Feedback — Advanced MaxEnt / POS Tagging / Parsing Intro

Help

You submitted this quiz on **Wed 18 Apr 2012 5:56 AM PDT**. You got a score of **5.00** out of **5.00**.

Question 1

Suppose we build a maxent model for part of speech tagging a word (x), over a set of just 3 parts of speech (y): $Noun$, $Verb$, and $Other$. Our model has just one feature:

```
f(x,y) = [x="breeze" & y="Noun"]
```

Our training data consists of 5 observations:

```
[x="breeze" & y="Verb"]
[x="breeze" & y="Verb"]
[x="breeze" & y="Verb"]
[x="breeze" & y="Noun"]
[x="breeze" & y="Noun"]
```

The maxent model will be trained in the usual way to give the feature f a weight λ, so that the model expectation for the feature matches its empirical expectation.

The weight of the feature will be $\log X$ (natural log).

What is X?

Your Answer		Score	Explanation
◉ $4/3$	✔	1.00	Correct!
○ 2			
○ $3/5$			
○ 15			
Total		1.00 / 1.00	

Question Explanation

The empirical expectation of f is 2 from 5 identical observed data x. So $P(Noun|breeze) = 0.4$.

$P(Noun|breeze) = 0.4 = e^{\lambda}/(e^{\lambda} + e^{0} + e^{0}) = e^{\lambda}/(e^{\lambda} + 2)$

So:

$0.4[e^{\lambda} + 2] = e^{\lambda}$
$0.8 = 0.6e^{\lambda}$
$4/3 = e^{\lambda}$
$\lambda = \log(4/3)$

Question 2

Suppose we build a maxent model for part of speech tagging a word (x), over a set of just 3 parts of speech (y): $Noun$, $Verb$, and $Other$. Our model has just one feature:

```
f(x,y) = [x="breeze" & y="Noun"]
```

Our training data consists of 5 observations:

```
[x="breeze" & y="Verb"]
[x="breeze" & y="Verb"]
[x="breeze" & y="Verb"]
[x="breeze" & y="Noun"]
[x="breeze" & y="Noun"]
```

The maxent model will be trained in the usual way to give the feature f a weight λ, so that the model expectation for the feature matches its empirical expectation.

What probability will the model give to $P(Other|breeze)$?

Your Answer		Score	Explanation
○ $1/10$			
◉ $3/10$	✔	1.00	Correct!
○ $3/7$			
○ $3/14$			
Total		1.00 / 1.00	

Question Explanation

$P(Noun|breeze) = 0.4$, and the rest is uniform, so 0.3 for each other class.
You can also confirm this from:
$P(Noun|make) = P(Other|make) = e^0/(e^{\log(4/3)} + e^0 + e^0) = 1/(4/3 + 1 + 1) = 3/10$

Question 3

Suppose we build a maxent model for part of speech tagging a word (x), over a set of just 3 parts of speech (y): $Noun$, $Verb$, and $Other$. Our model has just one feature:

```
f(x,y) = [x="make" & y="Verb"]
```

Our training data consists of 5 observations:

```
[x="make" & y="Verb"]
[x="make" & y="Verb"]
[x="make" & y="Verb"]
[x="make" & y="Verb"]
[x="make" & y="Noun"]
```

The maxent model will be trained in the usual way to give the feature f a weight λ, so that the model expectation for the feature matches its empirical expectation. Suppose we now put gaussian regularization into the same model over the same data with $\sigma = 1$.

The optimal weight of the feature will still be of the form $\log X$ (natural log). What is X?

Your Answer		Score	Explanation
○ $3/4$			
◉ $5/8$	✔	1.00	Incorrect.
○ 3			
○ 9			
Total		1.00 / 1.00	

Question Explanation

We have in terms of expectations that:

$4 - predicted\_count(\lambda) - \sigma = 0$

I.e., $4 - 5[e^\lambda/(e^\lambda + 2)] - 1 = 0$

So,

$3/5 = e^\lambda/(e^\lambda + 2)$
$3/5e^\lambda + 6/5 = e^\lambda$
$6/5 = 2/5e^\lambda$
$3 = e^\lambda$
$\lambda = \log 3$

Question 4

Do the part-of-speech tagging by hand for the following sentence using Penn Treebank POS tags:

Stanford/? and/? Open/? University/? have/? been/? downloaded/? 50/? million/? times/? on/? iTunes/? !/?

Your Answer		Score	Explanation
Stanford/NNP and/CC Open/NNP University/NNP have/VBP been/VBN downloaded/VBN 50/SYM million/CD times/NNS on/IN iTunes/NNS !/.			
Stanford/NNP and/CC Open/NNP University/NNP have/VBP been/VBN downloaded/VBN 50/CD million/CD times/NNS on/IN iTunes/NNP !/.	✔	1.00	Correct!
Stanford/NNP and/CC Open/NNP University/NNP have/VB been/VBN downloaded/VBN 50/CD million/CD times/NNS on/JJ iTunes/NNS !/.			
Stanford/NNP and/CC Open/JJ University/NN have/VBP been/VBN downloaded/VBN 50/CD million/CD times/NNS on/IN iTunes/NNP !/.			
Total		1.00 / 1.00	

Question 5

Which of the given bracket groups is a constituent in the following sentence?

Sentence: *When she was younger she would play outside with her friends until the bell rang for*

dinner.

[When she was younger] [[she would play] outside] [with her friends] [[until the bell rang] for dinner].

Your Answer		Score	Explanation
outside with her friends			
she would play outside	✔	1.00	Correct!
play outside			
she was younger			
Total		1.00 / 1.00	