Module Name - Software Development Group Project

Module Code - 5COSC009C.2

Module Leader - Banuka Athuraliya

CREDIT MART

Team Name - 50 Cent

Abstract

One of the main goals of a company is to make sure that their product is available at every place possible so that the consumer will not have to go through a hassle in finding the product. Failing to meet this goal can result in a loss of sale or at worst, a customer choosing another product due to ease of access. To achieve to this goal many companies sell their product <u>on credit</u> to a local retailer who will get a credit period to return the money for the items borrowed.

This process of finding local retailers and lending products usually end up in a situation where the retailer is unable to keep his promise of returning the money in given period of time or sometimes not paying at all which leads to bad debts of the company. Credit mart is an application which can potentially reduce the amount of bad debts a company would face in the future by financially rating other companies or so-called retailers. Since from the beginning of time business transactions have been done in good faith between the two parties but as an organization grows larger it is hard to keep track all their customers. Credit mart identifies this potential gap of trust and keeps it in terms of financial statistics of that company.

This research uses financial indicators of a company such as their profitability and past financial transactions to give the company a financial rating whereas the company willing to lend products will be able provide goods at a level the consumer will be able to meet his debt obligation and finally satisfying both parties in the process. The research will consider other existing credit analyzing products and models. These models will be studied in depth and analyzed in order to provide a good final product.

Team - 50 Cent

Acknowledgement

First and foremost, we would like to thank god for giving this team the strength and unity to keep moving forward no matter what the problem we faced. Even though the journey had barely begun the team had to face many challenges which we have resolved successfully so far.

Secondly, we would like to thank the module leader Mr. Banuka Athuraliya and the entire feedback panel for listening to our ideas with patience and providing feedback based on their knowledge in that area. Every feedback session attended gave the team more insight on the project. Special thanks to the domain experts and data science experts for giving the team specialized feedback on the project and how to approach technical problems. We appreciate the few companies who took the time to fill in our online surveys and contributing their ideas towards this project.

Finally, we would like to thank all our colleagues, friends and other SDGP teams who have helped us whenever we needed a help or information regarding the module.

Contents

Team - 50 Cent

Team - 50 Cent

Team - 50 Cent

List of figures

List of tables

Team - 50 Cent

Appendix

Team - 50 Cent

1.Introduction

1.1 Chapter Overview

The below chapter gives a detailed introduction to the system Credit-Mart. The reason it was created and the possible problems it can solve in the future for all companies transacting in credit basis. Hence the name Credit-Mart which emphasizes the main goal of the system. The chapter further elaborates on the real-life scenarios faced by businesses and available existing solutions to such problems, furthermore, explains the aims and objectives set to achieve by the system and finally the requirements needed to develop the system.

1.2 Project Background

1.2.1 Introduction to problem

1.2.1.1 Problem boundary

Businesses face major issues when they give on credit basis which occur for only credit translation. So, these businesses face a major problem when they carry the risk of their clients not meeting their debt obligations, so these companies face huge losses due to bad debt because of this they do not tend to do business with new customers on credit. These companies tend to create a trust with their client by transacting through cash at first and then they give their products on credit basis

The continuous risk of credit sales can result in even closing of a business. A company cannot only rely on cash sales because it will have some bad effects. Credit allows individuals to purchase large quantities of items on a credit basis. If companies did not offer such credit sales, they would not generate as many Profit.

Also, the business cannot rate a particular customer as a good customer or a bad customer by looking at their business name or reputations, so they tend to do transactions on credit by keeping curators as a security. This is the main problem companies face when they are doing credit sales.
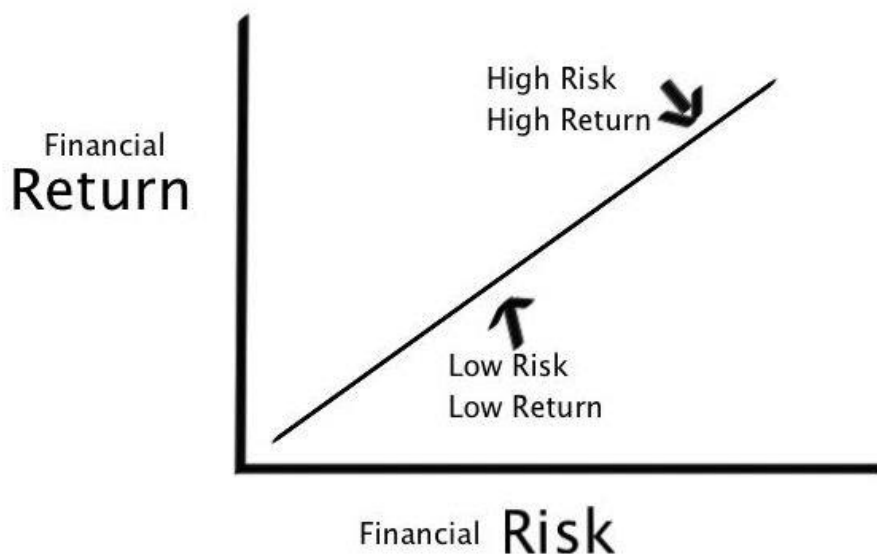
Figure 1: Financial risk vs Financial return graph

Team - 50 Cent

1.2.1.2 Example scenario of problem

Most of the supply chains are based on a post-paid method in which they supply the product first and receive money after the product reaches the ground level of the supply chain. If we investigate the journey of a particular product from its manufacture to consumer there can be many levels because manufactures cannot reach their consumers directly.
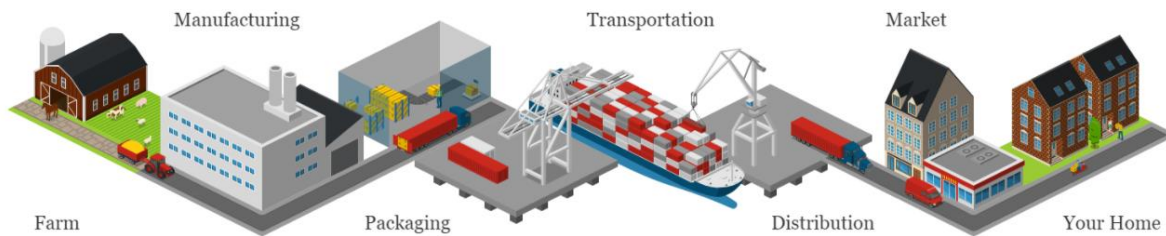
Figure 2: Supply chain example

Usually what happens is that manufacturers issue their products to middle entities which can deliver their product to actual consumers. This middleman of a supply chain can consist of many levels. After the product reaches the ground level of the supply-chain (actual consumer) middle entity will get paid and then only the manufactures get paid. It is more like a chain system. In a complex supply chain, there can be many parties involving this process. If the chain breaks, money flow will also break which leads the company or business to have bad debts. This bad debt problem is a common problem for every scale of a company whether they sell a product or a service.

Assets

ITEM	2016	2017
Cash & Short Term Investments	67.16B	74.18B
Cash & Short Term Investments Growth	-	10.46%
Cash Only	12.27B	14.52B
Short-Term Investments	54.89B	59.67B
Cash & ST Investments / Total Assets	20.88%	19.76%
Total Accounts Receivable	29.3B	35.67B
Total Accounts Receivable Growth	-	21.76%
Accounts Receivables, Net	15.75B	17.87B
Accounts Receivables, Gross	15.81B	17.93B
Bad Debt/Doubtful Accounts	(53M)	(58M)
Other Receivable	13.55B	17.8B
Accounts Receivable Turnover	7.31	6.41
Inventories	2.13B	4.86B

Figure 3: Bad debt example

Team - 50 Cent

As an example, above figure is abstracted from the financial states of APPL.inc which is a giant in the tech industry. They have nearly $ 53M bad debts yearly.

1.2.2 Problem definition- Define the problem

Credit is the ability to borrow money or access goods or services with the understanding that you will pay later. Simply it says buy now pay later. Lenders, merchants (Creditors) grant credit based on their belief that you can be trusted to pay back what you borrowed.
To get that creditors consider you worthy of their trust, Debtors are being addressed creditworthy, or to have "good credit." so by this nowadays and from the past decade's credibility is measured using trust.

In accounting, a credit is an entry recording a sum that has been received. Traditionally, credits appear on the right-hand side of the column with debits on the left. And every year a provision will be taken as an expenditure for the debt which cannot be collected. By this kind of debt company financially falls small amounts do not matter much but when it comes to large amounts of debt companies will face huge losses due to this type of debt.

There are small, medium, and large-scale companies but registered companies mean PLC companies show their financial report to the public because they collect capital from public shares. "The biggest advantage of forming a public limited company (PLC) is that it grants the ability to raise capital by issuing public shares" and, we can see that public limited companies do transactions with other companies and these transactions are also based on credit because of the huge amount of stock purchase.

1.2.3 Research questions

1. Default can be defined in multiple ways, considering the available data and how strict we want our model to be. Do we consider it a default if the client is only one day late on payments? Or do we let them be 30, 45, 60 days late before acting?
2. Are we going to consider a client defaulted if he owes 10 rupees? What if the client did not pay one bill but has been paying all the bills after?

We are required to make a judgment call and choose a definition that meets the creditor's needs best.

1.3 Aims and Objectives

1.3.1 Research Aim

The aim of this research is to design a web application to reduce the risk of bad debts. For small and medium companies which mainly rely on credit sales. To further elaborate on the aim, the proposed application uses the consumer's past transactions and their financial information to propose a predicted credit limit whereas the company can stick to the limit to reduce risk of selling more than which the customer would pay back. Also, additional features such as smart contracts reduce the amount of paperwork used for a transaction and to reduce the risk of them being altered.

1.4 Project Scope

1.4.1 In Scope

The main features that will be covered in the project are mentioned below:

- A web application that predicts credit affordability of business enterprises: based on past audited financial statements of customers, an estimated credit limit will be predicted. Therefore, traders can quantify the credit sales which can be offered to the customer.

Team - 50 Cent

- Providing information about customers' risk factor: The application has the ability to analyze customers' previous financial statistics and can provide information about whether they have enough assets to pay the debts or is it risky to have credit sales with them.

1.4.2 Out Scope

The following parts will not be covered in the project.
- The Application is mainly focused on business enterprises which share their financial details. Due to the above reason the application will not predict credit affordability of enterprises who are not willing to share their financial statements.
- The Application is mainly based on past financial statements. Therefore, the same will not cater to identify any unforeseen drawbacks of the companies which can happen due to the economy or market changes.
- Application will not provide any credit risk information of companies which have not been disclosed in their financial statements.

1.5 Objectives

1.5.1 Research Objectives

The main objective of CREDIT-MART system is to reduce the amount of bad depts and to minimize the risk factor which take place during a credit transaction.

The research objectives which were considered while undergoing the project are –

- To evaluate the possibility of using machine learning techniques to detect the bad debts.
- To evaluate the possibility of using machine learning techniques to detect and analyze the risk factor which arises during a credit transaction between two private limited companies.
- To identify new legible and advanced methods that can be used in the Machine Learning process.

1.5.2 Academic Objectives

- Improving time management skills and project management skills
- understanding about various approaches used in Machine learning and Deep learning and clustering techniques
- improving the critical thinking and improve the problem-solving capability
- Applying the theoretical principles to play
- Applying the knowledge and practice gained from modules in the degree program
- Getting an understanding on writing and publishing research paper
- Learning the ways of applying concept to real world problems

1.5.3 Operational Objectives

Objective	Objective Description
Objective 01 -Purpose	- Discuss the draw backs of credit sales with businesses who are involving with credit basis transactions. - Carry out an in-depth research about classification models in ML and how to apply them in financial sides.

Team - 50 Cent

Objective 02- Data Gathering	• Searching financial data of companies online.
Objective 3- Study and review existing work	• Evaluation an existing model that can use to analyses financial data of companies. • Performing a literature survey by evaluating pervious research in domain of PD models in the area of linear regression. • Identifying technologies that can be used to select appropriate technologies and algorithms. • Submit Literature Review document.
Objective 4- Design	• Gathering data from businesses about the nature of a supply chain and identifying functional and non-functional requirements. • Researching and selecting most suitable techniques and tools for project. • Preparing final SRS document and submit it.
Objective 5- Implementation	• Researching different techniques of software development methodologies and find the best suited model for the project. • Selecting testing methodologies for the system • Preparing the project management document and submitting it.
Objective 7- Evaluation	• Carefully analyzing the social ethical and legal background of are of project domain and identifying the legal, social, and ethical issues and discussing how to overcome them. • Finalizing the conclusion chapter and submit the full report.

