

CA4012: Statistical Machine Translation

Assignment 2: Language Modelling

Eoin Murphy 11487358

Training Data D :

<s> I am a student </s>

<s> he is a teacher </s>

<s> she is a student </s>

In the training data I add the start of sentence token <s> and the end of sentence token </s> this is just make it easier for calculating bi-grams.

Q1

Uni-gram model M_1 :

$$p(w) = \frac{\text{occurences of } w}{\text{number of tokens}}$$

Word w	$P(w)$
I	$1/18=0.05$
am	$1/18=0.05$
a	$3/18=0.16$
student	$2/18=0.1$
he	$1/18=0.05$
is	$2/18=0.1$
teacher	$1/18=0.05$
she	$1/18=0.05$
<s>	$3/18=0.16$
</s>	$3/18=0.16$

Bi-gram model M_2 :

$$p(y|x) = \frac{\text{occurences of 'x y'}}{\text{occurencs of x}}$$

w_x, w_y	$p(w_y w_x)$
<s> , I	$1/3=0.3$
I , am	$1/1=1$
am , a	$1/1=1$
a , student	$2/3=0.6$
student , </s>	$2/2=1$
<s> , he	$1/3=0.3$
he , is	$1/1=1$
is , a	$2/2=1$
a , teacher	$1/3=0.3$
teacher , </s>	$1/1=1$
<s> , she	$1/3=0.3$
she , is	$1/1=1$

Q2

Words in language = { I, am, a, student, he, is, teacher, she }

vocabulary size with sentence boundaries = 9

Smoothed count table with our alpha value applied:

	I	am	a	student	he	is	teacher	she	</s>	SUM
<s>	1+0.2=1.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	1+0.2=1.2	0+0.2=0.2	0+0.2=0.2	1+0.2=1.2	0+0.2=0.2	4.8
I	0+0.2=0.2	1+0.2=1.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	2.8
am	0+0.2=0.2	0+0.2=0.2	1+0.2=1.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	2.8
a	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	2+0.2=2.2	0+0.2=0.2	0+0.2=0.2	1+0.2=1.2	0+0.2=0.2	0+0.2=0.2	4.8
student	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	2+0.2=2.2	3.8
he	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	1+0.2=1.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	2.8
is	0+0.2=0.2	0+0.2=0.2	2+0.2=2.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	3.8
teacher	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	1+0.2=1.2	2.8
she	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	1+0.2=1.2	0+0.2=0.2	0+0.2=0.2	0+0.2=0.2	2.8

where:

add - α smoothing:

$$p = \frac{c + \alpha}{n + \alpha v}$$

$$p = p(w_n | w_1, \dots, w_{n-1})$$

c = count of n-gram in corpus

n = count of history in corpus

v = vocabulary size

$\alpha = 0.2$

Smoothed probability table:

	I	am	a	student	he	is	teacher	she	</s>
<s>	0.25	0.0416	0.0416	0.0416	0.25	0.0416	0.0416	0.25	0.0416
I	0.0714	0.4285	0.0714	0.0714	0.0714	0.0714	0.0714	0.0714	0.0714
am	0.0714	0.0714	0.4285	0.0714	0.0714	0.0714	0.0714	0.0714	0.0714
a	0.0416	0.0416	0.0416	0.4583	0.0416	0.0416	0.25	0.0416	0.0416
student	0.0526	0.0526	0.0526	0.0526	0.0526	0.0526	0.0526	0.0526	0.5789
he	0.0714	0.0714	0.0714	0.0714	0.0714	0.4285	0.0714	0.0714	0.0714
is	0.0526	0.0526	0.5789	0.0526	0.0526	0.0526	0.0526	0.0526	0.0526
teacher	0.0714	0.0714	0.0714	0.0714	0.0714	0.0714	0.0714	0.0714	0.4285
she	0.0714	0.0714	0.0714	0.0714	0.0714	0.4285	0.0714	0.0714	0.0714

Bag of Words B :

{ teacher, she, a, is }

Exhausting all the possible sentences using our bag of words B gives us 4! possible sentences to calculate.

