Capstone Report

On

PREDICTIVE MODEL ON STUDENT PERFORMANCE

School of Computer Science & Engineering

Under the Guidance of

Mr Kewal Krishan

Submitted by

AMRITESHWAR DWIVEDI (11505682)
JASHANPREET KAUR (11504725)
MADHAV (11506582)
Sv. RAJARSHI (11504761)
RAJNISH YADAV (11502996)

Declaration:-

We hereby declare that the capstone work entitled "Predictive Model on Student Performance" is an authentic record of our own work carried out in B.Tech degree in Computer Science and Engineering from Lovely Professional University, Phagwara, under the guidance of Mr. Kewal Krishan, during January to May 2018. All the information furnished in this capstone report is based on our own intensive work and is genuine.

CERTIFICATE

This is to certify that the capstone project entitled "Predictive Model on Student Performance" submitted by Amiteshwar Dwivedi, Jashanpreet Kaur, Madhav, Sv.Rajashri, Rajnish Yadav in partial fulfilment of the requirements for the award of Degree of Bachelor of Technology in Computer Science and Engineering at Lovely Professional University, Punjab is an authentic work carried out by them under my supervision and guidance. To the best of my knowledge, the matter embodied in the capstone has not been submitted to any other university institute for the award of any Degree.

Mr. Kewal Krishan
Head of Department
School of Computer Science and Engineering,
Lovely Professional University,
Phagwara, Punjab.

Date: 27/04/2018

Acknowledgment: -

1. Introduction

The students are future of our country and they will help to make the future lifestyle better and easier by enhancing the technology and research. Many students are distracted and fail to perform well academically as they have various things to learn and to attend during their growing years. So we have to intend to find the reasons and distractions by predicting and analyzing the same. We shall be comparing various records of the institute to help students in an effective manner we mainly focusing on the problems of the students.

1.1 Objective of the project:

The main objective of this project is to increase the student performance. The customer can be from any background that is from any college or a school or from any institution. The customer will be able to know the different problems faced by the students and then be able to help the students to increase their performance. We will gather views from the students related to problems in their area of specialisation, to analyse the data in order to predict the specific problem and to give suggestions to use our predictive model. Institute/Universities/Colleges for the benefit of overall development of the Student University or College.

1.2 Description of Project:

In this project we will make the predictive model of students problem which will analyse the day today aspects of students and to predict what all things can be done to improve his/her performance.

In this project we will make feedback forms, Google forms and conduct survey in Universities and will predict the problem why students lack in their performance.

After taking both the dataset we will Acquisitive dataset and by using tools we will analyse the whole dataset and predict the main cause, and then by using tableau we will predict and give valuable data which will help in improving student academics.

And later we will make certain Statistics which will help in tell on which aspect the students lack and on which things a University/College should focus on to improve the skillset and performance of students.

We mainly focusing on the problems of the students which are as follow

- Low CGPA
- Backlogs
- Lack of interest

1.3 Scope of Project:

This project will give a boost to the whole performance of the University/College which will l tell the weak points of students and how students can improve their performance. So, a University/College can increase the efficiency of the students and each and every students will get good grades. As we all know each and every student is important, so if knowledge is increasing then the overall ranking of University will increase. The University can improve ranking in comparison to other Universities globally.

As the overall learning experience and the obstacles of students are focused, hence the University can work on specific areas which will be helpful for them.

Developers Benefit:

We will get more projects and as overall learning of students is increasing, our country will also develop more efficiently, and a better job prospect can be offered.

2. System Description:

As the world is developing day to day we have new technologies and subjects emerging. These subjects and technologies are taught to the students as they are our future and key to the evolving world. It will be a problem they don't understand what is being taught to them. Data science plays an important role in analysing and predicting the outcome.