Table 1: Operational objectives

Team - 50 Cent

1.6 Proposed Solution

1.6.1 Rich picture diagram
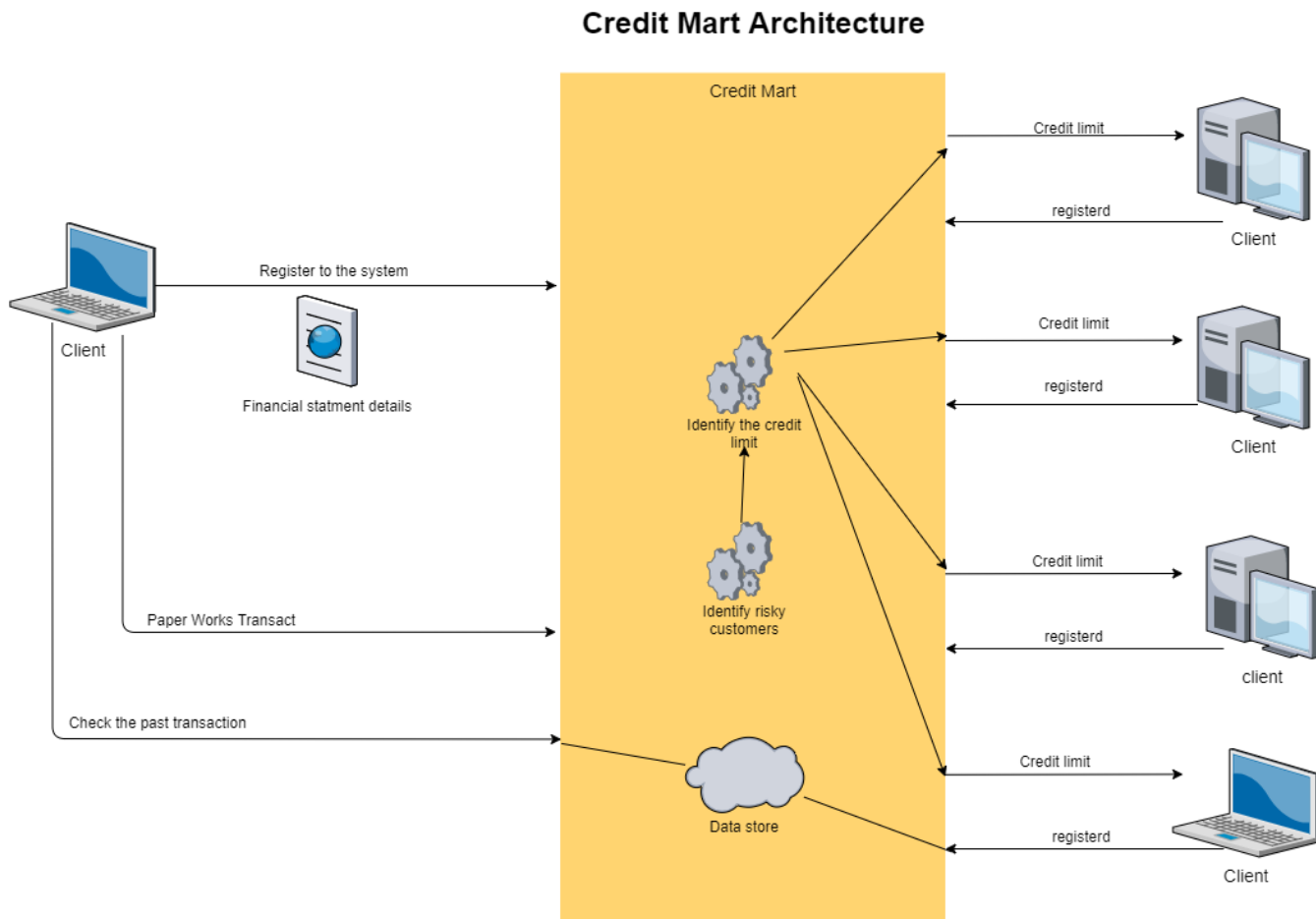
Figure 4: Rich picture diagram

1.6.2 Features of the prototype

- Predict the credit limit
  Based on the past financial statements undergone by the buyers or customers, an estimated credit limit will be predicted. Based on this, the buyer can decide on how far he could make his credit purchases in.
- Identify the risk factor of the customer
  Able to calculate if the customer is liable to repay debts according to his or her financial status. This is done by viewing the financial statements and records of his business. This feature will give out a prediction saying if the customer is good to go with the transaction or is it a risk for the sellers to have business and trade relationships with the customer.

1.7 Resource Requirements

1.7.1 Hardware requirements

Hardware Requirements	
Processor	**Intel Core I7 10700U CPU 2.40GHZ A processor** should be required for data processing purposes
RAM	**8GB RAM A RAM** is needed by most of the processors to perform efficient results with faster access memory.
USB Pen drive	**500gb portable hard drive** to store the source code and daily updates affiliated with project
Disk space	**50gb or above** to store the source code and testing results

Table 2: Hardware requirements description

1.7.2 Software requirements

Software Requirement	Description
Developing and Training model	To train the model using Python language (CNN). **Anaconda Navigator (Jupyter notebook).**
Source code Version Control and Backup	To keep back up on the source code. **GitHub**
Web application development tools	To build a web-based application. **Visual Studio Code / Pycharm**

Team - 50 Cent

Storing daily updates	To store the daily updates and project related documentations. **Google Drive**
Diagramming tools	Designing Use cases, Activity diagrams and sequence diagrams. **Star UML / Figma**
Reference Management tools	To keep journal articles, books and other project related references organized and auto formatting citations. **Zotero.org, Mybib.com**
Communication	To keep team communication within the project period. **Google meet, Discord**
Project management	To divide group work between group members and project management. **Trello**

Table 3: Software requirement description

1.7.3 Technology stack

This is the list of all the technologies that is needed to build the application. The team will use React to build the front end and Python to build our back end.

Users → Front End (React) → Back End (Python) → Data Set

1.8 Chapter Summary

The chapter involved in giving a clear idea on the system and the background overview of the idea. It was clearly mentioned as why this topic was selected as the main domain. A clear scope for the aims and objectives meant to be achieved was discussed and finally the technological feasibility for the system was briefly mentioned. The next chapter will discuss on the previous research that has been done towards this idea.

2. Literature Review

2.1 Chapter Overview

The previous chapter gave an overview idea of the credit mart system. This chapter discusses deeper into the problem domain, previous applications that were implemented to solve this problem, previous research that were related to this problem domain. The algorithms, tools and the special techniques that will be used in the Credit Mart system. The chapter provides a better proposed solution compared to other previously implemented programs to overcome the potential threat on the increasing bad debts in businesses.

2.2 Similar products and technologies

2.2.1 Features and its comparison

When businesses are done on credit basis, there are many threats and defaults which arise at the time of debt collections. This depends on perspective of the customer who is the buyer and the party involved in selling the goods. The main target audience of our system or application is to reduce the risk factor that would take place during a credit transaction. The 'Credit-Mart' application is focused with immense technology and key features providing ethical and legal solutions for the problems which arise during a credit business transaction.

- **Identify credit limits**

    Each customer is allowed to have a certain amount of credit depending on his or her companies past financial statements. There are some online tools which calculates credit limits when inserted the relevant data.
    o My DSO (days sales outstanding) Manager
        (What is the DSO report? 2020)
        This is an online software or tool which manages all sorts of receivables. It has a direct effect on cash flow of the business. It also maintains records of the number of days of sales invoiced and not paid yet. 'CREDIT TOOLS' is sub tool of My DSO Manager which calculates the credit limit accordingly.
        CREDIT TOOLS –
        (Credit Management tools and resources for SMEs and international groups, no date)
    o CGI (Credit Guru Inc)
        (How to Set Credit Limits - Credit Management Article -Credit Guru, no date)
        This site shows the methods and of calculating the credit limit for a credit company. It explains the risk factors and methods to be followed when giving goods on credit. It also shows issues and related matters to be consider when handling credit transactions.

- **Store Financial Statements**

    'CREDIT-MART' application identifies credit limits for the customer according to the financial statements provided by the customer. There are some applications which maintain the financial statements. 'CREDIT-MART' only stores the financial statements provided by the customer.
    The below application maintains the personal as well as company financial statements too. It links bank account, credit and debit cards and calculate the financial statement. This automatically updates each time a transaction is done.
    o FreshBooks
    o QuickBooks Online
    o Wave

o Xero
o Zoho Books

These are the top-rated accounting apps to track business expenses

Top-Rated Accounting Apps to Track Business Expenses

	FreshBooks	QuickBooks Online	Wave	Xero	Zoho Books
Overall rating (score out of 5)	4.5	4.4	4.6	4.3	4.4
Business expense tracking capabilities	Reconcile bank accounts, store expense receipts, organize and track expenses, assign recurring expenses.	Reconcile bank account, develop custom rules for expense categorization, create business reports, save expenses.	Reconcile bank and credit card accounts, scan and submit expense receipts, sync receipts on the app.	Reconcile expenses, get real-time reports, capture expenses, assign and label expenses.	Track, categorize, and bill expenses; automate expense generation; attach bills and credit notes to expenses; create reports.

GetApp

Figure 5: Top rated accounting apps

- **Calculate and show PD (Probability of Default)**

    Probability of Default is a result gained by predicting whether a customer could repay his dept or not within a period of time. It generates if the customer is reliable to fulfil his obligation on debts. (Herga et al., 2016)
    o CRIB (Credit Information Bureau of Sri Lanka) □
        This system is carried out by banks in-order to lend loan facilities to their customer.

- **Identify risk factor of customer**
    o Differentiate Good and Bad customer
        Good and bad customer are differentiated based on their dept repayments according to this project. PD model is used to in here in-order to calculate and differentiate if the customer is good or bad. There is no particular software or application which performs this action.

- **Sellers can identify their buyer**
    Sellers meeting the buyers could be seen in various E-commerce platforms. In 'CreditMart', the clients can identify their respective customer according to the PD and various other key features which would be available for their usage.

2.2.2 Feature comparison table

Similar Products Features	My DSO	FreshBooks / Zoho Books	CRIB	CREDIT MART
Identify credit limits	❌			❌
Store Financial Statements		❌		❌
Calculate and show PD (Probability of Default)			❌	❌
Identify risk factor of customer				❌
Sellers can identify their buyer				❌

Table 4: Feature Comparison Chart

2.3 Research Gap

Credit transactions are commonly done in business at any range from raw material purchases to final product sales. The creditor uses three types of approaches to collect and manage these so-called bad debts. The holistic approach, hard business approach and a one-size-fits-all approach.

- We mostly consider the Holistic approach when collecting the debt arrears. The primary objective is to maintain a good customer relationship in this approach the creditor uses the court to collect the debt.
- Hard businesses approaching the debt collections are widely considered to have a minimum cost for the debt recovery process. Customer relationship is not considered in this approach.
- In the One-size-fits-all approach, the creditors follow a standard set of procedure. In this approach we use the court system for the legal approach. One-size fits- all approaches tend to be drawn from more limited types of creditors. (Dominy and Kempson, 2003)

The approach which is focused to be carried out in Credit Mart would be a combination of all three approaches simplified by giving a credit limit and their credibility.

Team - 50 Cent

Credit liability in commercial bank loans is determined by using a machine learning approach along with logistic regression techniques. This process uses indicators such as age, number of children, previous loan amount and many more. This is carried out by identifying the common patterns of customer behavior individually (Dominy and Kempson, 2003). Credit Mart application and its related research will be mainly focusing on the Public limited companies and to find their credibility.

In the modern business world, machine learning is a very familiar technology. Although it is very much familiar, still in some key areas of business which helps management to make decisions are using different primitive methods. Analyzing credit risk of their customer in one such method according to research. Creditors carry the risk of their clients by not meeting their debt obligations. These kinds of events are often referred to as *default events*. This can be modeled for each company through a probability of default (PD) (Herger et al. 2016).

Finding the risk of a customer in banks is done through calculating PD – probability of default using supervised learning. The most common method they use is logistic regression (Yuan, 2015) but using random forest is more efficient to find the weighting factors of the PD model.

The Altman Z-score is a widely used credit scoring model. It is also said that the Altman method is very much possible to show the development of an internal rating model. This process was based on a limited number of financial ratios. This provides the credit decision with an immediate response. (Bazzi and Hasna, 2015).

According to several research and final outcomes, it has been shown that Altman Z score is not suitable to use in the context of small and medium scale enterprises (based on income) (R. C. Merton. 1947). Due to the complexity of these methods, most of the companies and even financial institutions do not use a sophisticated credit scoring model for their credit holders. Even the Sri Lanka bank institute uses CRIB (credit information bureau of Sri Lanka) to measure their customers' risk factors. In credit information systems they only gather information about liability details about most of the bank customers. Even though it is useful for financial institutes, it is hard to find a movement where businesses use these kinds of systems.

Therefore, an application that can be used by small and medium scale enterprises, can be considered as a useful approach to improve the health of the business by identifying their credit customers and analyzing their condition. This would enhance the profitability of the company and it would open new opportunities to deal with more and new suppliers with trustful nature

2.4 Research on Approaches and Techniques used in related domains
2.4.1 Credit Model

A credit model is a mathematical approach used to predict the probability of default (PD). The probability of default is a financial term which describes the likelihood of a debtor to repay the debt. Many financial institutions use various methods to calculate PD. The Altman Z-score (Altman, 1998) is a widely used credit-scoring model. It uses several financial indicators to predict a company's degree of PD.

Accurate predictions of scoring models help minimize the risk of bad debts. However, consumer behavior can change rapidly during different financial situations and season times. Therefore, the model must also be able to adapt to financial cycles and rapid changes the consumer world. Techniques used to create such models include:

- Logistic regression and linear regression.
- Machine learning and predictive analytics.
- Binning algorithm (i.e., monotone, equal frequency, and equal width).
- Random Forest and Decision Trees.

Above mentioned methods all have its pros and cons. Therefore, it is crucial to use a model which will work best for small and medium scale companies whereas that is the target audience in the credit mart system. The expected result from the credit scoring model is to rate the consumer according to their financial information and providing a credit limit where the creditor has the ability to lend items to the borrower based on this credit limit to reduce the risk of bad debts.

2.4.2 Probability of Default

As explained in previous chapters *probability of default* is a widely used term in case of credit risk analysis. In case of credit sales and loans there can be various factors involved such as time duration, loan amount etc. According to (Herga, et al., 2016) there are many ways to calculate a companies' degree of PD.

- Altman Z-score
- Merton's bond pricing model
- Predicting PD using artificial neural networks

Altman Z-score

Altman Z-score is calculated by a linear combination of five mostly used financial indicators. Initial formula of Altman Z-score as follows (Altman, 1968).

$Z = 1.2\ X_1 + 1.4\ X_2 + 3.3\ X_3 + 0.6\ X_4 + 1.0\ X_5$

X_1 = working capital / total assets

X_2 = retained earnings / total assets

X_3 = earnings before interest and taxes / total assets

X_4 = market value of equity / book value of total liabilities

X_5 = sales / total assets

But in many cases Altman Z-score is not suitable to calculate probability of default in the context of small and medium scale companies (D. Michala, et al., 2013).

