

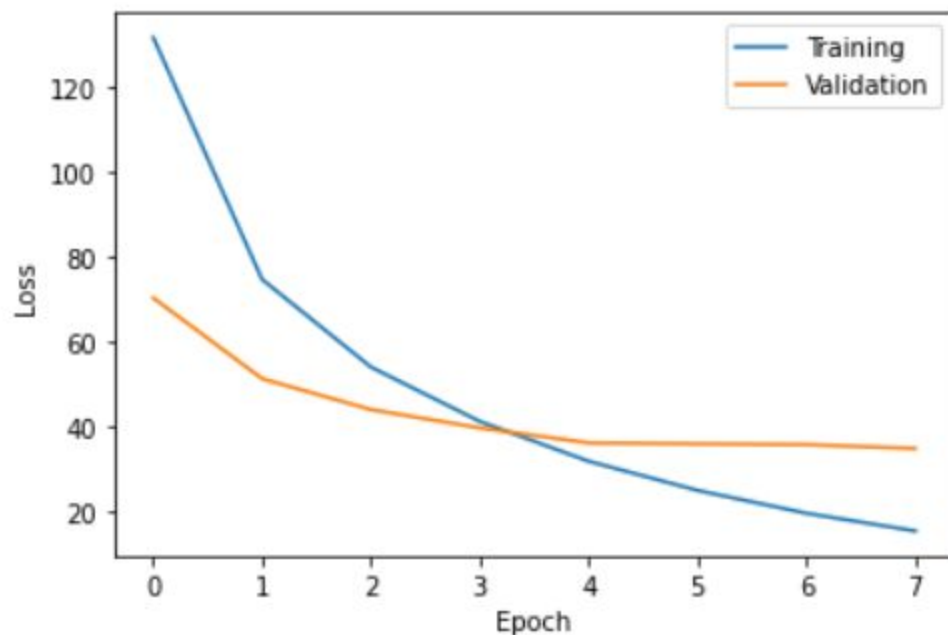
# Advanced NLP Assignment 2 Report

Tanish Lad (2018114005)

---

## 1. Encoder-Decoder Without Attention

The loss is taken as  $e^{\text{(actual\_loss)}}$  to get a better analysis of the models because the loss changed very minutely after epoch in the later part of the training and hence it was infeasible to compare without taking its exponent.



