21AI643

Sixth Semester B.E. Degree Examination, June/July 2024
Natural Language Processing

Time: 3 hrs. Max. Marks: 100

Note: *Answer any FIVE full questions, choosing ONE full question from each module.*

Module-1

1 a. Define NLP. List and explain different phases of analysis in NLP with an example for each. **(10 Marks)**

b. Explain the following with suitable example :
   i) $\overline{X}$ – theory
   ii) Theta theory. **(10 Marks)**

OR

2 a. Define n-gram model. Derive expression for n-gram model. Solve the following test case sentence :
   Training sentence :
      \<s> The Arabian Knights
      \<s> These are the fairy tales of the cast
      \<s>The stories of the Arabian Knights are translated in many languages.
   Test sentence :
      \<s> The Arabian Knights are the fairy tales of the cast
   Apply the bi-gram model for the test sentence and estimate the probability. **(10 Marks)**

b. What is LFG models? Explain the e-structure and f-structure concept used in language modeling. Find f-structure and c-structure for the following sentence'
   she saw stars in the sky
   CFG rules :
      $S \rightarrow NP\ VP$
      $VP \rightarrow V\ \{NP\}\ \{NP\}\ PP^*\ \{S'\}$
      $PP \rightarrow P\quad NP$
      $NP \rightarrow Det\quad N\{PP\}$
      $S' \rightarrow comp\ S$. **(10 Marks)**

Module-2

3 a. Draw the NFA for the language consisting of all string containing only a's and b's and ending with baa. Draw the state transition table. Find R.E for the above language. **(10 Marks)**

b. Write the algorithm for minimum edit distance and compute the minimum edit distance between "Peaceful" and "Peaceful". **(10 Marks)**

OR

4 a. Write CYK parsing algorithm, Find the sequence of states created by CYK algorithm while parsing the sentence : "The man read this book". Consider the following simplified grammar in CNF.

S → NP VP	Det → that/ this /a / the
S → VP	Noun → book / flight / meal / man
VP → Verb NP	Verb → include /read
NP → Det Noun	AUx → does

(10 Marks)

   b. Derive a top-down, depth – first, left – to – right parse tree for the given sentence :
'The angry bear chased the frightened little squirrel'.
Use the following grammar rules to create the parse tree :

S → NPVP	Det → the
NP → Det Nom	Adj → little / angry / frightened
VP → V NP	N → squirrel / bear
Nom → Adj Nom/N	V → chased

(10 Marks)

Module-3

5 a. With a neat diagram, explain the learning framework architecture. **(10 Marks)**
   b. Write a note on :
   i) Global security org
   ii) Infact system. **(10 Marks)**

OR

6 a. Explain shortest dependency path hypothesis. Show various shortest dependency path among the relations in the "Jellise created an atmosphere of terror in the camp by killing, abusing and threatening the detainess'. **(10 Marks)**
   b. What is Annotation? Explain the strategies used in active learning approach for acquiring labels using committee based classification scheme. **(10 Marks)**

Module-4

7 a. Explain latent semantic analysis feedback and topic models feedback system. **(10 Marks)**
   b. With the neat diagram explain the evolutionary model for KDT(Knowledge Discovery from Text). **(10 Marks)**

OR

8 a. Describe the following with example :
   i) iSTART
   ii) cohesion and cohesion matrix. **(10 Marks)**
   b. Explain SVM (Support Vector Machine) learning method in sequence model estimation. **(10 Marks)**

Module-5

9 a. Explain the architecture of an information retrieval system with a neat diagram. **(10 Marks)**

b. A user submitted a query to an IR system. Out of the first 15 documents returned by the system, those ranked 1, 2, 5, 8 and 12 were relevant compute non-interpolated average precision for this retrieval. Assume that the total number of relevant document is 6.

(10 Marks)

OR

10 a. Define the following with respect to IR
   i) wordNet
   ii) frameNet.

(10 Marks)

b. Explain the following classical model with example :
   i) Boolean model
   ii) Vector space model.

(10 Marks)

* * * * *