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:	5-March-20	Weight	5%
Section:	CS	Page(s):	2
Exam:	Quiz 2		

Q1) Given following test collection, compute the probability of test document for positive and negative class using normal Naïve Bayes classifier with Laplace smoothing. What will be class prediction of Naïve Bayes for doc 5?

Training	Doc	Words	Class
	1	I like this movie	**Positive**
	2	ordinary cast but great script	**Positive**
	3	interesting plot average film	**Negative**
	4	movie is interesting but long and slow paced	**Negative**
Test	5	great cast but average movie	**?**

Solution:

Vocabulary = V = 18
P(Positive) = 2/4 = 0.5
P(Negative) = 2/4 = 0.5

P (great | Positive) = (1+1)/(9 + 18) = 2/27
P (cast | Positive) = (1+1)/(9 + 18) = 2/27
P (but | Positive) = (1+1)/(9 + 18) = 2/27
P (average | Positive) = (0+1)/(9 + 18) = 1/27
P (movie | Positive) = (1+1)/(9 + 18) = 2/27

P (great | Negative) = (0+1)/(12 + 18) = 1/30
P (cast | Negative) = (0+1)/(12 + 18) = 1/30
P (but | Negative) = (1+1)/(12 + 18) = 2/30
P (average | Negative) = (1+1)/(12 + 18) = 2/30
P (movie | Negative) = (1+1)/(12 + 18) = 2/30

P (Positive | d5) = P(d5 | Positive) * P(Positive) = 2/27*2/27*2/27*1/27*2/27*1/2 = 5.575 e-7

P (Negative | d5) = P(d5 | Negative) * P(Negative) = 1/30*1/30*2/30*2/30*2/30*1/2 = 1.646 e-7

D5 belongs to Positive class

Name: _____ Reg #: _____ Section: _____

Q2) Following is Tf.Idf based vector representation of words “exam” and “pass”. What is cosine similarity of the two words?

Words	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5	Dim 6
Exam	0	2	0	2	1.2	2.6
Pass	1	2.5	0	0	3.4	0.5

Solution:

Do Product (Exam, Pass) = 0*1 + 2*2.5 + 0*0 + 2*0 + 1.2*3.4 + 2.6*0.5 = 10.38
Length (Exam) = sqrt (2*2 + 2*2 + 1.2*1.2 + 2.6*2.6) = 4
Length (Pass) = sqrt(1*1 + 2.5*2.5 + 3.4*3.4 + 0.5*0.5) = 4.36

Cosine (Exam, Pass) = 10.38 / (4*4.36) = 0.595