In this project we are going to analyse the data we gather from the student and the data that we receive from the institution regarding the student's performance and then find where the problem is and give suggestions to the institution on how to increase the performance of their students

This project mainly focuses on the following things

1) Gathering the data from the students regarding various issues that they are facing which is resulting in their performance degradation

2) Receiving the data from the institution and then analysing

3) Comparing both the data that we have gathered and then make a report why students' performance is low

4) Finally giving suggestion to the institutions on what ways and how to increase their student performance

2.1 Customers:

The main objective of this project is to increase the student performance. The client can be from any background that is from any collage or a school or from any institution. The client will be able to know the different problems faced by their students and then be able to help the students to increase their performance. So our customer are mainly the institute so that they can help there institute to help the students to increase their performance.

2.2 Functional Requirements:

A system with Tableau and a data mining software tool in it. We use Tableau for data visualization, creating reports, graphs and statistical analytics we use Data Mining for the data that we have gathered and make patterns from it and drawing inference that why students are facing problems in their academic performance. Data mining is used to take accurate decisions and also to predict the results of the student. With the results the institution can focus on what to teach and how to teach.
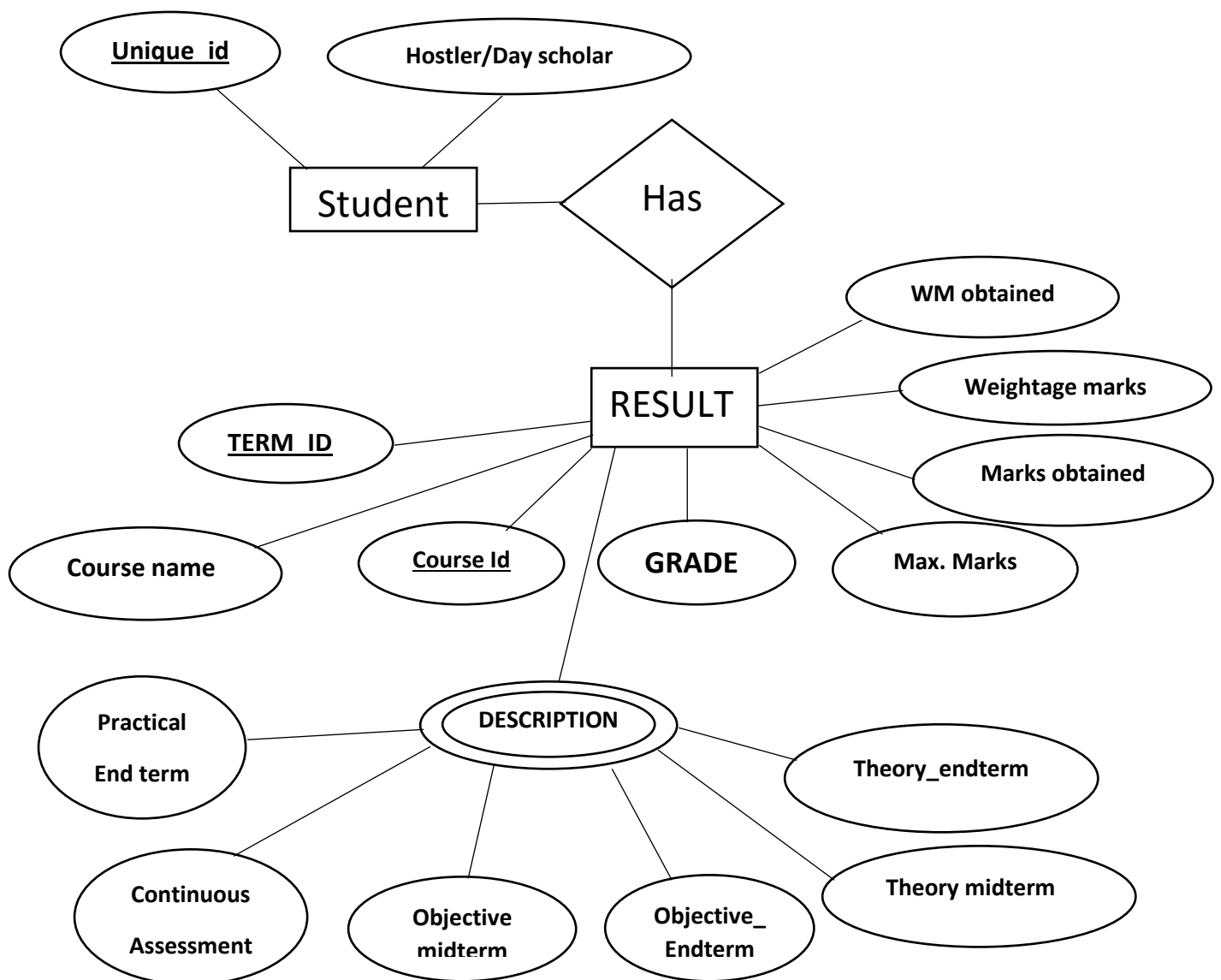
2.3 Non-Functional Requirements:

