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
An Effective Data Preprocessing Technique for Improved Data Management in a Distributed Environment

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 inturn 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

- Peng Jin, Yun-Long Zhu And Kun-Yuan Hu. August, 2007 A Clustering Algorithm For
  Data Mining Based On Swarm Intelligence Proceedings Of Sixth International Conference On
  Machine Learning Cybernetics, Hong Kong, 19-22
  San Francisco, CA, USA
- C., Lavrac, N., Moyle, S., Kavsek, B. 2001Integrating Aspects of Data Mining,
  Decision Support and Meta-Learning: Internal SolEuNet Session, ECML/PKDD&amp;apos;01
  workshop notes 43–52
- B. Liu and A. Tuzhilin 2008: Managing and Analyzing Large Collections of Data Mining
- Zulaiha Ali Othman, Azuraliza Abu Bakar, Abdul Razak Hamdan, Khairuddin Omar and
  Nor Liyana Mohd Shui, 2007 &quot;Agent based preprocessing,&quot; International Conference
  on Intelligent and Advanced Systems.
- Cristian Aflori and Florin Leon 2008: &quot;Efficient Distributed Data Mining using
  Intelligent Agents Authors&quot;,
- Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth 2005: &quot;From Data
  Mining to Knowledge Discovery in Databases&quot;
- I. A. Mohtar, 2006 &quot;Multiagent Approach to Stock Price Prediction,&quot; University Kebangsaan,
  Malaysia.
- P. Nurmi, M. Przybilski, G. Linden, and P. Floreen, 2005 &quot;An architecture for
distributed agent based data pre-processing. &quot; Pp. 122-132.
- C. Li, and Y Gao, 2006&quot;Agent-based pattern mining of discredited activities in
  public services,&quot; Proceedings of the 2006 IEEE/WI C/ACM International Conference on
  Web Intelligence and Intelligent Agent Technology.
- Dr. T. R. Gopalakrishnan Nair, Lakshmi Madhuri, Sharon Christa, Dr. V. Suma, 2012
  &quot;Data Preprocessing Model Using Intelligent Agents&quot; International Conference on
  Information Systems Design and Intelligent Applications.
- Agent Working Group, 2000 &quot;Agent technology,&quot; OMG Document
  ec/2000-08-01, Version 1.0.
- Stuart Russell and Peter Norvig 1995 &quot;Artificial Intelligence: A Modern
  Approach;&quot;, c Prentice-Hall, Inc.
- Sharon Christa, K. Lakshmi Madhuri and V. Suma 2012, &quot;A Comparative Analysis
  of Data Mining Tools in Agent Based Systems&quot;, International Conference on Systemics,
  Cybernetics and Informatics
- 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