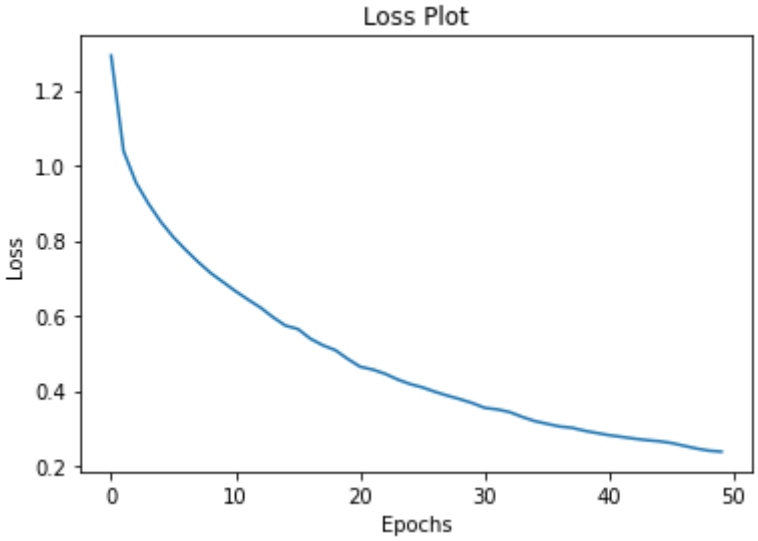
Assignment 4
Output

Part1:

We trained the model as given in the paper, following is the loss plot :

Loss Plot for 50 epochs :



Following are the results on **BLEU 1-4** AND **METEOR** Metric -

We evaluated the Validation Set and Test Set on **BLEU 1-4** AND **METEOR** Metric following are the results:

This table contains individual ***n-gram BLEU*** scores.

S.No.	BLEU-1	BLEU-2	BLEU-3	BLEU-4	METEOR
1. Validation Set	9.49	87.5	87.6	88.2	24.734
2. Test Set	8.9	87.6	87.9	86.9	24.787

This table contains ***cumulative n-gram BLEU*** scores with equal weightage.

S.No.	BLEU 1-2	BLEU 1-3	BLEU 1-4
Validation Set	27.79	40.332	49.179
Test Set	27.32	41.04	49.102

Image 1 :

Predicted Caption: dog playing in the field <end>

Image 2:

Predicted Caption: two people are walking towards a beach <end>

Image 3:

Predicted Caption: a brown dog is on a shallow water <end>

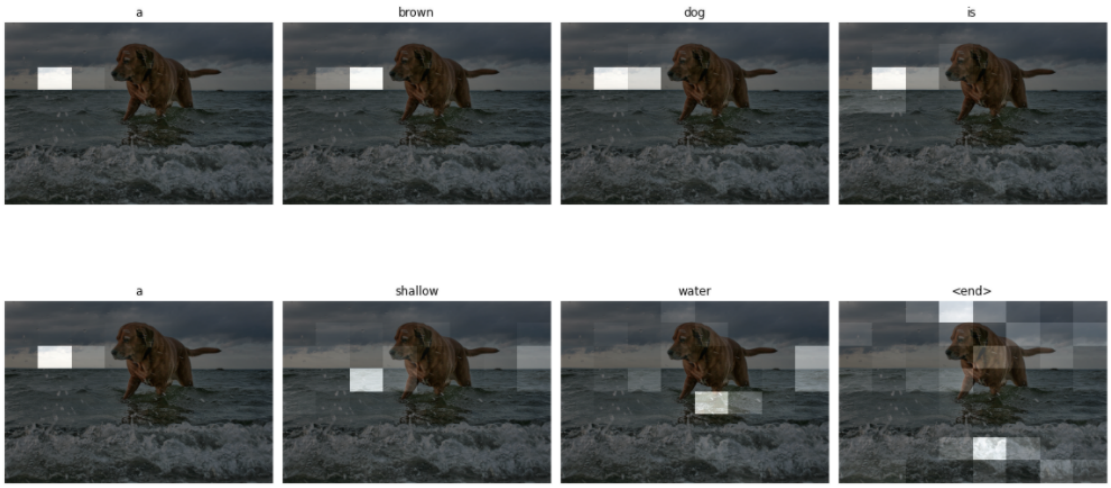
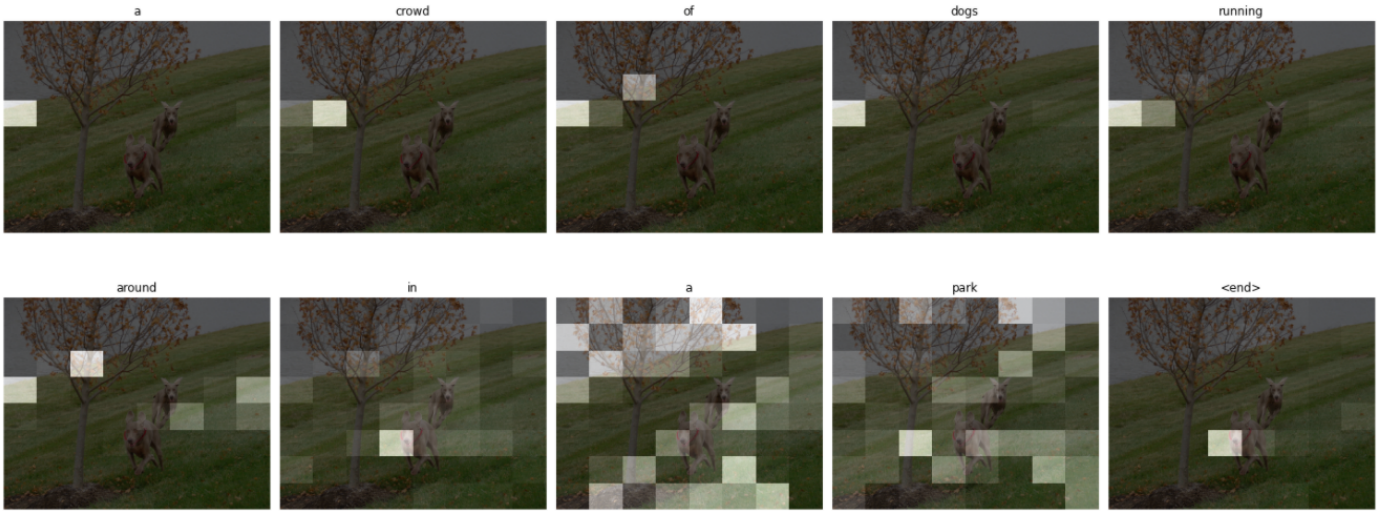


