

MES COLLEGE OF ENGINEERING, KUTTIPPURAM
DEPARTMENT OF COMPUTER APPLICATIONS
20MCA245 – MINI PROJECT

PRO FORMA FOR THE APPROVAL OF THE THIRD SEMESTER MINI PROJECT

(Note: All entries of the pro forma for approval should be filled up with appropriate and complete information. Incomplete Pro forma of approval in any respect will be rejected.)

Mini Project Proposal No: _____
(Filled by the Department)

Academic Year : 2021-2022

Year of Admission: 2020

1. Title of the Project : STUDENT PERFORMANCE PREDICTION

2. Name of the Guide : _____

3. Number of the Student: MESMCA20-2044

4. Student Details (in BLOCK LETTERS)

Name SARANYA.S

Roll Number 44



Signature

1. _____

Date: 7/12/2021

Approval Status : Approved / Not Approved

Signature of
Committee Members }

Comments of The Mini Project Guide

Dated Signature

Initial Submission :

First Review :

Second Review :

Comments of The Project Coordinator

Dated Signature

Initial Submission:

First Review

Second Review

Final Comments:

Dated Signature of HOD

STUDENT PERFORMANCE PREDICTION

SARANYA.S

Introduction:

In present educational systems, student performance prediction is getting worsen day by day. Predicting student performance in advance can help students and their teacher to keep track of progress of a student. This project intends to approach student achievement in schools using machine learning techniques. The real-world data (e.g. student grades, demographic, social and school related features) will be collected by using school reports and questionnaires. This is modelling under binary/five-level classification and regression tasks. Also, four DM models (i.e. Decision Trees, Random Forest, Logistic Regression and Support Vector Machines) and three input selections (e.g. with and without previous grades) will test. Although student achievement is highly influenced by past evaluations, an explanatory analysis has shown that there are also other relevant features (e.g. number of absences, parent's job and education, alcohol consumption). As a direct outcome of this project, more efficient student prediction tools can be developed, improving the quality of education and enhancing school resource management.

Objectives:

- The aim is to predict student achievement and if possible to identify the key variables that affect educational success/failure.
- This project will be modelled under three DM goals:
 - i) Binary classification (pass/fail);
 - ii) Classification with five levels (from I very good or excellent to V - insufficient);
 - iii) Regression, with a numeric output that ranges between zero (0%) and twenty (100%).
- For each of these approaches, three input setups (e.g. with and without the school period grades) and four DM algorithms (e.g. Decision Trees, Random Forest) will be tested. Moreover, an explanatory analysis will be performed over the best models, in order to identify the most relevant features.

Problem Definition:

In this project, we will analyze recent real-world data from two Portuguese secondary schools. Two different sources were used: mark reports and questionnaires. Since the former contained scarce information (i.e. only the grades and number of absences were available), it was complemented with the latter, which allowed the collection of several demographic, social and school related attributes (e.g. student's age, alcohol consumption, mother's education).

Basic functionalities:

Predicting student performance in advance can help students and their teacher to keep track of progress of a student. It learns iteratively over time by observing different examples, similarly to how students can perform in schools by using their marks, other activities and social behaviour. Using this application, we have to mould our children with a better and easy way. And it will be essential for a productive country.

Tools / Platform, Hardware and Software Requirements:

Hardware Requirements:

Processor: i3

Hard Disk: 500 GB

RAM: 4 GB

Software Requirements:

Language: Python

Front End: Python-django

Back end: SQLite

Dataset: Students' Academic Performance Dataset from Kaggle Website.

Techniques Used: Binary/five-level classification and regression tasks. Also, four DM models (i.e. Decision Trees, Random Forest, Logistic Regression and Support Vector Machines).

IDE: Visual Studio Code

OS: Windows/Linux