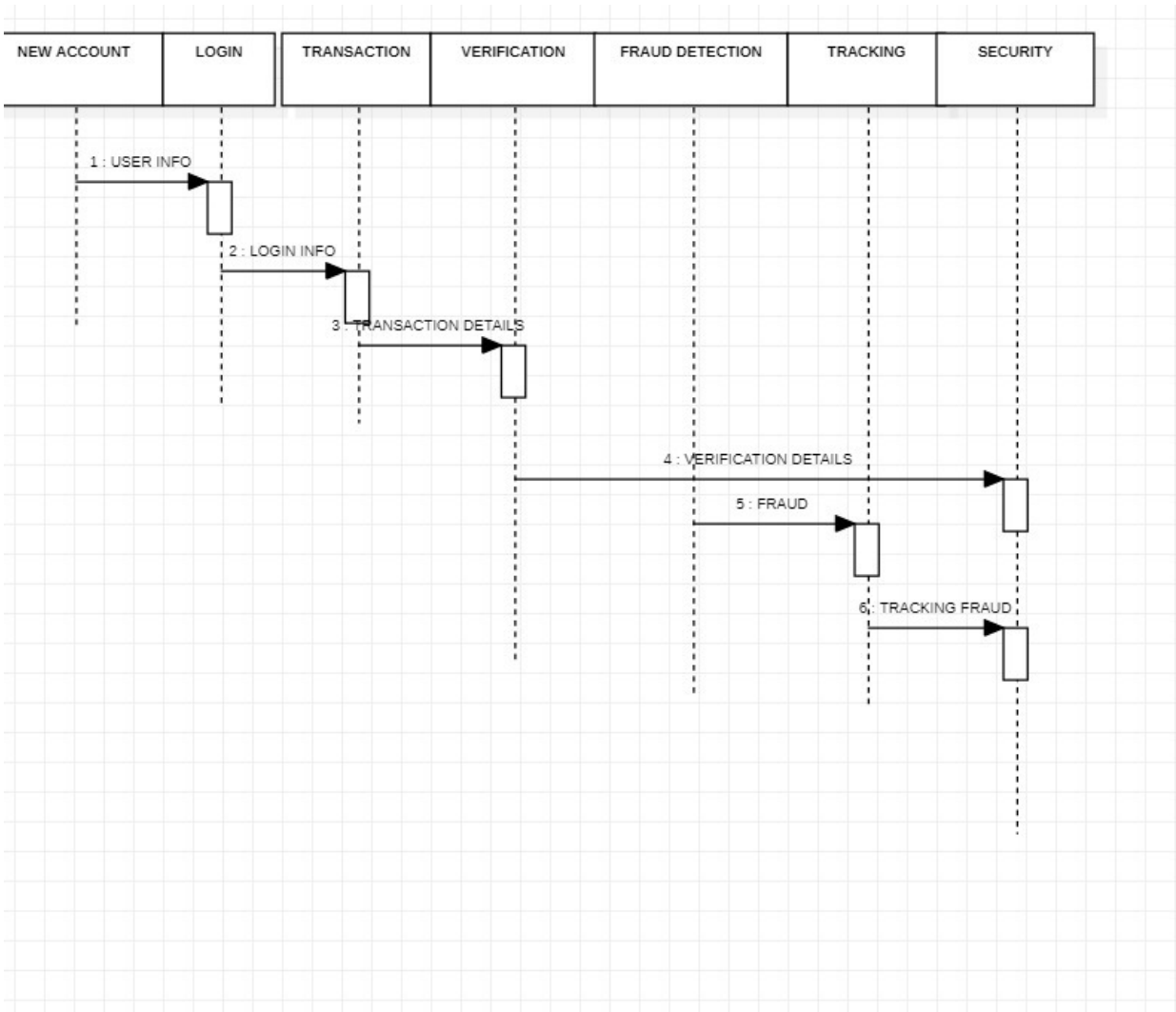
Aim

To create the sequence and collaboration diagram for the project fraud detection in Credit card system.

