CSCE 550 Project2

Due: October 15, 2019

1- Objective

In this assignment, we build a simple HMM model for a discrete dataset.

2- Dependencies

Python > 2.7
Pandas toolkit (If it is needed)
Numpy toolkit (If it is needed)

Note: Using Python's Scikit-learn tool **is not allowed** for this assignment.

3- Dataset

The dataset (Project2Data.txt) contains 1000 rows of past weather observations. The states (ω) are "sunny", "rainy" and "foggy". The emission states are "yes", "no" indicating if an umbrella was observed.

4- HMM program guideline

a- You may use your own way to break the problem in any format/function you like however; a skeleton of functions is in the file HMM.py.

b- Your HMM program will take an observation sequence (V^T) as an input and will outputs

I. The state matrix and all needed probabilities such as $a_{ij}\ and\ b_{jk}$ need to be calculated from given data

II. The probability of the given observation using ***Viterbi*** algorithm.

III. The most probable path to generate the given observations using ***Decoding*** algorithm

c- For initialization condition, assume that in the $T(0)$ system is in state sunny.

5- What we need to generate?

Your program accepts a sequence of visible states (V^T) and calculates the probability of this sequence occurs.

To achieve this goal you need to calculate transition probabilities and emission probabilities from dataset.

Outputs:

a- a_{ij} matrix

b- b_{jk} matrix

c- The probability of the HMM producing the given visible state (***Viterbi*** algorithm)

d- The sequence of hidden states that given visible states generates (***Decoding*** algorithm)