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

Feature Selection by Mining Optimized Association Rules based on Apriori Algorithm

- K. Z. Mao, &quot;Fast Orthogonal Forward Selection Algorithm for Feature Subset
- J. Jelonek, Jerzy S., &quot;Feature Subset Selection for Classification of Histological Images.
  Artificial Intelligence in Medicine&quot;, 1997. 9:22-239.
- B. Sahiner, H. P. Chan, N. Petrick, R. F. Wagner, and L. Hadjiiski, &quot;Feature
  Selection and Classifier Performance in Computer-Aided Diagnosis: The Effect of Finite Sample
  Size&quot; Medical Physics, 2000. 27(7): 1509-1522.
  Possible Subset Regression Models Using the {QR} Decomposition&quot; Parallel Computing,
  29, pp. 505-521.
- Gatu C. And Kontoghiorghes E. J. (2005). &quot;Efficient Strategies for Deriving the
  Subset {VAR} Models&quot; Computational Management Science, 2 (4):253-278.
  Computing the Best-Subset Regression Models&quot;. Journal of Computational and Graphical
- T. Joliffe, &quot;Principal Component Analysis&quot;, New York: Springer-Verlag,
  1986.
- K. L. Priddy et al., &quot;Bayesian selection of important features for feed-forward
- L. M. Belue and K. W. Bauer, &quot;Methods of determining input features for
- J. M. Steppe, K. W. Bauer Jr., and S. K. Rogers, &quot;Integrated feature and
- Q. Li and D. W. Tufts, &quot;Principal feature classification&quot;, IEEE Trans. Neural
- L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and
- Hanchuan Peng, Fuhui Long, Chris Ding, Feature Selection Based on Mutual
  Information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy, IEEE
  Transactions on Pattern Analysis and machine Intelligence, vol. 27, No. 8, August 2005.
- S Nojun Kwak and Chong-Ho Choi, Input Feature Selection by Mutual Information Based
  on Parzen Window, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24,
  No. 12, December 2002.
- Thomas Drugman, Mihai Gurban and jean-Philippe Thiran, &apos; Feature Selection and
  Bimodal Integration for Audio-Visual Speech Recognition&apos;, School of Engineering- STI
  Signal Processing Institute
- Georgia D. Tourassi, Erik D. Frederick, Mia K. Markey, Carey E., Floyd, Jr., &quot;Application
- Gang Wang, Frederick H. Lochovsky, Qiang Yang, &quot;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.