IST 707 Data Analytics

Student Grade Prediction Academic Project

(Group Project)

Course Learning Objectives:

* Document, analyze, and translate data mining needs into technical designs and solutions
* Apply data mining concepts, algorithms, and evaluation methods to real-world problems
* Employ data storytelling and dive into the data, find useful patterns, and articulate what patterns have been found, how they are found, and why they are valuable and trustworthy

Course skill sets:

* data mining, including data preparation, concept description, association rule mining, classification, clustering, evaluation and analysis
* Be able to apply data mining skills to business, science or other organizational problems

The goal for this project is to extract valuable insight from a data set through implementing the data mining skills, including Data Preprocessing, Visualization, and Machine Learning Algorithms.

In this project, I have learnt how to preprocess a dataset in R and transform the data through functions or packages to perform Machine Learning Algorithms in order to get desired predication outcome. The Machine Learning Algorithms implemented in the project include association rule mining, random forest, K-means clustering, Naïve Bayes and Support Vector Machine.

The Student Grade Prediction dataset is from Kaggle (<https://www.kaggle.com/dipam7/student-grade-prediction>). We have done preprocessing on the dataset, as well as the data transformation. We converted the dataset into desired format for each algorithm. For clustering analysis, we converted nominal and ordinal variables into numeric variables. And for association rule mining analysis, we converted all attributes into factor variables. We generated several plots, and used Machine Learning Techniques, such as clustering analysis and association rule mining analysis, to answer our research questions. And further, we successfully build a Support Vector Machine model with 96% of accuracy on predicting if a student is going to pass or fail the final exam.