Abstract

Most of the data mining algorithm focuses on frequent patterns, few algorithm emphases on rare items, but rare items [1] also have importance, for example, network intrusion detection, where among various normal connections we need to detect the rare malicious connections. Classification of such a non-uniform data set is a challenging issue. Most classifiers perform poorly in such a data set. Realizing the importance of rare class classification, in this paper we propose a classification algorithm (CBMR Algorithm) that is based on association rules mined by MSApriori approach [2] and is capable of classifying rare classes. The performance evaluation of the proposed algorithm has been done for different data sets [3] and in comparison with existing technique like [4], it is found that algorithm has efficient and superior performance for classifying rare cases.
Classification Algorithm based on MS Apriori for Rare Classes

- Coenen, F. (2003), The LUCS-KDD Discretized/normalised ARM and CARM data library http://www.csc.liv.ac.uk/~frans/KDD/Software/LUCS_KDD_DN/, Department of Computer Science, The University of Liverpool, UK.
- B. Liu, W. Hsu and Y. Ma; "Integrating classification and association rule mining," proceedings of the fourth international conference on knowledge discovery and data mining, 1998, pp. 80-86.
- M. S. Chen, J. Han, P. S. Yu; "Data mining: an overview from a database perspective," IEEE Transactions on Knowledge and Data Engineering, 1996, 8, pp. 866-883.
- W. Li, J. Han and J. Pei; "CMAR: Accurate and efficient classification based on multiple class-association Rules," In ICDM'01, San Jose, CA, Nov. 2001, pp. 369-376.

Index Terms

Computer Science
Artificial Intelligence
Keywords
Rare Classes  Msapriori Algorithm  Classification  Data Mining