ourdiu.com | Shared by Abir Mia -231

Daffodil International University

Faculty Of Science & Information Technology

Department Of Computer Science & Engineering

Course Code: CSE 432 **Course Title: NLP**

QUIZ 2 , Spring 2026 **Time:45 min**

1. A single neuron computes a weighted sum followed by a ReLU activation.

$$\hat{y} = \text{ReLU}(h)$$

Given the following values:

Term	Value
Bias input	1, with weight $w_0 = 0.05$
x_1	2, with weight $w_1 = 0.3$
x_2	1, with weight $w_2 = -0.2$
x_3	4, with weight $w_3 = 0.15$
Target y	2

(a) Compute h, $\hat{y}$, and $L(\hat{y}, y)$. 3

(b) Using the chain rule, write the general expression for $\partial L/\partial w_i$ in terms of $\partial h/\partial w_i$, $\partial \hat{y}/\partial h$, and $\partial L/\partial \hat{y}$, then compute $\partial L/\partial w_0$, $\partial L/\partial w_1$, $\partial L/\partial w_2$, and $\partial L/\partial w_3$. 3

(c) Hence write the full gradient vector $\nabla L(\hat{y}, y)$. 1

1. Describe the architecture of RNN model with diagram 5

2. Write down the limitation of Stochastic Gradient Descent (SCD) & How can overcome this limitation? 3

Scanned with CamScanner