Abstract

This paper presents a novel feature selection based on association rule mining using reduced dataset. The key idea of the proposed work is to find closely related features using association rule mining method. Apriori algorithm is used to find closely related attributes using support and confidence measures. From closely related attributes a number of association rules are mined. Among these rules, only few related with the desirable class label are needed for classification. We have implemented a novel technique to reduce the number of rules generated using reduced data set thereby improving the performance of Association Rule Mining (ARM) algorithm. Experimental results of proposed algorithm on datasets from standard university of California, Irvine (UCI) demonstrate that our algorithm is able to classify accurately with minimal attribute set when compared with other feature selection algorithms.

References

- Reunanen, J. (2003). &quot;Overfitting in making comparisons between variable
Feature Selection by Mining Optimized Association Rules based on Apriori Algorithm

- Thomas Drugman, Mihai Gurban and jean-Philippe Thiran, “Feature Selection and Bimodal Integration for Audio-Visual Speech Recognition”, School of Engineering-STI Signal Processing Institute.
- Gang Wang, Frederick H. Lochovsky, Qiang Yang, “Feature Selection with
Conditional Mutual Information MaxiMin in Text Categorization, Department of Computer Science, Hong Kong University of Science and technology, Kowloon, Hong Kong, 2004.
- Pang-Ning Tan, Michael Steinbach, Vipin Kumar, "Introduction to Data Mining," Addison Wesley.

Index Terms

Computer Science Algorithms

Keywords

Feature selection Association Rule Mining (ARM) Apriori Classification.