INDIAN INSTITUTE OF INFORMATION TECHNOLOGY KALYANI

Kalyani, Nadia-741235, West Bengal

(An institute of National Importance)

Course Title: Machine Learning Lab Paper Code: CSC612 Spring 2025

Assignment 10

Date: 31/03/2025 Due date: 31/03/2025 1 PM 5 Marks

An HMM consists of two types of variables: hidden states and observations.

The hidden states are the underlying variables that generate the observed data but are not directly observable. Here we consider NE tags as the hidden states.

The observations are the output variables that are observed. Here we consider POS tags as the observations.

Implement the Hidden Markov Model (HMM) algorithm for decoding problem using the following steps. Use first 47,700 sentences of Assignment10.csv for training and last 259 sentences for testing. (Dataset is attached in the classroom for Assignment 9).

1. Define the state space and observation space.
2. Calculate the initial state distribution (Pi), the state transition probabilities (A) and the state observation probabilities (B).
3. Calculate the probability of the best tag sequence of the test sentences.
4. Calculate the accuracy.
5. Show the accuracy in graphs.