KMV-Merton's bond pricing model

KMV-Merton model is based on equity approach introduced by Merthon (S. T. Bharath & T. Shumway, 2015).

But in some research, it is shown that merton model does not produce enough statistics for probability of default.

2.4.3 Non-machine learning Techniques used in related domains

Historically, businesses and banks have always faced problems with customers who are unwilling to pay their debts. Collecting debt can be a very traumatic experience for any business. Due to that reason, there is a significant amount of research work done for identifying customer risk factors, rating modules and credit risk management systems.

Prior to financial sector deregulation, banks were highly motivated to grant credit facility to clients who could easily express their wellness. Therefore, a high number of bad debts affected the commercial banks. Afterwards Credit risk management processes enforce the banks to establish a clear process for approving new credit as well as for the extension to existing credit. (Altman, 1998). In 2000, (Basel, 2000) announced several credit risk management principles that can be used in the banking sector. The document addresses establishing an appropriate credit risk environment and maintaining an appropriate credit administration system within the bank. Also, the

author says credit risk management systems may differ from bank to bank depending on their nature and complexity of credit activities. (Basel, 2000).

(Natalija Konovalova, 2016) conduct research about factors that can be affected to credit risk management such as average income of the borrower, the loan amount, and the loan term. In that research author defines a static method to identify common patterns of bank customers' economic behavior and formulate a set of differentiated requirements for borrowers in accordance with their economic behavior.

(Hasna, 2015) conduct research about static methods that predict credit limits of bank customers. The elements that author considered are:

- The company's margin is defined under the analytical accounting.
- The probability of default of the borrower.
- The forecast turnover to achieve by the borrower in a given time.

The author defines a formula to predict the credit limit of borrowers using the above elements. Also, the author recommends applying this method on customers that are reliable to the bank (Hasna, 2015).

2.4.4 Machine Learning approach

In the credit mart system, there are two main models. One is PD model, as explained in previous chapters probability of default shows the probability that a company might face financial distress in a given period of time.

Calculating the PD of a company can be done in various methods. The Altman Z-score which is a widely used method (Herga, et al., 2016). But results have shown this method does not give an accurate PD for small and medium scale enterprises. They show that artificial neural network produces the best result (I.-C. & C.-h., 2009).

2.4.5 Supervised Learning

Supervised learning is a process which operates under a specific supervision. Supervised learning techniques can be further grouped into regression and classification techniques.

Linear Regression – The Linear regression is one of the oldest supervised technique that is mainly used for predictions, forecasting and to find relationships "between quantitative data". (Rouse, n.d.)

Classification – The classification techniques is another supervised learning method which has the potential to analyze data while recognizing patterns in order to produce a qualitative response. Below are some of the widely used classification algorithms (Rouse, n.d.),

- · Support vector machines
- · Neural Networks
- · Logistic Regression
- · K-nearest neighbors
- · Decision Trees
- · Random Forest
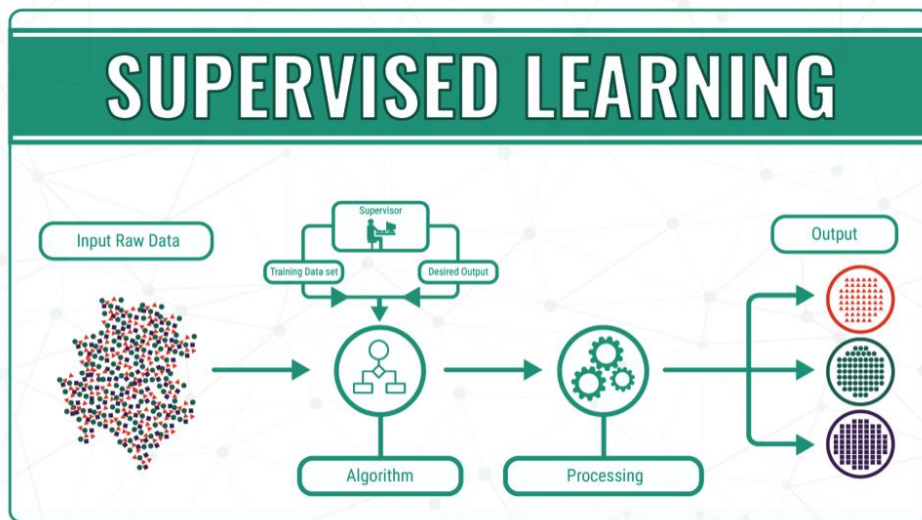
Figure 6: Supervised Learning

2.4.5.1 Choosing Classification Algorithms

Algorithm	Advantages	Limitations
Logistic Regression	● Easy to implement, interpret and very efficient to train.	● It constructs linear boundaries.
K – Nearest Neighbors	● This algorithm is faster than other classification algorithms. ● KNN is very easy to implement.	● Does not work well with large datasets. ● Need feature scaling. ● Does not work well with missing values and noisy data.
Support Vector Machine (SVM)	● SVM work well when there is a clear margin between classes.	● SVM algorithm is not suitable for large datasets. ● Does not perform well when the target classes are overlapping.
Random Forest	● Random forest can automatically handle missing values. ● Random forest can be used to solve both classification and regression problems.	● Takes long time to train. ● Since it creates lot of trees the algorithm is complex

Table 5: classification of algorithms

Team - 50 Cent

In the data set collected, there are 200 financial indicators of PLC companies which are available in the stock market. The first task is to classify the companies available to good or bad companies by the health of the company. There are multiple algorithms available for classifying tasks such as Logistic regression, K- nearest neighbors, support vector machine and Random forest. In the below figure we can identify the advantages and limitations of the following algorithms.

When considering the facts, using <u>logistic regression and random forest</u> are the most suitable algorithms for the expected model.

2.4.5.2 Indicators selection.

Indicators to calculate the probability of default covers the companies' liquidity, solvency, profitability, and economic structure. Also, includes two or three common indictors which are common to all the companies such as GDP etc.

Current assets / Current Liabilities	Liquidity
Working capital / Total assets	
Need Working capital / Working capital	
Current Liabilities / Total assets	
Cash / Current Liability	
Working capital / Current Liabilities	
(Current Assets - Stocks)/ Current Liability	
Current assets / Current Liabilities	
Net Income / Total Assets	Profitability
Net Income / Equity (ROE)	
Net Income / Turnover	
EBIT / Total Assets	
Total debt / Total equity	Solvency
Total debt / Total assets	
Operating income (or EBIT) / Interest expense	
Annual average balance	Behavior
Annual credit balance	
Annual debit balance	

Table 6: Indicator selection

2.4.5.3 Weight of Evidence

Using the dataset, indicators that covers aspects such as solvency profitability and liquidity of a company can be extracted. Then each indicator value can be taken and used to create *n* number of value ranges from min to max

for each indicator. Then it can be assigned each company to the above ranges and calculate number of good and bad companies in each range. Then the weight of evidence can calculate by below equation (Herga, et al., 2016).

$$WOE = log\frac{P(companies = Good)}{p(companies = Bad)}$$

Finally, the probability of default can be calculated using a logistic distribution function.

F(x) = 1 / 1 + exp−(β0+β1·woe(x1)+β2·woe(x2)+...+βn·woe(xn))

2.4.6 Credit Limit Model

The above models and methodologies give an explanation as of how to identify a good or bad customer by calculating their PD. The common question which arises after finding the status of the customer is how to manage them when doing a transaction with that party. The answer is to allocate a credit limit. A limit for the products a consumer can purchase in credit or borrow.

There are two points to consider when setting a credit limit.

1. If the limit is too high the customer may not be able to pay resulting in bad debts.
2. The limit being too low whereas the company faces a loss in terms of sale reduction.

The model which will be used to overcome the above problem is inspired by P.Krokhmal et al. which uses some standard financial terms. Value at risk (VaR) and Conditional Value at Risk (CVaR), as an example by taking VaR (95) = 3% this means that there is a 5% (100% - 95%) chance to lose 3% or more from a given transaction. If CVaR (95) = 4% we can consider that in 5% (100% - 95%) of worst returns, the average loss will be 4%. (P. Krokhmal, 2002)
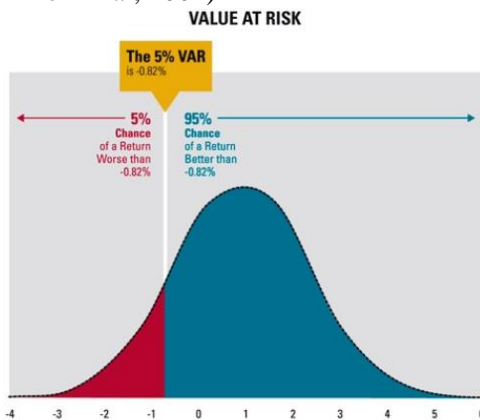
As a basic idea it can concluded that VaR gives a range of potential losses and CVaR gives an average expected loss.

Moving on with the above idea and equations provided in (Herga et al., 2016) the credit limit can be set by finding upper and lower bounds by using the equations.

Figure 7: Value at risk visualization

2.5 The final selected approach

After performing an in-depth Literature review on the approaches and techniques used in related domains it was clear to move on with logistic regression to calculate the PD and the use of Decision trees to classify each business on their credibility.

2.5.1 Logistic Regression

Logistic regression is the classic model for binary classifications. As it is easy to implement and efficiently train, it is the ideal method to use in calculation of PD.

Logistic regression aims to model the probability of an event occurring on the values of the independent variables. In this case the independent variables being the indicators which was explained in the 'indicators selection' section above. These independent variables are used as predictor variables to help model the probability of default.

It predicts the outcome of an occurrence by fitting data into a logical function which in our case has been explained in the 'Weight of evidence' section and provides the result. The expected final result will be a binary value (1 or 0) that is, 1 as in the borrower is default or 0 where the borrower is not in default. Therefore, we can conclude that logistic regression is a supervised classification algorithm used to predict a binary outcome.

2.5.2 Decision Trees

Decision trees can be used for binary classification challenges with supervised machine learning. Using this method, predictors can be set and matched to arrive at some outcome.

There are few steps to look forward to when creating a decision tree to predict whether a company is a healthy company or a non-healthy company. First, create three predictors as *Credit balance*, which records the credit balance if it is high, low or zero and *Debit balance*, which records debit balance as high or low and finally **Probability of Default**, which records as zero and one. It also creates a 'One outcome variable' called *Business type*. This is a binary classification problem. There are only two possibilities whether the business is healthy or unhealthy.

As the next step, this creates some training data based on some data collected about companies.

Predictors			Outcome
Credit balance	**Debit balance**	**Probability of Default**	**Business type**
High	Low	0	Non-Healthy
High	Low	1	Non-Healthy
Low	High	0	Healthy
High	High	0	Non-Healthy
High	Low	0	Non-Healthy
Zero	Low	1	Non-Healthy
High	Low	1	Non-Healthy
Zero	High	0	Healthy
Zero	High	0	Healthy
High	Low	0	Non-Healthy
High	Low	0	Non-Healthy
High	Low	1	Non-Healthy

Table 7: Decision tree prediction

The first row of data in the training set shows that the company is with high credit balance, low debit balance and PD is zero. Using decision trees, it predicts that the business type is unhealthy. If a company is with zero credit balance, high debit balance and if the PD is zero, then the business type is healthy. These combinations of predictors will create a different outcome.

To create a decision tree, the data set like the above mentioned should be taken and it should be split based on the predictors. The decision tree can be generated by any of the three above predictors.

As an example, *Credit balance* is chosen as the predictor for the root node. The root node should be chosen at the very trunk of the tree. Any other predictors can be chosen accordingly.

The next level of the tree after it has been split out, is called the decision node. Here, it displays all the options in the root node. So, in this case, there are three options: high, low and zero. At this level, it records the first set of outcomes. At this level outcome should be binary. This process is taken under in-order to check if the business is healthy or unhealthy.
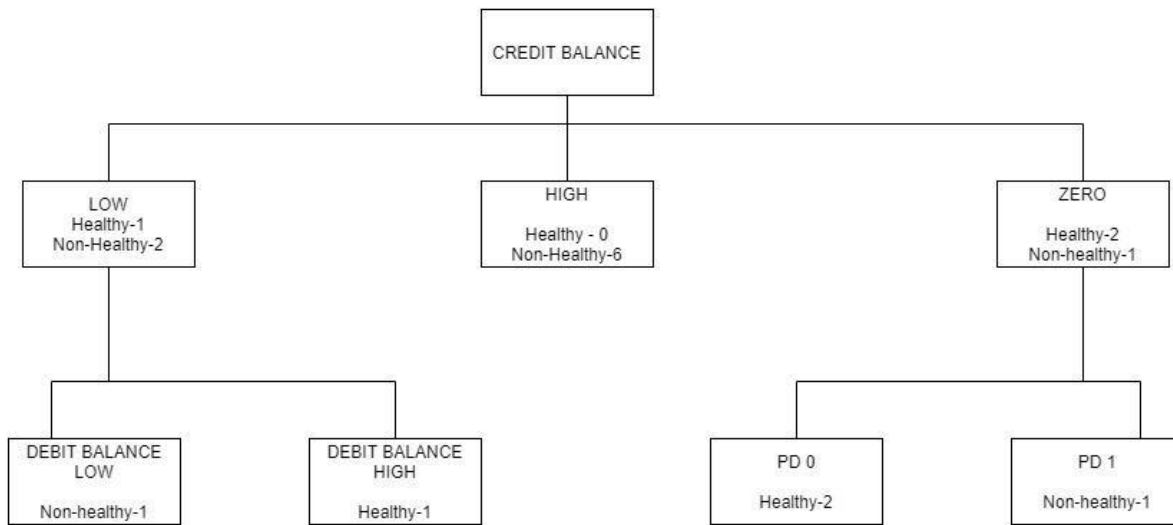
Figure 8: Decision tree example

In each of these decision nodes, it displays the total number of possible outcomes from the training data. There are three scenarios where companies can be featured as healthy and nine scenarios of it can be labelled as unhealthy companies. These different outcomes will let us break the decision node down into several different leaf nodes. This is the second set of predictors which show possible outcomes at the decision nodes. There are no leaf nodes when the decision node is having a high credit balance. This specifies that there are no outcomes with this combination. Since there are no other combinations, there is no need of breaking down the decision tree any further. When creating decision trees, it is compulsory to have a clear path to a yes or no outcome.