Structure of the data: datasets and the type of data that we have gathered

- Reliability: The outcome that we have given to the institution should be 100% reliable
- Appearance: presenting the report to the institution
- Size of the Data: Depends on the institution and number of students who we analysing
- Processing Time: 3-15 Seconds estimated, may also depend on the system we are working with
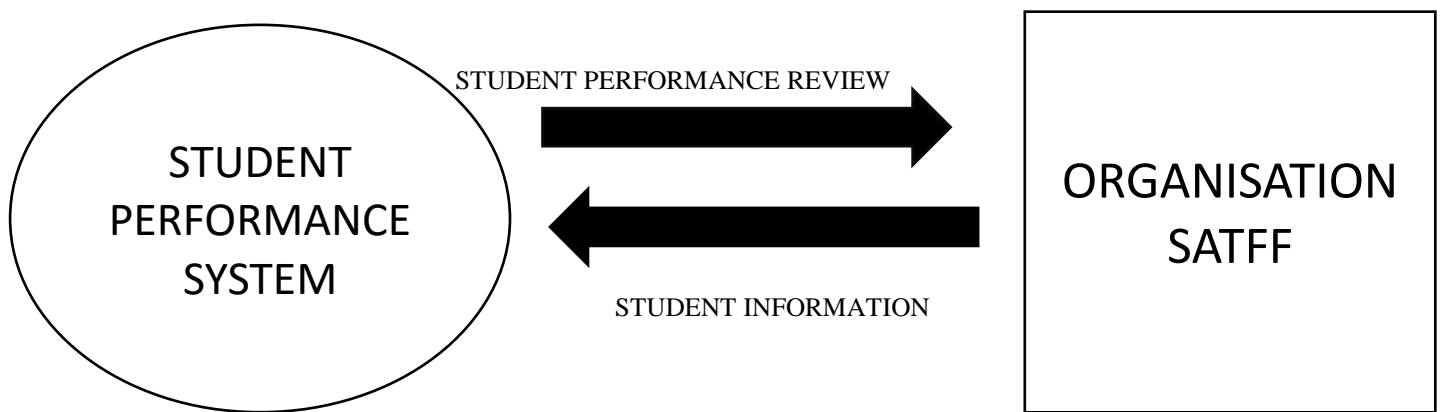
3. Design

