

NLP

Assignment-3

Abhinay Natti
-150438

Firstly I have done the preprocessing over the train dataset which had links commented .So I removed them and made sure that there are only 2 new line gaps between two sentences

For training the data I did unlabelled training . I made a dictionary which has heads associated to its children(head for a word is given by the 6th item in the file). For feature vector I have taken the top 2 elements of the stack(a b) and top 2 elements from the (c d)input list. For each word I have considered the word vector loaded from the glove.6B.300d.txt. If a word is not present in the glove model the sentence containing the word is not used for training.Left arc transition takes place if the 2nd element(a) in the stack has the its head as the topmost element(b) and if the there are no children left anymore for the 2nd element(a). Similarly right arc is done if the 1st element has no children left and has its head as the 2nd element . Finally shifting takes place either if the stack contains a single element (or) if the children of the top element exist even though the element in the counterpart is its head.

I have used MLPclassifier for training the set.

Accuracy : 80.61 %

Then I've made parser by just using the same method again instead of appending features I've used predict from the classifier and using transitions accordingly