Abstract

The growth of data-mining applications such as classification and clustering has shown the need for machine learning algorithms to be applied to large scale data. In this paper we present the comparison of different classification techniques using Waikato Environment for Knowledge Analysis or in short, WEKA. WEKA is open source software which consists of a collection of machine learning algorithms for data mining tasks. The aim of this paper is to examine the performance of different classification methods for a set of large data. The algorithm which have been tested are J48, SMO, Part, OneR, ZeroR and Navies Bayes Algorithm. The Syskill and webert we page rating data [11] with a total data of 1660 and a dimension of 332 rows and 5 columns will be used to test and validate the differences between the classification methods or algorithms.

References

- Classification at http://chem-eng. utoronto. ca/~datamining/dmc/classification. htm
- Jiawei Han and Micheline Kamber: Data Mining Concepts and Techniques, Elsevier Inc., Edition 2, 2006, ISBN no. 978-81-312-0535-8
Evaluation of Different Classification Techniques for WEB Data

- Decision Tree at http://chem-eng. utoronto. ca/~datamining/dmc/decision_tree. htm
- Navies Bayes at http://chem-eng. utoronto. ca/~datamining/dmc/naive_bayesian. htm
- http://chem-eng. utoronto. ca/~datamining/dmc/zeror. htm
- http://chem-eng. utoronto. ca/~datamining/dmc/oner. htm
- Data Mining Practical Machine Learning Tools and techniques, Author Ian H. Witten & Eibe frank [Part at page number 404, chapter explorer]
- Kappa at http://www. dmi. columbia. edu/homepages/chuangj/kappa
- Gabriel Fiol-Roig, Margaret Miró-Julià, Eduardo Herraiz: Data Mining Techniques for Web Page Classification. Math and Computer Science Department, University de les Illes Balears, SPAIN {biel. fiol@uib. es, margaret. miro@uib. es}
- A. Jebaraj Ratnakumar: An Implementation of Web Personalization using Web Mining Techniques. Professor and Head, Department of Computer Science and Engineering, Apollo Engineering College, Chennai, Tamil Nadu, India E-Mail: ajrk_jeba@yahoo. co. in
- Magdalini Eirinaki, Michalis Vazirgiannis: Web Mining for Web Personalization. Department of Informatics Athens, University of Economics and Business {eirinaki@auel. gr, mvazirg@auel. gr}
- Dimitrios Pierrakos, Georgios Paliouras, Charistos Papatheodorou and Constantine D. Spyropoulos: Web Usage Mining as a Tool for Personalization. The Institute of Informatics and Telecommunications, NCSR, email: paliourg@itdemokritos. gr

Index Terms

Computer Science

Artificial Intelligence

Keywords

Machine Learning Data Mining WEKA Classification Web data web mining