Abstract

The data mining can be defined as discovery of relationships in large databases automatically and in some cases it is used for predicting relationships based on the results discovered. Data mining plays a vital role in various applications such as business organizations, e-commerce, health care industry, scientific and engineering. In the health care industry, the data mining is mainly used for predicting the diseases from the datasets. Various data mining techniques are available for predicting diseases namely Classification, Clustering, Association rules and Regressions. This paper analyzes the classification tree techniques in data mining. The aim of this paper is to investigate the experimental results of the performance of different classification
techniques for a heart disease dataset. The classification tree algorithms used and tested in this work are Decision Stump, Random Forest, and LMT Tree algorithm. Comparative analysis is done by using Waikato Environment for Knowledge Analysis or in short, WEKA. It is open source software which consists of a collection of machine learning algorithms for data mining tasks.

References

- K. P Soman, Shyam Diwakar, V. Vijay; "Insight into Data mining theory and practice";
- Ruben D. Canlas Jr.; "DATA MINING IN HEALTHCARE: CURRENT APPLICATIONS AND ISSUES"; August 2009
- Esra Mahsereci Karabulut & Turgay ?brikçi; "Effective Diagnosis of Coronary Artery Disease Using The Rotation Forest Ensemble Method"; June 2011 / Accepted: 30 August 2011 / Published online: 13 September 2011 # Springer Science+Business Media, LLC 2011
- MA. JABBAR, Dr. PRITI CHANDRA, B. L. DEEKSHATULU; "Cluster Based Association Rule Mining For Heart Attack Prediction"; JTAIT Vol. 32 No. 2 October 2011.
- Milan Kumari, Sunila Godara; "Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction"; IJCST Vol. 2, Issue 2, June 2011
- Dr. D. Raghu. T. Srikanth, Ch. Raja Jacob; "Probability based Heart Disease Prediction using Data Mining Techniques"; IJCST Vol. 2, Issue 4, Oct - Dec. 2011, ISSN: 0976-8491 (Online) | ISSN: 2229-4333(Print)
- Dr. K. Usha Rani; "Analysis of Heart Diseases Dataset Using Neural Network Approach"; (IJKP) Vol. 1, No. 5, September 2011
- Niels Landwehr, Mark Hall and Eibe Frank, "Logistic Model Trees"; Institute for Computer Science, University of Freiburg, Freiburg, Germany. Department of Computer Science, University of Waikato, Hamilton, New Zealand.
- L. Breiman, Random Forests, Machine Learning (2001), 45, 5-32
Received 28 March 2006; received in revised form 20 October 2006; accepted 21 October 2006

- B. Nithyasri, K. Nandhini, Dr. E. Chandra "CLASSIFICATION TECHNIQES IN EDUCATION DOMAIN" (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 05, 2010, 1679-1684

Index Terms

Computer Science  Data Mining

Keywords

Classification  Decision Stump  Lmt  Random Forest  Heart Disease And Weka