V1V1V1V1V1V1V1V1V1V1V1V1V1V1V1V1V1V1V1V1V1V1V1V1V1

National University of Computer & Emerging Sciences
FAST-Karachi Campus
CS4051- Information Retrieval
Quiz#3

Dated: March 29, 2023 Marks: 20
Time: 20 min.
Std-ID: _____Sol_____

**Question No. 1**

What is meant by a language model? **[5]**

A language model is a function that puts a probability measure over strings drawn from some vocabulary. That is, for a language model M over an alphabet set lambda and set of all strings S: $\Sigma$ P(s) = 1

Language models analyze bodies of text data to provide a basis for their word predictions. Language modeling (LM) is the use of various statistical and probabilistic techniques to determine the probability of a given sequence of words occurring in a sentence. A simple definition of a Language Model is an AI model that has been trained to predict the next word or words in a text based on the preceding words, it assigns probablities to different sentences by using some dependency between sequence of word. For example: uni-gram model is the simplest one which compute $P_{uni}$(t1t2t3t4) = P(t1) P(t2) P(t3) P(t4) assuming all words are independent to each other.

**Question No.2**

Differentiate between the following pair of terms: [5]

| Probabilistic Information Retrieval Model | Language Model for Information Retrieval |
|---|---|
| - In this model we try to present document in the decreasing value of P(R=1/ q,d).<br>- All documents are evaluated as per degree of relevance.<br>- Very relax assumptions. | - In language model for IR we estimate the probability of a query generating from the same document model that is P(q/Md).<br>- Every document is treated as different model and the query generation process for each is estimated.<br>- Very restricted. |

**Question No.3**

Consider making a language model from the following training text: [5]

"the martian has landed on the latin pop sensation ricky martin"

a. Under a MLE-estimated unigram probability model, what are P(the) and P(martian)?

P(the) = 2-appearance of the in text / 11- total token in text = 2/11
P(martian) = 1/11

b. Under a MLE-estimated bigram model, what are P(sensation|pop) and P(pop|the)?

P(sensation | pop) = count(pop, sensation) / count(pop) = 1/1 = 1

P(pop| the) = count(pop, the) / count(the) = 0/2 = 0

**Question No. 4**

What is the difference between MAP and MLE estimation? [5]

Maximum A Posteriori (MAP) and Maximum Likelihood (MLE) are both approaches for making decisions from some observation or evidence. MAP takes into account the prior probability of the considered hypotheses. MLE does not.

$\Theta_{mle} = \text{argmax } \Pi\ p\ (x_i / \Theta)$

$\Theta_{map} = \text{argmax } \Pi\ p(x_i) * p\ (x_i / \Theta)$