

Question1

$$1.400$$

$$2.400^2=160000$$

$$3.400^3=64000000$$

Question2

$$p(i|<s>)=2/4=1/2$$

$$p(</s>|am)=1/1=1$$

$$p(\text{not}|here)=0/2=0$$

$$p(\text{not}|dog\ is)=1/1=1$$

Question3

1. $p(<s> \text{ I want English food } </s>)=$
 $p(i|<s>)p(\text{want}|i)p(\text{english}|\text{want})p(\text{food}|\text{english})p(<s>|\text{food})$
 $=0.25*0.33*0.0011*0.5*0.68=0.000031$
2. We have 6 words, $n=6$
 $\text{perplexity} = p(w_1 w_2 \dots w_n)^{-1/n}=0.000031^{-1/6}$

Question4

$$1/2+1/2=1 \text{ (we will interpolate bigram and unigram)}$$

$$\begin{aligned} p(\text{Sam}|am) &= 1/2 * p(\text{sam}) + 1/2 * p(\text{Sam}|am) \\ &= 0.5 * 4/25 + 0.5 * 2/3 \\ &= 31/75 \end{aligned}$$

Question5 (Use Add one smoothing instead of MLE)

New Doc : fast, couple, shoot, fly

Comedy1 :fun, couple, love, love

Comedy2: couple, fly, fast, fun, fun

Action1 : fast, furious, shoot

Action2 : furious, shoot, shoot, fun

Action3: fly, fast, shoot, love

$|V| = \text{vocabulary size} = \{\text{fun, couple, love, fly, fast, furious, shoot}\} = 7$

$$p(\text{fast}|\text{Comedy}) = (1+1)/(9+7) = 1/8$$

$$p(\text{couple}|\text{Comedy}) = (2+1)/(9+7) = 3/16$$

$$p(\text{shoot}|\text{Comedy}) = (0+1)/(9+7) = 1/16$$

$$p(\text{fly}|\text{Comedy}) = (1+1)/(9+7) = 1/8$$

$$p(\text{Comedy}|\text{Doc}) = (1/8) * (1/8) * (3/16) * (1/16) * (2/5) = 3/40960$$

$$p(\text{fast}|\text{Action}) = (2+1)/(11+7) = 3/18$$

$$p(\text{couple}|\text{Action}) = (0+1)/(11+7) = 1/18$$

$$p(\text{shoot}|\text{Action}) = (4+1)/(11+7) = 5/18$$

$$p(\text{fly}|\text{Action}) = (1+1)/(11+7) = 2/18$$

$$p(\text{Action}|\text{Doc}) = (3/18) * (1/18) * (5/18) * (2/18) * (3/5) = 1/5832$$

$$p(\text{Action}|\text{Doc}) > p(\text{Comedy}|\text{Doc})$$

The action class will be more likely

Question 6:

- $p(y=1|x) = \text{stimod}([1, 1, 0, 1] \cdot [3, -5, 1, 2] + 3)$
 $= \text{stimod}(3)$
 $= 1/(1 + e^{-3}) = 0.95$
 $p(y=0|x) = 1 - 0.95 = 0.05$

2. For $y=1$
answer $= -\log(\text{stimod Function}(w \cdot x + b))$
 $= -\log(0.95) = 0.05$
3. For $y=0$
answer $= -\log(0.05) = 2.996$

Question7.

We have two class EOS and not-EOS

I may simply define a lowercase word followed by a period as EOS (with positive weight).

Otherwise, it is defined as not-EOS. For example, U.S. is define as non-EOS because it is upper case that followed by a period