

START OF QUIZ

**Student ID: 60226669, Mo-
hammad, Waquas**

Question 1

Topic: Lecture 3

Source: Lecture 3

If our vocabulary consists of just symbols A and B and our corpus consists of the sequence: A B B A A B and we build a bigram language model by applying add-one smoothing to the maximum likelihood estimate from the corpus, what is the probability $P(B|A)$? Please show your work. (2)

Question 2

Topic: Lecture 2

Source: Lecture 2

Why is the Forgy initialization sub-optimal? (1)

Question 3

Topic: Lecture 1

Source: Lecture 1

Explain why edit distance (given our formulation) will always choose a substitution, if it can. (1)

Question 4

Topic: Lecture 3

Source: Lecture 3

In your own words, explain the Markov assumption, and how it is used in language modeling.

(1)

Question 5

Topic: Lecture 4

Source: Lecture 4

Why can we use logarithms for the Viterbi algorithm, but not the forward algorithm? (1)

Question 6

Topic: Lecture 4

Source: Lecture 4

Iterative algorithms often require a stopping condition. Briefly explain why this is necessary, and why perplexity is a metric to use for stopping HMMs. (2)

Question 7

Topic: Lecture 2

Source: Lecture 2

Why do outliers cause problems for clustering algorithms like k-means? How can we deal with them? (1)

Question 8

Topic: Lecture 1

Source: Lecture 1

Explain what modifications would need to be made to our dynamic edit distance algorithm to incorporate weighted edit distance. (2)

Question 9

Topic: Long

Source: Lecture 4

Please refer to the "Long" question from Lecture 4.

END OF QUIZ