Exploring Diverse Frequent Patterns in Classification

Mayank Gupta (201101004), Ankush Jain (201101010)

Problem

To extend the notion of diverse-frequent patterns and explore their applications in classification

Diversity - measure of the variety in a transaction

Intuitively, $\{Soap, Chocolate, Bucket\}$ is more diverse than $\{Bread, Butter, Milk\}$ since it contains items of more diverse categories.

Background

Frequent Patterns - patterns that occur frequently in a transaction. Example, dataset:

{A, B, C}
{A, D, E}
{A, F, G}
{A. B, D}
{A, C, D}

Here, {A, D} occurs very frequently in the above transactions, and is hence called a frequent pattern.

A threshold is defined for a pattern to be considered a frequent pattern. A property, *Support* is defined as

$$S = \frac{(No.\ of\ transactions\ containing\ pattern)}{Total\ Number\ of\ transactions}$$

A minSupport is defined by a user, and only those patterns that have a support > minSupport are considered.

We then derive association rules from frequent patterns. A rule is of the format

$$A \Rightarrow B$$

where $A \cup B$ is a frequent pattern

Since a large number of rules are generated, we filter rules by a parameter called confidence.

The confidence of an association rule is defined as

$$C = \frac{S(A \cup B)}{S(A)}$$

where $S(A)$ is support of A

Work Done

- Read the related research papers
    - Fast Algorithms For Mining Association Rules - Agarwal et al
    - Mining Frequent Patterns without Candidate Generation - Jiawei Han et al
    - Discovering Diverse-Frequent Patterns in Transactional Databases - Somya Srivastava et al

- Implemented the Apriori Algorithm

- Tested the implementation using a 100,000 line dataset from FIMI Repository and analysed results
- Compared results against standard Apriori implementation

Work Ongoing

- Read the research paper *Extracting Diverse-Frequent Patterns with Unbalanced Concept Hierarchy, M. Kumaraswamy et al*
- Write an implementation for mining Diverse-Frequent patterns using Unbalanced Concept Hierarchy