2.6 Chapter summary

In this literature review multiple PD modeling methods and implementations were covered. This chapter analyses the various types of approaches to collect and manage bad debts and how those approaches are applicable for the use of credit mart. How to identify credit limits using credit mart, how to store financial statements using credit mart and comparing features with some previously implemented similar products to credit mart are discussed deeply in this chapter. Before the discussion of the machine learning approach in this chapter, non-machine learning techniques which are based on the related domain are mainly focused and discussed. To analyze the credit risk, the probability of default has been used. Under the probability of defaults multiple algorithms such as 'logistic regression', 'K – nearest neighbors', 'support vector machines' and 'Random forest' are discussed briefly. To predict the probability of default, the credit model is further discussed in this chapter. Credit models are used to minimize the credit risk and it depicts the risk factor of the client or customer. The next chapter will discuss the project management phase of the project and will detail the methodologies used in this project.

3. Project management

3.1 Chapter Overview

This chapter will mainly focus on the life cycle of the project implementation, the methodologies in how project can be developed and the suitable method which can be used in starting to develop this project. Additionally, the chapter will focus on the project plan, possible risk and the mitigations for such risks will be discussed.

3.2 Methodologies

3.2.1 Research approach

Research is basically done in-order to overcome any sort of issues or problems which would arise at the time of conducting the project. There are some categories which research approaches are conducted.

1. Inductive approach
   - In this process, the researchers are focused with the generation of new theory. This approach starts with observation. As a result of the observations, the theories are generated towards the end. The researcher is free to change the nature of the research after the beginning of the development process. (Dudovskiy, 2011)
2. Deductive approach
   - In this process, the research is based on an existing theory. It is mostly used in proving and testing an existing hypothesis. By then, the research strategy is designed in-order to test the problem and hypothesis. *"Deductive means reasoning from the particular to the general."* (Dudovskiy, 2010)

In general, a deductive approach starts with a hypothesis or a problem, whereas an inductive approach typically uses research questions to limit the study's reach (Gabriel, 2013). (Dudovskiy, 2010) says, "Deduction begins with an expected pattern "that is tested against observations, whereas induction begins with observations and seeks to find a pattern within them.""

CREDIT-MART is typically a new theory, where the credit limit is analyzed to start-up a transaction and recommends if its wise or not to continue with the current selected transaction. This is purely done using the financial data provided when the customer registers into the application. Based on the above said and verified explanation, CREDIT-MART is applicable to follow Inductive approach.

3.2.2 Process Model

Credit mart is a research base project. Some of key areas of credit mart consist of requirements that need to be changed based on new findings. So that in summary we can conclude that a credit mart project needs to carry out based on a prototype and should develop the requirements on iterations based on the stakeholders' feedback and new findings.

Iterative Model

The idea behind the iterative model is to start with a simple implementation of the software and iteratively improve the developing adaptations until the full framework is actualized. In each iteration the system is developed, and modifications are made in each stage.

Iterative models are mostly used if the requirements of the complete system are clearly defined and understood.
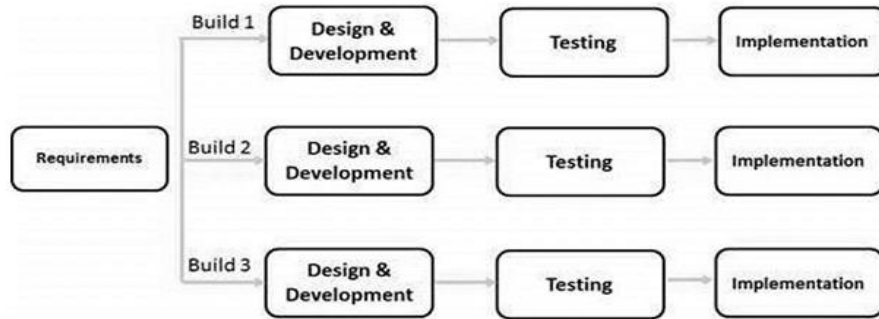
Figure 9: Iterative model

Agile Model

In the agile model every project needs to be handled differently and the existing method needs to be adjusted to best suit the project requirements. Agile model follows an iterative approach and working prototypes are built after each iteration. In each build, a prototype is developed with requirements and the final build will contain all the features required by the customer.

Agile method uses an adaptive approach, where there is no detailed planning and clarity on future tasks. Also, in agile there is feature driven development and the credit mart team can adapt to the changing product requirements dynamically.
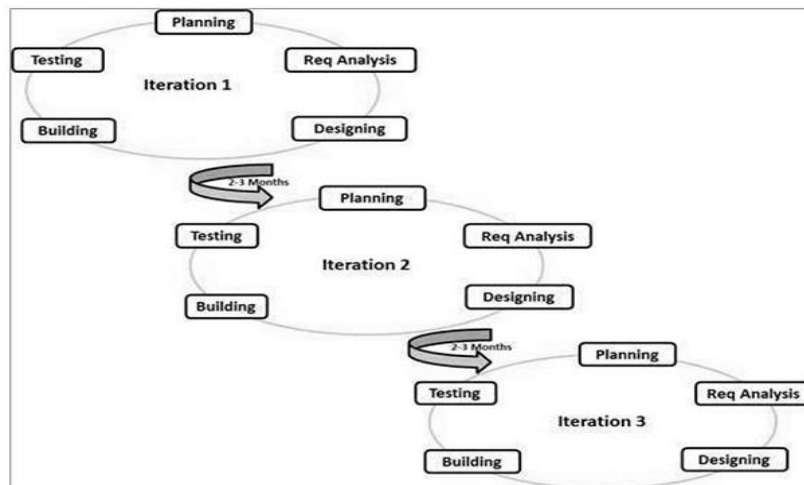
Figure 10: Agile model

RAD Model

Rapid application development is a software development model that produces a prototype that contains features that are functionally equivalent to a component of the product. In RAD models, the functional models are developed in parallel as prototypes.

RAD models can be applied effectively to the ventures where clear modularization is conceivable. In the event that task cannot be broken into modules RAD may fail. Also, RAD models should not be chosen if domain experts are not available with relevant business knowledge.

Team - 50 Cent
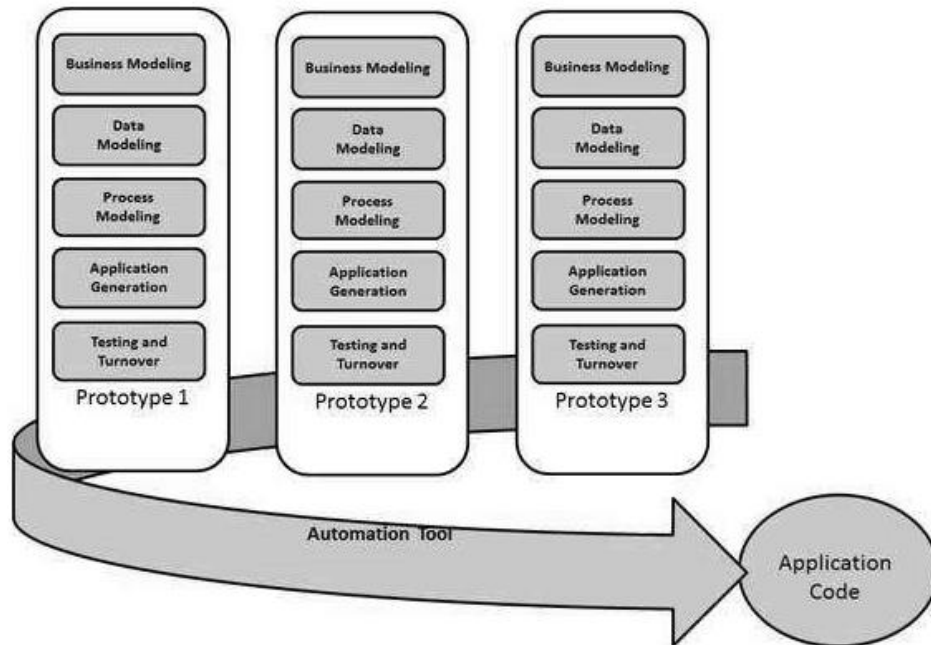
Figure 11: RAD model

According to the above points Best suited model for credit mart project is Agile model due to the above-mentioned facts.

3.2.3 Analysis and design approach

After conducting an analysis of the two foremost Structured System analysis design approaches in between Structured system analysis and design approach (SSAD) and object-oriented analysis design approach (OOAD), the object-oriented analysis design approach (OOAD) was selected as the most suitable design methodology for the project by considering the following reasons.

- The System does contain classes as required by (OOAD) design methodology.
- This approach decreases the complication of the system and increases the readability of the code.

3.2.4 Testing methodology

Testing is a process where the system is checked if it is working according to the required manner. The main aim of testing is to verify if the system is working accurately as needed. It is very much likely for bugs and errors to arise while working on the project. This is the main reason for testing. When a testing methodology is followed, the errors could be identified and fixed thereby.

There are some testing methodologies which can be followed in-order to overcome such possible errors.

1. Functional Testing
   Functional testing requires testing the system against the needs of financial institution.
   - Unit testing
   - Integration testing
   - System testing

- Acceptance testing
2. <u>Non-Functional Testing</u>
   Non-functional testing requires all the types based on the operational concepts of a software program.
   - Quality testing
   - Performance testing
   - Security testing

(Software Testing Methodologies, 2019)

The testing methodology which are supposed to be followed while developing the system for CREDIT-MART are specified below.

<u>During Development</u>

- Unit testing
- Integration testing

<u>After Development</u>

- Quality testing
- Performance testing
- Security testing

3.2.5 Project management method

Among many project management methods Agile is selected to handle the credit mart project.
Credit mart project is a research base project. Agile model is a flexible method. Which means it gives the change to experiment and make incremental changes to the project.

3.2.6 Data Gathering Methods

Gathering data is a necessary requirement in-order to conduct project in a well-mannered and structured way. By gathering necessary data, it would be helpful to overcome all sorts of research questions and problems that arise during the implementation process too.

1. Observation
   - Analyzing the data and projects on previously made models.
2. Interviews
   - Interviews are conducted in-order to gain ideas from domain experts and relevant people who work in the same domain as CREDIT-MART
3. Questionnaires
   - Questionnaires are sent in-order to get the ideas and overview about the businessman on the project CREDIT-MART
4. Brainstorming
   - Brainstorming was an implicit requirement which was very useful to come to a conclusion. This process was helpful to discuss the technical aspects of the project.
5. Proto typing
   - In order to identify how the system design can be changed, prototypes will be built to get ideas from target audience.

3.3 Risks and mitigations

The possible risk in involved with the Project and mitigation instances

NO	Description	Probability	Impact	Mitigation action
N1	Lack of knowledge in technologies and techniques and the System domain	High	High	Getting advice from experts and constant and doing a background study on the businesses and their way of transaction self-studying on the related research domain.
N2	Scope change when adding new requirements to the system and contacting domain experts.	Medium	High	Prioritize the requirements. Work for the Given scope if any changes are to be made first completing the main component and then making other changes.
N3	Hardware/software product failures	Medium	High	Make a cloud backup using Google drive and git hub and commit our work on a daily basis to avoid data loss.
N4	Missing deadlines due to complexity of the project, Unable to complete the task	High	High	Prioritize the tasks and meet the given deadline by the leader. Finishing the main task first
N5	Problems with implementation	High	High	Important the requirements and implement what is most important first. Look for alternatives if some parts are implemented
N6	Getting sick	Low	High	Planning the work ahead

Table 8: Risk and mitigation

3.4 Activity Schedule

The activity schedule along with the project deliverables are detailed in table below.

Date	Activity	Time Frame
29/09/2020	Finalizing the project idea	1 Week
24/10/2020	Submission of the project proposal	2 Weeks
30/11/2020	Submission of the LR	1 Week
20/12/2020	Completed the SRS	1 Week
22/12/2020	Completed the Project Management	1 Week
31/12/2020	Completed the Conclusion Chapter	1 Week
08/02/2021	Re-submission of the report	2 Weeks
01/04/2021	Finalized the report	12 Weeks

Table 9: Activity Schedule

3.5 Work breakdown structure

Appendix 1: Work breakdown structure

3.6 Gantt chart diagram

Appendix 2: Gantt chart

Team - 50 Cent

3.7 Chapter Summary

The above chapter focused on creating a detailed plan as how the project needs to be proceeded. The chapter identified various system development methods and discussed on the most suitable method for the current project. Also, the project elaborated on the various risks which may arise and the required mitigations for such situations. Finally, a clear work breakdown structure was shown. The next chapter will discuss on the stakeholder analysis requirement gathering techniques and other system requirements.

4. System Requirements Specification (SRS)

4.1 Chapter Overview

This chapter introduce the requirement specification process of the research project. The stakeholders of the project are introduced in the stakeholder analysis. Furthermore, this chapter will be covering the sections including use case diagrams, use case descriptions, Functional and Non-functional requirements of the system.

