Student Number: Student Name: Date

CS5483 Lecture Review Question 7

What is the rational of having various data mining technique? In other words, how can one decide which technique of the following to select in data mining?

Association rules

Clustering

Decision Tree

Neural network

Web Mining

Genetic programming

What are the major difference between Apriori algorithm and Frequent Pattern Tree (FP-tree) with respect to performance? Justify your answer.

Model answer

Different data mining techniques are applied to meet various user application requirements as follows:

Data Mining Technique	Rationale	Applications example
Association rules	To predict the probability of two events occurrences together	Cross selling
Clustering	To categorize each data instance into different groups	High/Medium/Low Risky slopes categorization
Decision Tree	To rank the priorities of important factors for decision making	Important factors for credit card fault
Neural network	To weight a profile of factors on a prediction	Customer profile analysis
Web Mining	Data mining on the web	Web pages/content design
Genetic programming	To use machine learning program for prediction	Computer Chess Game

The major difference between Apriori and Frequent Pattern Tree is that even though they both aim to derive frequent pattern data set. However, Apriori algorithm includes all source data set (i.e. non-candidate dataset) as frequent data pattern for data analysis while Frequent Pattern Tree (FP Tree) only considers candidate data set (valid data set according to users requirements) as frequent data pattern for data analysis. As a result, FP tree can perform much better than Apriori algorithm to derive Frequent Pattern data.