Assignment 2: Word Sense Disambiguation

CS626: Speech and Natural Language Processing and the Web

Problem statement

- Modify the modules created for POS tagging using Viterbi to perform **HMM based word sense labeling.**
- Dataset: **SemCor (Sense-tagged Corpus)**
- Input: **A sentence**
- Output**: WordNet sense ids for the words in the sentence**
- Create a document which reports the following
  - P, R, F1-scores
  - Compare the performance of your sense-tagger against MFS baseline
  - Perform detailed error analysis
  - Write a short paragraph on your learning.

Note: Use 5-fold cross-validation for reporting all accuracies

Submission instructions

- The assignment is to be submitted in groups of 3 (Same group for every assignment and project)
- The submission link will be created on moodle to submit the assignment
- Only one person from the group with the lowest id is supposed to make the submission
- The name of the folder should be <id1_id2_id3>_Assignment2.zip
  - The uncompressed folder should contain code, readme and the slides used for presentation
  - The readme should contain details about the tools, versions, pre-requisites if any, and how to run the code for both approaches.

Deadline

- No-Hard deadline (Continuous Evaluation). Evaluation date will be announced soon