COMP-SCI 5542 (SP17) - Big Data Analytics and Applications

Tutorial 11 Assignment (Due 04/12/17 by 11:59 PM)

***Dayu Wang* (45)**

1. TensorFlow Programming

1.1. Training Dataset

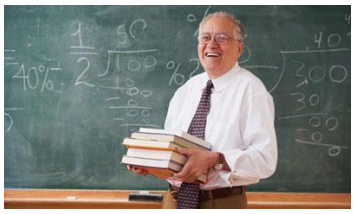
In this project, the **Flickr30k** dataset was chosen to serve as the training dataset, since the Flickr30k dataset has become a standard benchmark for sentence-based image description[1]. Since the dataset is too big (way over the file size limit of GitHub), the dataset itself was not uploaded to the Git repository.

1.2. Number of Captions for Each Image: **4**

```
def __init__(self,
             model,
             vocab,

             beam_size = 4,

             max_caption_length = 20,
             length_normalization_factor = 0.0):

  self.vocab = vocab
  self.model = model

  self.beam_size = beam_size
  self.max_caption_length = max_caption_length
  self.length_normalization_factor = length_normalization_factor
```

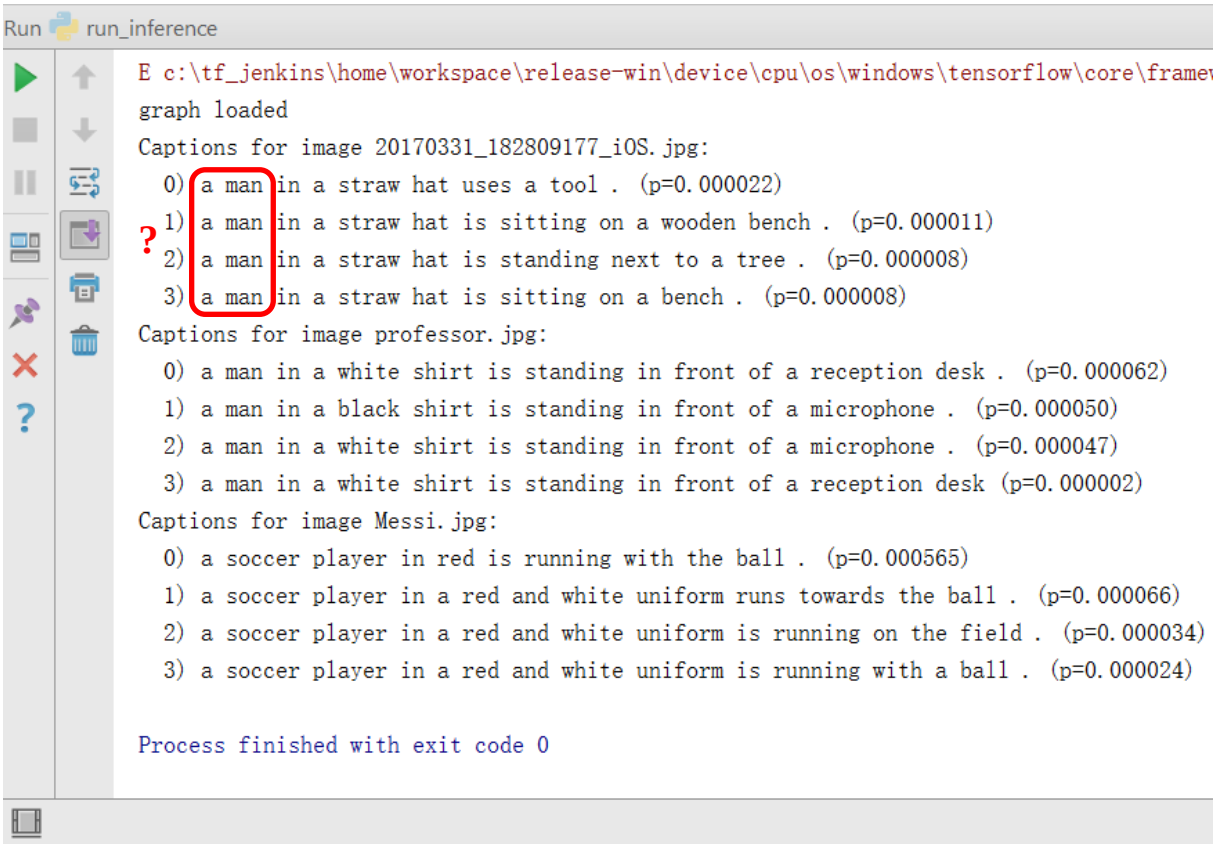
1.3. Screenshots of Results

20170331_182809177_iOS.jpg

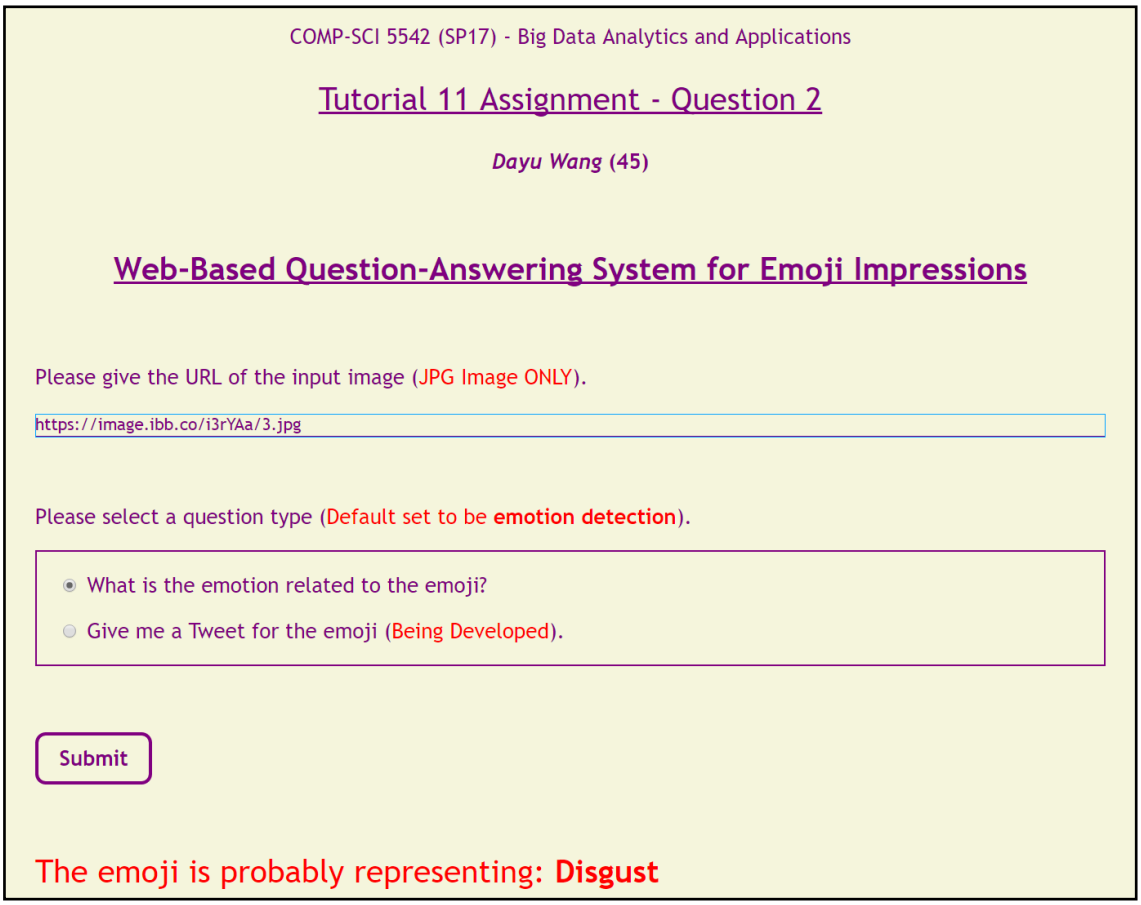


professor.jpg

Messi.jpg



2. Web-Based Question Answering (Screenshot)



3. Reference

[1] Plummer, B. A., Wang, L., Cervantes, C. M., Caicedo, J. C., Hockenmaier, J., & Lazebnik, S. (2015). Flickr30k Entities: Collecting Region-to-Phrase Correspondences for Richer Image-to-Sentence Models. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2641-2649). View Original Paper