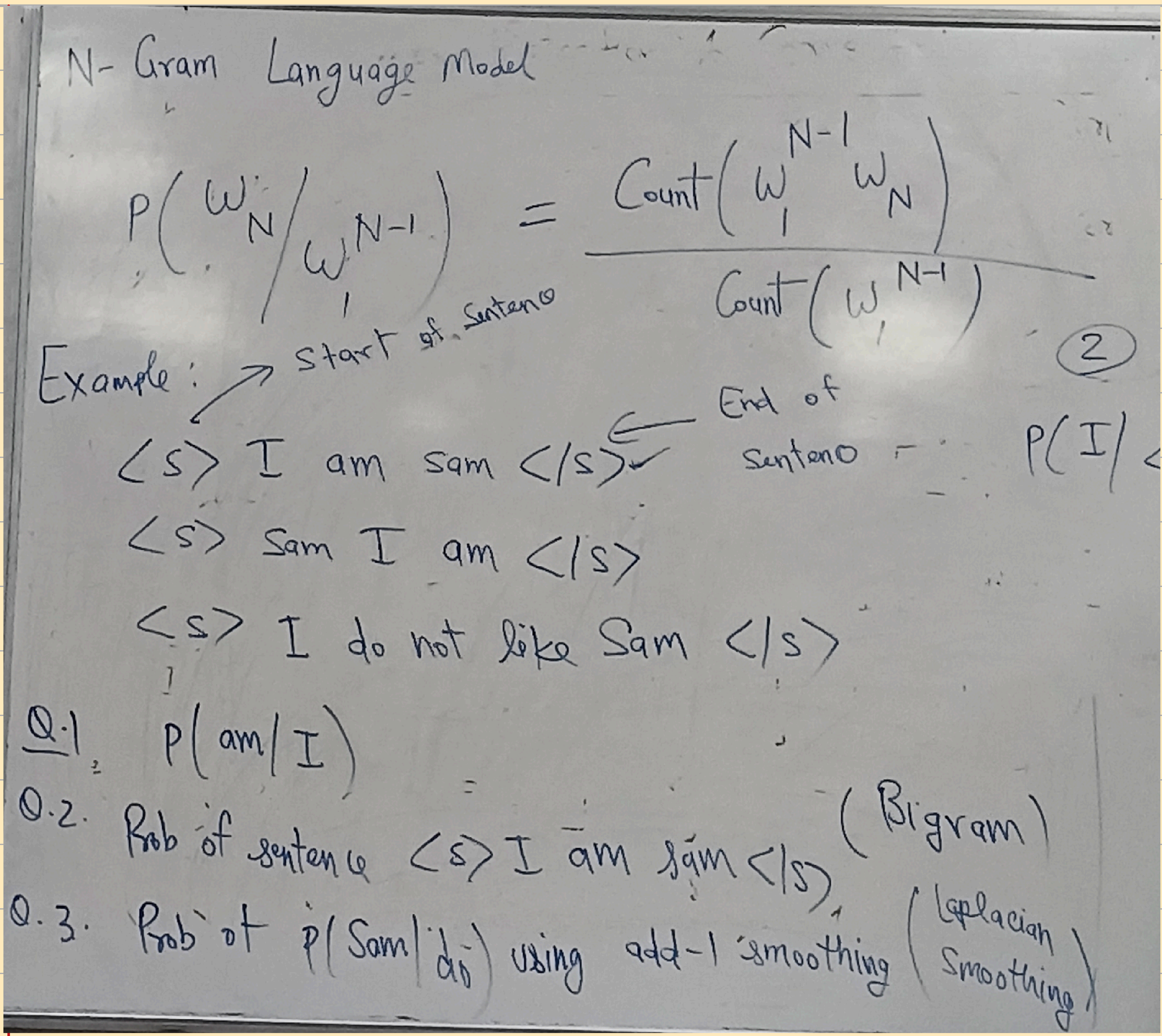
10/02/26



N- Gram Language Model

$$P\left(w_N / w_1^{N-1}\right) = \frac{Count\left(w_1^{N-1} w_N\right)}{Count\left(w_1^{N-1}\right)}$$

