

## ***Abstract***

The main goal of this project work was to develop, implement and evaluate a fast and efficient algorithm that enables mining association rules from datasets that contain items with real values. This was accomplished by extending and enhancing the Apriori Algorithm, which is the standard algorithm to mine association rules. Here, we deal with real attributes by fine partitioning the values of the attribute. We also introduce measures to determine the number of partitions for each attribute. The major drawback of basic data mining techniques is that after applying data mining on the quantitative data, the number of extracted association rules is very huge. When many association rules are obtained, the usefulness of them will be reduced. In this report, we introduce an approach to reduce and summarize the extracted association rules. We use an iterative approach, in which first with the help of a training dataset the set of all possible association rules is mined. From this set we keep only a subset of rules with very high confidence and which can be used to identify a class uniquely. This set of rules is termed as strong rules. Next with the help of a testing dataset the strength of the rules is measured. The strength is calculated using misclassification error. If this measure crosses an user given threshold then we again repeat the whole process by resetting the minimum support and minimum confidence values. Finally, these rules are partitioned into different classes according to their ability to identify the respective class. Similar rules within a class are joined. Within a class these rules are arranged systematically to minimize the searching complexity.

***Keywords:*** data mining, apriori, association rule mining, rule reduction, rule classification.