## MES COLLEGE OF ENGINEERING-KUTTIPPURAM

DEPARTMENT OF COMPUTER APPLICATIONS

RLMCA 352 PROJECT AND VIVA-VOCE

## PRO FORMA FOR THE APPROVAL OF THE FINAL SEMESTER PROJECT

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

|  |  |  |
| --- | --- | --- |
| |  | | --- | | Project Proposal Number:  *(Filled by the Department)* | | Academic    Year    :  2018-2021    Year    of    Admission  :    2018        Admission    Number        :      15204    Roll    Number    :  18MCA11013    Register    Number    :        MES18MCA013 |
| |  |  | | --- | --- | | E-Mail | : jincykv1998@gmail.com | | Mobile No. | : 9020334773 | |

* 1. Name of the Student (in BLOCK LETTERS) : JINCY K V
  2. Title of the Project : Academic Performance Prediction Based on Multisource, Multi feature Behavioural Data

Date : 15-03-2021 Signature of the Student:

**Comments of The Project Guide**

1. Name of the Guide Assigned :

(Internal-Department)

1. Thrust Area of the Project :

Initial Submission :

|  |  |  |  |
| --- | --- | --- | --- |
| Approval Status : Approved / Not Approved      First Review :    Second Review : | Dated Signature of | Guide | HOD |
| **Comments of The Project Coordinator**  Initial Submission:  First Review  Second Review | Dated Signature of | Projet Coordinator: | or: |

ABSTRACT

## Academic Performance Prediction Based on Multisource, Multifeature Behavioural Data

## JINCY K V, MES18MCA013, 18MCA11013

**Introduction:** Digital data trails from disparate sources covering different aspects of student life are stored daily in most modern university campuses. However, it remains challenging to (i) combine these data to obtain a holistic view of a student, (ii) use these data to accurately predict academic performance, and (iii) use such predictions to promote positive student engagement with the university. To initially alleviate this problem, in this project, a model is proposed. In our project, (1) first, an experiment is conducted based on a real-world campus dataset of college students that aggregates multisource behavioural data covering not only online and offline learning but also behaviours inside and outside of the classroom. Specifically, to gain in-depth insight into the features leading to excellent or poor performance, metrics measuring the linear and nonlinear behavioural changes (e.g., regularity and stability) of campus lifestyles are estimated; furthermore, features representing dynamic changes in temporal lifestyle patterns are extracted by the means of long short-term memory (LSTM). (2) Second, machine learning-based classification algorithms are developed to predict academic performance. (3) Finally, visualized feedback enabling students (especially at-risk students) to potentially optimize their interactions with the university and achieve a study-life balance is designed.

**Objectives**: Many academic performance prediction systems have been developed for college students, the following challenges persist: (i) capturing a sufficiently rich profile of a student and integrating these data to obtain a holistic view; (ii) exploring the factors affecting students’ academic performance and using this information to develop a robust prediction model with high accuracy; and (iii) taking advantage of the prediction model to deliver personalized services that potentially enable students to drive behavioural change and optimize their study-life balance. To address these challenges, four representative prediction systems (including one online system and three offline systems) are used.

**Motivation or Relevance**: As an important issue in the education data mining field, academic performance prediction has been studied by many researchers. However, due to lack of richness and diversity in both data sources and features, there still exist a lot of challenges in prediction accuracy and interpretability. To initially alleviate this problem, our project aims at developing a robust academic performance prediction model, to gain an in-depth insight into student behavioural patterns and potentially help students to optimize their interactions with the university. In our project, a model is proposed to predict the academic performance of college students.

**Problem Definition and Draft Product Backlog**: A model is proposed to predict the academic performance of college students. Our contributions in this project are related to three sources. First, regarding data fusion, to the best of our knowledge, this work is the first to capture, analyse and use multisource data covering not only online and offline learning but also campus-life behaviours inside and outside of the classroom for academic performance prediction. Based on these multisource data, a rich profile of a student is obtained. Second, regarding the feature evaluation, behavioural change is evaluated by linear, nonlinear, and deep learning (LSTM) methods respectively, which provides a systematical view of students’ behavioural patterns. Specifically, it is the first time that three novel nonlinear metrics (LyE, HurstE, and DFA) and LSTM are applied in students’ behavioural time series analysis. Third, our experimental results demonstrate that this model can predict academic performance with quite high accuracy, which help to formulate personalized feedback for at-risk (or unself-disciplined) students.

**Tools / Platform, Hardware and Software Requirements**:

HARDWARE REQUIREMENTS**:**

* System : x86 64-bit CPU (Intel / AMD architecture)
* Hard Disk : 5 GB.
* Input Devices : Keyboard, Mouse
* Ram : 4 GB

SOFTWARE REQUIREMENTS**:**

* Operating system : Windows 10.
* Coding Language : Python