Image 4:

Predicted Caption: a small child walking on the grass <end>

Image 5:

Predicted Caption: a crowd of dogs running around in a park <end>



Part2:

We tried the model architecture of the paper and used this architecture, WithoutAttention and with attention (IAN). Results are:

Laptop:

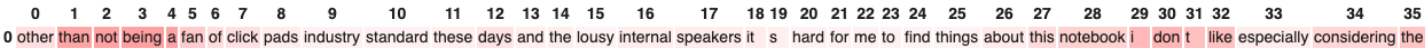
Without Attention: Test-Accuracy → 66.3%

With IAN: Test-Accuracy → 71.5%

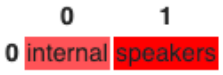
Attention weights for inputs:

Input 1: `other than not being a fan of click pads  industry standard these days  and the lousy internal speakers  it s hard for me to find things about this notebook i don t like  especially considering the  350 price tag`
`==> internal speakers`
`==> negative`

Context →

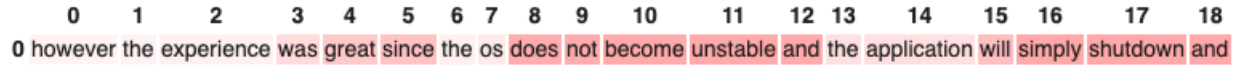


Target →



Input 2: `however  the experience was great since the os does not become unstable and the application will simply shutdown and reopen`
`==> os`
`==> positive`

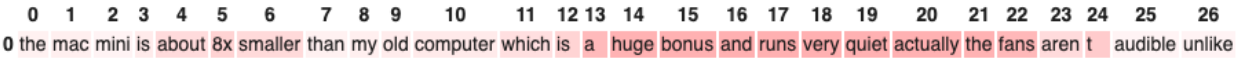
Context →



Target →

Input 3: `the mac mini is about 8x smaller than my old computer which is a huge bonus and runs very quiet  actually the fans aren t audible unlike my old pc`
`==> runs`
`==> positive`

Context →



Target →

Restaurant:

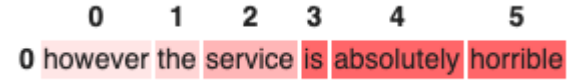
Without Attention: Test-Accuracy → 76.4%

With IAN: Test-Accuracy → 78.4%

Attention weights for inputs:

Input 1: `however  the service is absolutely horrible`
`==> service`
`==> negative`

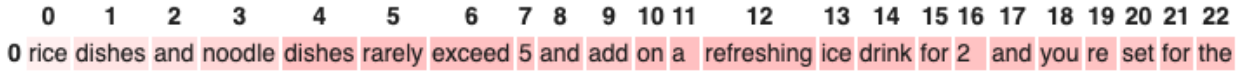
Context →



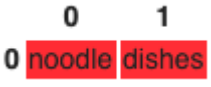
Target →

Input 2: `rice dishes and noodle dishes rarely exceed  5 and add on a refreshing ice drink for  2 and you re set for the night`
`==> noodle dishes`
`==> positive`

Context →



Target →



Input 3: `the service is really fast and friendly  and the value is great`
`==> value`
`==> positive`

Conext →

	0	1	2	3	4	5	6	7	8	9	10	11
0	the	service	is	really	fast	and	friendly	and	the	value	is	great

Target →

	0
0	value