Example: → start of sentence

<s> I am sam </s> ← End of sentence

<s> Sam I am </s>

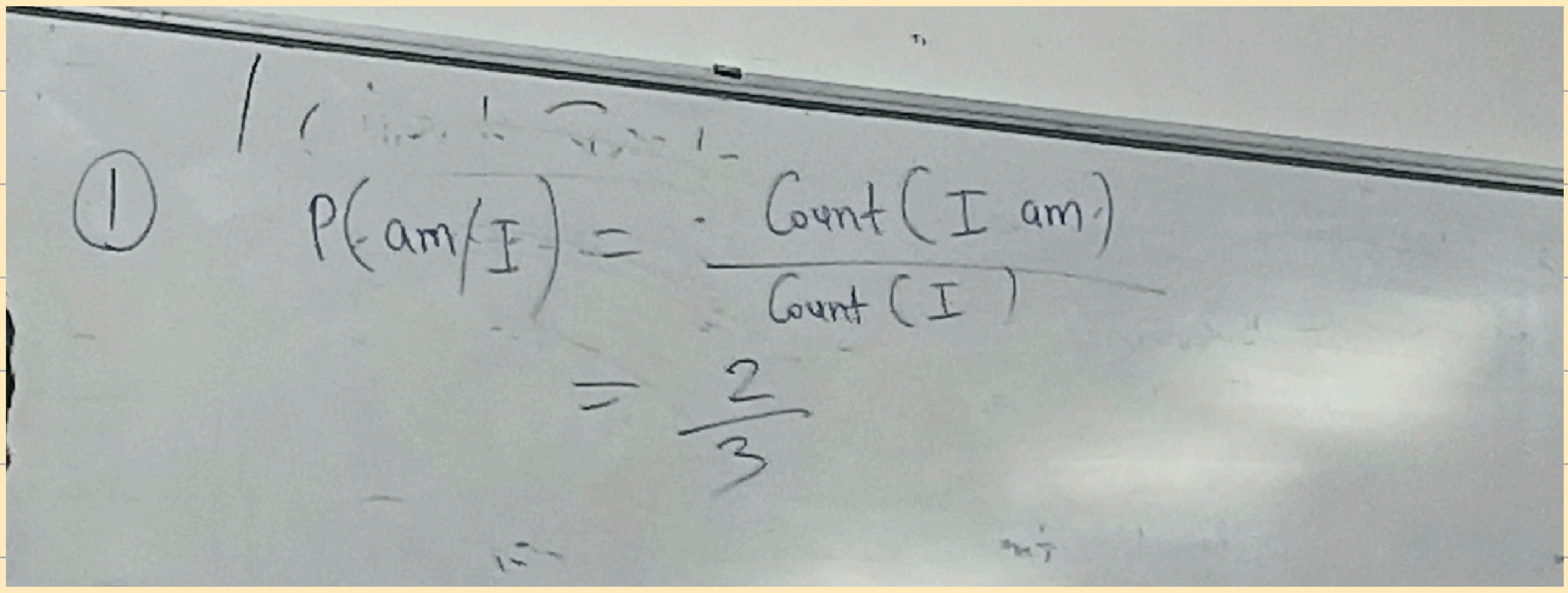
<s> I do not like Sam </s>

(2) P(I/<

Q.1 P(am/I)

Q.2. Prob of sentence <s> I am sam </s> (Bigram)

Q.3. Prob of p(Sam/do) using add-1 smoothing (Laplacian Smoothing)



(1)

$$P(am/I) = \frac{Count(I\ am)}{Count(I)} = \frac{2}{3}$$

N-Gram Language Model

$$P\left(w_N / w_1^{N-1}\right) = \frac{Count\left(w_1^{N-1} w_N\right)}{Count\left(w_1^{N-1}\right)}$$

Example: → start of sentence

<s> I am sam </s> ← End of sentence

<s> Sam I am </s>

<s> I do not like Sam </s>

(2) P(I/<

Q.1 $P(am/I)$

Q.2. Prob of sentence <s> I am sam </s> (Bigram)

