Homework 4

BT3102 - Computational Methods for Business Analytics
Due: 22-03-2022 (11:59 PM)

Question 1. (5 points) Assume that we have the Hidden Markov Model (HMM) with the following properties - each of the states can take on k different values, and a total of m different observations are possible across all states. Further, consider that the observation sequence X is of length n: $X = \{x_1, x_2, ..., x_n\}$, and x_{n+1} is the stop state. Given this information, answer the following questions:

(a) How many total number of parameters (not the minimum) are required to fully define this HMM? Justify your answer. *(4 points)*

(b) Does your answer to the above question depend on the length of the HMM? Explain. *(1 point)*

Question 2. (15 points)

Consider a small town where people are either healthy or feverish. Only the doctor can differentiate or identify the people having fever from healthy people. He does so by inquiring about their symptoms. People can have one of the following symptoms: {*'cold', 'dizzy', 'normal'*}.

The doctor believes that the health condition of his/her patients obeys the markov property. However, the health states: (*'Healthy', 'Fever'*), are hidden from the doctor. The doctor is now asking for help from the students in BT-3102 (who are known to be experts in bayes networks). The doctor wants to identify the most likely sequence of states and its corresponding probability for a given sequence of symptoms.

The doctor has also provided us the information on the parameters of Hidden Markov Model: initial probabilities (π), transition probabilities (A), and emission probabilities (B). Also, the doctor says it is equally likely for all the states to reach the stop state, hence you can just ignore the transition to the *stop* state. These parameters are included in the file $viterbi - hw4\_rename\_before\_submission.py$. Using these parameters and the observation sequence (which is also included in the file), help the doctor using Viterbi algorithm. While you are equipped to calculate this path by hand, the doctor is asking for the program to reuse your solution in the future. Hence, please implement the Viterbi algorithm.

Note that your code should return the **mostly likely state sequence (or the path with highest probability)** and the **Viterbi matrix (V)**, which contains the maximum probability value of the paths ending at each state across time. We are expecting you to return the V matrix instead of just the maximum probability value of most likely sequence to

check the correctness of your code. You are also allowed to add any additional intermediate functions if necessary.

Testing your code: Please calculate the most likely sequence of states using probabilities by hand. Further, check if the output of your code aligns with your calculation.

Grading: Your codes will be run automatically, so it is important that you name your submission file based on your matric-number (ex: A0222374L.py). Full points will be awarded if the code runs and outputs the correct sequence of states and its corresponding probability value. 12 points if either one of them is correct. 5 points if it runs but the output is wrong. Zero marks if the code doesn't run or doesn't follow the naming convention.