$$\begin{aligned} p(<s> \text{ teacher she a is } </s>) &= p(\text{teacher}|<s>) \times p(\text{she}|\text{teacher}) \times p(\text{a}|\text{she}) \times p(\text{is}|\text{a}) \times p(</s>|\text{is}) \\ &= 0.0416 \times 0.0714 \times 0.0714 \times 0.0416 \times 0.0526 \\ &= 0.0000004640 \end{aligned}$$

$$\begin{aligned} p(<s> \text{ teacher she is a } </s>) &= p(\text{teacher}|<s>) \times p(\text{she}|\text{teacher}) \times p(\text{is}|\text{she}) \times p(\text{a}|\text{is}) \times p(</s>|\text{a}) \\ &= 0.0416 \times 0.0714 \times 0.4285 \times 0.5789 \times 0.0416 \\ &= 0.00003062 \end{aligned}$$

$$\begin{aligned} p(<s> \text{ teacher is she a } </s>) &= p(\text{teacher}|<s>) \times p(\text{is}|\text{teacher}) \times p(\text{she}|\text{is}) \times p(\text{a}|\text{she}) \times p(</s>|\text{a}) \\ &= 0.0416 \times 0.0714 \times 0.0526 \times 0.0714 \times 0.0416 \\ &= 0.0000004640 \end{aligned}$$

$$\begin{aligned} p(<s> \text{ teacher a she is } </s>) &= p(\text{teacher}|<s>) \times p(\text{a}|\text{teacher}) \times p(\text{she}|\text{a}) \times p(\text{is}|\text{she}) \times p(</s>|\text{is}) \\ &= 0.0416 \times 0.0714 \times 0.0416 \times 0.4285 \times 0.0526 \\ &= 0.000002784 \end{aligned}$$

$$\begin{aligned} p(<s> \text{ teacher is a she } </s>) &= p(\text{teacher}|<s>) \times p(\text{is}|\text{teacher}) \times p(\text{a}|\text{is}) \times p(\text{she}|\text{a}) \times p(</s>|\text{she}) \\ &= 0.0416 \times 0.0714 \times 0.5789 \times 0.0416 \times 0.0714 \\ &= 0.000005107 \end{aligned}$$

$$\begin{aligned} p(<s> \text{ teacher a is she } </s>) &= p(\text{teacher}|<s>) \times p(\text{a}|\text{teacher}) \times p(\text{is}|\text{a}) \times p(\text{she}|\text{is}) \times p(</s>|\text{she}) \\ &= 0.0416 \times 0.0714 \times 0.0416 \times 0.0526 \times 0.0714 \\ &= 0.0000004640 \end{aligned}$$

$$\begin{aligned} p(<s> \text{ she teacher a is } </s>) &= p(\text{she}|<s>) \times p(\text{teacher}|\text{she}) \times p(\text{a}|\text{teacher}) \times p(\text{is}|\text{a}) \times p(</s>|\text{is}) \\ &= 0.25 \times 0.0714 \times 0.0714 \times 0.0416 \times 0.0526 \\ &= 0.000002788 \end{aligned}$$

$$\begin{aligned} p(<s> \text{ she a teacher is } </s>) &= p(\text{she}|<s>) \times p(\text{a}|\text{she}) \times p(\text{teacher}|\text{a}) \times p(\text{is}|\text{teacher}) \times p(</s>|\text{is}) \\ &= 0.25 \times 0.0714 \times 0.25 \times 0.0714 \times 0.0526 \\ &= 0.00001675 \end{aligned}$$

$$\begin{aligned} p(<s> \text{ she a is teacher } </s>) &= p(\text{she}|<s>) \times p(\text{a}|\text{she}) \times p(\text{is}|\text{a}) \times p(\text{teacher}|\text{is}) \times p(</s>|\text{teacher}) \\ &= 0.25 \times 0.0714 \times 0.0416 \times 0.0526 \times 0.4285 \\ &= 0.00001673 \end{aligned}$$

$$\begin{aligned} p(<s> \text{ she is a teacher } </s>) &= p(\text{she}|<s>) \times p(\text{is}|\text{she}) \times p(\text{a}|\text{is}) \times p(\text{teacher}|\text{a}) \times p(</s>|\text{teacher}) \\ &= 0.25 \times 0.4285 \times 0.5789 \times 0.25 \times 0.4285 \\ &= 0.006643 \end{aligned}$$

$$\begin{aligned} p(<s> \text{ she teacher is a } </s>) &= p(\text{she}|<s>) \times p(\text{teacher}|\text{she}) \times p(\text{is}|\text{teacher}) \times p(\text{a}|\text{is}) \times p(</s>|\text{a}) \\ &= 0.25 \times 0.0714 \times 0.0714 \times 0.5789 \times 0.0416 \\ &= 0.00003069 \end{aligned}$$