Q.3. Prob of $P(Sam/do)$ using add-1 smoothing (Laplacian Smoothing)

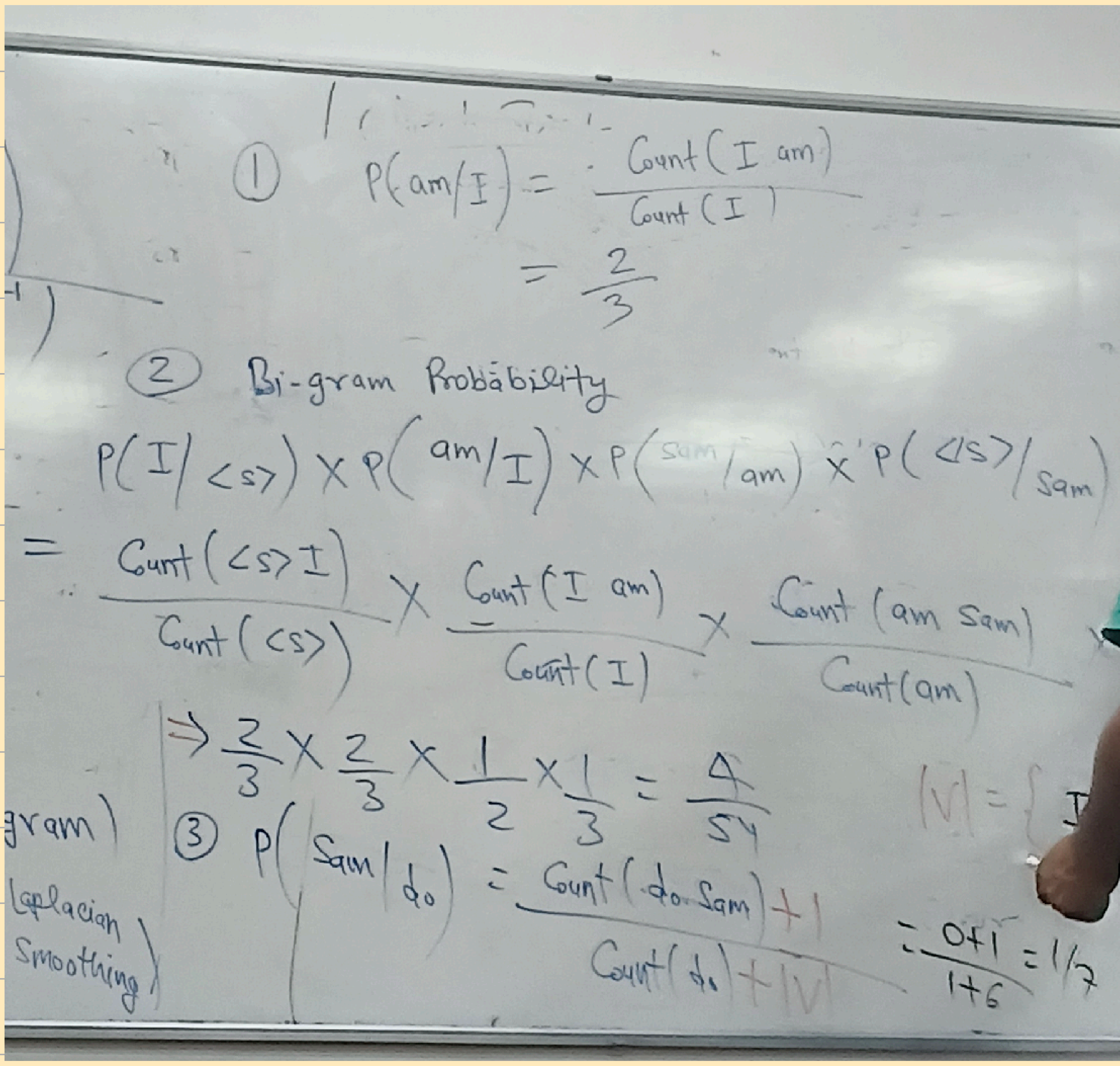
(1) $$P(am/I) = \frac{Count(I\ am)}{Count(I)} = \frac{2}{3}$$

