Assignment-2 Report

Implementation of Baum-Welch Algorithm:

Select initial values for the matrices A, B and π, where π is 1 × N, while A = {aij} is N × N and B = {bj (k)} is N × M (where M is the vocabulary size), and all three matrices are row-stochastic. Initialised A,B,pi with random initial parameters. Made sure that each row sums to 1 and the elements of each matrix are not uniform.

Difficulties faced:

As the length of observation sequence increases and the number of observations increases, the values of all the matrices tends to zero.
So, I tried handling them using scaling and manual heuristics. Log probabilities in this case will not work and I tried it out.

Observations:

After running certain number of iterations for training the HMM on the entire brown corpus, I got the output emission matrix. Many of the top 10 words emmited for each tag are same. The top 10 words for each tag get varied, as we run the algorithm again & again or you can say i.e a 10 fold validation. Such pattern could also be explained on the basis of transition probabilities between different tags as in the case of POS tagging also. Because the occurrence of one tag is dependent on the other tags too which is incorporated in A matrix.