Natural Language Processing
Assignment 11
Type of Question: MCQ

Number of Questions: 10   **Total Marks: 10×1=10**

Question 1: Your teacher recommended you to read the book 'Natural Language Processing with Python'. After reading the book, you want to summarize it. What kind of summarization method would you use for this purpose?

1. Abstractive single document summarization
2. Abstractive multi document summarization
3. Extractive single document summarization
4. Extractive multi document summarization

a. 1, 2
b. 3, 4
c. 1, 3
d. 2, 4

Answer: c
Solution:

Question 2: What kind of summarization approach is lexrank?

a. Extractive multi document generic
b. Extractive multi document query specific
c. Abstactive multi document query specific
d. Abstractive multi document generic

Answer: a
Solution

.

Question 3: Identify whether the following statements are True or False.

1. Maximum Marginal Relevance strives to reduce redundancy while maintaining query relevance.
2. Query-focused summarization can be thought of as a complex question-answering system.

a. True, False
b. True, True
c. False, True
d. False, False

Answer: b
Solution:

For question 4-8, use the data given in Table 1.

Suppose you have trained a image classifier with 5 classes - cat, dog, lion, tiger, and deer. Consider the confusion matrix shown in Table 1.

		Gold Labels				
		cat	dog	lion	tiger	deer
Predicted Labels	cat	130	17	9	7	40
	dog	15	150	25	10	7
	lion	10	45	150	23	5
	tiger	15	15	20	120	30
	deer	40	30	20	10	155

Table 1

Question 4: What is the macro averaged precision?

    a. 0.6696

    b. 0.6078

    c. 0.6433

    d. None of the above

Answer: c

Solution

Question 5: What is the macro averaged recall?

    a. 0.6464

    b. 0.6540

    c. 0.6190

    d. None of the above

Answer: a

Solution:

Question 6: What is the accuracy of your classifier?

    a. 0.6421

    b. 0.6536

    c. 0.6319

    d. None of the above

Answer: a

Solution:

Question 7: What is the micro averaged precision?

    a. 0.6915

    b. 0.6421

    c. 0.6245

    d. None of the above

Answer: b

Solution: Separate confusion matrix for each class is as follows:

class

TP	FP
FN	TN

cat

130	73
80	815

dog

150	57
107	784

lion

150	83
74	791

tiger

120	80
50	848

deer

155	100
82	761

$$\text{recall} = \frac{TP}{TP+FN}$$
$$\text{precision} = \frac{TP}{TP+FP}$$
$$\text{accuracy} = \frac{TP+TN}{TP+FP+TN+FN} = \frac{\text{number of correct predictions}}{\text{the total number of predictions}} = \frac{705}{1098} \approx 0.6421$$

macro averaged precision $= (0.6404 + 0.7246 + 0.6438 + 0.6 + 0.6078)/5 \approx 0.6433$

macro averaged recall $= (0.6190 + 0.5837 + 0.6696 + 0.7059 + 0.6540)/5 \approx 0.6464$

For micro averaged results, create pooled confusion matrix from all the classes.

micro averaged precision $=$ micro averaged recall $= \frac{705}{705+393} \approx 0.6421$

.

Question 8: What is the micro averaged recall?

   a. 0.6190
   b. 0.6535
   c. 0.6421
   d. None of the above

Answer: c
Solution:

Question 9: It is estimated that 20% of ChatGPT generated texts are fake. Some AI system has been applied to filter these fake contents. An AI system claims that it can detect 98% of fake contents, and the probability for a false positive (a real content detected as fake) is 3%. Now if a content is detected as fake, then what is the probability that it is in fact a real content?

   a. 0.084
   b. 0.109
   c. 0.119
   d. None of the above

Answer: b

Solution:

Let, A = Event that a content is detected as fake
B = Event that a generated text is fake
$P(B) = 0.2$
$P(B') = 0.8$
$P(A|B) = 0.98$
$P(A|B') = 0.03$

$$P(B'|A) = P(A|B')P(B')/P(A)$$

$$= P(A|B')P(B')/(P(A|B)P(B) + P(A|B')P(B'))$$

$$= (0.03 \times 0.8)/(0.98 \times 0.2 + 0.03 \times 0.8)$$

$$\approx 0.109$$

Question 10: Consider the system generated summary (S) and the reference sum- mary as follows:

S: ChatGPT is powered by deep learning, a technique that involves training a neural network with extensive data.
R: ChatGPT is deep learning model that uses a neural network to understand lan- guage patterns.
What is the ROUGE-1 recall for the given summary with respect to the reference?

   a. 0.500

   b. 0.571

   c. 0.470

   d. None of the above

Answer: b   **Solution:** ROUGE-1 recall = 8/14 = 0.5714