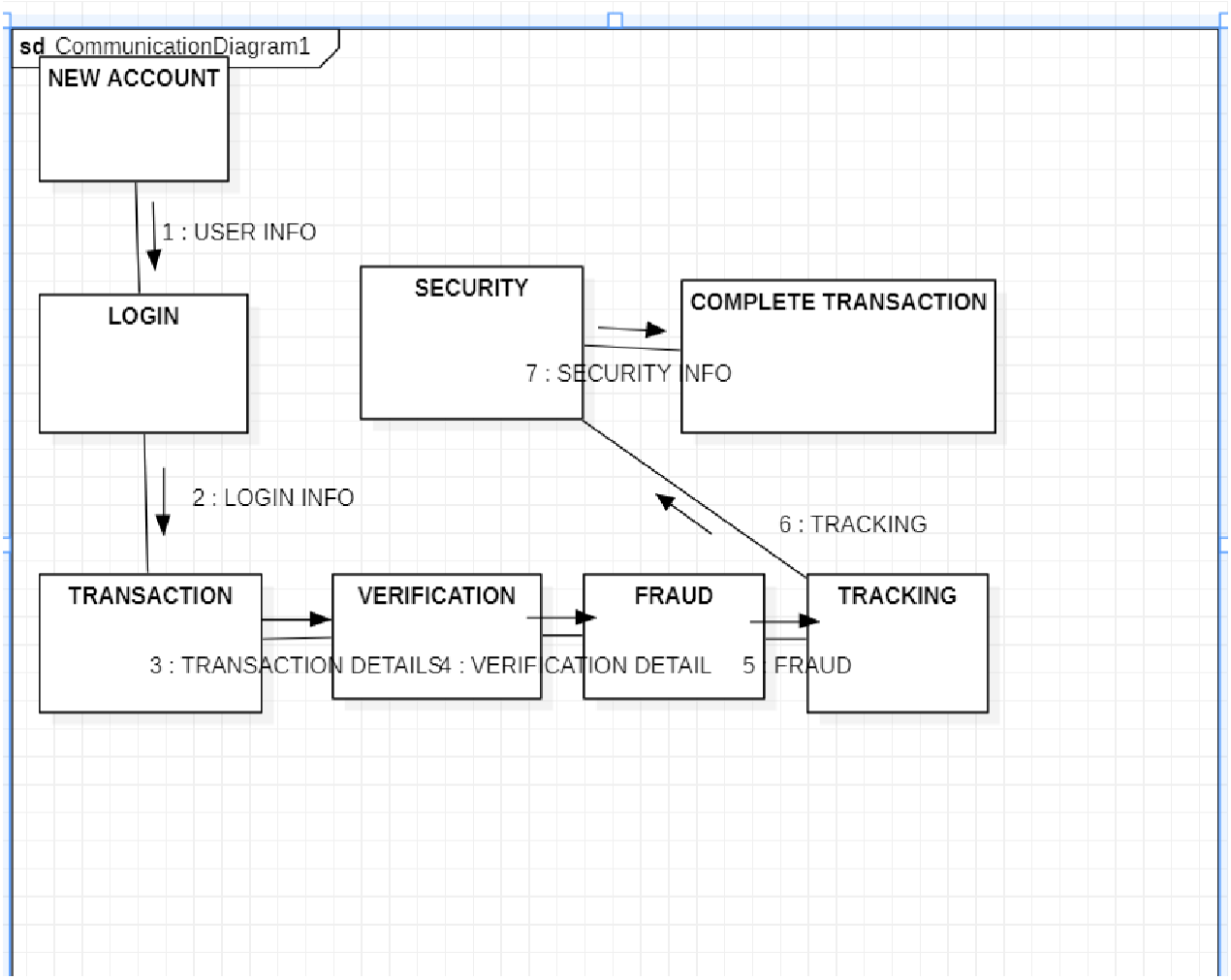
Team Members:

S No	Register No	Name	Role
1	RA2011003010363	SAKTHI MAHALAKSHMI S	**Rep/Member**
2	RA2011003010355	KASHISH KEDIA	**Member**
3	RA2011003010372	KETSI DEBORAH P	**Member**

SEQUENCE DIAGRAM:



COLLABORATION DIAGRAM:



Result:

Thus, the sequence and collaboration diagrams were created for the project fraud detection in credit card system.

