Abstract

With the evolution of distributed computing, the databases are inherently distributed across the globe and therefore data analysis from various data sources is very essential in decision making. The core need in the current industrial environment is hence to extract information from the huge, complex and dynamic data through data mining techniques. Integrating data from multiple data sources and analysing the large, complex dynamic data is a tedious and
complex work. Additionally, database consists of inconsistent and noisy data. Further, with the decrease in quality of data to be mined the quality of knowledge model obtained from it also decrease which in turn affects the decision making process. However optimization of data preprocessing can resolve the aforementioned issues. This paper provides design and development of data preprocessing software, based on intelligent agents. This software enables data preprocessing operations to be performed in an automated mode, and gives accurate results in lesser time when compared to manual data preprocessing.

References

- Pyle, D. 1999 Data Preparation for Data Mining. Morgan Kaufmann Publishers, Inc. , San Francisco, CA, USA
- C., Lavrac, N., Moyle, S., Kavsek, B. 2001 Integrating Aspects of Data Mining, Decision Support and Meta-Learning: Internal SolEurope Session, ECML/PKDD’01 workshop notes 43–52
- Cristian Aflori and Florin Leon 2008: “Efficient Distributed Data Mining using Intelligent Agents Authors”
- Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth 2005: “From Data Mining to Knowledge Discovery in Databases”
- Ranjit Bose, Vijayan Sugumaran 1998, "IDM: An Intelligent Software Agent Based Data Mining Environment," IEEE.
- Dr. Joseph P. Bigus, Jennifer Bigus "Constructing Intelligent Agents with JAVA;"

**Index Terms**

Computer Science  
Hpc Applications

**Keywords**

Discretization Agent  
Transformation Agent