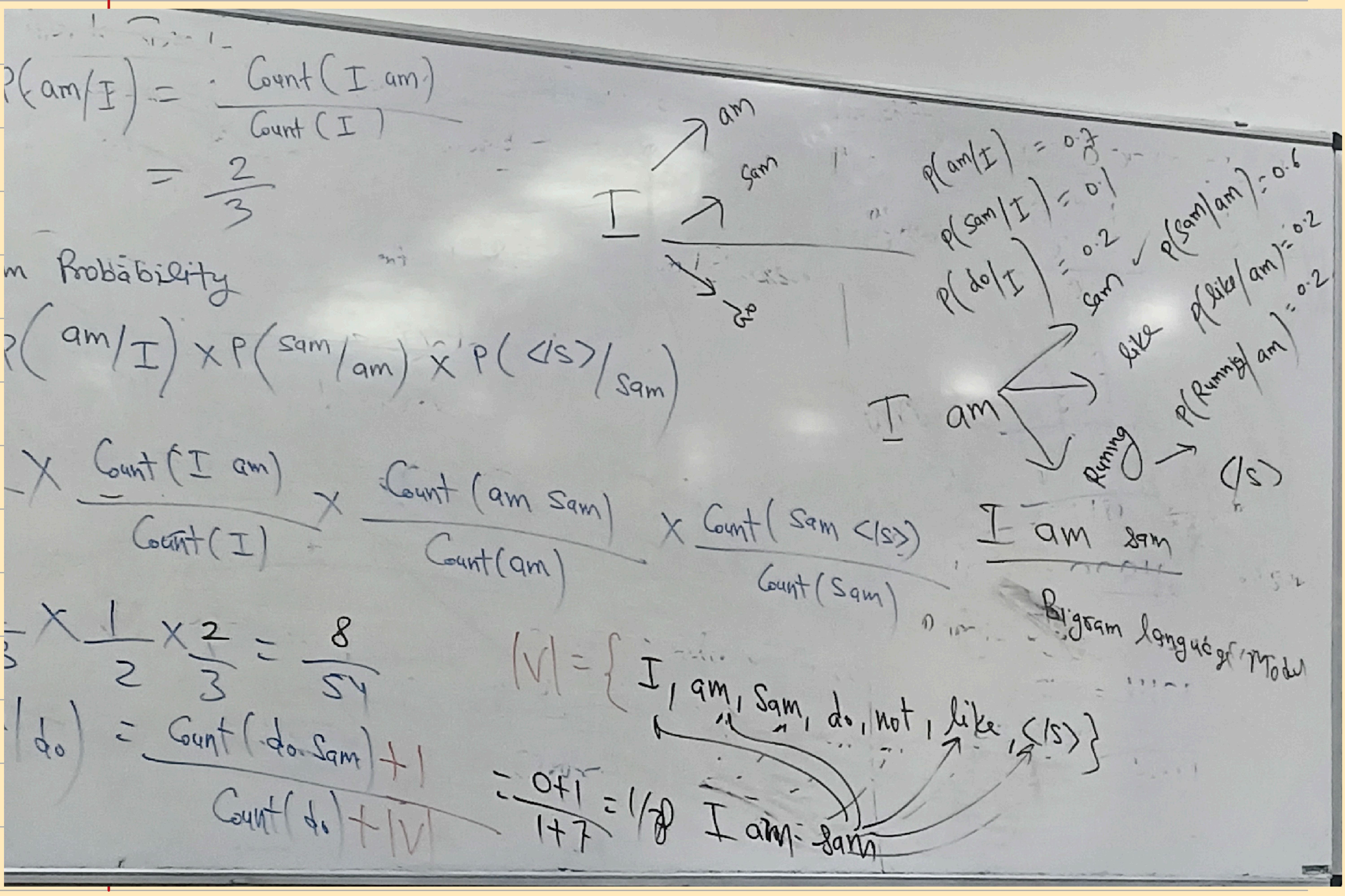
(2) Bi-gram Probability

$$P(I/\langle s\rangle) \times P(am/I) \times P(sam/am) \times P(\langle/s\rangle/sam)$$