School of Computing

SRM IST, Kattankulathur – 603 203

Course Code: 18CSC206J

Course Name: Software Engineering and Project Management

Experiment No	10
Title of Experiment	Develop a Testing Framework/User Interface
Name of the candidate	P. KETSI DEBORAH
Team Members	1) KASHISH KEDIA 2) SAKTHI MAHALAKSHMI S
Register Number	RA2011003010372
Date of Experiment	19-05-2022

Mark Split Up

S. No	Description	Maximum Mark	Mark Obtained
1	Exercise	5	
2	Viva	5	
	Total	**10**	

Staff Signature with date

Aim

To develop the testing framework and/or user interface framework for the fraud detection in credit card system.

Team Members:

S No	Register No	Name	Role
1	RA2011003010363	SAKTHI MAHALAKSHMI S	**Rep/Member**
2	RA2011003010355	KASHISH KEDIA	**Member**
3	RA2011003010372	KETSI DEBORAH P	**Member**

CATEGORY	METHODOLOGY	TOOLS REQUIRED
Functional Requirements	Anomaly detection Artificial Neural Network Data Mining User Profiling	Word Template

ANOMALY DETECTION :

It involves the comparison of an individual details, recognising unsual data entries that appear irregular in a current dataset. This method typically highlights entries that fail to follow an expected pattern or format, which previously seemed comparable and alike, which has proven beneficial in the credit card sector. Anomaly detection can be a difficult technique to initiate as it can be challenging to define a region of activity which is deemed to be normal. Therefore, identifying a boundary between normal and abnormal datasets can prove problematic.

ARTIFICIAL NEURAL NETWORKS:

These are information processing systems, which is modelled on the biological nervous system and the way in which the human brain processes information.These networks are capable of pattern recognition and machine learning but require extensive training and learning prior to use. Using numerous background information i.e. individuals income, previous transactions or current fraud affecting other banks. Artificial Neural Networks use algorithms to predict future transactions, whilst determining if current transactionsare genuine or fraudulent

DATA MINING:

The main goal of Data Mining is to predict future actions, through the searching of consistent and familiar patterns in current data sets. The most common form of data mining is predictive modelling, which involves the production of a data model, through various statistical techniques, to predict and highlight the probability of a future or otherwise unknown outcome. Through predictive modelling, risk levels associated with individuals can be determined, providing guidance as to whether the individual is reliable and capable to make future repayments when expected.

USER PROFILING:

It can be used to access an individual point of application, in order to draw a conclusion as to whether they possess a suspicious background. It is possible
to construct a typical user profile, summarising what is deemed to be the users' typical behaviour. The created user profile will then be referred to in future transactions, allowing the comparison of current and historic actions. This comparison
perceived normal behaviour, any deviations in behaviour comparison can help to build a suspicious case against the individual, helping an organisation to reach their final decision

CATEGORY	METHODOLOGY	TOOLS REQUIRED
Non Functional Requirements	1. Assess user expectation	Word Template
	2. Recognize the market demand. ...	
	3. Analyze the performance	

ASSESS USER EXPECTATION:

Think about what level of quality your target audience is seeking. For example, they may expect enhanced security options or high-speed performance. Meeting their expectations can help improve the public perception of your application's performance.

RECOGNISE MARKWET DEMAND:

Consider conducting market research to determine the demand for specific nonfunctional requirements and comparing them to what your competitors already offer. Then you can innovate designs that other companies haven't released yet, such as crack-resistant screens.

ANALYSE THE PERFORMANCE:

Try testing how your nonfunctional requirements work to assess their quality. You can gather a focus group to provide feedback on the usability or review the number of shutdowns to find out how quickly an application can overcome technical errors.

Result:

Thus, the testing framework/user interface framework has been created for the Fraud detection in credit card system.

School of Computing

SRM IST, Kattankulathur – 603 203

Course Code: 18CSC206J

Course Name: Software Engineering and Project Management

Experiment No	11
Title of Experiment	Test Cases
Name of the candidate	P. KETSI DEBORAH
Team Members	1) KASHISH KEDIA 2) SATHI MAHALAKSHMI S
Register Number	RA2011003010372
Date of Experiment	26-05-2022

Mark Split Up

S. No	Description	Maximum Mark	Mark Obtained
1	Exercise	5	
2	Viva	5	
	Total	**10**	

Staff Signature with date

Aim

To develop the test cases manual for the <project name>

Team Members:

S No	Register No	Name	Role
1	RA2011003010363	SAKTHI MAHALAKSHMI S	**Rep**
2	RA2011003010355	KASHISH KEDIA	**Member**
3	RA2011003010372	KETSI DEBORAH P	**Member**

