

ESCORT – UNIVERSITY STUDENTS GUIDANCE SYSTEM BASED ON NLP

2022-179

Project Proposal Report

Shanghavi.R

B.Sc. (Hons) Degree in Information Technology
(Specialization in Software Engineering)

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

February 2022

ESCORT – UNIVERSITY STUDENTS GUIDANCE SYSTEM BASED ON NLP

2022-179

Project Proposal Report

Shanghavi.R

B.Sc. (Hons) Degree in Information Technology
(Specialization in Software Engineering)

Department of Information Technology


Sri Lanka Institute of Information Technology

Sri Lanka

February 2022

Declaration

I declare that this is my own work, and this proposal does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any other university or institute of higher learning, and to the best of my knowledge and belief, it does not contain any material previously published or written by another person except where the acknowledgment is made in the text.

Name	Student ID	Signature
Shanghavi.R	IT19155708	

The supervisor/s should certify the proposal report with the following declaration. The above candidate is carrying out research for the undergraduate Dissertation under my supervision.

Signature of the supervisor :

Date:

Abstract

The ultimate goal of any university is to offer the best educational experience and knowledge to its students. Identifying the students who need extra support and recommending the appropriate solution to enhance their performance plays an important role in achieving that goal. Students' performance plays an important role in educational institutions and the economic growth of society by producing graduates [1].

Predicting students' performance is very important in matters related to higher education [2]. Early detection of students at risk, along with preventive measures, can drastically improve their success. That is essential in order to help at-risk students and assure their retention while providing excellent learning resources and experience and improving the university's ranking and reputation [3]. Assignments, lab tests, mid exams, and final exams all have an impact on a student's academic success or failure.

Predicting academic performance and recommending E-learning resources are critical considerations for students who want to enhance their grades. One of the benefits of the student's prediction is that it reduces the official warning signs as well as expels students because of their inefficiency. Effective performance prediction approaches help to recommend solutions and resources more accurately. The existing literature is primarily focused on predicting academic results and how effectively a teacher might design a specific course to help students improve their academic performance.

This research will assist students to identify learning gaps and determine ways to resolve such gaps. Having a Natural Language Processing and Machine Learning system will effectively help university students. Our study's main goal is to use classification approaches to uncover students' performance and help recommend e-learning resources. The proposed system will offer student performance prediction through the rules generated via data mining techniques. The data mining technique used in this project is classification, which classifies the students based on their transcripts.

Keywords - Machine Learning, Data mining technique, Performance, classification

Table of Contents

Declaration	iii
Abstract	iv
Table of content	v
List of Figures	vi
List of Tables	vii
List of Appendices	viii
1. Introduction	01
1.1. Background	01
1.2. Literature survey	02
1.3. Research Gap	03
1.4. Research Problem	04
2. Objectives	05
2.1. Main Objective	05
2.2. Sub Objectives	05
3. Methodology	06
4. Description of personal facilities	08
5. Budget and Budget Justification	09
References	10
Appendices	
Appendix A: Work Breakdown Structure	11
Appendix B: Gantt Chart	11

List of figures

	Page
Figure 3.1 System Overview Diagram	7

List of tables

	Page
Table 1.3.1 Comparison Table	3
Table 4.1 Personal Facilities	8

List of appendices

Appendix	Description	Page
Appendix - A	Work Breakdown Structure	11
Appendix – B	Gantt Chart	11

1. Introduction

1.1. Background

Education is one of the fundamental needs and a right of all human beings. Even though universities are considered excellent centers for education, some students are unable to gain the full benefits of the university for several reasons [4]. There are many common factors causing students in institutions of higher education to be distressed and perform poorly or drop out without completing their courses [4]. They face a lot of pressures and challenges that present various physical, social, and emotional challenges. It is important to provide better service to university students to address their concerns. We provide a solution for students with communication problems with the administration, who struggle with psychological problems, who struggle to choose the right career path, and who are unable to identify the problems and weaknesses in their education.

Student academic performance in higher education is extensively researched to tackle academic underachievement, increased university dropout rates, and graduation delays, among other tenacious challenges [5]. Simply put, student achievement refers to the extent to which students meet short- and long-term educational objectives. The academic performance of students must be assessed to promote growth and improve the quality of higher education.

