Predicting CSE Undergraduate Students’ Careers Using Machine Learning Algorithms in the Context of Bangladesh

Israt Jahan Reshma

Bangabandhu Sheikh Mujibur Rahman Science and Technology University

CSE 488

Author Note

Include any grant/funding information and a complete correspondence address.

Abstract

Keywords: Add keywords here.

# INTRODUCTION

Computer Science and Engineering (CSE) has emerged as one of the most popular and rapidly evolving fields among Bangladeshi students. Alongside other engineering disciplines such as electrical and civil engineering, CSE is often considered a prestigious and high-potential career path. CSE is a highly sought-after discipline, with students driven by various factors such as job security, high earning potential, and social prestige. However, the rapid advancement of technology, diverse career options, and intense competition in the job market make this decision increasingly complex. Many students find themselves uncertain about which career path aligns best with their skills, interests, and personal values.

Students often grapple with significant external pressures, particularly from family expectations. In a society where job security and financial stability are highly valued, parents often encourage their children to pursue careers in engineering, and CSE stands out as a prime choice. The rapidly changing tech landscape adds another layer of complexity, as students must continually update their knowledge and skills to stay relevant. This creates a challenging environment where career decisions are sometimes made based on external influences rather than personal interest, resulting in a disconnect between one’s job and passions. Consequently, even students who secure prestigious positions at well-known companies may experience dissatisfaction if their roles do not align with their genuine interests.

Despite these hurdles, a shift is occurring among the new generation of Bangladeshi CSE students. More and more are choosing career paths that resonate with their personal interests and strengths, whether in software development, data science, artificial intelligence, or other emerging tech fields. This growing awareness emphasizes the importance of aligning one's career with their skills and passions, not just external expectations. As a result, students are becoming more mindful of making informed and fulfilling career choices, reflecting a broader, more thoughtful approach to professional development in the Bangladeshi context.

< ???????????????????????????????????????>

This study explores the development of a machine learning model designed to predict the most suitable career paths for Bangladeshi CSE undergraduate students upon graduation. The model leverages a comprehensive data collection form that gathers information on students’ personal and academic profiles. It includes details such as CGPA, academic skills, proficiency in key technical areas, soft skills, number of research publications and projects, and individual career preferences. Additionally, the model considers external factors, like family expectations and the desire for a balanced work-life, which can heavily influence career choices.

By analyzing these factors, the machine learning model aims to provide personalized career recommendations that align with each student’s strengths and aspirations. The objective is to assist students in making informed and strategic career decisions, reducing the risk of career dissatisfaction and skill mismatch. Furthermore, this research offers insights into the current trends and challenges faced by CSE students in Bangladesh, contributing to the broader discourse on career planning and academic counseling in the region.

This study aims to determine whether students' academic performance in the field of Computer Science and Engineering (CSE) is influenced more by aptitude or personality traits. Additionally, it seeks to develop a predictive model to analyze student performance effectively. Machine learning is a powerful approach to automatically analyzing data and using it to make accurate predictions. By processing large datasets, patterns and rules can be identified, which are then used to characterize new data automatically. This process assists in continuously refining data analysis, ultimately supporting students in improving their learning activities and helping them better understand potential career paths.

Various classification techniques, such as K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Stochastic Gradient Descent (SGD), Logistic Regression, Decision Trees, and Neural Networks, are employed to evaluate educational performance on a broader scale. The model developed focuses on student-related variables, including skills, interests, hobbies, and academic achievements, to predict future career opportunities. The learned rules and patterns are presented using decision trees to make the analysis interpretable.

In this study, the SVM algorithm is emphasized to enhance prediction accuracy. The application of machine learning techniques encompasses data analysis, visualization, performance prediction, feedback provision, and student grouping. Student performance data is pre-processed through feature extraction and selection, which involves data cleaning, tokenization, sentiment analysis, and the removal of irrelevant words. The culmination of these processes yields a comprehensive understanding of students' academic performance, guiding them toward the most suitable career options in engineering and technology after graduation.

