Unit 6 - Week 4

Assignment 4

The due date for submitting this assignment has passed.
As per our records you have not submitted this assignment.

Due on 2019-08-28, 23:59 IST.

1) Let us define an HMM Model with K classes for hidden states and T data points as observations. The dataset is defined as $X = \{x_1, x_2, \dots, x_T\}$ and the corresponding hidden states are $Z = \{z_1, z_2, \dots, z_T\}$. Please note that each x_i is an observed variable and each z_i can belong to one of classes for hidden state. What will be the size of the state transition matrix, and the emission matrix, respectively for this example. *1 point*

1. $K \times K, K \times T$
2. $K \times T, K \times T$
3. $K \times K, K \times K$
4. $K \times T, K \times K$

1.
2.
3.
4.

No, the answer is incorrect.
Score: 0
Accepted Answers:
1.

2) You are building a model distribution for an infinite stream of word tokens. You know that source of this stream has a vocabulary of size 1000. Out of these 1000 words you know of 100 words to be stop words each of which has a probability of 0.0019. With only this knowledge what is the maximum possible entropy of the modeled distribution. (Use log base 10 for entropy calculation) *1 point*

1. 0.0079
2. 1
3. 2.9841
4. 3

1.
2.
3.
4.

No, the answer is incorrect.
Score: 0
Accepted Answers:
3.

3) In beam search decoding approach for POS tagging, what is the maximum size of search space at every time step for a beam size of K? (V: Size of Word vocab, P: Number of possible POS Tags) *1 point*

1. K^2
2. $K * V$
3. $K * P$
4. $K * P * V$

1.
2.
3.
4.

No, the answer is incorrect.
Score: 0
Accepted Answers:
3.

4) For an HMM model with N hidden states, V observable states, what are the dimensions of parameter matrices A,B and π? A: Transition matrix, B: Emission matrix, π: Initial Probability matrix. *1 point*

1. $N \times V, N \times V, N \times N$
2. $N \times N, N \times V, N \times 1$
3. $N \times N, V \times V, N \times 1$
4. $N \times V, V \times V, V \times 1$

1.
2.
3.
4.

No, the answer is incorrect.
Score: 0
Accepted Answers:
2.

5) Which of the following is not a property of Hidden Markov Models? H and V denotes the sets of possible values for hidden and visible states respectively. "t" denotes time. $t \in [1, T]$ *1 point*

1. $P(v_T|h_{T-1}) = P(v_T)$
2. $P(v_1, v_2) = \sum_{h_1 \in H} \left(P(v_1, h_1) \sum_{h_2 \in H} P(v_2, h_2|h_1)\right)$
3. $P(v_T|v_{T-1}) = \sum_{h_T \in H} P(v_T|h_T) P(h_T|v_{T-1})$
4. $P(v_1, \dots, v_T|h_1, \dots, h_T) = P(v_1|h_1)P(v_2|h_2) \dots P(v_T|h_T)$

1.
2.
3.
4.

No, the answer is incorrect.
Score: 0
Accepted Answers:
1.

6) Suppose you have the input sentence "Death Note is a great anime". And you know the possible tags each of the words in the sentence can take. *1 point*

- Death: NN, NNS, NNP, NNPS
- Note: VB, VBD, VBZ
- is: VB
- a: DT
- great: ADJ
- anime: NN, NNS, NNP

How many possible hidden state sequences are possible for the above sentence and states?

1. $4 \times 3 \times 3$
2. 4^{3^3}
3. $2^4 \times 2^3 \times 2^3$
4. $2^{4 \times 3 \times 3}$

1.
2.
3.
4.

No, the answer is incorrect.
Score: 0
Accepted Answers:
1.

7) In Hidden Markov Models or HMMs, the joint likelihood of an observed sequence O with a hidden state sequence Q, is written as $P(O, Q; \theta)$. In many applications, like POS tagging, one is interested in finding the hidden state sequence Q, for a given observation sequence, that maximizes $P(O, Q; \theta)$. What is the time required to compute the most likely Q using an exhaustive search? The required notations are, N: possible number of hidden states, T: length of the observed sequence. *1 point*

1. Of the order of TN^T
2. Of the order of N^2T
3. Of the order of T^N
4. Of the order of N^2

1.
2.
3.
4.

