

Reg. No.:

Name :



Model Question Paper

Programme	: M.Tech Computer Science & Engineering	Semester	: Winter 2015
Course	: Natural Language Processing	Code	: CSE528
Faculty	: Prof. Tulasi Prasad Sariki	Slot	: F2
Time	: Three Hours	Max. Marks	: 100

Answer ALL Questions

Q.No.	Sub. Sec.	Question Description	Marks
1.	a)	Identify the Referent for the Underlined word in the following sentences. 1. The monkey ate the banana. <u>It</u> was hungry. 2. Susan dropped the plate. <u>It</u> shattered loudly. 3. The music stopped, and <u>that</u> upset everyone. 4. Rama saw the girl. <u>She</u> asked him for help. 5. Carol told Bob to attend the party. <u>They</u> arrived together.	05
	b)	Deduce the existing morphemes from each of the following words. 1. Unbreakable 2. Computerization 3. Antiintellectualism 4. Conformity 5. Etymologically	05
2.		Find out the thematic roles for the underlined phrases in the following sentences. 1. <u>They</u> filled <u>the pool</u> with <u>water</u> . 2. <u>John</u> drove <u>Mary</u> from <u>Austin</u> to Dallas in <u>his Toyota Prius</u> . 3. <u>Jack</u> ran with <u>crutch</u> for <u>Susan</u> . 4. <u>Sam</u> <u>destroyed</u> the car. 5. <u>Sujan</u> <u>put</u> the car <u>through the wall</u> .	15
3.		Solve the problem of Word Sense Disambiguation with any Machine Learning approach.	05
4.		A dot com company is feeling difficult to identify the fake reviews from the customers. Deduce a methodology for that.	10

5.	<p>Identify the phrases in the following sentences according to CONLL 2000 chunking</p> <ol style="list-style-type: none"> 1) The government plans to raise taxes were defeated. 2) Hence the activation of P600 marks the parser's attempt to revise the sentence's structural mis-match or ambiguity. 3) Depending on its location within a sentence, a disfluency either aids in the computation of a sentence or forces the parser to linger on the sentence for a longer period of time. 4) Garden path sentences are less common in spoken communication because the prosodic qualities of speech often serve to resolve ambiguities in the written text. 5) Time flies like an arrow; fruit flies like a banana. 	20																																																																																				
6.	<p>The table A given below shows the text corpus with words and their frequency of occurrence in each document. The table B shows the test document in terms of words and their frequency of occurrence. Can you demonstrate a method that can be used to classify the test document (Doc7) into one of the classes given in table A?</p> <p>A:</p> <table border="1" data-bbox="336 808 1382 1115"> <thead> <tr> <th rowspan="2">Document Name</th> <th colspan="6">Key Words</th> <th rowspan="2">Class Name</th> </tr> <tr> <th>Cost</th> <th>Rates</th> <th>Free</th> <th>Lic</th> <th>Irctc</th> <th>Tatasky</th> </tr> </thead> <tbody> <tr> <td>Doc1</td> <td>2</td> <td>1</td> <td>3</td> <td>0</td> <td>0</td> <td>1</td> <td>Spam</td> </tr> <tr> <td>Doc2</td> <td>1</td> <td>1</td> <td>1</td> <td>0</td> <td>0</td> <td>0</td> <td>Spam</td> </tr> <tr> <td>Doc3</td> <td>1</td> <td>1</td> <td>2</td> <td>0</td> <td>1</td> <td>0</td> <td>Spam</td> </tr> <tr> <td>Doc4</td> <td>0</td> <td>1</td> <td>0</td> <td>2</td> <td>1</td> <td>1</td> <td>Inbox</td> </tr> <tr> <td>Doc5</td> <td>0</td> <td>0</td> <td>1</td> <td>1</td> <td>1</td> <td>0</td> <td>Inbox</td> </tr> <tr> <td>Doc6</td> <td>0</td> <td>0</td> <td>0</td> <td>2</td> <td>2</td> <td>0</td> <td>Inbox</td> </tr> </tbody> </table> <p>B:</p> <table border="1" data-bbox="336 1189 1382 1301"> <thead> <tr> <th rowspan="2">Document Name</th> <th colspan="6">Key Words</th> <th rowspan="2">Class Name</th> </tr> <tr> <th>Cost</th> <th>Rates</th> <th>Free</th> <th>Lic</th> <th>Irctc</th> <th>Tatasky</th> </tr> </thead> <tbody> <tr> <td>Doc7</td> <td>2</td> <td>1</td> <td>1</td> <td>0</td> <td>0</td> <td>1</td> <td>?</td> </tr> </tbody> </table>	Document Name	Key Words						Class Name	Cost	Rates	Free	Lic	Irctc	Tatasky	Doc1	2	1	3	0	0	1	Spam	Doc2	1	1	1	0	0	0	Spam	Doc3	1	1	2	0	1	0	Spam	Doc4	0	1	0	2	1	1	Inbox	Doc5	0	0	1	1	1	0	Inbox	Doc6	0	0	0	2	2	0	Inbox	Document Name	Key Words						Class Name	Cost	Rates	Free	Lic	Irctc	Tatasky	Doc7	2	1	1	0	0	1	?	10
Document Name	Key Words						Class Name																																																																															
	Cost	Rates	Free	Lic	Irctc	Tatasky																																																																																
Doc1	2	1	3	0	0	1	Spam																																																																															
Doc2	1	1	1	0	0	0	Spam																																																																															
Doc3	1	1	2	0	1	0	Spam																																																																															
Doc4	0	1	0	2	1	1	Inbox																																																																															
Doc5	0	0	1	1	1	0	Inbox																																																																															
Doc6	0	0	0	2	2	0	Inbox																																																																															
Document Name	Key Words						Class Name																																																																															
	Cost	Rates	Free	Lic	Irctc	Tatasky																																																																																
Doc7	2	1	1	0	0	1	?																																																																															
7.	Suggest a suitable methodology for sentence boundary disambiguation	10																																																																																				
8.	You are given a task to find out the sequences of words automatically from a given corpus whose syntactic structure and semantics are not compositional. Can you give examples of such sequences of words and demonstrate a method to find them out automatically?	10																																																																																				
9.	<p>Your Manager asked you to create a machine translation tool for your Company by giving the following options. Which option will you choose and Why?</p> <ol style="list-style-type: none"> 1. Rule based Machine Translation 2. Example based Machine Translation 3. Interlingua Machine Translation 	10																																																																																				

