Machine Learning with Python-From Linear Models to Deep Learning

Course   Progress   Dates   Discussion   Resources

Course / Unit 4. Unsupervised Learning (2 weeks) / Homework 4

< Previous

2. Maximum Likelihood Estimation

☐ Bookmark this page

Homework due Apr 19, 2023 08:59 -03   Completed

Consider a general multinomial distribution with parameters θ. Recall that the likelihood
by:

$$P\left(\mathcal{D}; \theta\right) = \prod_{i=1}^{|\theta|} \theta_i^{c_i}$$

where c_i is the occurrence count of the i-th event.

The MLE of θ is the setting of θ that maximizes $P\left(\mathcal{D}; \theta\right)$. In lecture we derived this to b

$$\theta_i^* = \frac{c_i}{\sum_{j=1}^{|\theta^*|} c_j}$$

Unigram Model

4/4 points (graded)
Consider the sequence:

> A B A B B C A B A A B C A C

A unigram model considers just one character at a time and calculates $p\left(w\right)$ for $w \in \{$

What is the MLE estimate of θ? Give your result to three decimal places.

θ_A^* | 0.429 | ✔

θ_B^* | 0.357 | ✔

θ_C^* | 0.214 | ✔

Using the MLE estimate of θ on $\mathcal{D}$, which of the following sequences is most likely?

1.0/1 point (graded)
A bigram model computes the probability   as:

where   is the first word, and   is a pair of consecutive words in the documer

This is also a multinomial model. Assume the vocab size is   . How many parameters a

N^2-1 ✔

?   STANDARD NOTATION

Submit   You have used 2 of 3 attempts

Bigram Model 2

1/1 point (graded)
Which of the following represents the MLE for the **conditional probability**

○

○

◉

○

✔

Submit   You have used 2 of 3 attempts

Previous   **Next**

Submit   You have used 1 of 3 attempts

Discussion

Topic: Unit 4. Unsupervised Learning (2 weeks) :Homework 4 / 2. Maximum Likelihood Estimation

Show all posts   ⌄

? **Staff: Bigram Model 1 Solution is incorrect**

? **Confused by Bigram Model1 (Have seen staff hint)**
P(w) is a table of numbers of size N thus need N-1 Parameters. P(w2|w1) is of size NxN, where P(*|w1) will ha

💬 **Assistance**
Hello good day, Please i got this response "Could not format HTML for problem. Contact course staff in the d

? **Question on Bigram Model 1**
For Bigram Model 1, I provided my answer as a sum with 2 components (corresponding to the 2 sets of param

? **Bigram Model 1**
I have spent a lot of time thinking about how to solve the problem, however I am not able to get the correct a

💬 **P(null | B)**
Bigram Model 3 If I just us the MLE through counting, I get P(null | C) = 1/3 P(null | B) = 0 but if I plug in P(E

? **Bigram model 1 - what is actually given?**
Bigram model AFAIK breaks sentences into pairs of consequent words. Here it looks like 3 words - w0, w1, w

? **What is the difference between w1 and w'1 or w2 and w'2 for Bigram Model 2 supposed to me**
I do not get what is the difference between w1 and w'1 or w2 and w'2?

💬 **There is a typo at Bigram model3**
"Consider the same sequence from the unigram model:" should be "Consider the same sequence from the bi

? **Bigram Model 3. Starting probabilities are uniform?**
Uniform as in 1/3 ? and do we need the ending probabilities to be uniform too?

Legal

Terms of Service & Honor Code

Privacy Policy

Accessibility Policy

Trademark Policy

Sitemap

Cookie Policy

Do Not Sell My Personal Information

Connect

Blog

Contact Us

Help Center

Security

Media Kit