We have to types of design and are as follow:
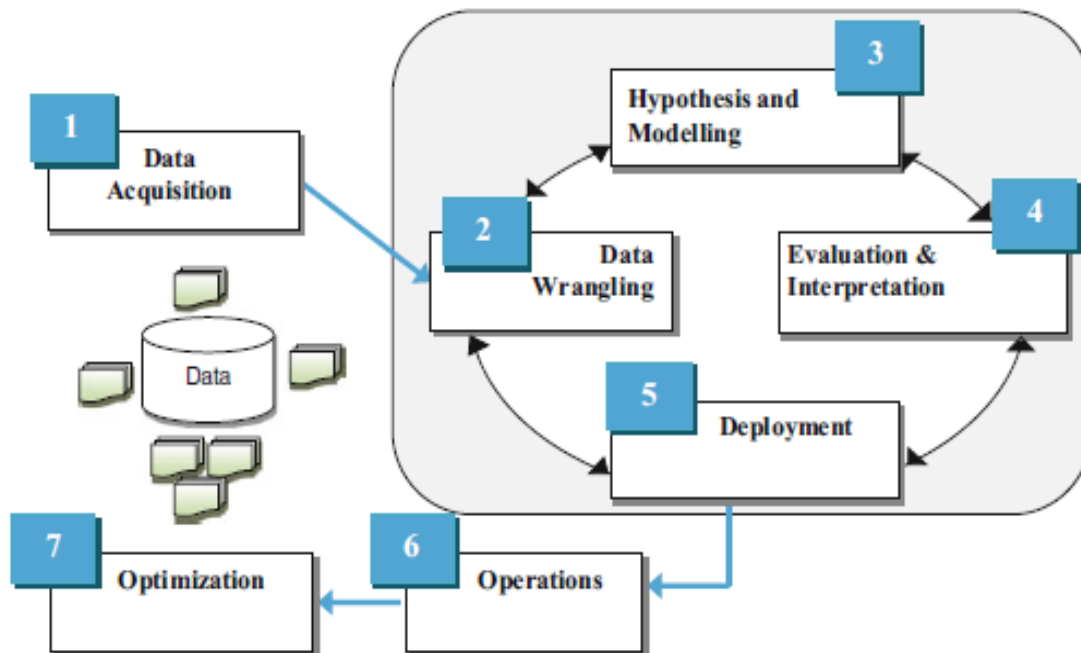
3.1 Entity Relationship Diagram:

3.2 Data Flow Diagram:

Since we are in the level 0 of our project so following is the level 0 DFD:-

STUDENT PERFORMANCE REVIEW

STUDENT PERFORMANCE SYSTEM

STUDENT INFORMATION

ORGANISATION SATFF

4. Scheduling and Estimation:

- **March-**
  Searched for the idea on which we have to work, the idea which has to be real life problem and defining the objective and scope of the idea for the project.

- **April-**
  Made the report and set the limits of the project on which we have to work.

- **May & June-**
  Data Acquisition comprises obtaining data from both internal and external sources, including social media or web scraping. Data may be generated by devices, experiments, sensors, or supercomputer simulations. We are gathering the data from the students and also requesting the data from the university.

- **July-**
  Data wrangling includes cleaning the data and reshaping it into a readily usable form for performing data science. This phase covers various activities to construct the final data set (data that will be fed into the modelling tools) from the initial raw data. This activity can be performed multiple times, and not in any prescribed order. Tasks also include table, record, and attribute selection, as well as transformation and cleaning of data for modelling tools.

- **August-**

Hypothesis and Modelling are the standard data mining steps—but in a data science project, these are not limited to statistical samples. In this phase, various modeling techniques are selected and applied, and their parameters are calibrated to optimal values. There are several techniques for the same data mining problem type. Some techniques have specific requirements on the form of data. Therefore, stepping back to the data preparation phase is often needed. A vital sub-phase is achieved here for model selection. This contains the partition of a training set for training the candidate models, and validation/test sets for evaluating model performances and selecting the best performing model, gauging model accuracy and preventing over fitting.

- **September-**
  Evaluation and Interpretation: Evaluating the Modelling phase again so that if any mistake or if we want some modification we can do that in this time also interpreting the model we have done in the previous phase.

- **October-**
  The moment the model has been deployed in production, it is time for regular maintenance and operations. This operations phase could also follow a target model which gets well with the continuous deployment model, given the rapid time-to market requirements in data-driven projects. The deployment contains performance tests to measure model performance, and can trigger alerts when the model performance lowers beyond a certain standard threshold.

- **November-**
  The optimization phase is the final step in the data science project life cycle. This could be triggered by inadequate performance, or the need to add new data sources and retraining the model, or even to deploy a better model based on enhanced algorithms.

- **December-**
  This is the time when we will be ready to present our analysed and predicted model for the performance of the student.