Agents are software programs that perform tasks on behalf of others and they are used to clean the text data with their characteristics. Agents are task oriented with the ability to learn by themselves and they react to the situation. Learning characteristics of an agent is done by verifying its previous experience from its knowledgebase. An agent concept is a complementary approach to the Object Oriented paradigm with respect to the design and implementation of the autonomous entities driven by beliefs, goals and plans. Preference based text data cleaning is based on the selection issue. Preferences are given by the user in the form of alphabets, numbers and special characters. Preference based Text data cleaning process transforms the given text data into structured database and extracts the required information using the given keyword. Agents incorporated in the architectural design of a Text data cleaning process combines the features of Multi-Agent System (MAS) Framework, MAS
with Learning (MAS-L) Framework. MAS framework reduces the development time and the complexity of implementing the software agents. MAS-L framework incorporates the intelligence and learning properties of agents present in the system. MAS-L Framework makes use of the Decision Tree learning and an evaluation function to decide the next best decision that applies to the machine learning technique. This paper proposes the design for Multi-Agent based Data Cleaning Architecture that incorporates the structural design of agents into object model. The design of an architectural model for an Intelligent Multi-Agent based Data Cleaning inherits the features of the Multi-Agent System (MAS) and uses the MAS-L framework to design the intelligence and learning characteristics.

Reference

- Margaret H Dunham, Sridhar S "Data Mining – Introductory and Advanced Topics"
- Raymond J Mooney, Razvan Bunescu, "Mining knowledge from text using information extraction", SIGKDD Explorations, 2003, vol 7 issue 1, pp 3-10.
- Zhang Jin, "Research on Data Cleaning in Data Acquisition", conference on data mining,
Index Terms

Computer Science          Data Mining

Key words

Text data          Preference          Agents          MAS

MAS-L

Architecture

Intelligent

Data Cleaning

IMASDC