FUNCTIONAL TEST CASE:

Test ID (#)	Test Scenario	Test Case	Execution Steps	Expected Outcome	Actual Outcome	Status	Remarks
test 1	Verify User Login credentials	Accept user id and password	1.User enters the bank portal 2.Enter user id and password 3.Click login button	User should be Logged in to the bank account	User logged in	Pass	test is successful
test 2	verify transaction details	checking transaction details	1. retrieving transaction data set 2.checking transaction history	abnormal transaction not detected	transaction complete	Pass	test is successful
test 3	verifying suspicious activity	checking transaction details	1. retrieving transaction data set 2.checking transaction history	abnormal transaction detected	warning message displayed with fraud detected	Pass	test is successful

NON-FUNCTIONAL TEST CASE:

Test ID (#)	Test Scenario	Test Case	Execution Steps	Expected Outcome	Actual Outcome	Status	Remarks
1	Performance Testing	Mean Response Time	1.User enters the app 2.User waits for the app to load completely	App should load with the mean response time of less than 5 second	Response time was around 5 second	pass	test is successful
2	Load Testing	Application Should not crash while loading	1.Open application and record the number of test cases in which the app fails to load completely on a stable metered connection. 2.The sample space is 1000	Failed Test cases should be less than 5%	Failed Test cases should be less than 0.5%	pass	test is successful
3	Serviceability	Application should be available 24*7	1.User should be able to use the app whenever in need	Application is serviceable	Application was found to be user friendly	pass	test is successful
4	Maintainability	Application should be updated frequently	1.bugs and glitches should be fixed	Application is updated and maintained	Application was found to be updated with new features	pass	test is successful

Result:

Thus, the test case manual has been created for the project fraud detection in credit card system.

School of Computing

SRM IST, Kattankulathur – 603 203

Course Code: 18CSC206J

Course Name: Software Engineering and Project Management

Experiment No	12
Title of Experiment	Manual Test Case Reporting
Name of the candidate	P. KETSI DEBORAH
Team Members	1) KASHISH KEDIA 2) SAKTHI MAHALAKSHMI S
Register Number	RA2011003010372
Date of Experiment	02-06-2022

Mark Split Up

S. No	Description	Maximum Mark	Mark Obtained
1	Exercise	5	
2	Viva	5	
	Total	**10**	

Staff Signature with date

Aim

To prepare the manual test case report for the project Fraud Detection In Credit Card System.

Team Members:

S No	Register No	Name	Role
1	RA2011003010363	SAKTHI MAHALAKSHMI S	**Rep/Member**
2	RA2011003010355	KASHISH KEDIA	**Member**
3	RA2011003010372	KETSI DEBORAH P	**Member**

Test Report :

Testing was successful. Obstacles were presented to the stakeholders and were further looked into. Obstacles were removed

Category	Progress Against Plan	Status
Functional Testing	Green	Completed
Non-Functional Testing	Green	Completed

Functional	Test Case Coverage (%)	Status
Module 1	100%	Completed
Module 2	100%	Completed
Module 3	100%	Completed
Module 4	100%	Completed
Module 5	100%	Completed
Module 6	100%	Completed

Result:

Thus, the test case report has been created for the project Fraud Detection In Credit Card System.

School of Computing

SRM IST, Kattankulathur – 603 203

Course Code: 18CSC206J

Course Name: Software Engineering and Project Management

Experiment No	13
Title of Experiment	Provide the details of Architecture Design/Framework/Implementation
Name of the candidate	P. KETSI DEBORAH
Team Members	1) KASHISH KEDIA 2) SAKTHI MAHALAKSHMI S
Register Numbers	RA2011003010372
Date of Experiment	09- 06-2022

Mark Split Up

S. No	Description	Maximum Mark	Mark Obtained
1	Exercise	5	
2	Viva	5	
	Total	**10**	

Staff Signature with date

Aim

To provide the details of architectural design/framework/implementation

Team Members:

S No	Register No	Name	Role
1	RA2011003010363	SAKTHI MAHALAKSHMI S	**Rep/Member**
2	RA2011003010355	KASHISH KEDIA	**Member**
3	RA2011003010372	KETSI DEBORAH	**Member**

Code: Importing all the necessary Libraries

```
# import the necessary packages
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from matplotlib import gridspec
```