Academic achievement is the level of achievement of the student's academic goal that can be measured and assessed through examinations, assignments, and other forms of measurement. Early prediction of performance is essential for maintaining students on a progressive path. If a student's performance is predicted well in advance, it can help to maintain or improve the quality of teaching by predicting students' interests in subjects, student-level activities, and contributing to their performance in universities.

The analysis of student performance and the availability of e-learning resources can be well planned during their institutional study period [3]. Timely forecasting of student performance makes it possible to detect underperforming students and recommend appropriate learning materials. Students could also improve their learning activities on their own.

1.2. Literature Survey

Many studies have already been carried out in this field. Academicians measure student success from a variety of perspectives, ranging from students' final grades, grade point average (GPA), to future employment prospects [5]. According to a recent comprehensive survey, nearly 70% of the examined work looked into predicting student performance using grades and GPAs, whereas only 10% of the studies looked into predicting student achievement using learning outcomes [3].

A basic study was carried out to examine similar existing systems used to perform the analysis of student performance. Three existing systems were chosen because they are similar to our system.

(a) Faculty Support System (FSS)

Shana and Venkatachalam have proposed a framework named Faculty Support System (FSS), which is low in cost as it uses cost-effective open-source analysis software, WEKA, to analyze student performance in a course provided by the Coimbatore Institute of Technology of Anna University [7]. Data mining techniques like the classification technique are used in this search. The classification technique is used for forecasting student performance. In addition, FSS focuses on identifying factors that contribute to student performance in a specific course.

(b) Student Performance Analyzer (SPA)

The SPA is a secure online web-based software that allows teachers to view student performance and track school data. SPA is a tool designed to analyze, display, store, and collect feedback on student assessment data [8]. It is a powerful analysis tool used by schools around the world to perform analyses and display the analysis data once the raw data of the students is uploaded into the system. The analysis is done by tracking the student or class to get the overall performance of the student or class. It helps identify the performance of students that is below the expected level, at the expected level, or above the expected level. This would allow teachers or staff to easily assess the performance of existing students. Apart from that, it enables the generation of several types of student performance reports, such as progress reports and achievement reports.

(c) Intelligent Mining and Decision Support System (InMinds)

InMinds helps the University Malaysia Sarawak (UNIMAS) to monitor the performance of various fields in each department of UNIMA [9]. The system enables top and mid-management in UNIMAS to have a clear look at the areas that need attention by looking at the figures, revenues, and risks. The characteristics, ease of use, and flexibility of the system make it ideal to analyze performance in UNIMAS. Charts are provided by the system to aid in the interpretation of student achievement.

1.3. Research Gap

Predicting students' performance eventually became a difficult endeavor. Prediction algorithms now in use are insufficient to forecast students' performance in higher education institutions. Many studies have been published to address the challenge of predicting student performance, however they do not provide enough recommendations for students who receive low grades.

As a result, there was a clear need for more advanced approaches for predicting at-risk students and determining what factors influence their outcomes, which inspired us to conduct this research. This study contributes to the existing knowledge base by predicting student academic performance and assisting in the identification of students with poor grades who can then be reviewed and given new learning materials and strategies to help them improve their marks.

When we compare our research ideas to existing ones. Most of them are designed to predict school students' performance, but we will design for university students. Most of these systems predict performance using GPA, CGPA, grades, student demographics, and psychological attributes. Our system predicts students' performance based on transcripts, which include assignment marks, lab test marks, and mid exam marks, so it helps the students improve their final exam marks. In particular, we recommend appropriate online learning materials based on performance. This will help them concentrate on each subject and improve their performance.

The goal of this project is to create a system that can predict student academic achievement in a specific module by evaluating the students' performance using data mining classification algorithms. The purpose of this research is to forecast student performance, with a particular focus on identifying students who may fail to satisfy course requirements [11]. As a result, the system will predict students' performance and based on it, recommend learning resources. It was created to help students by allowing them to view their previous performance in a certain course and recommend learning materials to help them overcome their weak subjects.

