Natural Language Processing

Assignment- 3

TYPE OF QUESTION: MCQ

Number of questions: 10 **Total mark: 10 X 1 = 10**

Question 1. Morphemes attached at the front and back of stem are called -

1. Prefixes
2. Infixes
3. Circumfixes
4. Suffixes

Answer: 3
Solution: Circumfixes (also called discontinuous morphemes) both precede and follow the root/stem.

Question 2: Find one tagging error in each of the following sentences that are tagged with the Penn Treebank tagset:

- **I/PRP need/VBP a/DT flight/NN from/IN Atlanta/NN**
- **Does/VBZ this/DT flight/NN serve/VB dinner/NNS**

1. Atlanta/NNS, dinner/NNP
2. Atlanta/NN, dinner/NN
3. from/JJ, dinner/NN
4. flight/NNS,Atlanta/FW

Answer - 2

Solution: The word "Atlanta" should be tagged with NNP instead of NN. Word "dinner" should be tagged with "NN" instead of "NNS".

Question 3: Which one of the following is an example of the discriminative model?

1. Naive Bayes
2. Bayesian Networks
3. Hidden Markov models
4. Logistic Regression

Answer - 4

Solution: Others are generative model

Question 4: Which one of the following is an example of the Generative model?

1. Conditional Random Fields
2. Naive Bayes
3. Support Vector Machine
4. Logistic Regression

Answer- 2

Solution: Others model in the option are discriminative model

---

Question 5. Natural language processing is essentially the study of the meaning of the words a human says or writes. Natural language processing is all around us all the time, but it also happens to be a way to improve the chatbot or product we interact with on a regular basis. Natural language processing is all about mimicking our own language patterns. Natural language processing can also improve the efficiency of business transactions and customer care. Natural language processing is the application of computer technology.

Suppose we want to check the probabilities of the *final words* that succeed the *string* language processing in the above paragraph. Assume d= 0; it is also given that no of unigrams = 78, no of bigrams = 122, no of trigrams = 130,, Question 6 and Question 7 are related to Question 5 corpus.

Solve the question with the help of **Kneser-Ney backoff technique.**

What is the continuation probability of *"is"* ?

1. 0.0078
2. 0.0076
3. 0.0307
4. 0.0081

Answer: 2

Solution: Refer week 3 lecture 12

Question 6: What will be the value of P(is| language processing) using Kneser-Ney backoff technique and choose the correct answer below. . Please follow the paragraph in Question .

1. 0.5
2. 0.6
3. 0.8
4. 0.7

Answer: 3

Solution: Refer week 3 lecture 12

Question 7. What is the value of P(can| language processing)? Please follow the paragraph in Question 5

1. 0.1
2. 0.02
3. 0.3
4. 0.2

Answer: 4

Solution: Refer week 3 lecture 12

Question 8: How does the state of the process is described in HMM?

1. Literal
2. Single random variable
3. Single discrete random variable
4. None of the mentioned

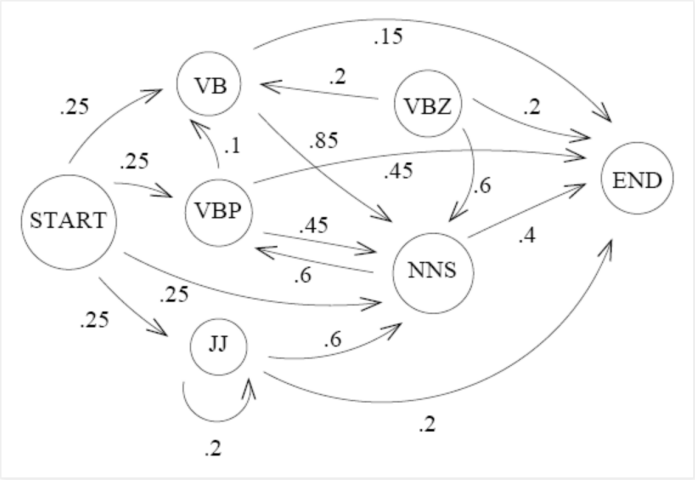
Answer: 3

Solution:
The state in HMM is defined as discrete random variable

Question 9: Consider the HMM given below to solve the sequence labeling problem of POS tagging. With that HMM, calculate the probability that the sequence of words "free workers" will be assigned the following parts of speech;

VB NNS

	free	*workers*
JJ	**0.00158**	**0**
NNS	**0**	**0.000185**
VB	**0.00115**	**0**
VBP	**0.00081**	**0**
VBZ	**0**	**0.00005**



The above table contains emission probability and the figure contains transition probability

1. $4.80 * 10^{-8}$
2. $0.80 * 10^{-8}$
3. $1.80 * 10^{-7}$
4. $1.80 * 10^{-8}$

Answer: 4

Solution:

P(free workers, VB NNS)

= P(VB|*start*) * P(*free*|VB) * P(NNS|VB) * P(*workers*|NNS)

* P(end|NNS)

= 0.25 * 0.00115 * 0.85 * 0.000185 * 0.4

= $1.80 * 10^{-8}$

Question 10: Which of the following is/are true?

1. FSA (Finite State Automaton) is always non-deterministic
2. FSA recognises regular language
3. Closed class words are mostly functional words
4. We need FSA for morphological analyzers

Answer: 2, 3

Solution:

FSA can be deterministic also, we need transducers to build morphological analyzers

---