$$\begin{aligned} p(<s> \text{ she is teacher a } </s>) &= p(\text{she}|<s>) \times p(\text{is}|\text{she}) \times p(\text{teacher}|\text{is}) \times p(\text{a}|\text{teacher}) \times p(</s>|\text{a}) \\ &= 0.25 \times 0.4285 \times 0.0526 \times 0.0714 \times 0.0416 \\ &= 0.00001673 \end{aligned}$$

$$\begin{aligned} p(<s> \text{ a teacher she is } </s>) &= p(\text{a}|<s>) \times p(\text{teacher}|\text{a}) \times p(\text{she}|\text{teacher}) \times p(\text{is}|\text{she}) \times p(</s>|\text{is}) \\ &= 0.0416 \times 0.25 \times 0.0714 \times 0.4285 \times 0.0526 \\ &= 0.00001673 \end{aligned}$$

$$\begin{aligned}
p(<s> a she teacher is </s>) &= p(a|<s>) \times p(she|a) \times p(teacher|she) \times p(is|teacher) \times p(</s>|is) \\
&= 0.0416 \times 0.0416 \times 0.0714 \times 0.0714 \times 0.0526 \\
&= 0.000004640 \\
p(<s> a she is teacher </s>) &= p(a|<s>) \times p(she|a) \times p(is|she) \times p(teacher|is) \times p(</s>|teacher) \\
&= 0.0416 \times 0.0416 \times 0.4285 \times 0.0526 \times 0.4285 \\
&= 0.00001671 \\
p(<s> a is she teacher </s>) &= p(a|<s>) \times p(is|a) \times p(she|is) \times p(teacher|she) \times p(</s>|teacher) \\
&= 0.0416 \times 0.0416 \times 0.0526 \times 0.0714 \times 0.4285 \\
&= 0.000002784 \\
p(<s> a teacher is she </s>) &= p(a|<s>) \times p(teacher|a) \times p(is|teacher) \times p(she|is) \times p(</s>|she) \\
&= 0.0416 \times 0.25 \times 0.0714 \times 0.0526 \times 0.0714 \\
&= 0.000002788 \\
p(<s> a is teacher she </s>) &= p(a|<s>) \times p(is|a) \times p(teacher|is) \times p(she|teacher) \times p(</s>|she) \\
&= 0.0416 \times 0.0416 \times 0.0526 \times 0.0714 \times 0.0714 \\
&= 0.000006405 \\
p(<s> is teacher she a </s>) &= p(is|<s>) \times p(teacher|is) \times p(she|teacher) \times p(a|she) \times p(</s>|a) \\
&= 0.0416 \times 0.0526 \times 0.0714 \times 0.0714 \times 0.0416 \\
&= 0.0000004640 \\
p(<s> is she teacher a </s>) &= p(is|<s>) \times p(she|is) \times p(teacher|she) \times p(a|teacher) \times p(</s>|a) \\
&= 0.0416 \times 0.0526 \times 0.0714 \times 0.0714 \times 0.0416 \\
&= 0.000006499 \\
p(<s> is she a teacher </s>) &= p(is|<s>) \times p(she|is) \times p(a|she) \times p(teacher|a) \times p(</s>|teacher) \\
&= 0.0416 \times 0.0526 \times 0.0714 \times 0.25 \times 0.4285 \\
&= 0.00001673 \\
p(<s> is a she teacher </s>) &= p(is|<s>) \times p(a|is) \times p(she|a) \times p(teacher|she) \times p(</s>|teacher) \\
&= 0.0416 \times 0.5789 \times 0.0416 \times 0.0714 \times 0.4285 \\
&= 0.00003065 \\
p(<s> is a teacher she </s>) &= p(is|<s>) \times p(a|is) \times p(teacher|a) \times p(she|teacher) \times p(</s>|she) \\
&= 0.0416 \times 0.5789 \times 0.4285 \times 0.0714 \times 0.0714 \\
&= 0.00005260 \\
p(<s> is teacher a she </s>) &= p(is|<s>) \times p(a|is) \times p(teacher|a) \times p(she|teacher) \times p(</s>|she) \\
&= 0.0416 \times 0.5789 \times 0.25 \times 0.0714 \times 0.0714 \\
&= 0.00003069
\end{aligned}$$

Q3

Perplexity

Prediction	P_{LM}	$-\log_2 P_{LM}$
(she <s>)	0.25	2.0
(is she)	0.4285	1.2226
(a is)	0.5789	0.7886
(teacher a)	0.25	2.0
(</s> teacher)	0.4285	1.2226
	Average	1.44676