The structure of this paper is as follows:

* **Section 2** literature review
* **Section 3** proposed work
* **Section 4** methodology
* **Section 5** results and discussions
* **Section 6** conclusion
* **Section 7** reference

# Literature Review

Ade & Deshmukh (2015) suggested an incremental ensemble learning approach for forecasting student career paths. They used voting to merge three classifiers (Naïve Bayes (NB), K-Star, and SVM). From these three classifiers, the accuracy percentage was respectively 89.6%, 89.2%, and 89.2%. They highlighted the importance of considering interest, talent, and projected growth in career choice.

Iqbal et al. (2017) explored various machine learning techniques, such as matrix factorization, classification, and regression, to predict student grades across different courses. Using data collected from ITU Pakistan, the study evaluated the performance of these methods and concluded that RPM is the most effective among the models analyzed.

Madhan and Reddy (2021a) have implemented a system for predicting suitable careers for computer science students based on students' skills using XGBoost and Decision Trees. They collected over 30 thousand records from students in the form of ratings in various required fields of computer science student. However, they did not report the accuracy of the algorithms used, nor did they provide an analysis of the model's performance metrics, such as the confusion matrix and other evaluative components.

Vaidu et al. have implemented machine learning techniques based on student performance to predict their employability skills. They have used KNN and Naïve Bayes models to classify the students into numerous groups. The prediction of the students’ employability from the KNN algorithm is 95.33% accurate, which is all for the Naïve Bayes is 67.67% accurate. (vadiu, 2017)

predicting the career an engineering student can select after their graduation using machine learning classification techniques and find the factors that can affect students’ decision to choose the right career path

References

Last Name, A. B. (Year). Article Title. Journal Title, Pages #-#. URL. URL.

Last Name, C. D. (Year). Book TitleBook Title URL.

Last Name, D. E., Last Name, F. G. (Year). Report TitleReport Title URL.

Last Name, H. I. (Year, Month Day). Article Title/Headline. Periodical.Periodical.

Organization Name. (Year, Month Day).Webpage Title. URL.

Footnotes

1For APA reports, add footnotes manually on their own page following references. Do not use the **Insert** **Footnotes** method on the **References** tab as they will not be formatted correctly. For APA formatting requirements, it’s easier to type your own footnote references and notes. To format a footnote reference, select the number and then, on the **Home** tab, in the **Styles** gallery, click **Footnote Reference**. The body of a footnote, such as this example, uses the **Normal** text style. If you delete this sample footnote, don’t forget to delete its in-text reference at the end of the sample Heading 2 paragraph on the first page of body content in this template.

Tables

Table 1

Table Title

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Column Head | Column Head | Column Head | Column Head | Column Head |
| Row Head | 123 | 123 | 123 | 123 |
| Row Head | 456 | 456 | 456 | 456 |
| Row Head | 789 | 789 | 789 | 789 |
| Row Head | 123 | 123 | 123 | 123 |
| Row Head | 456 | 456 | 456 | 456 |
| Row Head | 789 | 789 | 789 | 789 |

Note: Place all tables for your paper in a tables section, following references and footnotes. Start a new page for each table, include a table number and table title for each, as shown. All explanatory text appears in a table note that follows the table, like this one. Use the **Table/Figure** style, available on the **Home** tab, in the **Styles** gallery, to get the spacing between table and note. Tables in APA format can use single or 1.5 line spacing. Include a heading for every row and column, even if the content seems obvious. A default table style has been set up for this template that fits APA guidelines. To insert a table, on the **Insert** tab, click **Table**.

Figures Title

Figure 1.

Include all figures in their own section, following references, footnotes, and tables. Include a numbered caption for each figure. Use the Table/Figure style for easy spacing between figure and caption.



*For additional information on APA Style formatting, please consult the* [*APA Style Manual, 7th Edition*](https://apastyle.apa.org/style-grammar-guidelines)*.*