Code: Loading the Data

```
# Load the dataset from the csv file using pandas
# best way is to mount the drive on colab and
# copy the path for the csv file
data = pd.read_csv("credit.csv")
```

Code: Understanding the Data

```
# Grab a peek at the data
data.head()
```

	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14
0	0.0	-1.359807	-0.072781	2.536347	1.378155	-0.338321	0.462388	0.239599	0.098698	0.363787	0.090794	-0.551600	-0.617801	-0.991390	-0.311169
1	0.0	1.191857	0.266151	0.166480	0.448154	0.060018	-0.082361	-0.078803	0.085102	-0.255425	-0.166974	1.612727	1.065235	0.489095	-0.143772
2	1.0	-1.358354	-1.340163	1.773209	0.379780	-0.503198	1.800499	0.791461	0.247676	-1.514654	0.207643	0.624501	0.066084	0.717293	-0.165946
3	1.0	-0.966272	-0.185226	1.792993	-0.863291	-0.010309	1.247203	0.237609	0.377436	-1.387024	-0.054952	-0.226487	0.178228	0.507757	-0.287924
4	2.0	-1.158233	0.877737	1.548718	0.403034	-0.407193	0.095921	0.592941	-0.270533	0.817739	0.753074	-0.822843	0.538196	1.345852	-1.119670

Code: Describing the Data

```
# Print the shape of the data
# data = data.sample(frac = 0.1, random_state = 48)
print(data.shape)
print(data.describe())
```

Output:

```
(284807, 31)
```

	Time	V1	...	Amount	Class
count	284807.000000	2.848070e+05	...	284807.000000	284807.000000
mean	94813.859575	3.919560e-15	...	88.349619	0.001727
std	47488.145955	1.958696e+00	...	250.120109	0.041527
min	0.000000	-5.640751e+01	...	0.000000	0.000000
25%	54201.500000	-9.203734e-01	...	5.600000	0.000000
50%	84692.000000	1.810880e-02	...	22.000000	0.000000
75%	139320.500000	1.315642e+00	...	77.165000	0.000000
max	172792.000000	2.454930e+00	...	25691.160000	1.000000

```
[8 rows x 31 columns]
```

Code: Imbalance in the data

Time to explain the data we are dealing with.

```
# Determine number of fraud cases in dataset
fraud = data[data['Class'] == 1]
valid = data[data['Class'] == 0]
outlierFraction = len(fraud)/float(len(valid))
print(outlierFraction)
print('Fraud Cases: {}'.format(len(data[data['Class'] == 1])))
print('Valid Transactions: {}'.format(len(data[data['Class'] == 0])))
```

```
0.0017304750013189597
Fraud Cases: 492
Valid Transactions: 284315
```

Only *0.17%* fraudulent transaction out all the transactions. The data is highly Unbalanced. Lets first apply our models without balancing it and if we don't get a good accuracy then we can find a way to balance this dataset. But first, let's implement the model without it and will balance the data only if needed.

Code: Print the amount details for Fraudulent Transaction

```
print("Amount details of the fraudulent transaction")
fraud.Amount.describe()
```

Output:

```
Amount details of the fraudulent transaction
count     492.000000
mean      122.211321
std       256.683288
min         0.000000

25%         1.000000
50%         9.250000
75%       105.890000
max      2125.870000
Name: Amount, dtype: float64
```

Code: Print the amount details for Normal Transaction

```
print("details of valid transaction")
valid.Amount.describe()
```

Output:

