

Chapter 1

Introduction

Bangladesh University of Engineering and Technology(*BUET*) is a renowned Technological University of *Bangladesh*. Brilliant students around the country get admitted here for quality education. The huge amount of BIIS data represent the states of all the students of BUET in academic education. For a large educational institute like public university which generates large volumes of data, it requires an efficient way to apply data mining techniques for obtaining knowledge on the development and performance improvement of academic activities. The knowledge acquired from BIIS can be sufficient to look for answers to such questions as: which factors determine better or worse academic performance of students? Which factors in a Department can be crucial for quality education? Concepts and techniques of data mining are essential to discover the hidden knowledge from large datasets [1] .

1.1 Problem Definition

Every year BUET enrolls almost 1000 brilliant Engineering minded student into 11 different departments from all over the country and simultaneously almost 1000 students graduate from BUET. The BIIS data holds all the academic information of individual student in detail. Statistics from BIIS data shows that academic performance of a student varies from department to department. After close observation, it has been revealed that department might not be the only factor behind the performance of a student. There might be some other reasons such as CGPA, Hall Status, Gender, Class Test Marks, Attendance Status which affect a student's performance. Only statistical analysis is not sufficient for finding out the knowledge from the BIIS data. The hidden knowledge inside the institutional academic and personal data of students is necessary to find out the possible effects on any students

academic performance. That is why knowledge discovery and data mining from academic data is essential for educational institution like BUET to improve academic performance of students as well as reshaping the decision makings for the betterment of the institution. Data mining techniques are very effective for discovering the hidden knowledge from educational data and applying it properly for the decision making. But all the data mining techniques can not be applied directly on academic data because of complex structure. This requires rigorous preprocessing. Bringing all the relevant data in a useful scope and applying classification methods on them are other problems of this research.

1.2 Scope of the Work

The BIIS data represents the detailed academic performance of an individual student throughout the undergrad period. Every year almost 1000 students are enrolled into BUET whereas almost 1000 students are graduated simultaneously. These students are under 5 different faculties. These five faculties are Faculty of Architecture & Planning, Faculty of Civil Engineering, Faculty of Mechanical Engineering, Faculty of Engineering and Faculty of Electrical & Electronic Engineering. There are 11 different departments under these faculties which are Department of Electrical & Electronic Engineering, Department of Computer Science & Engineering, Department of Civil Engineering, Department of Mechanical Engineering, Department of Architecture, Department of Water Resource Engineering, Department of Naval Architecture & Marine Engineering, Department of Chemical Engineering, Department of Materials & Metallurgical Engineering, Department of Urban & Regional Planning and Department of Industrial & Production Engineering. [2] The scope of knowledge discovery from BIIS data is immense in context of undergraduate students. In this research BIIS data of 10 already graduated batches have been used which represents almost 10 thousand students. No current student has been brought under this research.

1.3 Objectives

Every year, a number of brilliant minded students are admitted into BUET, but at the end of their academic period, all of them can not utilize the full potential, it affects their academic performance. The main objectives of this research study are

- To find out the effect of different factors to the performance of a student.

- To discover knowledge of students academic performance and personal statistics through the impact of different assessment and factors e.g. Class Test, Attendance, CGPA, Credit Completion etc.
- Compare the impact of different factors e.g. Hall Status, Gender, Class Test, Attendance, Departments on the academic performances of a student

1.4 Thesis organization

We have developed a technique to discover knowledge using ID3 classification algorithm from institutional data of students who have completed their undergraduate in the department of CSE, BUET. All the literature studies e.g., preprocessing of raw data, preliminaries of Knowledge Discovery and Data Mining, Building decision tree, ID3 algorithm and related works have been elaborated in chapter 2. Description of the raw academic and analysis have been described in chapter 3. Basically, the raw BIIS data is not suitable for implementation of necessary procedures, the status of the data has been elaborated in chapter 3. The existing raw data needs to be rigorously preprocessed before the implementation of classification. The step by step methodologies and the result of this preprocess has been described in chapter 4. On chapter 5, the implementation of ID3 classification algorithm on the test data using training set has been elaborated. The result of the procedures taken has also been depicted in this section in detail. Finally, we have illustrated the summary of the findings along with quantitative analysis. We have also encompassed the scope of the extension of this research work by illustrating some significant future works in chapter 6.

Bibliography

- [1] Han J., Kamber M. and Pei J. *Data Mining: Concepts and Techniques*. Morgan Kauffman Publishers, San Francisco, 2011.
- [2] Bangladesh University of Engineering and Technology, General Information on: <http://www.buet.ac.bd>