4.2 Stakeholder analysis

4.2.1 Onion model
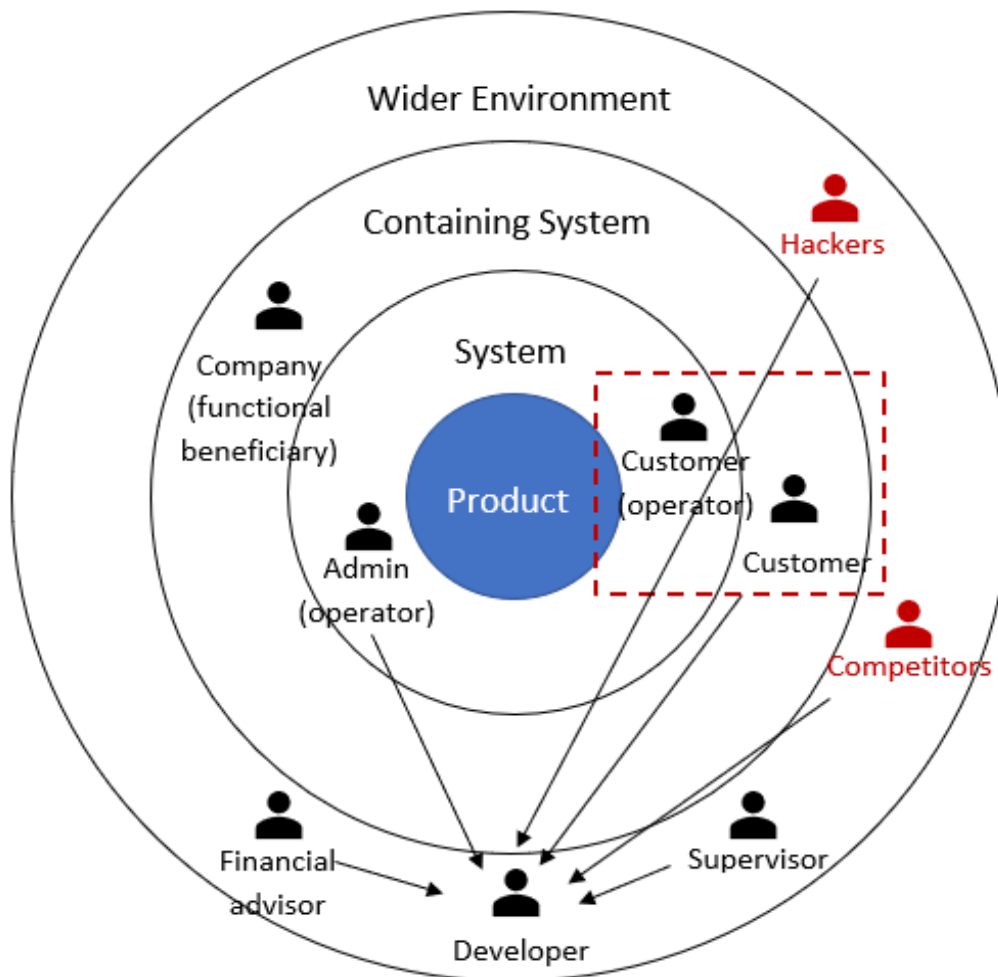
Figure 12: Onion Model

4.2.2 Stakeholder descriptions

Stakeholder Role	Stakeholder	Description
Direct Engagers	Admin	The company administrators directly involved in the system to find safe customers in the network and do transaction.
	Customer	Customers directly engage to the system to purchase products from registered companies.
Functional Beneficiary	Customer	Wants to be able to find companies to engage in credit transactions.
	Company	Wants to benefit from credit worthy transaction to avoid bad debts.
Negative	Hackers	Will disrupt the program and cause security breaches.
	Competitors	Wants to build a better system provided by Credit Mart.
Experts	Supervisor	Wants researcher to complete the program with proper standards
	Financial Advisor	Gives the developer crucial financial advice in keeping the system safe for users.

Table 10: Onion model description

4.3 Requirements gathering

The most important factor in the Software Development Life Cycle is 'Requirement Gathering'. This process can be also called as 'Elicitation Process' which literally means to the ways of gathering data and information to carry out a project successfully.

There are several techniques and ways, in-order to carry out this process. This will be further discussed in this chapter.

4.3.1 Techniques for Requirement Gathering

There are various types of techniques used for requirement gathering. The most common and widely used techniques are 'Interview', 'Questionnaire', 'Brainstorming', 'Storyboard', 'Prototyping', 'Use cases' and 'JAD (Joint Application Development)' (Mushtaq, 2016).

This section of the report explains the data gathering methods used to improve the 'CREDIT-MART' project.

No.	Method	Objective
01	Observation of existing systems	According to this technique, it was able to identify the similar types of project systems. Gaining some knowledge of previously created

Team - 50 Cent

		systems and research are much needed to carry out the project meaningfully. This strategy was extremely effective to identify the strengths and drawbacks of the existing approaches in the financial credit identification system.
02	Interviews with Domain Experts	Interviews are basically focused to get an overview opinion from the domain experts. The questions which are questioned out are mostly open-ended. This process gave a great outcome in continuing with project. The domain experts were very much enthusiastic about the project 'Credit-Mart'. This process also gave the opportunity to increase the depth of the project with various minor solutions. This discussion could sort out various threats and difficulties which would arise at the time continuing with the project.
03	Interviews with Businessman	This interview is targeted to businessman based on different aspects about credit sales. Businessman from Private Limited Companies are the primary target audience of the project Credit-Mart. By this interview, it can understand and get ideas about what these businessmen think on the aspect of credit businesses. This interview could be entangled on the view as a buyer and a seller too. There can be several minor points which could be analyzed deeply based on their aspect. This could help in building the project with more efficiency.
04	Questionnaires for Sellers (Private Limited Companies)	The questionnaire was targeted mainly to the primary target-audience of Credit-Mart. This questionnaire was mainly set-out to check if its legible to continue and perform our task appropriately. Explicit and data could be gathered by this approach. By this is was able to identify the ideas of which businessman had regarding credit sales and their thoughts regarding doing credit businesses with new customers.
05	Questionnaires for Buyers (Private Limited Companies)	The questionnaire was targeted mainly to the primary target-audience of Credit-Mart. This questionnaire was mainly set-out to check if its legible to continue and to know the opinions of the businessman on this field. Explicit and data could be gathered by this approach. By this it was able to identify the ideas which businessmen had regarding credit sales and their thoughts regarding doing credit businesses with new customers.
06	Brainstorming	Brainstorming process is a very useful process while engaging in this project. Gathering necessary information and other useful data by making a collective decision within friends, colleagues, and domain experts would lead this project to a better ending. This is particularly to share ideas and thoughts since this is a totally new concept. This session is explicitly engaged by the people who have a better knowledge about the topic.

07	Proto typing	By going ahead with this method, it would be more reliable for the consumer with the best deign. By processing on this, we can make pre-decisions in-order to make it look and feel to its best. When having gone through the proto-typing process it is easier to identify what to be included and what not to be included. At the stage of implementation, it is less likely to face errors and it would not be much of a confusion.

Table 11: Requirement gathering technique

4.3.2 Questionnaire Design

In-order to carry with development of the project, the developers sent out two questionnaires to gather necessary information.

1. Questionnaires for sellers (PLC)
2. Questionnaires for buyers (PLC)

Both these parties are the ones who are involved in buying and selling which can be also called as parties involved in the process of transactions. Since the primary target audience of Credit-Mart are the businessmen of the Private Limited Companies, questionnaires were directly sent to them. The responses that were attained by the questionnaires were very much useful in-order to understand what the people in the private sector have in the firm of credit sales. By developing Credit-Mart, it would be easier and fruitful for them to carry out their transactions. The summary of the questionnaire is briefly discussed in the next chapter (Requirement Analysis).

4.3.2.1 Questionnaires for sellers (PLC)

No.	Goals	Questions
01	To check the preference of the businessman if they are reliable to do businesses with new customer.	Would you prefer doing credit sales with new customers? • Yes • No
02	To determine if there is any process to determine the credit limit.	Do you determine the credit affordability of your customers before moving on with credit transaction? • Yes • No
		What is the method in which you calculate the credit limit for a certain customer?

03	To identify the risk factors that take place during a credit transaction.	How often do you experience bad debts from your customers? • 1 • 2 • 3 • 4 • 5
		What are the risk factors you find when taking part in credit transactions? • Bad debts. • Delaying payments. • Debtor being unavailable to contact after sales. • Other…

Table 12: Questionnaire description for sellers

4.3.2.2 Questionnaires for buyers (PLC)

No.	Goals	Questions
01	To identify the problems that arise when approaching new suppliers.	What kind of problems do you face when approaching a new supplier? • Trust issues • Under supplying requested items
02	To know the possibility of getting the financial data in-order to calculate the credit limits.	Do you wish to have a method to prove your credibility to your supplier? • Yes • No
		If yes would you like to share your financial data to maintain your financial profile? • Yes • No • Maybe
03	To check the preference of approaching new suppliers.	How often do you approach new suppliers? • 1 • 2 • 3 • 4 • 5

Team - 50 Cent

04	Identify the probability of using the software.	Would you be satisfied if there is an application which you can maintain your financial profile and find new suppliers? • Yes • No • Maybe

Table 13: Questionnaire description for buyers

4.3.2.3 Formal Interviews with Domain Experts

The interview session was conducted with a data science expert. We conducted this session in-order to gain necessary information regarding the project.
The aim of the interview was to resolve the technical issues with the data science component and on the available dataset. It was also a great opportunity to clarify about our project and to gain necessary information regarding implementation.
- **Interviewee** - Mohamed Nazeem Ayoob
- **Via** - Google Meets
- **Date of meeting** - 1st of February 2021

4.3.3 Analysis of gathered data

Questionnaire were sent out on 2020.11.13. Questionnaire is consisting of two main parts. One part is businesses which are already doing their sales on credit basis and other part is for people who is not doing credit sales.

4.3.3.1 Credit seller's questionnaire

Do you engage in "Credit Sales" with your customer?
9 responses
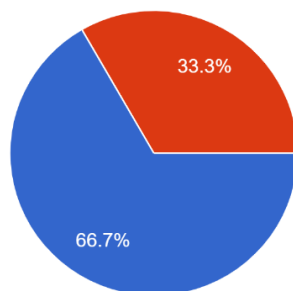
66.7% of businessmen who have responded to the questionnaire have said yes to this question. Which concludes that most people or experienced businessmen are commonly doing their trades on credit basis.

Team - 50 Cent

Figure 13: Credit seller's questionnaire response 1

Would you prefer doing credit sales with new customers?
8 responses

- Yes
- No
- Maybe

25%

75%
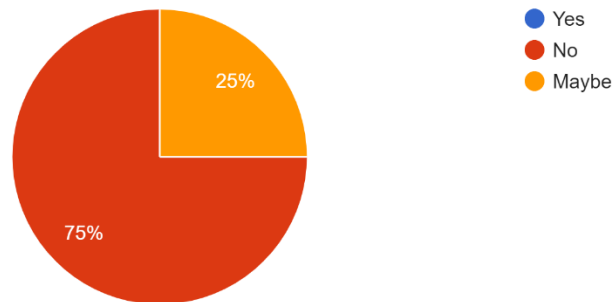
75% of the sellers who are engaging with credit sales have voted no for doing credit sales with new customers. It means that even though they are experienced in the field they do not trade on credit basis for new customers because of the risk they have to face. On the other hand, they are losing several new customers because of that fact.

Figure 14: Credit seller's questionnaire response 2

How often do you experience bad debts from your customers?
8 responses
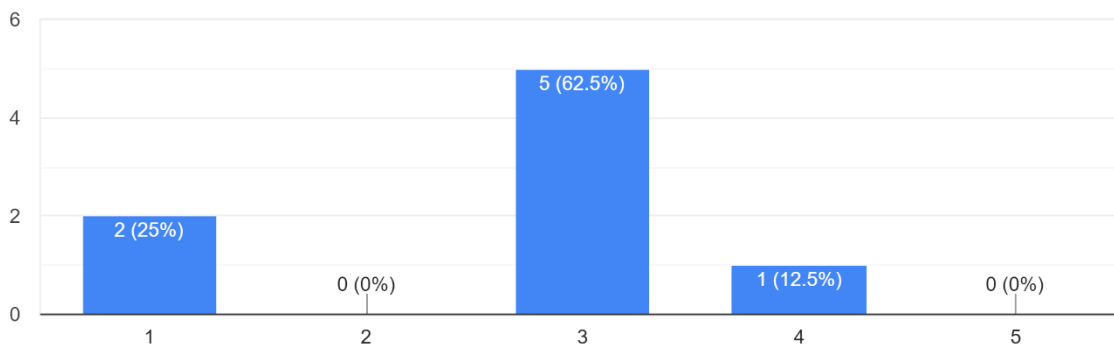
Figure 15: Credit seller's questionnaire response 3

From the total amount of people who are engaging with credit sales with their customers 75% of businesses have said they receive bad debts more often or sometimes. Which means most of the businesses are experiencing considerable amounts of bad debts.

4.3.3.2 Formal Interviews with Domain Experts

During the formal interview credit mart team explained the basic idea of credit mart and explained the technical methods that team is going to follow up. Several questions regarding the technical side of credit mart were formulated for the formal interview.

The domain expert explained certain areas that we can improve furthermore. One of them was the weight of the evidence method. Suggested to study about correlation methods and implement that.

Overall, his opinion was to start implementation as soon as possible and invest more time on training the model and tuning it to best fit.