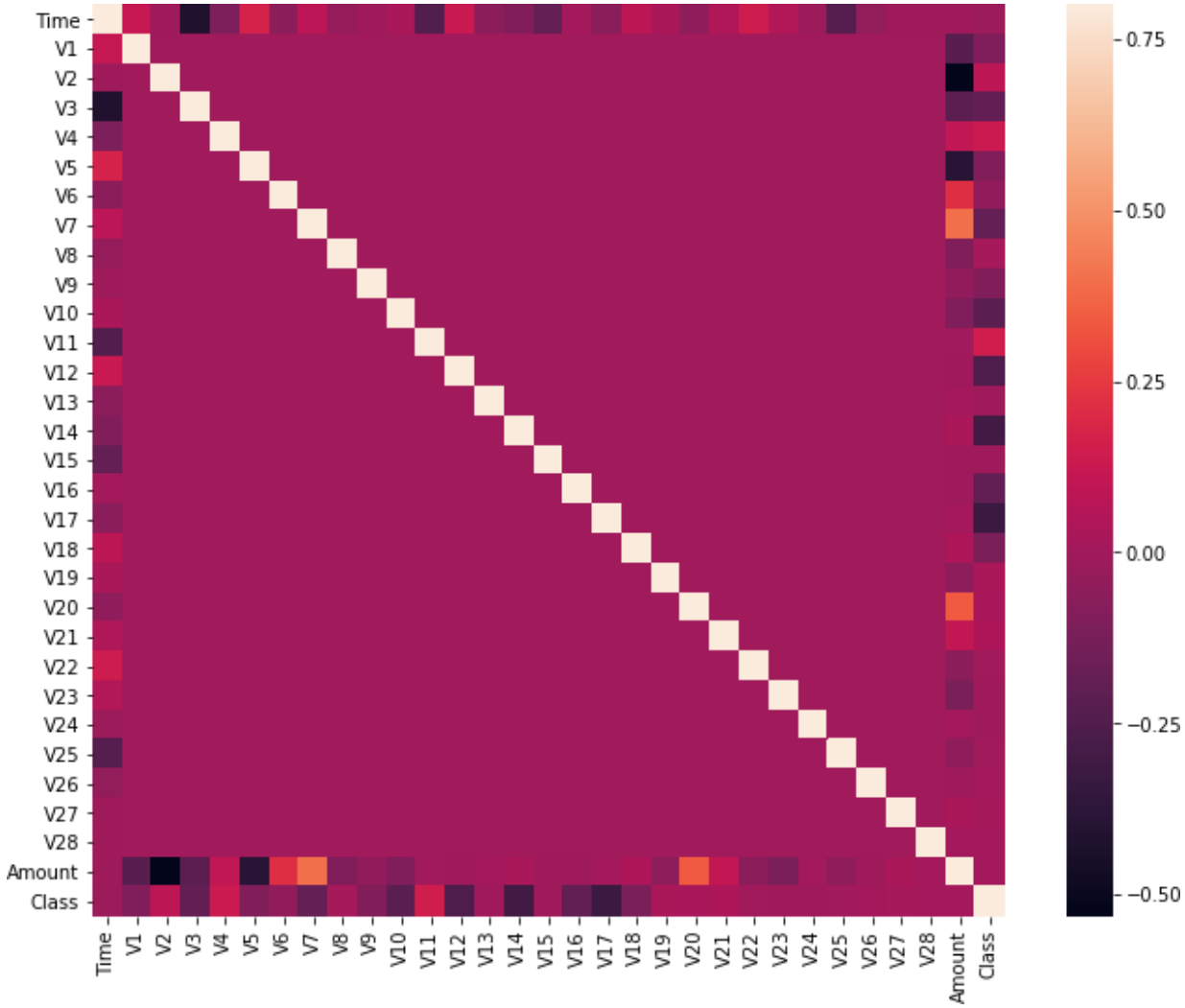
```
Amount details of valid transaction
count    284315.000000
mean         88.291022
std         250.105092
min           0.000000
25%           5.650000
50%          22.000000
75%          77.050000
max       25691.160000
Name: Amount, dtype: float64
```

As we can clearly notice from this, the average Money transaction for the fraudulent ones is more. This makes this problem crucial to deal with.

Code: Plotting the Correlation Matrix

The correlation matrix graphically gives us an idea of how features correlate with each other and can help us predict what are the features that are most relevant for the prediction.

```
# Correlation matrix
corrmat = data.corr()
fig = plt.figure(figsize = (12, 9))
sns.heatmap(corrmat, vmax = .8, square = True)
plt.show()
```



In the HeatMap we can clearly see that most of the features do not correlate to other features but there are some features that either has a positive or a negative correlation with each other. For example, *V2* and *V5* are highly negatively correlated with the feature called *Amount*. We also see some correlation with *V20* and *Amount*. This gives us a deeper understanding of the Data available to us.

