POS Tagging provides important grammatical as well as contextual information for each word in
the corpus. POS Tagging enables various companies to be able to track user reviews and can even be used for Speech Synthesis. In this paper, different POS Tagging Algorithms, namely, Memory-Based Learning Algorithm, Multi-Domain Web Based Algorithm and the Hybrid Model, will be compared on the basis of their execution time as well as efficiency. In Memory-Based Learning algorithm, the word to be tagged is searched in the lexicon using weighted similarity matrix, if an exact match is found, its lexical representation is retrieved, but, if it is not found, the lexical representation of its nearest neighbor is retrieved. Thus, the algorithm will not work efficiently for sparse data. On the other hand, Multi-Domain Web Based Algorithm is used to tag unknown words. The word is searched over the web for its possible tags. Due to the web search, runtime overhead is induced for each word. The tag with highest occurring probability is assigned to the word. The Hybrid Model executes Memory-Based Learning algorithm for known words and Multi-Domain Web Based Algorithm for unknown words.

References

- Aastha Gupta, Rachna Rajput, Richa Gupta & Monika Arora, 2014,'Hybrid Model to Improve Time Complexity of Words Search in POS Tagging'. Paper presented at International Conference on Data Mining and Intelligent Computing,IEEE, Delhi, India
- Jakub Zavrel & walter Daelemans, "Recent Advances in Memory Based Part of Speech Tagging." VI Simposio Internacional de Comunicacion Social, Santiago de Cuba
Comparative Study of POS Taggers

pp. 590-597, 1999
- Lars Bungum, Bjorn Gamback, "Evolutionary Algorithms in Natural Language Processing", Norwegian Artificial Intelligence Symposium, Gjøvik, 22 November 2010

Index Terms

Computer Science
Algorithms

Keywords
Pos Tagging Multi-domain Web Based Algorithm Memory Based Learning Algorithm Hybrid Model