Team - 50 Cent

4.4 Models
4.4.1 Use case diagram
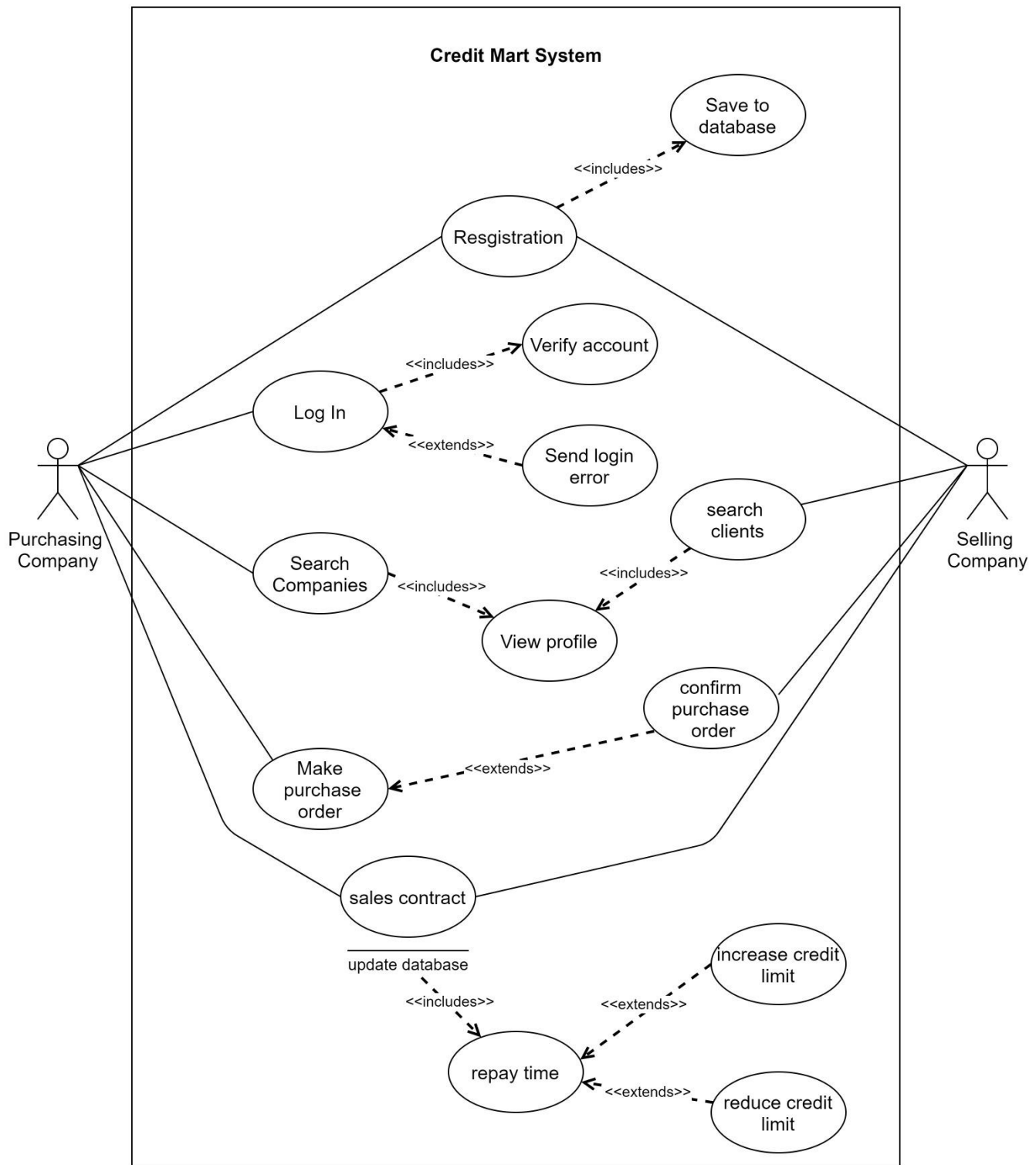
Figure 16: Use case diagram

Team - 50 Cent

4.4.2 Use case description

Use case Id	CREDITMART_UC_1
Use case Name	Sales contract
Priority	High
Primary Actors	Purchasing company Manager, Selling company Manger
Description	The contract given by Selling company Manger to the transaction process.
Pre-conditions	• Should have an active internet connection. • purchase order should be accepted by Selling company Manger
Main Flow	1. Use case begins once the purchase order is confirmed 2. Then the Selling company Manger and Purchasing company Manager will have to agree with the contract 3. Use case ends once both parties agree with the contract
Alternative flow/ Extension Flow	1.a. If the purchase order is rejected by the Selling company Manager - the Transaction will not continue 2.a. If the contract is rejected by the purchasing company - the Transaction will not continue

Table 14: Use case description 1

Appendix 6: Use case description 2

Appendix 7: Use case description 3

Appendix 8: Use case description 4

Appendix 9: Use case description 5

Appendix 10: Use case description 6

Team - 50 Cent
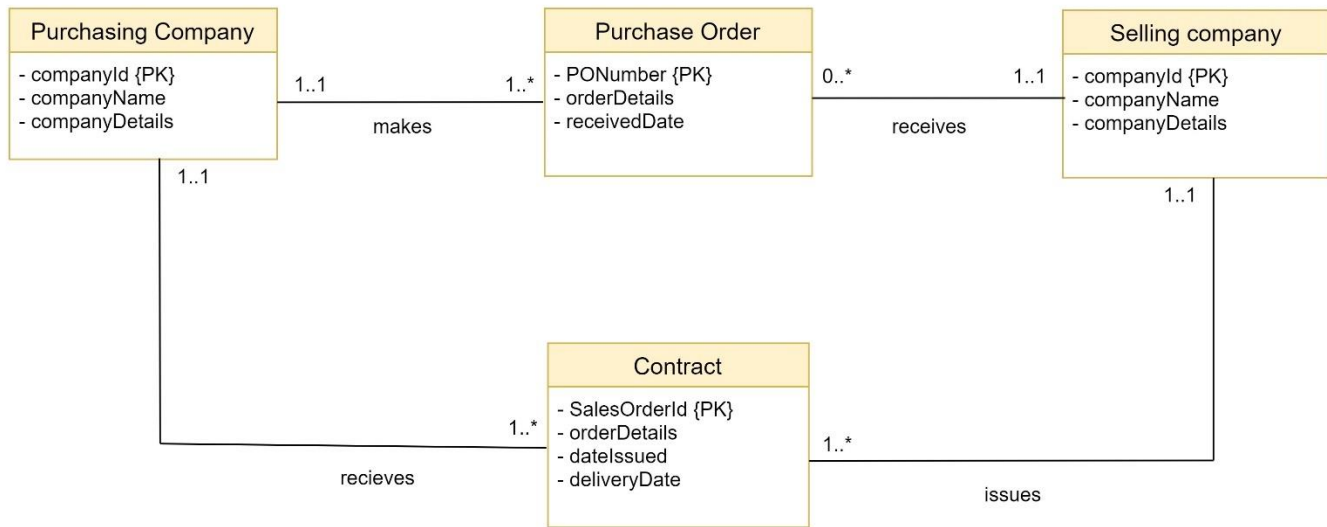
4.4.3 Domain Model

Figure 17: Domain model

4.5 Functional requirements

The below table illustrates the functional requirements affiliated with the project. The priority level of the requirements is defined. The priority level was defined based on the limited time frame and resources. Since the use cases are identified, the requirements are also mapped with use cases.

ID	Requirement	Priority	Use case mapping
FR1	The system should be able to get user input (Financial data).	**High**	**Registration**
FR2	The system should be able to identify the company's risk factor.	**High**	**View Profile**
FR3	The system should be able to predict credit limits for the customers.	**High**	**Search clients**

FR4	The system should be able to generate sales contract between buyer and seller	**Optional**	**Sales contract**

Table 15: Functional requirements

4.6 Non-Functional requirements

The non-functional requirements of the Credit Mart are listed below.

Requirements		**Description**
NR1	Accurate	Credit mart should be accurate when featuring upkeep issues.
NR2	Reliability	The system should be reliable for the credit management to address the issues featured.
NR3	Usability	The system user design and user experience must be good and user friendly.
NR4	Performance	Credit mart should function admirably with no extensive.
NR5	Security	With many users using the services of the credit mart, user data should remain accessible to only the authorized parties without the risk of breach.

Table 16: Non – functional requirements

4.7 Chapter Summary

The chapter commences by providing an analysis of the stakeholders that are associated with the project while defining their roles and viewpoints. The requirement elicitation techniques including interviews, questionnaires and observations of exiting systems were presented by illustrating the strengths and weaknesses of each technique and finally, all the findings of requirement elicitation process were summarized and presented in a table. Next, the use case diagrams and use case descriptions affiliated with the project are explained. Finally, the functional and non-functional requirements associated with the project are defined

5. Design

5.1 Chapter Overview

The previous chapter was focused on identifying the stakeholders of the system. This chapter will focus on providing high level designs to the proposed solution and what the application's main flow would look like. The chapter will also provide class diagram and a high level architecture diagram using UML notations and UI mockups.

5.2 High-Level architecture Diagram
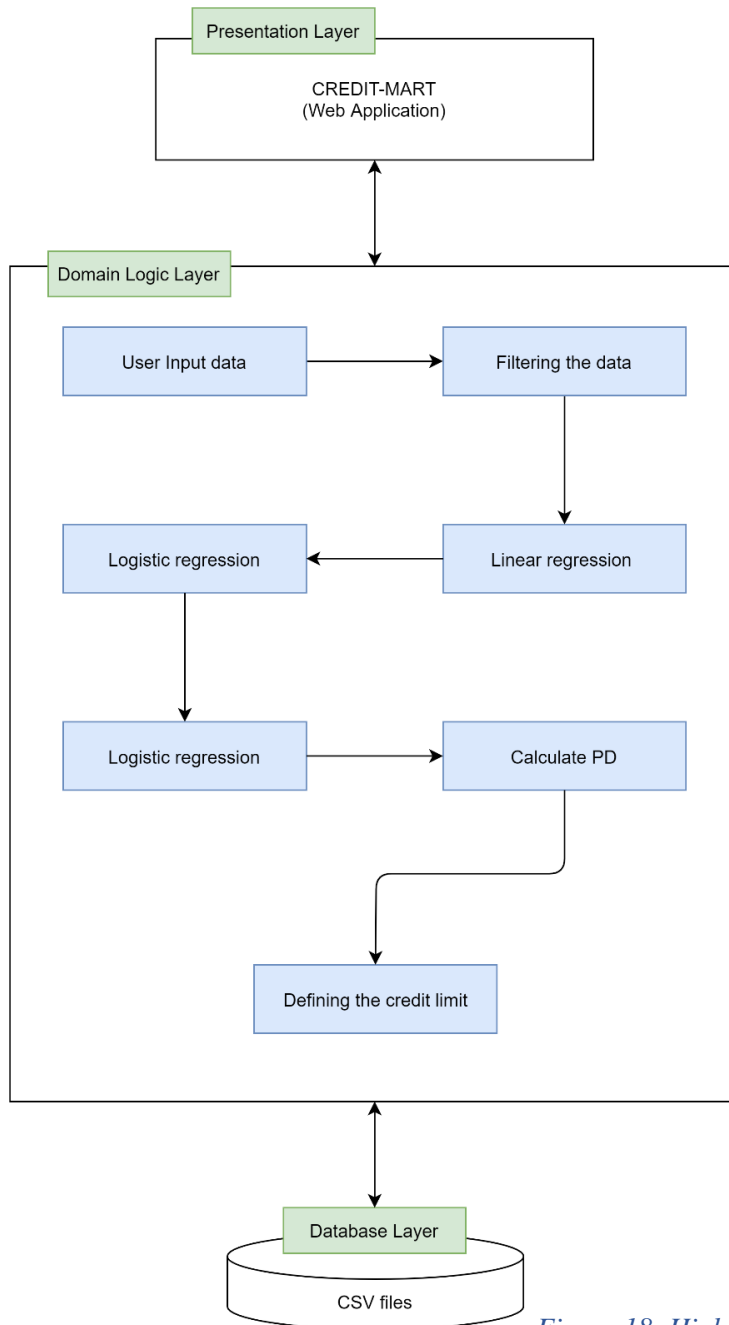
The diagram shown within this context is the High-Level Architecture diagram of the system CREDIT-MART. The system architecture is elaborated in the 3 different tiers.

- Presentation Layer

- Domain logic Layer

- Database Layer

The datasets which would be used in-order to calculate the credit limit as per the objective of CREDIT-MART is stored in the Database Layer.

The Domain Logic Layer includes the methodology of Credit-backend Mart's logic to evaluate the credit limit. The different kinds of models which were used in implementation of CREDIT-MART is marked up in this layer graphically.

The Presentation Layer shows the front end of the web application.

Figure 18: High level architecture diagram

Team - 50 Cent

5.3 Class Diagram

Appendix 3: Class diagram

Class	Responsibility	Collaboration
Company	Contain the company details	FinancialData and Controller
Controller	Handling backend controls and intercommunication methods.	Company, PDModel and DatabaseManager
DatabaseManager	Handling database	Controller, PDModel
PDModel	Calculating the probability of default of a company	CreditModel and DatabaseManager
CreditModel	Calculating a credit limit of a company	Controller and PDModel
UserInterface	Home view and control front end	Controller
FinancialData	Containing Financial Data of a company	Company

Table 17: Class description

5.4 Sequence Diagram
Appendix 4: Sequence diagram

The above figure illustrates the sequence diagram for the core function of Credit mart. The user company must first sign up to the credit mart system. When the required data is provided, it is saved in the database and a reply will be sent to the user stating that the company has been successfully signed up. If the user has already signed up and created a credit mart account, the user simply needs to enter credentials and sign in.

When updating financial data, the company must provide the required data which will be stored to the database, then the pd value is calculated using entered financial data using the machine learning model. Then the calculated pd value is validated and stored in the database. The credit limit is then calculated using the pd model and the financial data and the company profile will be updated as per entered data and other companies will be able to view your profile.