Code: Separating the X and the Y values

Dividing the data into inputs parameters and outputs value format

```
# dividing the X and the Y from the dataset
X = data.drop(['Class'], axis = 1)
Y = data["Class"]
print(X.shape)
print(Y.shape)
# getting just the values for the sake of processing
# (its a numpy array with no columns)
```

```
xData = X.values
yData = Y.values
```

Output:

```
(284807, 30)
(284807, )
```

Training and Testing Data Bifurcation

We will be dividing the dataset into two main groups. One for training the model and the other for Testing our trained model's performance.

```
# Using Skicit-learn to split data into training and testing sets
from sklearn.model_selection import train_test_split
# Split the data into training and testing sets
xTrain, xTest, yTrain, yTest = train_test_split(
        xData, yData, test_size = 0.2, random_state = 42)
```

Code: Building a Random Forest Model using skicit learn

```
# Building the Random Forest Classifier (RANDOM FOREST)
from sklearn.ensemble import RandomForestClassifier
# random forest model creation
rfc = RandomForestClassifier()
rfc.fit(xTrain, yTrain)
# predictions
yPred = rfc.predict(xTest)
```

Code: Building all kinds of evaluating parameters

```
# Evaluating the classifier
# printing every score of the classifier
# scoring in anything
from sklearn.metrics import classification_report, accuracy_score
from sklearn.metrics import precision_score, recall_score
from sklearn.metrics import f1_score, matthews_corrcoef
from sklearn.metrics import confusion_matrix

n_outliers = len(fraud)
n_errors = (yPred != yTest).sum()
print("The model used is Random Forest classifier")

acc = accuracy_score(yTest, yPred)
print("The accuracy is {}".format(acc))
```

```
prec = precision_score(yTest, yPred)
print("The precision is {}".format(prec))

rec = recall_score(yTest, yPred)
print("The recall is {}".format(rec))

f1 =  f1_score(yTest, yPred)
print("The F1-Score is {}".format(f1))

MCC = matthews_corrcoef(yTest, yPred)
print("The Matthews correlation coefficient is{}".format(MCC))
```

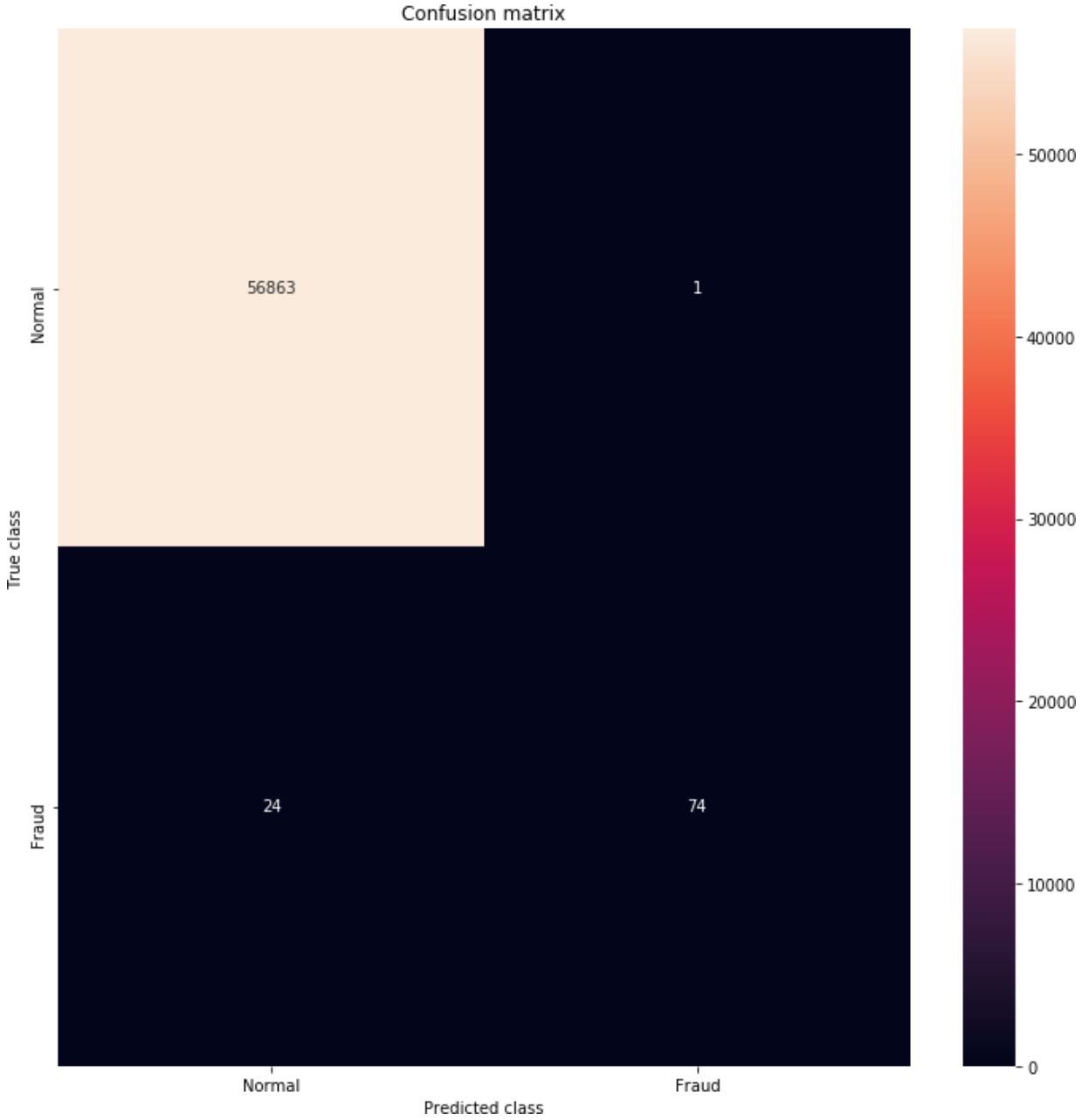
Output:

```
The model used is Random Forest classifier

The accuracy is 0.9995611109160493

The precision is 0.9866666666666667

The recall is 0.7551020408163265

The F1-Score is 0.8554913294797689

The Matthews correlation coefficient is0.8629589216367891
```

