Abstract

Pre-processing in data mining played essential role for enhancing data quality. The basic concept behind is that, learning with accurate and high quality data may provide more efficient classification results as compared to learning with poor quality of data. In this presented paper a pre-processing technique is implemented with slight modification which is based on the technique given in [1]. In this paper a promising approach of data pre-processing is provided which utilizes a fuzzy technique in order to improve the data quality. The implementation of available technique is performed using MATLAB. Additionally, the improved fuzzy technique is also implemented with it. The results demonstrate the effectiveness in classification accuracy after implementation of both techniques. Finally, the obtained results favour the proposed model for enhancing the performance of classifiers in both manners supervised and unsupervised manner.

References

- Der-Chiang Li and Chiao-Wen Liu, "Extending Attribute Information for Small Data Set Classification", IEEE Transactions on Knowledge and Data Engineering,
Vol. 24, No. 3, March 2012
- Yifei Ren, "Data Preprocessing for Data Mining," Bachelor's Thesis (UAS) Degree Program in Information Technology Information Technology 2013.
- X. Sun, S. Z. Sun, J. Tian and J. Han, "Sparse Kernel Principal Component Analysis on Seismic Denoising and Fluid Identification," 10 June 2013 DOI: 10.3997/2214-4609.20130642

Index Terms
Computer Science Information Sciences

Keywords
Data mining  pre-processing  data quality enhancement  classification performance improvement