Name: _____ Reg #: _____ Section: _____

National University of Computer and Emerging Sciences, Lahore Campus

Course:	**Natural Language Processing**	**Course Code:**	**CS 535**
Program:	**MS(Computer Science)**	**Semester:**	**Spring 2020**
Duration:	**20 Minutes**	**Total Marks:**	**12**
Paper Date:	**11-Feb-20**	**Weight**	**5%**
Section:	**CS**	**Page(s):**	**2**
Exam:	**Quiz 1**		

Q1) You are given the following training corpus: [1 + 1 + 1 + 3 + 2 = 8 Marks]

<s> I like cars </s>
<s> cars like I </s>
<s> We like bikes </s>
<s> I do not like bikes and cars </s>

a) Calculate the probability of following test sentence. Include </s> in your counts just like any other token. λ_1 = trigram weight, λ_2 = bigram weight, λ_3 = unigram weight, λ_1 = 0.5, λ_2 = 0.3, λ_3 = 0.2

<s> I like bikes </s>

i. Unigram Model

ii. Bigram Model

iii. Trigram Model

iv. Trigram language model with linear interpolation.

Name: _____ Reg #: _____ Section: _____

b) Calculate perplexity of test sentence using trigram model with linear interpolation

Q2) Give some example of how language modeling helps in task in speech. [2 Marks]

Q3) A dictionary can be used to identify spelling errors. How does language modeling help in task of spelling correction? [2 Marks]