Unit 4 - Week 3

Course outline

How to access the portal?

Week 1

Week 2

Week 3

- Examples for word prediction
- Introduction to Probability in the context of NLP
- Joint and conditional probabilities, independence with examples
- The definition of probabilistic language model
- Chain rule and Markov assumption
- Generative Models
- Bigram and Trigram Language models -peeking indide the model building
- Out of vocabulary words and curse of dimensionality
- Exercise
- Naive-Bayes Algorithm for classification
- Week 3 Lecture Materials
- Quiz : Assignment 3
- Assignment 3 : Programming Exercise for Self Assessment
- Week 3 feedback : Applied Natural Language Processing

Week 4

Week 5

Week 6

Week 7

Week 8

Week 9

Week 10

Week 11

Week 12

DOWNLOAD VIDEOS

Assignment 3

The due date for submitting this assignment has passed. **Due on 2019-08-21, 23:59 IST.**
As per our records you have not submitted this assignment.

1) Assume that the chain rule is used to compute the joint probability of the sentence " $I\ got\ this\ one$ ". The products of probabilities are represented by $P(got|I) \times P(this|I, got) \times P(one|I, got, this)$ *1 point*

- True
- False

No, the answer is incorrect.
Score: 0
Accepted Answers:
False

2) Assume that the language model is evaluated as given below *1 point*

$$\phi(W) = \sqrt{\frac{1}{P(w_1, w_2, \ldots, w_n)}}$$

$Note:\ n$ is the number of words in the sentence.
Smoothing will be used if the denominator $\to 0$.
Is the statement "$Minimizing\ \phi(W)\ is\ same\ as\ maximizing\ the\ probability\ P(w_1, w_2, \ldots, w_n)\ of\ the\ sentence$" true?

- True
- False

No, the answer is incorrect.
Score: 0
Accepted Answers:
True

3) Select one of the following bigram probabilities that represents the sentence *1 point*
`I love dogs`

(i) `<S>I love dogs</S>` `P (I) · P (love | I) · P (dogs | I love)`
(ii) `P (<S>) · P (I | <S>) · P (love | <S> I) · P (dogs | I love) · P (</S> | love dogs)`
(iii) `P (I | <S>) · P (love | I) · P (dogs | love) · P (</S> | dogs)`

- a
- b
- c
- All of the above

No, the answer is incorrect.
Score: 0
Accepted Answers:
c

4) The table given below contains some of the bigram frequencies of $(determine, w_i)$ where w_i represents every word in the column *1 point*

	the	how	this	a	his
determine	0.115	0	0.0125	0.006	0.0013

What is the conditional probability of $P(his|determine)$ if the probability of $determine$ as the starting word is 0.6?

- 0.0031
- 0.0022
- 0.0122
- 0.0128

No, the answer is incorrect.
Score: 0
Accepted Answers:
0.0022

5) Assuming that a language model assigns the following conditional probabilities to a 4-word sentence (S)=0.01212. What is the perplexity? *1 point*
Note: Perplexity is defined in question 2.

- 2.41
- 3.14
- 4.35
- 3.014

No, the answer is incorrect.
Score: 0
Accepted Answers:
3.014

6) Consider the following three sentences *1 point*
`Ram read a novel`
`Raj read a journal`
`Rai read a book`
What is the bigram probability of the sentence **`Ram read a book`**?
Include start and end symbols in your calculations

- 0.06
- 0.2222
- 0.1111
- 0.0556

No, the answer is incorrect.
Score: 0
Accepted Answers:
0.1111

7) Consider the following three sentences *1 point*
`Ram read a novel`
`Raj read a journal`
`Rai read a book`
What is the trigram probability of the sentence **`Ram read a book`**?
Include start and end symbols in your calculations

- 0.06
- 0.2222
- 0.1111
- 0.0556
- None of the above

No, the answer is incorrect.
Score: 0
Accepted Answers:
0.1111

8) In Naive Bayes classification, Posterior probability is estimated for predicting the class? *1 point*

- True
- False

No, the answer is incorrect.
Score: 0
Accepted Answers:
True

9) The following table contains the extracted features of emails and their classes (Spam and not Spam). The prior probabilities of the two classes (Spam,not Spam) = (0.57, 0.43) *1 point*

# of words=# of capitalized words	Subject in all capital letters?	# URLs> 3	Spam
No	No	0	No
No	No	5	No
Yes	Yes	4	Yes
Yes	No	4	Yes
Yes	No	0	No
No	Yes	7	Yes
Yes	No	0	No

- Yes
- No

No, the answer is incorrect.
Score: 0
Accepted Answers:
No