

Part of Speech Tagging Using Statistical Approach for Nepali Text

Authors : Archit Yajnik

Abstract : Part of Speech Tagging has always been a challenging task in the era of Natural Language Processing. This article presents POS tagging for Nepali text using Hidden Markov Model and Viterbi algorithm. From the Nepali text, annotated corpus training and testing data set are randomly separated. Both methods are employed on the data sets. Viterbi algorithm is found to be computationally faster and accurate as compared to HMM. The accuracy of 95.43% is achieved using Viterbi algorithm. Error analysis where the mismatches took place is elaborately discussed.

Keywords : hidden markov model, natural language processing, POS tagging, viterbi algorithm

Conference Title : ICNLPC 2017 : International Conference on Natural Language Processing and Cognitive Computing

Conference Location : Singapore, Singapore

Conference Dates : January 08-09, 2017