5.5 Activity Diagram
Appendix 5: Activity diagram

Team - 50 Cent

5.6 Wireframes

Before the implementation of the system, wireframes were sketched, and feedback was gathered by showing different Credit management systems and domain and some UI/UX designers. The figure below shows the wireframe of the entire Credit mart
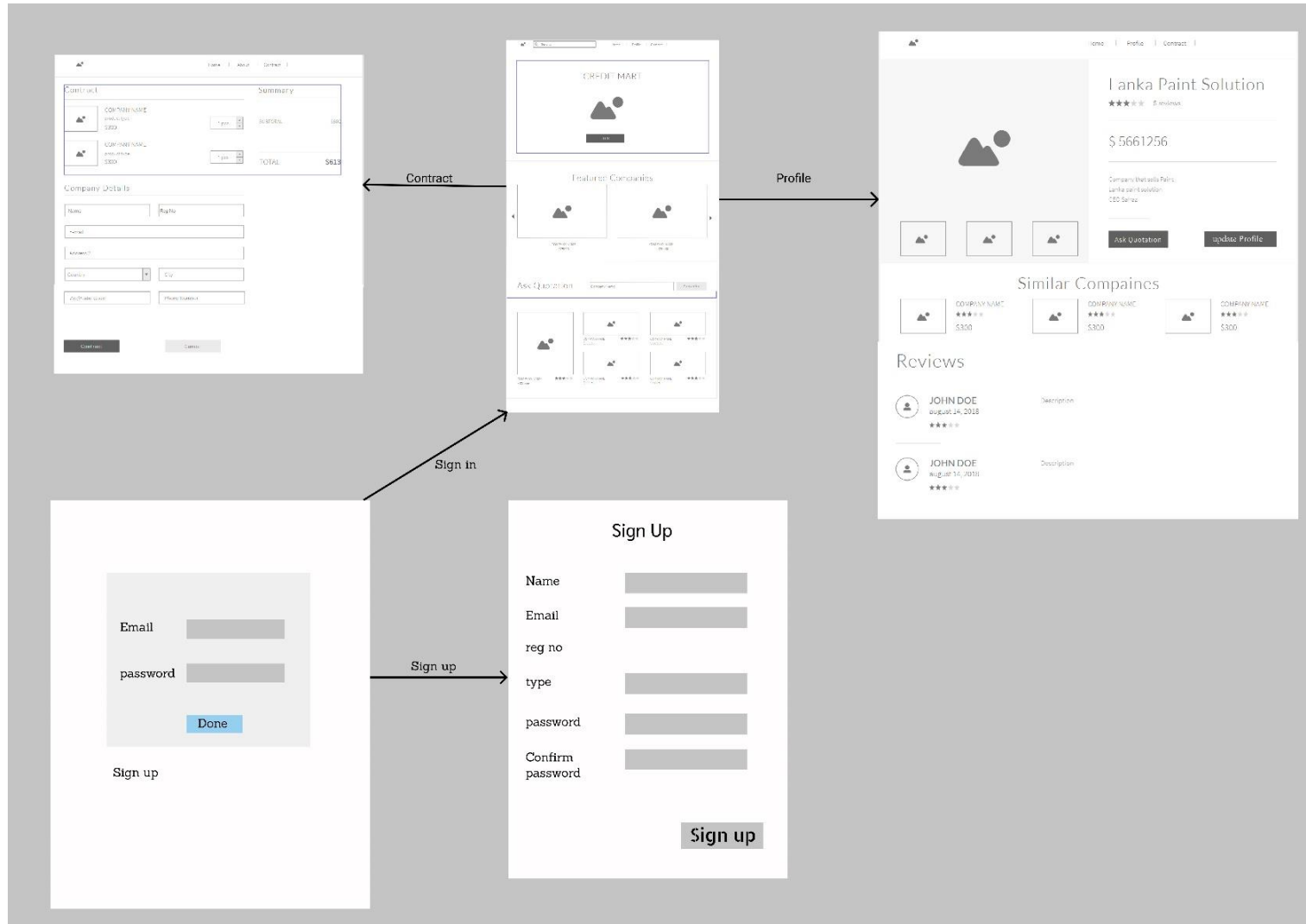
Figure 19: Wireframe

6. Conclusion

6.1 Chapter Overview

This chapter will focus on the data set that will be used to move forward with the data science component of the system. The legal, social, ethical, and professional issues that may arise and the how the Credit-Mart system will handle these related aspects and finally the plans on how the proposed system will be implemented.

6.2 Dataset

200+ Financial Indicators of US stocks (2014-2018) dataset is initially used to develop a stock market trading model. This data set consists of five .csv data files which contain more than 200+ financial indicators about each company based in the USA. This data set is a public dataset which is available in Kaggle.com with ethical clearance.

Financial Indicators of US stocks data set consists of nearly 200 financial indicators which covers the aspects of solvency, liquidity, debt, and profitability of companies. Those aspects are the key features to calm whether a company has good financial health or not.
Data file contains following files,
● 2014_Financial_Data.csv
● 2015_Financial_Data.csv
● 2016_Financial_Data.csv
● 2017_Financial_Data.csv
● 2018_Financial_Data.csv

In this data set some financial indicator values are missing so that dataset needs to be cleaned using data clearing techniques (dropna, fillna, etc).

6.3 Legal, social, ethical, and professional issues

6.3.1 Legal aspects

When developing CREDIT-MART, a high emphasizes was given to Data protection act and the Law of Sri Lanka regarding the business law. The data set was extracted from the Kaggle for credit mart and the terms and condition of use the dataset was read carefully and it has Creative Commons public, so we downloaded it and started using it. The data set was not changed and altered in any, but it was cleaned. It was not used for any illegal activities which disobey the law.
The Sri Lankan law on Credit sales and its legal aspects are considered in the context of the law. There are no amendments made till now, but England have made amendments that addresses the new concepts like Credit cards, Online purchases etc. (Chandrawansa, 2020)
Act, No. 28 of 72 The Need for Reforming the Sale of Goods Ordinance of Sri Lanka 2005 and Payment Devices Frauds Act, No. 30 of 2006 are the new laws introduced to this area. Section 5 of the Sale of Goods Ordinance states that, "the sale can be enforceable if the buyer has accepted the goods or part of it, paid the price or part of it or unless some note or memorandum in writing signed". (Chandrawansa, 2020)
The Contract which the credit mart has will be able to use as a legal agreement which can be used to file a case if the purchasing party did not repay and it is a contract that both parties have agreed for the credit sale.
Software License was also prioritized when developing CREDIT-MART. Credit Mart uses some legal software to sketch the Prototype such as Adobe Photoshop, Figma and other open-source tools.

The questionnaires sent out during the requirement gathering phase, did not gather any personal information of the response users. The data gathered were treated confidentially and the privacy of the responded users were protected, and no personal details were collected. Details of the users who answered the questionnaire were kept anonymous. The guidelines provided in the module handbook were strictly followed in the project and 1.5 spacing was used.

6.3.2 Professional aspects

Questionnaires were sent out on a pre specified and well-organized manner with all instructions and deadlines mentioned perfectly as the target audience of CREDIT-MART are the Private Limited Companies. It was a great achievement of the developers of CREDIT-MART to receive a minimum response due to heavy and tight busy schedule of the businessman. Although the responses were limited, it was handled explicitly without misleading private and confidential data. The businessmen were informed about the usage of their data by time.

The interview which was organized in-order to gain necessary data to continue the improvement of the project was handled in a well-organized way. The interviewer was happy to share his thoughts and ideas about the topic.

Trello was a well improvised tool, where all the updates the tasks were managed. This application is highly professionalized tool to improve the stability of the developers. Due to the current situation, all the developers were connected online via Discord. This application was very much helpful to make us stay connected. All sorts of development processes were updated in Git-Hub. This is a universal application where all developers use in-order to store and save their development works. This maintains the professionalism to a well standard.

6.3.3 Social aspects

The project Credit mart has some social effect on intermediate businessmen and also debt collection companies. The intermediaries who do act as the agent will be eliminated which has an impact to their business and the society. The companies registered under our system will have to maintain a credit limit which the companies so the society will be aware about the companies trust and repayable ability. The project will be implemented in only English and it might affect the people who do not understand the English language. Multi Language support will be added in later as a future development process. By using this web application, small companies can expand the business safely without a credit risk. By managing the credit risk and timely payments from credit mart it leads to better cash flow planning in society. Credit mart also works as a credit certificate for buyers. From credit mart creditors can also manage credit risks by performing periodic reviews of active customers. This help to analyze the social perceptions of indebts in society.

6.3.4 Ethical aspects

What is ethics? Ethics is a system of moral principles. They affect how people make decisions and lead their lives. Ethics is concerned with what is good for individuals and society and is also described as moral philosophy (2014, n.d.). The project credit mart has ethical advantages as well as ethical disadvantages. The main purpose of project credit mart is reducing bad debts of business enterprises. That reason helps society to maintain their lifestyle. That is the main ethical advantage of project credit mart.
In credit mart users should share financial data with the application. That reason can be considered as the ethical disadvantage of Credit mart, because in society no one likes to share private information with 3rd party applications.

6.4 Plans for implementation

As the necessary research regarding the entire system has been done, the plan for implementation will be to use the agile model to start development. As the agile model is an iterative feature driven methodology it suits best for the current system as the credit mart system involves in many features. The first step would be to build the base product with all the functional requirements such as calculating the company pd, estimating a credit limit and other functional requirements mentioned in chapter (4.5 Functional requirements). If the functional requirements are achieved before the deadline the team will focus on deploying new features to the system.

Appendix

Appendix 1: Work breakdown structure

Team - 50 Cent
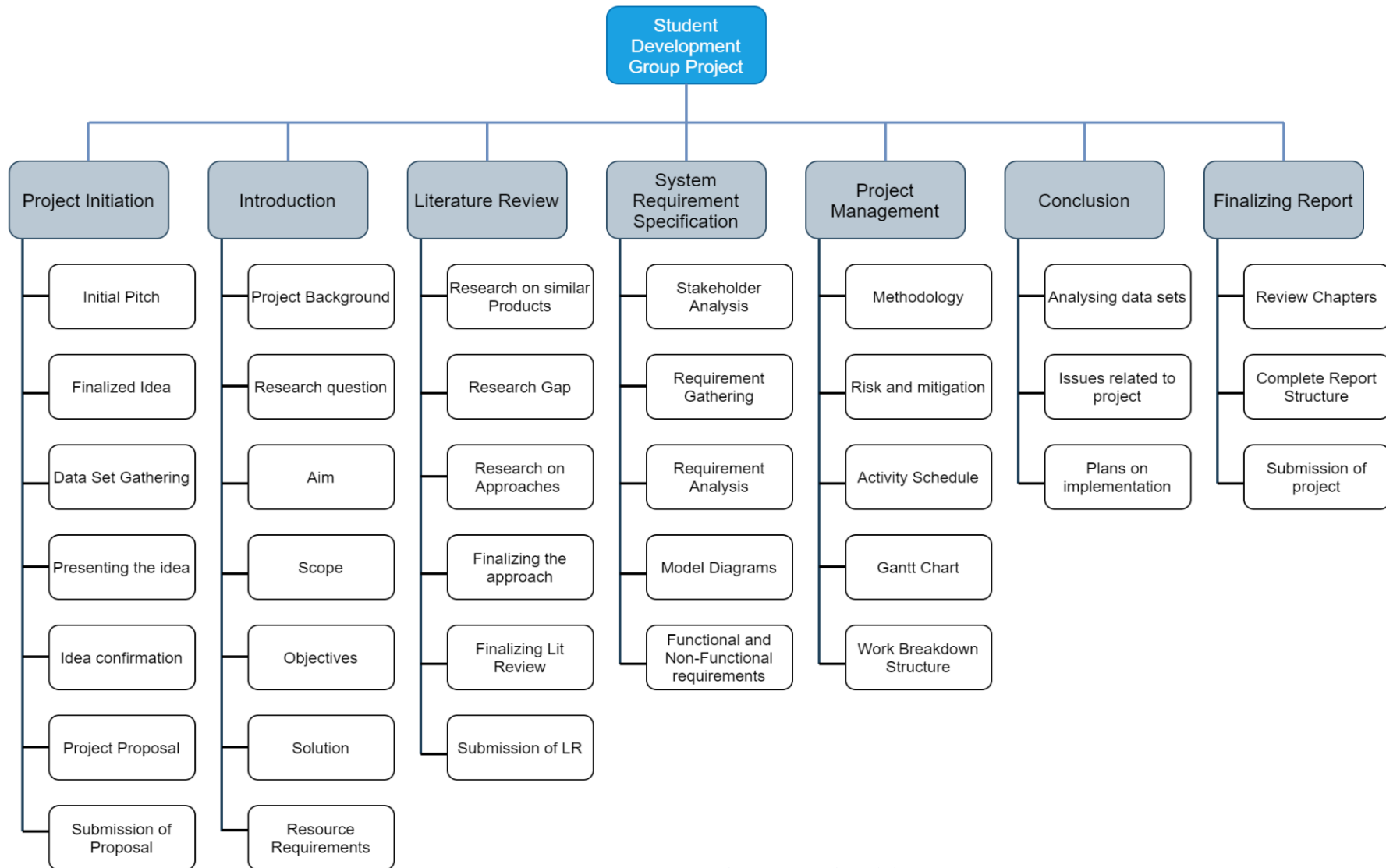
50 Cent (Gantt chart)

