Abstract

A web page which is a source of information consists of many parts among which only a part of the information is useful for a particular application and the remaining information are noises. An effective technique for users to extract the useful information from the total information is urgently required. Hence by removing those noise patterns from the web page, the efficiency of the web data extraction can be improved. This research work proposes an approach for removing the local noise from a given web page based on n x 1 table and XSL display method with filter feature for improving the efficiency of web data extraction.

References

- G. Poonkuzhali, K. Thiagarajan, K. Sarukesi and G. V. Uma, "Signed Approach for Mining Web Content Outliers", World Academy of Science, Engineering and
A Noise Reduction Approach based on n x 1 Table and XSL Display Method for Efficient Web Data Extraction


Index Terms

Computer Science

Artificial Intelligence
Keywords
  noise reduction  web data mining  data extraction  XML  XSL