Features	Faculty Support System (FSS) [7]	Student Performance Analyzer (SPA) [8]	Intelligent Mining and Decision Support System(InMinds) [9]	Student Performance Analysis System (SPAS) [10]	Escort
Designed for University Students	✗	✗	✓	✓	✓
Based on transcript	✗	✗	✗	✓	✓
Accurate performance prediction	✓	✓	✗	✗	✓
Categorizing before posting	✗	✗	✗	✗	✓
Recommending E-Learning Resources	✗	✗	✗	✗	✓

Table - 1.3.1

Comparison Table

1.4. Research Problem

Universities make it possible to improve knowledge, skills, and a way of life. However, nowadays, university students face many challenges [12]. They have difficulty contacting the university administration to address their educational needs. As a result, students are having difficulty obtaining administration-related information, such as payment problems, course details, IC status, and more.

More than that, students have difficulties in communicating language while sending emails and also when raising their problems about academic needs. Students who face psychological problems are increasing daily and also many students are facing problems in choosing an appropriate career path [13].

In addition, many students are not able to identify issues and weaknesses in their studies. They have a problem in referring to appropriate tutorials, videos, and documentation [14]. Most problems experienced by undergraduate students are due to a lack of knowledge about the content of lectures, unable to communicate with others regarding academic matters, and not having the appropriate time management for each topic. On the other hand, the continued failure of examinations by university students is the main issue for both teachers and students.

The lack of motivation with regard to academic career can also lead to a lower grade of undergraduate students. A good study design is an essential requirement for students to improve their academic performance. Students still have difficulty finding a suitable solution for their study plan. Even if they can develop a plan, it may not be efficient enough to improve their academic performance. Sometimes they do not give equal weight to all subjects. A good study plan and learning materials will give them the motivation to be responsible for their study and eventually will help them improve their academic performance.

2. Objectives

2.1. Main Objective

Our system assists students who are having difficulty with their grades and performance. It helps those who are struggling to find learning materials that meet their needs. Furthermore, most students find it difficult to recognize their weak subjects, so our method assists in identifying those modules as well as providing solutions for how to overcome them.

The main objective of this study is to identify and analyze students' academic performance based on their transcripts and suggest appropriate learning resources. This study is more valuable for identifying weak students and helping these students individually so that their performance improves in the future. Overall, this study met its goals of improving student performance, predicting at-risk and dropout students, and helping to enhance their final grades.

2.2. Sub Objectives

- Analyze assignment marks and mid exam marks in the transcript.
- Predicting performance level.
- Categorize wise performance prediction.
- Students' weakness identification.
- Recommending Learning materials based on performance.
- Helping to enhance final grade.

3. Methodology

Predicting students' performance in undergraduate studies is crucial for any university. Understanding students' performance in each module ahead of time is essential for supporting at-risk students in reducing the challenges they confront in their learning journeys and assisting them in excelling in the learning process.

The data collection was based on attributes such as assignment marks, mid-marks, and lab test marks to predict academic performance at the higher level of education. For this study, recent real data will be collected through a physical and online survey of university students.

Initially, when we upload a student's transcript, the system will be stored in the system database.

(a) Data Pre-processing and Balancing

The first step in data pre-processing. Data pre-processing is a major step in Machine Learning because the quality of data and the valuable knowledge that can be extracted from it directly affects our model's ability to learn. Its goal is to convert raw data into a format that can be used by mining algorithms. During this process, the following tasks will be completed.

- Data Integration.
- Data Cleaning.
- Discretization.

The data balancing approach is used to handle the class imbalance problem after data pre-processing. When the number of instances in one class is significantly less than the number of instances in another class or classes, this is known as a class imbalanced problem.

(b) Data Mining Approaches:

Data mining can be used as an effective method to identify and predict these dropouts. In educational data mining methods, predictive modeling is usually used in predicting student performance. In order to build predictive modeling, there are several tasks used, which are classification, regression, and categorization. The most popular task to predict students' performance is classification. There are several algorithms under classification tasks that have been applied to predict students' performance. Among the algorithms used are Decision tree, Artificial Neural Networks, Naive Bayes, K-Nearest Neighbor, and Support Vector Machine. We will use some of these data mining algorithms in our system for predicting student performance more effectively.

(c) Recommendation:

After predicting the student performance, the system will analyze the student weak area particularly and make recommendations like e- learning materials or resources to help them improve their academic achievement. The prediction results assist students in developing a good understanding of how well or poorly they will perform in a module and then taking appropriate steps.