Code: Visualizing the Confusion Matrix

```
# printing the confusion matrix
LABELS = ['Normal', 'Fraud']
conf_matrix = confusion_matrix(yTest, yPred)
plt.figure(figsize =(12, 12))
sns.heatmap(conf_matrix, xticklabels = LABELS,
            yticklabels = LABELS, annot = True, fmt ="d");
plt.title("Confusion matrix")
plt.ylabel('True class')
plt.xlabel('Predicted class')
plt.show()
```

Output:



Comparison with other algorithms without dealing with the imbalancing of the data.

What other Data Scientists got

Method Used	Frauds	Genuines	MCC
Naïve Bayes	83.130	97.730	0.219
Decision Tree	81.098	99.951	0.775
Random Forest	42.683	99.988	0.604
Gradient Boosted Tree	81.098	99.936	0.746
Decision Stump	66.870	99.963	0.711
Random Tree	32.520	99.982	0.497
Deep Learning	81.504	99.956	0.787
Neural Network	82.317	99.966	0.812
Multi Layer Perceptron	80.894	99.966	0.806
Linear Regression	54.065	99.985	0.683
Logistic Regression	79.065	99.962	0.786
Support Vector Machine	79.878	99.972	0.813

CONCLUSION

Credit card fraud is most common problem resulting in loss of lot money for people and loss for some banks and credit card company. This project want to help the peoples from their wealth loss and also for the banked company and trying to develop the model which more efficiently separate the fraud and fraud less transaction by using the time and amount feature in data set given. We build the model using some machine learning algorithms such as logistic regression, decisiontree, support vector machine, this all are supervised machine learning algorithm in machine learning.

In feature solving this problem statement using another part of artificial intelligence that is time series analysis, in our present project we used both and time and amount feature mainly for predicting the weather the transaction is fraud or Nonfraud transaction, in time series analysis we can reduce the number of parameters that is feature required for the model and we can achieve this model by using average method ,moving average or window method, naive method and sessional naive methods but all this method have some advantages and disadvantages.

REFERENCES

- ✓ https://www.geeksforgeeks.org/ml-credit-card-fraud-detection/
- ✓ https://cppsecrets.com/users/52711141051151049798104461141074955485253641031099710510846991111 09/Advanced-Python-Project-Credit-Card-Fraud-Detection.php
- ✓ https://github.com/topics/credit-card-fraud-detection
- ✓ https://www.researchgate.net/publication/336800562_Credit_Card_Fraud_Detection_using_Machine_Learning_and_Data_Science
- ✓ https://www.freeprojectz.com/entity-relationship/credit-card-approval-system-er-diagram