Abstract

The spreading amount of data usually generates interesting demand for the data analysis tools that spot regularities in these data. Data mining has turned up as great domain that contributes mechanism for data analysis, to find out the hidden knowledge, and self-ruling decision making in many operation domains. Supervised machine learning is using to find out the search for algorithms that reason from clearly supplied instances to produce general interpretation, which then makes predictions about future scenario or events. In other words, the goal of supervised learning is to make a small model of the distribution of class labels (distribution or classification) in terms of finding (predictor) features. The resulting classifier is then used to assign class labels (attributes) to the testing instances where the values of the predictor (attributes or properties) features are known, but the value of the class label is unknown. This paper explains various supervised machine learning classification techniques.

In this paper, we have discussed the about the classification algorithm which are available today, how they works, and what are their advantages and disadvantages. The algorithms
which we will discuss are Naïve Bayes, SVM, random forest, decision tree and logistic regression.

References

2. Jiawei Han and Micheline Kamber (2006), Data Mining Concepts and Techniques, published by Morgan Kauffman, 2nd ed.

Index Terms

Computer Science Algorithms

Keywords

Classification, Naïve Bayes, Random forest, Multiple regression dependent variable, independent variables, predictor variable, response variable