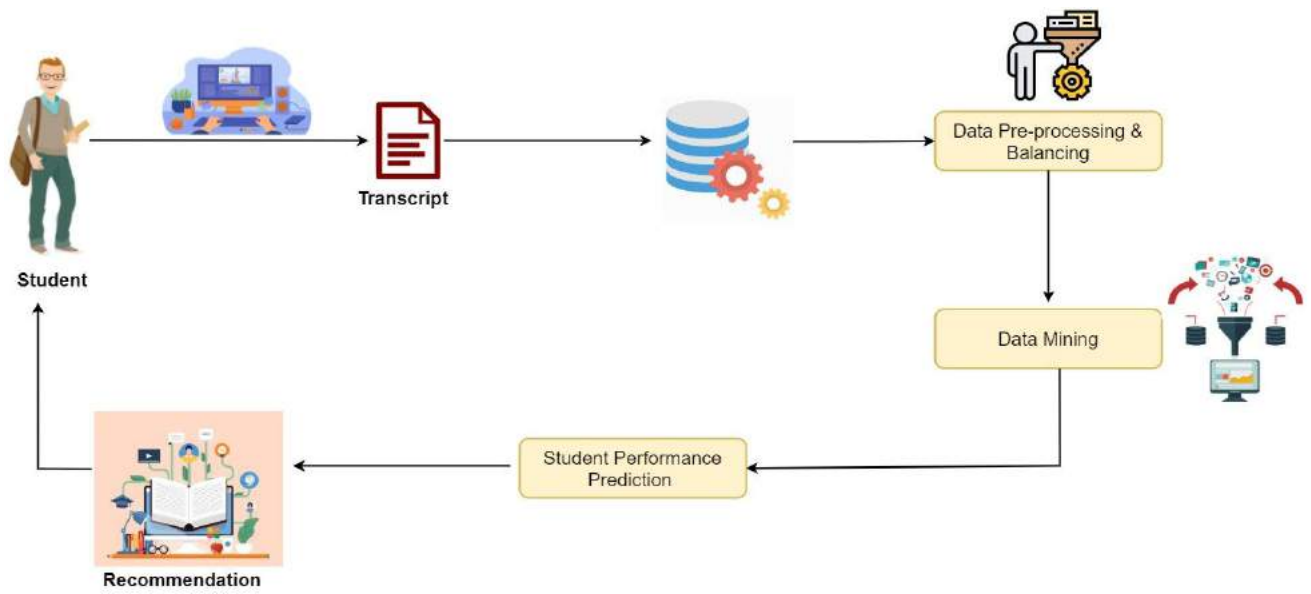


Figure - 3.1
System Overview Diagram

4. Description of personal facilities

Member	Component	Task
Shanghavi.R	Predict and analyze students' performance and recommend suitable Learning resources to the students.	<ol style="list-style-type: none">1. Information gathering and Analysis2. System Design3. ER Mapping and Database design4. Function Implementation<ul style="list-style-type: none">• Insert students' Transcript• Data Preprocessing & Balancing• Use data mining techniques• Predict students' performance• Output (recommended E-learning material links/ resources)5. Unit, Integration, and system testing Documentation

Table - 4.1
Personal Facilities

5. Budget and Budget Justification

The estimated cost of the proposed system for developing and testing the accuracy of the research can be categorized as follows.

