

Both models reduce loss from 9 to 2 in 1 epoch

THIS IS ORIGINAL-----

985

<start> a **red** double decker bus parked in a parking lot . <end> **actually yellow**

962

<start> a man holding a tennis racquet on a tennis court . <end>

285

<start> a bear is standing in the middle of **a field** . <end> **not a field probably**

836

<start> a person on a snowboard jumping over a snow covered slope . <end>

1089

<start> a man in a suit and tie sitting on a **couch** . <end> **desk**

1063

<start> a giraffe standing in a field of grass . <end>

133

<start> a wooden table topped with a variety of books . <end> **ignores bed nearby**

1205

<start> a living room with a **couch** , chair and a window . <end> **bed**

Bleu_1: 0.647

Bleu_2: 0.463

Bleu_3: 0.320

Bleu_4: 0.220

computing METEOR score...

METEOR: 0.206

computing Rouge score...

ROUGE_L: 0.483

computing CIDEr score...

CIDEr: 0.739

computing SPICE score...

SPICE: 0.142

THIS IS SELF IMPLEMENTED

Bleu_1: 0.645

Bleu_2: 0.460

Bleu_3: 0.314

Bleu_4: 0.209

computing METEOR score...

METEOR: 0.208

computing Rouge score...

ROUGE_L: 0.472

computing CIDEr score...

CIDEr: 0.743

computing SPICE score...

SPICE: 0.141

397

<start> a pizza sitting on top of a **white** plate . <end>**brown paper plate**

328

<start> a group of men **standing** next to each other . <end>**sitting**

1029

<start> a large airplane flying in the sky with clouds . <end>

-428

<start> a little girl is holding a cake with a **candle** in her hand . <end>**balloon**

3

online RNN(probably LSTM or more evolved LSTM versions) for stock prediction, especially high freq trading. Currently stocks are getting huger and huger datasets since the intro of high frequency trading by miliseconds. I interned at Credit Suisse last summer and their data lake was impressive and of all kinds:

stock, trades, books, news, client info, client chats... A combination of numerics and also text. I think LSTM would do well as stock prediction will heavily rely on past patterns and "memory" and do well with texts.

The data should be a combination of stock related numerics, news and social media, and client info and chats. We should encode the text data into embeddings, while adding vectors that indicate their lag, importance, relativity to the stock.

4

Attention distill information on image down to most salient objects, also enabling visualization of what model sees.

By introducing the context vector Z_t , enhanced model consider the right place to produce the text place. Also the next place network looks depends on the already generated words.

Author tested both soft and hard attention Rnn on Flickr8k and 30k and found that it has state of art performance, and also gives better explainability.

I didn't do any hyperparameter tuning on the model, so tuning on epochs, #hidden, #layers would have effects on the models

We can use transferred learning. Our model tends to misclassify the table beds and chairs as couches as the model was not seeing a lot of table and chairs like this in trainset. By using the transferred learning, we hope that our model can self-taught to classify such less frequent item in the training set correctly.