Data Challenge - Executive Summary

CS F415 - Data Mining

Abhay Koushik --- 2015A7PS0056P

Gautham V --- 2014B4A70637P

Problem 1:

Objective: To find a dynamic pricing scheme to increase profits, while at the same time minimizing a given penalty value.

After preprocessing, the following techniques were used to obtain inferences for decisions:

- **Decision Tree Classifiers**:
    - **Done with**: IBM SPSS Modeler
    - **Important Parameters** : Rule set, Non-partitioned data, no cross-validation, expert, minimum records per branch = 30
    - **Inferences Obtained** : Item no. 128, followed by 154, 76, 169, 134 are most frequent, and rules for them were found
- **Association Rule Mining**:
    - **Done with**: R with the *apriori* package
    - **Important Parameters** : support = 0.1, maxlen = 5
    - **Inferences Obtained** : Association rules were derived along with frequent itemsets, and after sorting with respect to support value and penalty weight, the best rules were combined to put together the result.

In addition to the above, certain statistical measures about the training data were found.

Decision Made: The updated *newPrices.csv* has been attached with the decisions made from these classifiers.

- **Profit** : A profit of **6.95%** was obtained.
- **Penalty** : The computed penalty value was **259996**.