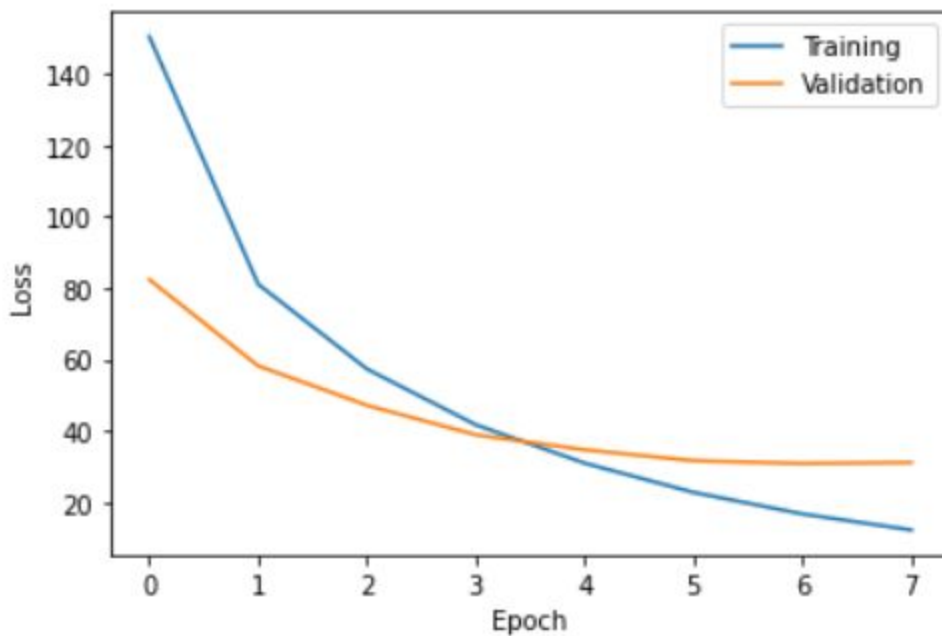
BLEU Score: 0.33565255962128826

## Analysis of 20 Sentences

Predicted	Actual
मैं तुम्हें साथ आ रहा .	-मैं तुम्हारे साथ चलूँगी।
चलो , चलो । चलो .	चलो, चलो चलते हैं
नहीं , हम हम सकते सकते हैं ।	ठीक है, हम उसे कर सकते हैं..
तुम मुझे के लिए कर रहे हैं	तुम मुझे मदद करने के लिए जा रहे हैं।.
" जानेमन मैं तुम पर कसम खाता । "	"जानेमन मैं तुम पर कसम खाता हूँ।".
नहीं , वह कहा नहीं नहीं था	ठीक है, वह कभी नहीं कहा..
क्या नहीं नहीं नहीं सकता	तुम ऐसा नहीं कर सकते!.
क्या क्या कर सकते हैं ?	तुम क्या कर सकते हो?.
नहीं , तुम नहीं सकते सकते हैं ।	नहीं है, तुम मर नहीं हो सकता।.
नहीं , नहीं , नहीं ,	- नहीं, नहीं, नहीं..
पुरुष अपनी 2 में रहा है	पुरुष अपनी मातृभाषा में बोल रहा है:.
तुम क्या कर रहे हैं	क्या कर रहे हो?.
तुम क्या कर रहे हो ?	- आप क्या कह रहे हैं?.
यही क्या चाह रहे हो ? ? ?	यही कह रहे हो?.
चलो से से बाहर .	चलो अब चलो यहाँ से ..
और वे वे नहीं कर सकते .	आप वे नहीं कर सकते पता है..
आप सिर्फ है कि <unk> .	आप बस लटक गया पता..
तुम क्या हो रहे हो	तुम कहाँ जा रहे हो?.
<unk> , यह नहीं नहीं कर सकते हैं .	. = नहीं, मैं ऐसा नहीं कर सकते..
" , पागलपन पागलपन के लिए नहीं है । "	"मैं तुम्हारे बिना शांति नहीं है।".

## 2.Encoder-Decoder With Attention

The loss is taken as  $e^{\text{(actual\_loss)}}$  to get a better analysis of the models because the loss changed very minutely after epoch in the later part of the training and hence it was infeasible to compare without taking its exponent.



BLEU Score: 0.3390682622909464

## Analysis of 20 Sentences

Predicted	Actual
<unk> तुम क्या कर रहे हो	आप क्या कर रहे हैं?.
<unk> गलती नहीं है .	यह उसकी गलती नहीं है।.
<unk> ठीक ठीक हो जाएगा .	मैं सब ठीक हो जाएगा..
<unk> यह क्या है रहा है	- सब कैसा चल रहा है?.
<unk> क्या क्या है ?	क्या मिला?.
मैं बारे में चिंता नहीं करो	उसपे ध्यान मत दीजिए।.
<unk> हम एक मिला है .	और वह हमें मिल गया।.
, एक मिनट मिनट ..	Saru, एक मिनट रुको।.
<unk> तुम क्या कर रहे हो	- क्या कर रहे हो?.
लक्ष्य के के लिए ..	तुम धीरे करोगे?.
<unk> <unk> है , है ठीक है ..	ठीक है..
एक अच्छा एक है .	आप एक अच्छा एक है..
<unk> इस इस साथ साथ हैं ?	आप इस के साथ ठीक हो?.
इस इस साथ साथ हैं ?	- नहीं। चलो।.
नाम नाम नाम है ?	- तुम्हारा नाम क्या है?.
<unk> यह पत्नी है है ?	जहां इस आदमी पर है?.
हम कहाँ जा जा रहे ?	हम कहां जा रहे हैं ?.
तुम आप जा कर कर सकते	- हलो, क्या मैं आपकी मदद कर सकता हूँ?.

unk> बंद बंद ..	May get warded off....
<unk> कुछ कुछ है .	वहाँ कुछ गड़बड़ है..

## Analysis

- Attention outputs are better overall.
- Lesser data leads to bad results
- Even the dataset is not perfect and it does not have perfect translations.  
This is one of the reasons along with lesser data for the model having bad results.
- Some of the predicted sentences were perfect acceptable translations of the input sentence (marked in bold for model without attention)
- Both the models correctly predict the punctuation marks with very high accuracy.
- Comparing both the models, the loss of model with attention is less and also the bleu score is higher. Hence the model with attention is definitely and obviously a better model.

### **Model Checkpoints are present at the below drive link**

- <https://drive.google.com/drive/folders/1Rl9JYtybVDtR96JdgVNDRKk5VOV1P513?usp=sharing>