Abstract

Decision tree technology has proven to be a valuable way of capturing human decision making within a computer. As ID3 select those attribute as splitting attributes which have different values whether it classify dataset properly or not. There is another drawback of ID3 which repeat sub tree many times and select same attribute many times. These drawbacks can be
removed by the proposed algorithm. By the proposed algorithm firstly large volume of dataset which contain redundant instance are reduced. These redundant instance doesn’t make any contribution to take decision hence can be deleted from the dataset. After reducing the volume of the dataset decision tree is constructed through rough set. The main concept of rough set theory is degree of dependency which is used in the proposed algorithm to select splitting attribute on the compressed data. Thus the proposed algorithm reduces the complexity of tree and in addition increases the accuracy. We have used some UCI machine learning repository. By the experimental result it is shown that proposed algorithm gives better accuracy and diminishes the complexity of tree.

Reference

- Gökhan Silahtarolu, An Attribute-Centre Based Decision Tree Classification Algorithm, World Academy of Science, Engineering and Technology 2009, pp: 302-306.
- Jinmao Wei and Dao Huang, Rough Set Based Decision Tree, Proceedings of the 4 World Congress on Intelligent Control and Automation, IEEE 2002.
- Zhang Li-Juan and Li Zhou-Jun, A Novel Rough Set Approach for Classification, 1999.
- Longjun Huang, Minghe Huang, Bin Guo, Zhiming Zhang, A New Method for Constructing Decision Tree Based on Rough Set Theory, IEEE 2007, pp: 241-245.
- Baowei Song Chunxue Wei, Algorithm of Constructing Decision Tree Based on Rough Set Theory, International Conference on Computer and Communication Technologies in Agriculture Engineering, 2010.

Index Terms

Computer Science     Data Mining
Key words

Classification

Rough Set

Decision Tree

Dimensionality Reduction

ID3