"Deep Learning Report HW2"

Title: Deep Learning Course

Instructor: Dr. Feng Luo

Student: Sayed Pedram Haeri Boroujeni

Date: 03. March. 2023

Step 1: Details of my model_S2VT:

```
S2VT(
    (drop): Dropout(p=0.5, inplace=False)
    (linear1): Linear(in_features=4096, out_features=500, bias=True)
    (linear2): Linear(in_features=500, out_features=4096, bias=True)
    (lstm1): LSTM(500, 500, batch_first=True, dropout=0.5)
    (lstm2): LSTM(1000, 500, batch_first=True, dropout=0.5)
    (embedding): Embedding(4096, 500)
)
```

Step 2: Details of the data (train):

```
['<BOS> A cat rubs up against an infant <EOS>',
 "<BOS> A black and white colored cat is snuggling itself against a baby's face. <EOS>",
 '<BOS> A black-and-white-cat is snuggling heads with a baby. <EOS>',
 '<BOS> A cat is nuzzling a baby. <EOS>',
 '<BOS> A cat is rubbing against a baby. <EOS>',
 '<BOS> A cat is rubbing against a baby. <EOS>',
 "<BOS> A cat is rubbing against baby's face. <EOS>",
 '<BOS> A cat is rubbing its face on a small child. <EOS>',
 "<BOS> A cat is rubbing it's head against a young child. <EOS>",
 '<BOS> A cat is snuggling a baby. <EOS>',
 '<BOS> A cat is snuggling with a baby. <EOS>',
 '<BOS> A cat plays with a baby. <EOS>',
 '<BOS> A cat rubs its head against a baby. <EOS>',
 '<BOS> A cat rubs up against a baby. <EOS>',
 '<BOS> The cat cuddled with the baby. <EOS>',
 '<BOS> The cat is caressing the baby. <EOS>',
 '<BOS> The cat snuggled with the baby. <EOS>']
```

Step 3: Details of the data (test):

```
['<BOS> Someone is cutting a stack of crisp tortillas into quarters. <EOS>',
 '<BOS> A person is slicing a stack of tortillas. <EOS>',
 '<BOS> A person is slicing a stack of tortillas. <EOS>',
 '<BOS> Someone cuts a stack of tortillas into quarters and then they start to slice a tomato. <EOS>',
 '<BOS> A person is slicing roties  into pieces. <EOS>',
 '<BOS> A person is slicing flour tortillas with a knife. <EOS>',
 '<BOS> A man is cutting vegetables. <EOS>',
 '<BOS> Someone is slicing something. <EOS>',
 '<BOS> Someone is cutting a stack of tortillas. <EOS>',
 "<BOS> Someone is slicing tortila's. <EOS>",
 '<BOS> A person is slicing some tortillas. <EOS>',
 '<BOS> Someone cut tortillas into quarters. <EOS>',
 '<BOS> The cook is slicing tortillas. <EOS>',
 '<BOS> Tortillas are cut. <EOS>',
 '<BOS> A person cuts up some tortillas. <EOS>',
 '<BOS> The lady cut through the tortillas. <EOS>']
```

Step 4: Details of the vocabularies and their values:

	Word	Value
0	multi-serve	1
1	accident.	1
2	asphalt	1
3	barrier.	1
4	skidded.	1
5	rhino	12
6	rhino.	2
7	rhinos	2
8	rhinoceros	1
9	mom.	2
10	Yellow	1
11	capped	2
12	Mt.	1
13	Fuji.	2
14	Mount	1
15	countryside.	2

Step 5: Details of the training procedure:

No.	Information	Detail
1	Dataset	MSVD
2	Number of Epochs	30
3	Number of Iterations	2000
4	Batch Size	10
5	LR	0.001

Step 6: Some examples during the training procedure:

Epoch1: Generated Output

```
A man is playing a guitar.
A man is slicing a bowl.
A man is slicing a piece of a bowl.
A woman is slicing a piece of a bowl.
A woman is slicing a piece of a bowl.
A man is playing a guitar.
A man is playing a guitar.
A man is playing a guitar.
A man is playing a guitar.
A woman is slicing a piece of a bowl.
```

Epoch1: Ground Truth

```
<BOS> A gymnast is doing back flips then falls.
<BOS> The cat swatted someone from inside a box.
<BOS> A woman is putting a dish of liquid in a microwave.
<BOS> A woman is skewering shrimp.
<BOS> A man empties rice roni from a small carton box to a pan containing melted butter.
<BOS> A band is playing instruments.
<BOS> A small car explodes in a field.
<BOS> A man shoots another man on a horse.
<BOS> A boy is kicking a ball.
<BOS> The lady cut up shrimp and put it in the salad.
```

Epoch1: Images

Epoch15: Generated Output

```
A woman is putting ingredients in a bowl.
A person is adding a piece of meat in a bowl.
A woman is putting a piece of bread.
A man is slicing a piece of bread.
A man is riding a horse.
A man is pouring oil into a pot.
A man is pouring oil into a pot.
A cat is playing with a ball.
A woman is talking on a phone.
A girl is playing a flute.
```

Epoch15: Ground Truth

```
<BOS> A woman is cracking eggs.
<BOS> A man is beating two eggs in a bowl with a fork.
<BOS> A girl is cutting a piece of plastic.
<BOS> A man is slicing a cucumber.
<BOS> The boy hit a sign on his bike and fell off.
<BOS> The man put water into the tomato sauce can.
<BOS> A man is adding oil to a bowl.
<BOS> A cat is eating out of a bowl and a puppy is biting the cats ear.
<BOS> The man ate a banana.
<BOS> A girl plays the violin.
```

Epoch15: Images

Epoch30: Generated Output

```
A boy is looking out a window.
A toy fish is moving in the water.
A man is staring at his hand.
Two girls are talking on the phone.
A baby is sneezing and other baby.
A man is playing the guitar.
A person is frying something.
A man is mixing ingredients in a bowl.
Two cats are fighting.
A panda is lying on logs.
```

Epoch30: Ground Truth

```
<BOS> A small boy has placed his hands on the window pane and
<BOS> Bath toys are swimming.
<BOS> A bright triangle shape appears on the hand of a man.
<BOS> Two girls are talking in telephone.
<BOS> The babies are sitting down.
<BOS> A man is playing a guitar.
<BOS> The woman is frying vegetables.
<BOS> The man mixed the dough in a bowl.
<BOS> Two cats are playing while sitting on stools.
<BOS> A panda is lounging around.
```

Epoch30: Images

Step 7: Details of the score output:

It is higher the 0.6 (BL)

```
(base) pedramhaeri@Pedrams-MacBook-Pro P2 % python bleu_eval.py test_output.txt ]
Average bleu score is 0.6077167702578283
```

Step 8: Final Loss Figure