$$\frac{Count(\langle s\rangle I)}{Count(\langle s\rangle)} \times \frac{Count(I\ am)}{Count(I)} \times \frac{Count(am\ sam)}{Count(am)} \times \frac{Count(sam\ \langle/s\rangle)}{Count(sam)}$$

$$= \frac{2}{3} \times \frac{2}{3} \times \frac{1}{2} \times \frac{2}{3} = \frac{8}{54}$$


① P(am/I) = Count(I am) / Count(I)
= 2/3
② Bi-gram Probability
P(I/<s>) x P(am/I) x P(sam/am) x P(</s>/sam)
= Count(<s> I) / Count(<s>) x Count(I am) / Count(I) x Count(am sam) / Count(am)
⇒ 2/3 x 2/3 x 1/2 x 1/3 = 4/54
gram)
Laplacian Smoothing)
③ P(sam/do) = Count(do sam)+1 / Count(do)+|V|
= 0+1 / 1+6 = 1/7
|V| = {


P(am/I) = Count(I am) / Count(I)
= 2/3
Probability
P(am/I) x P(sam/am) x P(</s>/sam)
X Count(I am) / Count(I) x Count(am sam) / Count(am) x Count(sam </s>) / Count(sam)
X 1/2 x 2/3 = 8/54
(do) = Count(do sam)+1 / Count(do)+|V|
= 0+1 / 1+7 = 1/8
I am sam
|V| = { I, am, sam, do, not, like, </s> }
I → am
sam
do
P(am/I) = 0.7
P(sam/I) = 0.1
P(do/I) = 0.2
I am
sam
like
Running
</s>
P(sam/am) = 0.6
P(like/am) = 0.2
P(Running/am) = 0.2
I am sam
Bigram language Model

Perplexity: is a standard metric used to evaluate how well a language model predict a sentence.

Perplexity is inverse probability of test set normalized by the no of words. for a sentence W with N words

$$W = \{w_1, w_2 \ldots w_N\}$$

$$\text{Perplexity} = \left[P(w_1 w_2 \ldots w_N)\right]^{-\frac{1}{N}} = \sqrt[N]{\frac{1}{P(w_2/w_1) \times P(w_3/w_2) \ldots P(w_N/w_{N-1})}}$$

$P(w_1 w_2 \ldots w_N)$: Bigram or Trigram or N-gram

(For Bigram only)

ample:

<s> I am sam </s> ← End of sentence

<s> Sam I am </s>

<s> I do not like Sam </s>

$P(\text{I like Sam } \langle/s\rangle) \Rightarrow$ Bigram

$P(I/\langle s\rangle) \times P(like/I) \times P(Sam/like) \times P(\langle/s\rangle/Sam)$

Perplexity Score

Perplexity is inverse probability of test set normalized by the no of words. for a sentence W with N words

$$W = \{w_1, w_2 \ldots w_N\}$$

$$\sqrt[N]{\frac{1}{P(s)}}$$

$N = 4$

$$\text{Perplexity} = \left[P(w_1 w_2 \ldots w_N)\right]^{-\frac{1}{N}} = \sqrt[N]{\frac{1}{P(w_2/w_1) \times P(w_3/w_2) \ldots P(w_N/w_{N-1})}}$$

Bigram or Trigram or N-gram

(For Bigram only)

(1) $P(am/I) = \frac{Count(I\ am)}{Count(I)} = \frac{2}{3}$

Probability

$P(am/I) \times P(Sam/am) \times \frac{Count(I\ am)}{Count(I)} \times$

$= \frac{?}{3} \times \frac{1}{2} \times \frac{2}{3} = \frac{8}{5?}$

$= \frac{Count(do\ Sam)}{Count(do)}$