Indraprastha Institute of Information Technology Delhi

CSE641 - Deep Learning
Assignment 04

Date:05-04-2021 **Total Marks: [80+70]**
Deadline: 17-04-2021 [11:59 PM]

Instructions:

1. The assignment can be submitted in a group of maximum of three members. Group details can be found here.
2. For Plagiarism, institute policies will be followed strictly.
3. Extension policy as discussed in class can be found here. Spamming +1s on Classroom will lead to usage of your group's one extension.
4. Make sure to use Pickle or any other library to save all your trained models. There will not be enough time during the demo to retrain your model. This is a strict requirement. You must upload your best models on Classroom to reproduce your results during the demo. If you are not able to reproduce your results during the demo then no marks will be given.
5. You need to submit README.pdf, Code files (it should include both .py files and .ipynb files), Output.pdf, and models dumped after training.
6. Mention methodology, helper functions, preprocessing steps, any assumptions you may have, and contribution of each member in README.pdf.
7. Mention your sample outputs in the output.pdf.
8. You are advised to prepare a well-documented code file.
9. Submit code, models, readme, and output files in ZIP format with the following name: A4_Member1_Member2_Member3.zip
10. Use classroom discussion for any doubt.

****

Q1. Implement the paper: Show, Attend and Tell: Neural Image Caption Generation with Visual Attention using **soft attention** on this dataset. The train, validation, and test samples are stored in each folder separately. The captions are stored in .pkl files within each folder. The .pkl file is a dictionary which takes image name with extension as the key and has value as a list of five captions associated with the current image. Use the pretrained VGG model [PyTorch, Tensorflow] as an image encoder. You can fine-tune the encoder, if you want.
Show attention weights for each word in 5 images.

Input: An image **Output:** Caption **Evaluation:** BLEU[1-4] and METEOR

Note: Depending upon the availability of computing resources, you are allowed to reduce the size of the train set. However, use the complete validation set and test sets for evaluations.

Q2. Implement the paper: Interactive Attention Networks for Aspect-Level Sentiment Classification using this dataset. You can find the details of the dataset here. Train two models for each domain (Laptop and Restaurant) - one with the interactive attention and another without any attention mechanism (concatenate the pooled vector directly and predict the sentiment). Use GloVe embeddings.
Show attention weights for 5 samples, as depicted in Figure 2 of the paper.

Input: <Context, Target>

1. <I liked the *service* and the *staff*, but not the *food,*
2. <I liked the *service* and the *staff*, but not the *food,*
3. <I liked the *service* and the *staff*, but not the *food,*

Output: [positive, negative, or neutral]
service> ⇒ *positive*
staff> ⇒ *positive*
food> ⇒ *negative*

Note: You can ignore all instances with conflict class.

Evaluation: Accuracy