National University of Computer and Emerging Sciences 0030

Lahore Campus

Data Analysis and Visualization (DS3001)

Sessional II

Date: November 2nd 2024

Total Time (Hrs):

Total Marks:

Total Questions 2

Course Instructor(s)

Ms. Umme Ammarah

Ms. Esha T...

Do not write below this line

- Show complete working of all the questions on **Answer sheet**. And fill the tables in **Question Paper.**
- Attempt all the parts of a question together and in order.

Attach question paper with answer book.

CLO #1:

Q1 (a): Disease data set: **[10 Marks]**

Patient ID	Fever	Cough	Fatigue	Disease (0 = No, 1 = Yes)
1	1	1	1	1
2	1	1	0	1
3	1	0	1	1
4	0	1	1	0
5	0	0	1	0
6	1	0	0	1
7	0	1	0	0
8	1	1	1	1
9	0	0	0	0
10	1	1	0	1

Patient ID: Identifier for each patient. For all other attributes 1 and 0 corresponds to Yes and No respectively. Given the following testing data of patients, for each determine if the patient have disease (1 for Yes) or not (0 for No) using **Naïve Bayes**. Show complete working on answer sheet for all the 3 patients.

Fever	Cough	Fatigue	Disease
1	0	1	?
0	1	1	?
1	1	1	?

Q1(b): Find the cosine similarity of the following two sentences. 2/ **[5 Marks]**

Document 1: **"I love programming and data science."**

Document 2: **"Data science and programming are fun."**

Formula:	$\frac{a \cdot b}{\|a\| \cdot \|b\|}$
Similarity Value	$\frac{1}{2}$

CLO # 1

Q2. Imagine you have a simple RNN that predicts the weather for the next day based on the temperature observations over the past three days. Each observation consists of two temperature readings for a given day. Based on these readings, the model will predict the weather for the next day as one of three possible conditions: **Sunny**, **Rainy**, or **Cloudy**. Input Layer size is 2 as for each observation there are 2 values, hidden layer has 2 neurons. For output layer 1st neuron represents sunny, 2nd represent rainy and 3th represents cloudy class. You only need to get the output at time=3. Transposed weights are given below, W^{xh} are weights between input and hidden layer. W^{hh} are weights between hidden layer to previous hidden layer and W^{hy} are weights between hidden and output layer. b^h is bias for hidden neurons and b^o is bias for output layer. 4/ **[15 Marks]**

$$W^{xh} = \begin{bmatrix} 0.2 & 0.1 \\ 0.3 & 0.4 \end{bmatrix}, W^{hh} = \begin{bmatrix} 0.5 & 0.2 \\ 0.3 & 0.6 \end{bmatrix}, W^{hy} = \begin{bmatrix} 0.3 & 0.5 \\ 0.1 & 0.4 \\ 0.2 & 0.3 \end{bmatrix}, h^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Input Sequence:

$$x^{(1)} = \begin{bmatrix} 20 \\ 15 \end{bmatrix} \quad x^{(2)} = \begin{bmatrix} 22 \\ 18 \end{bmatrix} \quad x^{(3)} = \begin{bmatrix} 19 \\ 16 \end{bmatrix}$$

I waited for bias value.
Please check rough work!

Softmax: $\tau(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$

$$\tanh(x) = \frac{2}{1+e^{-2x}} - 1$$

Output of Hidden State $h^{(1)}$	Output of Hidden State $h^{(2)}$	Output of Hidden State $h^{(3)}$
?	?	?

Output vector $Y^{(3)}$	?
Predicted Weather condition	

$b^h = \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix}$, $b^o = \begin{bmatrix} 0.1 \\ 0.1 \\ 0.1 \end{bmatrix}$