Abstract

Efficient Human Computer Interaction (HCI) is an absolute necessary for many applications these days. Computational Linguistics supports HCI to make computers to understand human languages. Advanced Computational models can be built using many technologies to provide easy communication between human and computers. Data mining has emerged to address problems of understanding ever-growing volumes of information for structured data. Data mining is a process to extract hidden knowledge from huge amount of data which can be used to build computational model. The usage of Association Rules (AR), one of the data mining techniques, to build an effective communication between human and computers is elucidated in this paper. The comparative performance of two different Association rule algorithms is illuminated in building a model to legalize semantics of sentences in linguistics domain. The sequence of operations to build the model is explored with necessary constraints at each stage. The model is to verify the meaning of English sentences which are syntactically correct using Apriori and Frequent-pattern tree growth algorithm in a limited domain. As a prerequisite, syntax verification of the sentence is also done and as a follow up, it also provides an interface which can be used for interaction between human and computer. The association rules, a data mining concept is employed in semantic analysis in a distinct way. Since the natural language understanding is an endless process, this work opens the door for the usage of association
rules in semantic analysis of natural language sentences in a defined domain.

References

- Yuhua Li, Zuhair A. Bandar and David McLean. "An approach for measuring semantic similarity between words using multiple information sources," IEEE Transactions on Knowledge and Data Engineering, July/August 2003, Volume 15, No. 4.
- Jiawei Han and Micheline Kamber. Data Mining-Concepts and Techniques. Morgan Kaufmann Publishers.
- Sam Y. Sung, Zhao Li, Chew L Tan and Peter A. Ng. "Forecasting Association Rules using Existing Datasets," IEEE Transactions on Knowledge and Data Engineering, Volume 15, No. 6.

Index Terms

Computer Science

Information Sciences
Keywords
Syntax Analysis  Semantic analysis  Apriori algorithm  Question Answering System