Project Start Date – 22-09-2020

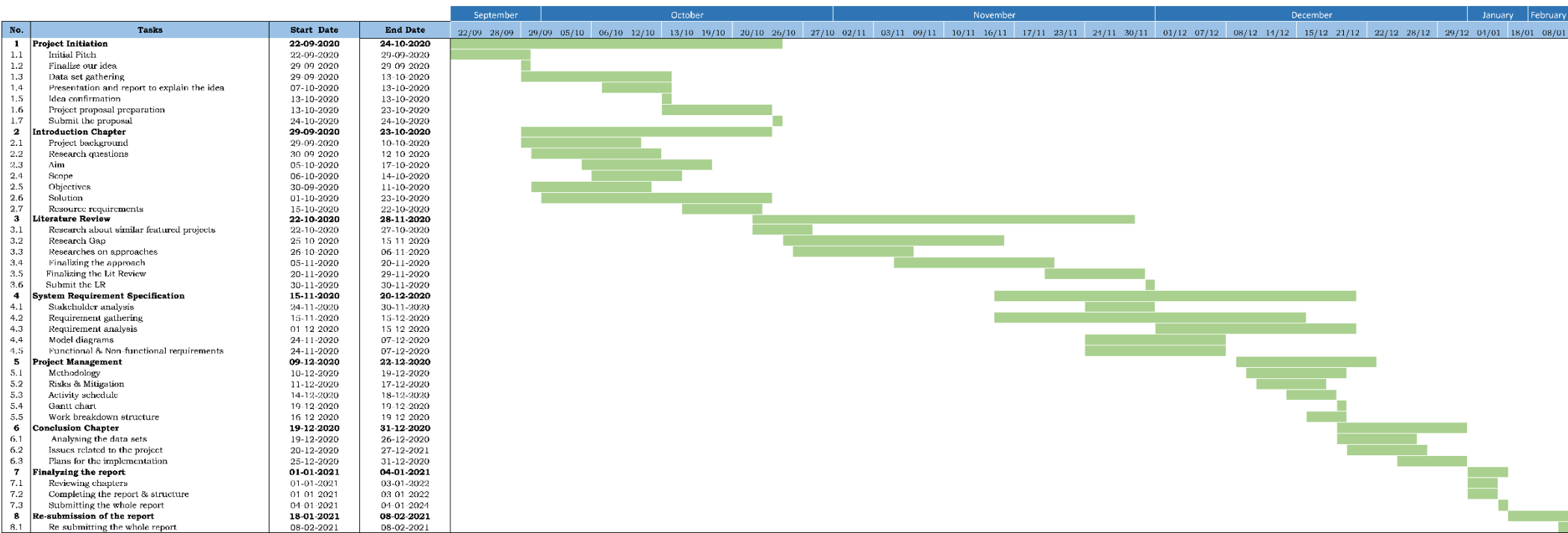| No.	Tasks	Start Date	End Date	September		October						November						December				January	February
				22/09 28/09	29/09 05/10	06/10 12/10	13/10 19/10	20/10 26/10	27/10 02/11	03/11 09/11	10/11 16/11	17/11 23/11	21/11 30/11	01/12 07/12	08/12 14/12	15/12 21/12	22/12 28/12	29/12 01/01	18/01 08/01
1	**Project Initiation**	**22-09-2020**	**24-10-2020**																
1.1	Initial Pitch	22-09-2020	29-09-2020																
1.2	Finalize our idea	29-09-2020	29-09-2020																
1.3	Data set gathering	29-09-2020	13-10-2020																
1.4	Presentation and report to explain the idea	07-10-2020	13-10-2020																
1.5	Idea confirmation	13-10-2020	13-10-2020																
1.6	Project proposal preparation	13-10-2020	23-10-2020																
1.7	Submit the proposal	24-10-2020	24-10-2020																
2	**Introduction Chapter**	**29-09-2020**	**23-10-2020**																
2.1	Project background	29-09-2020	10-10-2020																
2.2	Research questions	30-09-2020	12-10-2020																
2.3	Aim	05-10-2020	17-10-2020																
2.4	Scope	06-10-2020	14-10-2020																
2.5	Objectives	30-09-2020	11-10-2020																
2.6	Solution	01-10-2020	23-10-2020																
2.7	Resource requirements	15-10-2020	22-10-2020																
3	**Literature Review**	**22-10-2020**	**28-11-2020**																
3.1	Research about similar featured projects	22-10-2020	27-10-2020																
3.2	Research Gap	25-10-2020	15-11-2020																
3.3	Researches on approaches	26-10-2020	06-11-2020																
3.4	Finalizing the approach	05-11-2020	20-11-2020																
3.5	Finalizing the Lit Review	20-11-2020	29-11-2020																
3.6	Submit the LR	30-11-2020	30-11-2020																
4	**System Requirement Specification**	**15-11-2020**	**20-12-2020**																
4.1	Stakeholder analysis	24-11-2020	30-11-2020																
4.2	Requirement gathering	15-11-2020	15-12-2020																
4.3	Requirement analysis	01-12-2020	15-12-2020																
4.4	Model diagrams	24-11-2020	07-12-2020																
4.5	Functional & Non-functional requirements	24-11-2020	07-12-2020																
5	**Project Management**	**09-12-2020**	**22-12-2020**																
5.1	Methodology	10-12-2020	19-12-2020																
5.2	Risks & Mitigation	11-12-2020	17-12-2020																
5.3	Activity schedule	14-12-2020	18-12-2020																
5.4	Gantt chart	19-12-2020	19-12-2020																
5.5	Work breakdown structure	16-12-2020	19-12-2020																
6	**Conclusion Chapter**	**19-12-2020**	**31-12-2020**																
6.1	Analysing the data sets	19-12-2020	26-12-2020																
6.2	Issues related to the project	20-12-2020	27-12-2021																
6.3	Plans for the implementation	25-12-2020	31-12-2020																
7	**Finalyzing the report**	**01-01-2021**	**04-01-2021**																
7.1	Reviewing chapters	01-01-2021	03-01-2022																
7.2	Completing the report & structure	01-01-2021	03-01-2022																
7.3	Submitting the whole report	04-01-2021	04-01-2024																
8	**Re-submission of the report**	**18-01-2021**	**08-02-2021**																
8.1	Re-submitting the whole report	08-02-2021	08-02-2021																

Appendix 2: Gantt chart

Team - 50 Cent

FinancialData

+currentAssets
+currentLiabilities
+workingCapital
+totalAssets
+netIncome
+equityROE
+turnOverReceiverbales
+ebit
+totalDebt
+interestExpense
+totalLiabilities
+equityShareHolders
+longTermDebt

company

+companyUserName: String
+companyName: String
+contactNo: String
+email: String

+getCompanyName(): String
+setCompanyName(companyName)
+getCompanyUserName(): String
+setCompanyUserName(userName)

controller

+addNewCompany()
+deleteData()
+updateFinancialData()
+calculatePD()
+calculatecCreditLimit()

creditLimit

+targetCompany: Company
+pdValue: float

+calculateCreditLimit()

+uses

UserInterface

+searchCompany()
+addDetails()
+viewPD()
+logIn()
+SignUp()

DatabaseManager

+companyUserName: String

+addCompany()
+searchCompany()
+updateDetails()

PDModel

+targetCompany: Company

+calculatePD()

Team - 50 Cent   *Appendix 3: Class diagram*

Appendix 4: Sequence diagram

Team - 50 Cent

Appendix 5: Activity diagram

Use case Id	CREDITMART_UC_2	
Use case Name	save and update database	
Priority	High	
Primary Actors	-	
Description	Company registration details are saved to the database and the sales and purchases are updated to the database	
Pre-conditions	• Should have an active internet connection. • Should have past credit details	
	Save database	Update database
Main Flow	1. Use case begins once the companies are registered to the system 2. The user should give the financial statement for the system 3. Use case ends when it has done and submitted correctly	1. Use case begins when the contract is accepted by both Managers 2. Its updates to the table automatically 3. The use case end when the table is done updating
Alternative flow/ Extension Flow	Save database	Update Database
	1.a. If the user enters wrong details - ask for validations 2.a. If the contract is rejected by the purchasing company - the credit limit will not be predicted	1.a. If the contract is canceled - No effect to the system database

Appendix 6: Use case description 2

Team - 50 Cent

Use case Id	CREDITMART_UC_7
Use case Name	Registration
Priority	High
Primary Actors	Purchasing company Manager, Selling company Manger
Description	The registration when the companies want to use the system
Pre-conditions	1. should have internet connection 2. Should show the financial statement 3. Should be a PLC company
Main Flow	4. Use case begins one the user clicks the register button 5. The time is asking for the details and the financial stamen of the company 6. Use case ends when it submits the registration
Alternative flow/ Extension Flow	2.a. If the company is not willing to share financial statement or status - cannot register for the system

Appendix 7: Use case description 3

Use case Id	CREDITMART_UC_6
Use case Name	Repay time
Priority	High
Primary Actors	Purchasing company Manager, Selling company Manger
Description	The repay time will be given for the Purchasing company Manager to pay the debt
Pre-conditions	1. should have done with the contract
Main Flow	1. Use case begins once the contract is confirmed 2. The time is determined by the Selling company Manger for Purchasing company Manager 3. Use case ends the repay time is notified to the selling company manager

	1.a. If the purchase order is rejected by the Selling company Manger
Alternative flow/ Extension Flow	- the Transaction will not continue
	2.a. If time is not determined by the selling company
	- the time is automatically set to 1 month and 15 days

Appendix 8: Use case description 4

Use case Id	CREDITMART_UC_8
Use case Name	Search Company
Priority	High
Primary Actors	Purchasing company Manager
Description	The purchasing company will be able to search for the related selling companies according to their wish
Pre-conditions	- Should have an active internet connection. - Should be registered into the system - Related company needs to be registered into the system.
Main Flow	1. Use case begins once the company starts searching for a seller company 2. Then the Purchasing company Manager will be able to request for an order as their wish. 3. Use case ends once the notification is sent to the related to the companies.
Alternative flow/ Extension Flow	1.a. If the purchase company doesn't have a stable connection The company will not be search for seller companies

Appendix 9: Use case description 5

Use case Id	CREDITMART_UC_10
Use case Name	Purchase Order
Priority	High
Primary Actors	Selling company Manager, Purchasing company Manager

Team - 50 Cent

Description	The purchasing company manager will make a purchase order which will be notified to the selling company manager
Pre-conditions	● Should have an active internet connection. ● Should be registered into the system ● Related company needs to be registered into the system.
Main Flow	4. Use case begins once the company starts searching for a seller company 5. Then the Purchasing company Manager can view the selling company credit limit and company details 6. Then the company makes a purchase order request 7. The selling company manager then should accept the request to procedure 8. Use case ends once the user moves from the details page
Alternative flow/ Extension Flow	1.a. If the purchase company does not have a stable connection ● The company will not be search for seller companies 4a. if the selling company manager reject ● It will not proceed further

Appendix 10: Use case description 6

Bibliography

- Software Testing Methodologies. (2019). *smartbear.com*. Available from https://smartbear.com/learn/automated-testing/software-testing-methodologies/ [Accessed 13 December 2020].

- Gabriel, D. (2013). Inductive and deductive approaches to research | | Dr Deborah Gabriel. *deborahgabriel.com*. Available from https://deborahgabriel.com/2013/03/17/inductive-and-deductive-approaches-to-research/#:~:text=The%20main%20difference%20between%20inductive [Accessed 13 December 2020].

- Dudovskiy, J. (2011). Inductive Approach (Inductive Reasoning) - Research-Methodology. *Research-Methodology*. Available from https://research-methodology.net/research-methodology/research-approach/inductive-approach-2/ [Accessed 13 December 2020].

- Dudovskiy, J. (2010). Deductive Approach (Deductive Reasoning) - Research-Methodology. *Research-Methodology*. Available from https://research-methodology.net/research-methodology/research-approach/deductive-approach-2/ [Accessed 13 December 2020].

- P. Krokhmal, J. P. a. S. U., 2002. *Portfolio optimization with onditional value at risk objetives and onstraints.*. s.l.:s.n.

- What is the DSO report? (2020). *My DSO Manager*. Available from https://www.mydsomanager.com/online-help.php?question=2 [Accessed 16 November 2020].

- Credit Management tools and resources for SMEs and international groups. (no date). Available from https://www.creditmanagement-tools.com/ [Accessed 16 November 2020].

- How to Set Credit Limits - Credit Management Article -Credit Guru. (no date). Available from https://www.creditguru.com/index.php/credit-management/commercial-credit-management-articles/40-how-to-set-credit-limits [Accessed 16 November 2020].

- 5 Best Accounting Apps for Keeping Track of Business Expenses. (2019). *GetApp*. Available from https://www.getapp.com/resources/top-rated-accounting-apps-for-expense-tracking/ [Accessed 16 November 2020].

- Herga, Z. et al. (2016). Modeling probability of default and credit limits. *ResearchGate*. Available from https://www.researchgate.net/publication/320041913_Modeling_probability_of_default_and_credit_limits[Accessed 16 November 2020].

- Dominy, N. and Kempson, E., 2003. Cannot pay or will not pay? A review of creditor and debtor approaches to the non-payment of bills. *Department for Constitutional Affairs Research Unit*, p.82.

- Konovalova, N., Kristovska, I. and Kudinska, M., 2016. CREDIT RISK MANAGEMENT IN COMMERCIAL BANKS. *Polish Journal of Management Studies*, 13(2), pp.90-100.

- Altman, E. C. J. &. N. P., 1998. Credit-risk Measurement and Management.

- Basel, 2000. *Principles for the Management of Credit Risk.* Basel, Basel Committee on Banking Supervision.

Team - 50 Cent

- Natalija Konovalova, I. K. K., 2016. Credit risk management in commercial banks. *Polish Journal of Management Studies*, june.
- Yuan, D., 2015. Applications of Machine Learning: Consumer Credit Risk Analysis. *Applications of Machine Learning: Consumer Credit Risk Analysis*, [online] Available at: <http://hdl.handle.net/1721.1/100614> [Accessed 13 November 2020].
- 2014, n.d. Ethics: a general introduction. *Ethics guide, BBC*

Topics	Student	Student ID
C1. Aims and Objectives, C2. Logistic Regression, Credit Limit Model C3. Work breakdown structure C4. Stakeholder analysis C6. Plans for implementation	Safraz	20191066
C1. Project Background C2. PD-model, Machine Learning Supervised Learning C3. Methodologies C4. Requirement Gathering C5. Class diagram C6. Dataset	Oshan	20191052
C1. Proposed Solution C2. Research Gap C3. Risk and mitigation C4. Models C5. Wireframe C6. Legal aspects	Abdul Baasith	2019566
C1. Project Scope, C2. Supervised Learning, Non-machine Learning approach C3. Methodologies C4. Functional and non-functional requirements C5. Activity Diagram C6. Social aspects	Navindu	20191031
C1. Resource requirements C2. Decision tree C3. Activity Schedule C4. Requirement Gathering C5. Sequence Diagram C6. Ethical aspects	Kavindu	20191067
C1. Objectives C2. Feature Comparison C3. Methodologies C4. Requirement Gathering C5. High Level Architecture Diagram C6. Professional aspects	Ashfaaq	2019394

Table 18: Work breakdown

Team - 50 Cent