

Natural Language Processing (CS4063)

Course Instructor(s):

Dr Ahmad Raza Shahid, Dr Mehreen Alam

Sections: C and D

Sessional-I Exam

Total Time (Hrs): 1

Total Marks: 50

Total Questions: 3

Date: Sep 24, 2024

Course Section

Do not write below this line.

Signature

Attempt all the questions.

Question 1:

Points 5 + 5 + 5 + 5 = [20 marks]

Given the following corpus, calculate the probability and perplexity of the test sentence "Efforts to combat this environmental issue are crucial". Show working for unigram and bigram models. Show your work to get the credit.

Corpus:

Deforestation is a major environmental issue. The rate of deforestation has increased significantly over the past few decades. This environmental issue leads to loss of biodiversity and disruption of ecosystems. Efforts to combat this issue include reforestation and sustainable land management. The impact of this environmental issue on climate change is profound. Addressing this issue requires global cooperation and innovative solutions.

Question 2:

Points = [20 marks]

Consider the dataset given below of sentences with positive and negative sentiment:

Train Sentences:

Positive: "I love the vibrant colors of the sunset." 8

Positive: "The movie had vibrant characters and a fantastic plot." 9

Positive: "Her work was praised for its vibrant creativity." 8

Negative: "The food at the restaurant was vibrant but disappointing." 9

Negative: "He felt very sad after the disappointing meal." 9

Test Sentence:

"Vibrant sunset was disappointing."

Identify the class of the test sentence using naïve Bayesian classifier. Show your work to get the credit.

Question 3:

Points 4 + 3 + 3 = 10 marks]

Answer the questions based on the following corpus. Write code and display output for the following using the nltk library.

Elon Musk is both an innovator and an entrepreneur. As an innovator, he has developed new ideas and technologies through companies like SpaceX, which makes rockets, and Tesla, which makes electric cars. Musk's work with SpaceX led to the first commercially built spacecraft to carry astronauts to the International Space Station. As an entrepreneur, he has started and grown these companies into successful businesses. He is known for taking on big challenges and industries, like automotive and aerospace, and finding ways to improve and change them. His success with multiple companies shows his skill as an entrepreneur, while his contributions to technology show his strength as an innovator.

- remove stop words from the first two sentences. Make a reasonable assumption for stop word list.
- perform lemmatization on the first sentence.
- find concordance for the word 'he'. Apply for only 1 word on the left and 3 words to the right.