Itamar Trainin 315425967

Ex. 2 part 2 description file
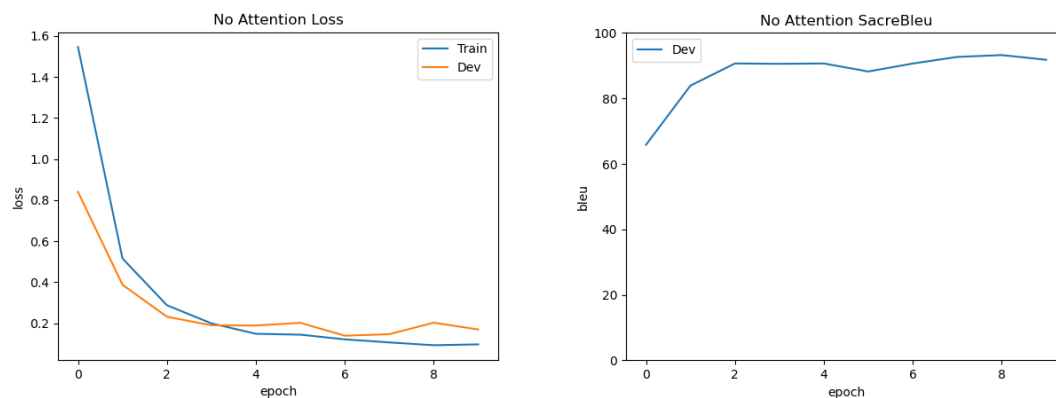
Note: In this part the Attention mechanism architecture was according to the architecture presented in the paper attacked to the assignment: `*Effective Approaches to Attention-based Neural Machine Translation*`, *by Luong et al*.

Hyperparameters:

```
embd_dim = 50
hidden_dim = 200
output_dim = 500
num_layers = 1

lr = 0.001
num_epochs = 10
```

Dev Bleu/Loss vs. Epoch Graph:

Results Comparison:

It is possible to see that the results were greatly improved when the attention was added (from around bleu score of 75 on the test set in part one to blue score of around 95. In addition the total running time was almost the same (about 10-15 seconds longer for each epoch).

Test results:

test-loss: 0.1122, test-bleu: 95.5989

opts: {'test_src': '../data/test.src', 'test_trg': '../data/test.trg', 'model_file': '../models/model_attention_93_50_200_500_0.001'}

\<s> o f \</s>	\<s> M r . \</s>	\<s> S a u n d e r s , \</s>
\<s> f o g . \</s>	\<s> S h e \</s>	\<s> w h o r l s \</s>
\<s> o f \</s>	\<s> p l a s t i z u l \</s>	\<s> n e x t \</s>
\<s> B o n d ' s \</s>	\<s> M A L E \</s>	\<s> T V \</s>
\<s> t h e \</s>	\<s> O F \</s>	\<s> h e r e . \</s>
\<s> h e r e , \</s>	\<s> U p \</s>	\<s> 1) \</s>
\<s> w i t h \</s>	\<s> c a m e r a , \</s>	\<s> t u r n e d \</s>
\<s> U z i . \</s>	\<s> H A L L W A Y \</s>	\<s> u p \</s>

<s> s a f e s t </s>
<s> h e a d </s>
<s> a t </s>
<s> h e </s>
<s> c u r r e n t l n </s>
<s> c l a i m i n g </s>
<s> d e c k - t h o u t e </s>
<s> i n </s>
<s> a f t e r o n o o </s>
<s> i n </s>
<s> D o e s </s>
<s> d o w n </s>
<s> H A R M S W A Y </s>
<s> I ' v e </s>
<s> T h e </s>
<s> s h o u l l </s>
<s> S I D N E Y </s>
<s> V I D E O </s>
<s> o f </s>
<s> B o n d </s>
<s> a n d </s>
<s> h e r e </s>
<s> Y e s . </s>
<s> h i m </s>
<s> B r i t i s h </s>
<s> h e r </s>
<s> h a p p y </s>
<s> l o n g </s>
<s> t h e </s>
<s> c a l l s </s>
<s> W O R M </s>
<s> c o p y </s>
<s> f e e t . </s>
<s> A </s>
<s> i n t o </s>
<s> - </s>
<s> h i m . </s>
<s> t h e </s>
<s> t h e </s>
<s> k n o w </s>
<s> m a n ' s </s>
<s> u n i v e r s e </s>
<s> - </s>
<s> H A R M S W A Y </s>
<s> W O R M ')) </s>
<s> t h e </s>
<s> M e n </s>
<s> E X T . </s>
<s> o f </s>
<s> T H E </s>
<s> m a r k s </s>
<s> k e e p i n g </s>
<s> W U . </s>
<s> A P P E A M S . </s>
<s> c a l l s </s>
<s> O b l i v i o n s , </s>
<s> a </s>
<s> C i t i b a n d . </s>

<s> . . . </s>
<s> t h r o u g h </s>
<s> w h o </s>
<s> t h e </s>
<s> l u r c h e s </s>
<s> o n </s>
<s> t h e m . </s>
<s> A t t e n t i n . </s>
<s> N o r t h </s>
<s> a l i v e i </s>
<s> t o </s>
<s> t h a t </s>
<s> W i t h </s>
<s> G A <NULL> A G E </s>
<s> C h i n a </s>
<s> t h i r d </s>
<s> a n d </s>
<s> r a d i o o </s>