No, the answer is incorrect.
Score: 0
Accepted Answers:
1.

8) What will be the runtime for above problem, if we use forward algorithm for finding the most likely Q? What is the speed up compared to Question 8? *1 point*

1. N^T, T
2. TN^2, N^{T-2}
3. T^2, T^{N-2}
4. N^2, TN^{T-2}

1.
2.
3.
4.

No, the answer is incorrect.
Score: 0
Accepted Answers:
2.

9) Find the Viterbi Decoding of the sequence "ki fin yeni!". Possible POS tags are {T1, T2, T3, T4}. Assume all POS tags are equally likely to be at the starting of a sequence. Emmision probabilities are, *1 point*

	ki	fin	yeni
T1	0.1	0.1	0.8
T2	0.8	0.1	0.1
T3	0.2	0.2	0.6
T4	0.8	0.1	0.1

Table 1: Output Symbol probabilities

Transition matrix is

	T1	T2	T3	T4
T1	0.18	0.01	0.8	0.01
T2	0.9	0.0	0.05	0.05
T3	0.4	0.5	0.05	0.05
T4	0.4	0.5	0.05	0.05

Table 2: Hidden State transition matrix

Calculate $P(x_1 = \text{"}ki\text{"}, x_2 = \text{"}fin\text{"}, y_1 = \text{"}T1\text{"}, y_2 = \text{"}T2\text{"})$.

1. 0.000025
2. 0.0001
3. 0.0025
4. None of the above

1.
2.
3.
4.

No, the answer is incorrect.
Score: 0
Accepted Answers:
1.

10) For the question above what is the most probable POS tag sequence for the given input. *1 point*

1. T_2, T_1, T_3
2. T_4, T_3, T_1
3. T_2, T_3, T_1
4. T_4, T_2, T_1

1.
2.
3.
4.

No, the answer is incorrect.
Score: 0
Accepted Answers:
1.

11) What is the joint probability of most probable sequence found above? *1 point*

1. $P(\text{"ki"}, \text{"fin"}, \text{"yeni"}, T_{t=1}, T_{t=2}, T_{t=3}) = 0.0072$
2. $P(\text{"ki"}, \text{"fin"}, \text{"yeni"}, T_{t=1}, T_{t=2}, T_{t=3}) = 0.00864$
3. $P(\text{"ki"}, \text{"fin"}, \text{"yeni"}, T_{t=1}, T_{t=2}, T_{t=3}) = 0.00972$
4. $P(\text{"ki"}, \text{"fin"}, \text{"yeni"}, T_{t=1}, T_{t=2}, T_{t=3}) = 0.01152$

1.
2.
3.
4.

No, the answer is incorrect.
Score: 0
Accepted Answers:
2.

12) For an HMM, The initial state distribution is given as $P(z_1) = \pi_{z_1}$. The emission distribution is a vector parameterized by ϕ_{*k}, where $\phi_{xk} = P(X = x|Z = k)$, and Φ is the emission matrix that contains all the emission vectors for all possible values of Z. ($Z \in \{1, ..., K\}$). The transition probabilities are given by $a_{kj} = P(Z_t = k|Z_{t-1} = j)$. Let $\theta = \{A, \Phi, \pi\}$. What is the full expression for $P_\theta(X = x_1, x_2, \dots, x_T)$ *1 point*

1. $\left(\sum_{z_T}(\sum_{z_{T-1}} \cdots (\sum_{z_2}(\sum_{z_1} \pi_{z_1}\phi_{x_1,z_1}a_{z_2,z_1})\phi_{x_2,z_2}a_{z_3,z_2}) \dots \phi_{x_{T-1},z_{T-1}}a_{z_T,z_{T-1}})\phi_{x_T,z_T}\right)$
2. $\pi_{z_1}\phi_{x_1,z_1}\prod_{t=2}^{T}\phi_{x_t,z_t}a_{z_t,z_{t-1}}$
3. $\prod_{t=1}^{T}\pi_{z_t}\phi_{x_t,z_t}a_{z_t,z_{t-1}}$
4. None of the above

1.
2.
3.
4.

No, the answer is incorrect.
Score: 0
Accepted Answers:
1.