CS671
Assignment 3
Jayant Agrawal
14282

Preparing Training Data
1. Create a transition sequence along with operations from the Treebank using the transition rules given in the slides.
2. Each transition state along with the corresponding operation serves as a training pair.

Features
1. Use word vectors for 3 words from top of the stack and top of the buffer.
2. Each word vector contains word embedding from GLoVE/Word2Vec concatenated with one-hot encoding of the POS tag for that sentence.
3. Concatenate all 6 vectors to form the feature for a configuration.

Results

Training Sentences: 1000
Test Sentences: 100
Final Accuracy: **91.14 %**

Training Sentences: 10000
Test Sentences: 1000
Final Accuracy: **94.64 %**

Sources
1. TreeBank: https://github.com/UniversalDependencies/UD_English-EWT
2. GLoVE Vectors: http://nlp.stanford.edu/data/glove.6B.zip