PROGRESS REPORT

Dynamic Lexicon Generation for Natural Scene Images

Team No - 28

Mentor : Sarthak Sharma

Analysing the Data:

we take the MS-COCO data set , it has captions with corresponding images_ids.

And we maintain the dictionary of {Image_id : Text Corpus of corresponding images } as key value pair

example:

Text Corpus Associated with this Image:

```
[ ‘A black Honda motorcycle parked in front of a garage.’,
'A Honda motorcycle parked in a grass driveway',
'A black Honda motorcycle with a dark burgundy seat.',
'Ma motorcycle parked on the gravel in front of a garage',
'A motorcycle with its brake extended standing outside']
```

Our Target is to apply LDA topic Model to learn the latent topics of images from text corpus.

LDA will learn : Prob(word/topic) and Prob(topic/document)

1.First we apply the tokenisation and iterate through each word.

2. Remove the stop words and Lemmatise the each word.

Result of above example after preprocessing :

```
['black', 'honda', 'motorcycle', 'park', 'garage']
['honda', 'motorcycle', 'park', 'grass', 'driveway']
['black', 'honda', 'motorcycle', 'dark', 'burgundy', 'seat']
['motorcycle', 'park', 'gravel', 'garage']
['motorcycle', 'brake', 'extend', 'stand', 'outside']
```

like this , we send each text corpus to preprocessing and to form a documents of relevant words. So that we can form the bag of words easily.

we pass our dictionary into doc2bow() vector to get bag of words.
To implement LDA we use the gensim module , we can in which we pass no of passes and no of topics as hyperparameters.

Screen shot of result after LDA: with Prob(word/topic) are in written in decreasing order.

```
Topic: 0
Words: 0.047*"bathroom" + 0.039*"white" + 0.035*"toilet" + 0.028*"window" + 0.027*"sit" +
rror" + 0.018*"black" + 0.016*"small" + 0.015*"elephants"
Topic: 1
Words: 0.092*"clock" + 0.070*"build" + 0.053*"fly" + 0.045*"kite" + 0.043*"large" + 0.041*
0.021*"kit" + 0.017*"tall" + 0.016*"people"
Topic: 2
Words: 0.065*"sit" + 0.042*"woman" + 0.037*"room" + 0.035*"phone" + 0.034*"hold" + 0.032*"
+ 0.028*"pizza" + 0.023*"desk" + 0.023*"live"
```

It shows that document 0 -> is mixture of topic0(11.99%), topic4(6.33%), topic7(41.98%) and topic9(37.37%)

```
In [92]: lda_model[bow_corpus[0]]
Out[92]: [(0, 0.11993622), (4, 0.06336602), (7, 0.41989934), (9, 0.37370506)]
```

CNN Model to predict probability distributions over LDA’s topics.

1.For this we have to generate a set of training set.

2. We obtain a set of M training examples of the form { (x1,y1), ... ,(xM,yM) } such that xi is an image and yi is the probability distribution over topics obtained by projecting its associated textual information into the LDA topic space.