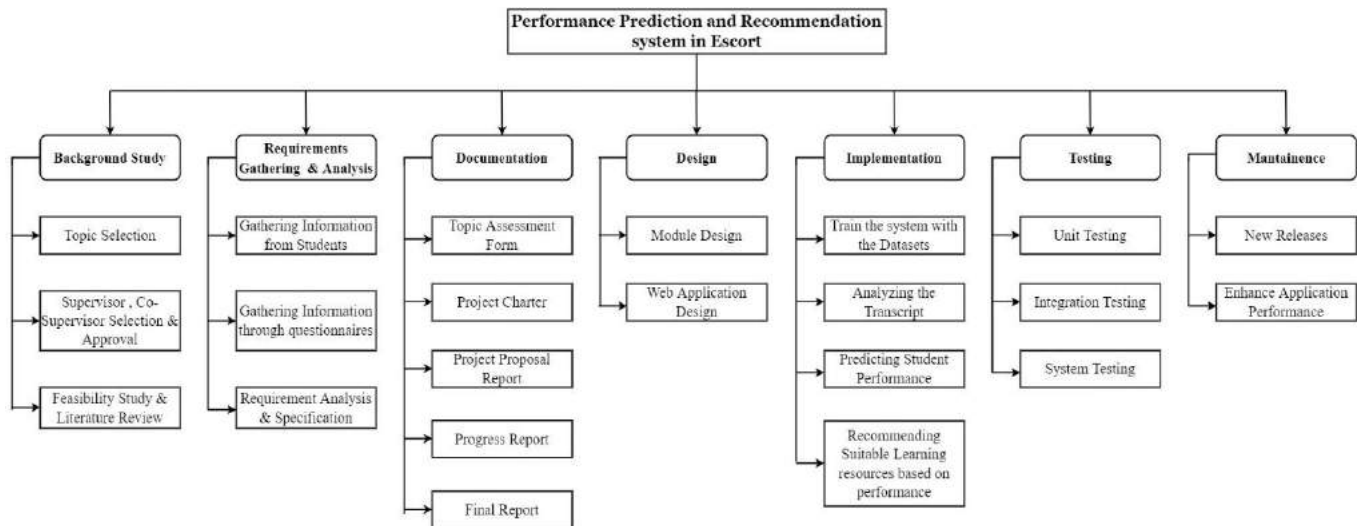
Internet usage for research	- Rs. 5000
Domain name registration (annual)	- Rs. 1500
Hosting (annual)	- Rs. 10000
Preparation for report and printing	- Rs. 1500
Other costs (Data Collection, Travelling Expenses)	- Rs. 10000
Total	- Rs. 28000

References

- [1] S. JothiLakshmi and M. Thangaraj, "Recommender System for Student Performance Using EDM," vol. 7, 3 December 2018.
- [2] S. Li and T. Liu, "Performance Prediction for Higher Education Students Using Deep Learning," 14 July 2021.
- [3] E. Alyahyan and D. Düşteğör, "Predicting academic success in higher education: literature review and best practices," 10 February 2020.
- [4] M. A. N. R. Perera, "Problems Faced by Undergraduates in the Learning Environment: Some Evidence from a Sri Lanka University," vol. 3, June 2013.
- [5] A. Namoun and d. A. Alshanqiti, "Predicting Student Performance Using Data Mining and Learning Analytics Techniques: A Systematic Literature Review," 29 December 2020.
- [6] A. Siddique, A. Jan, F. M. A. I. Qahmash, N. N. Quadri and d. M. O. A. Wahab, "Predicting Academic Performance Using an Efficient Model Based on Fusion of Classifiers," 13 December 2021.
- [7] J. Shana, and T. Venkatacalam, "A framework for dynamic Faculty Support System to analyse student course data", International Journal of Emerging Technology and Advanced Engineering, Vol. 2, No. 7, 2012, pp.478-482.
- [8] M. SRRICHARAN, M. SANKEPALLY, M. K. SAI, M. LAKKIREDDY and M. M. KUMAR, "STUDENT PERFORMANCE ANALYSER USING SUPERVISED LEARNING ALGORITHMS," 7 June 2021.
- [9] H. M. Ibrahim, S. M. Z. Harun, M. a. Latifah, L. A. a. F. Bungan, B. a. Hazrinda and H. Ali, "Intelligent Mining and Decision Support System (InMinds)," 12 April 2016.
- [10] C. L. Sa, D. H. b. A. Ibrahim, E. D. Hossain, and M. b. Hossin, "Student Performance Analysis System (SPAS)," January 2015.
- [11] H. Bydžovská, "A Comparative Analysis of Techniques for Predicting Student Performance," 2015.
- [12] Rohana Bandara Weerasooriya, "Major problems and issues in Sri Lankan university system- study focus on the student perspective", 2013
- [13] Wasantha Kalyani and Nirmal Chathuranga, "Career Choice Determinants of the Management Students: A Case of a Sri Lankan State University", 2021, Journal of Accounting Research, Business and Finance Management, Volume-2, Issue-1 (July-December 2021), e-ISSN:2582-8851
- [14] Innocent B. Mboya, Beatrice john, Eneck S. Kibopile, Lisbeth Mhando, Johnston George, and James S. Ngocho, "Factors associated with mental distress among undergraduate students in northern Tanzania", 2020

Appendices

Appendix A: Work Breakdown Structure



Appendix B: Gantt Chart

