

1. You are given an English file `wsj_pos_tagged_en.txt` that contains 3914 Part-of-Speech (POS) tagged sentences where each word in a sentence is associated with a tag.

Perform the following tasks.

- a. Split the data into K folds ( $K \geq 3$ ).
- b. Find the emission and transition probabilities from the training data.
- c. Implement the Viterbi decoding to predict the tag sequence given a sentence.
- d. Evaluate the performance of the tagger using precision, recall, and F1-score across all folds.