Abstract

Knowledge is not only power, but also has significant competitive advantage. Data warehousing is not a new idea. The use of corporate data for strategic decision making, as opposed to the use of data for tracking and enabling operations, has gone on for a computing itself. As the business these days contain huge amounts of data and the users connected to these databases across the globe and round the clock have the necessity for maintaining a separate database for the sake of analysis. This paper proposes one method of feature selection of NB Tree Algorithm. The Proposed algorithm (NB Tree) gives an effective Classification Algorithm for reducing computational time and gives better accuracy results compare with another algorithms. In many applications, however, an accurate ranking of instances based on the class probability is more desirable. The dependence between two attributes is determined based on the probabilities of their joint values that contribute to true and false classification decisions. The paper also evaluates the approach by comparing it with existing feature selection algorithms over 8 datasets from University of California, Irvine (UCI) machine learning databases. The proposed method shows better results in terms of number of selected features, classification accuracy, and running time than most existing algorithms.
References


Index Terms

Computer Science

Information Sciences

Keywords

Feature selection  Data mining  Classification  J48  Decision Tree  and NB Tree