Abstract

Today, several tools are available for solving data mining problems, both in open source and commercial category. For solving classification problems these tools provide variety of strategies such as decision tree, neural networks, lazy classifiers etc. For each strategy, the tools allow the user to select specific values for large number parameters[1] for e. g. in case of a neural network classifier, user needs to provide values for parameters such as epochs, learning rate, momentum etc. Although default setting for such parameters is provided by tools, it is often found that the classifier performance (accuracy) may be enhanced by making series of experiments with different values for these parameters. Thus, for a novice user it is difficult to guess proper values for these parameters and the only option is to try with series of experiments which is time consuming. This paper aims at developing a database to record the nature of data such as number and type of attributes, presence or absence of missing values etc along with various values for building classifier models and the accuracy of the classifier. Such a data is then can be made available to novice users to build a model based on past experience. The work also aims at developing required forms reports
User Assistance for Effective Data Mining

References

- WEKA manual.
- Zdravko Markov, Ingrid Russell, An Introduction to the Weka DataMining System

Index Terms

Computer Science Information Sciences

Keywords

Data Mining classification classifier filter WEKA