ourdiu.com | Shared by Md. Fahmid Bin Mostafa

Daffodil International University

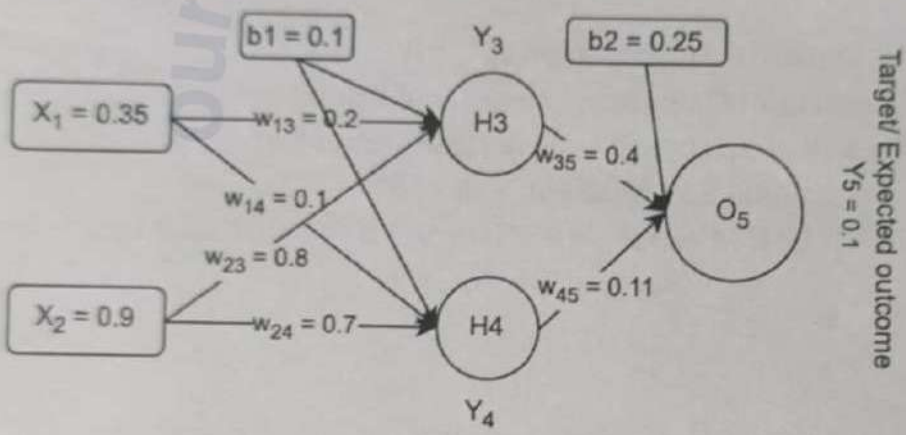
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Final Examination, Fall - 2024
Course Code: CSE445, Course Title: Natural Language Processing
Section: Re_A

Time: 2 Hours **Marks:40**

Answer All Questions
[The figures in the right margin indicate the full marks and corresponding course outcomes. All portions of each question must be answered sequentially.]

1. Explain one-hot encoding as a method for representing words as vectors. Discuss its limitations why this poses a challenge in natural language processing.	5	**CO2**
2. (a) Explain the concept of Latent Dirichlet Allocation (LDA) in topic modeling. **(b)** Describe how Collapsed Gibbs Sampling (CGS) is applied in LDA to infer topic distributions with example. Create 2 topics on your own to illustrate the CGS.	2 3	
3. Imagine you're typing on your phone, and it suggests the next word based on what you've already written. A recurrent network is working behind this suggestion. It processes sequential data one step at a time, maintaining a hidden state that captures information about previous inputs to influence future outputs. Now, your task is to **draw** that model's architecture and explain its **advantages** and **disadvantages**.	5	
4. (a) Define and analyze 3 types of activation function (sigmoid, ReLu, Tanh) with graphical representation and required equations.	5	

ourdiu.com | Shared by Md. Fahmid Bin Mostafa

(b) Assume that the neuron has a sigmoid activation function, apply a Backpropagation algorithm to perform a forward pass and a backward pass to update the value of **w35** and **w45** on the network. Assume that the learning rate is 1.  	10	
5. A team is working on developing a speech recognition system for a voice-activated assistant. To ensure the system performs accurately, they need to create a robust dataset, preprocess the speech data, and implement appropriate analysis techniques. **(a)** Explain two preprocessing techniques that can be applied to the speech data to improve its quality before analysis. Provide examples of tools or methods that could be used. **(b)** Write down the importance of speech dataset for the system is being developed and the possible challenges that they may face while collecting data.	4 6	**CO2, CO3**

Best of luck!