

CEP Course on Deep Learning for NLP

Assignment on Multi-Task Learning for Emotion Detection and Intensity Prediction

Task Description: Existing emotion datasets are mainly annotated categorically without an indication of degree of emotion. In contrast, it is often useful for applications to know the degree to which an emotion is expressed in text.

Given a tweet and an emotion X, determine the intensity or degree of emotion X felt by the speaker -- a real-valued score between 0 and 1. The maximum possible score 1 stands for feeling the maximum amount of emotion X (or having a mental state maximally inclined towards feeling emotion X). The minimum possible score 0 stands for feeling the least amount of emotion X (or having a mental state maximally away from feeling emotion X). The tweet along with the emotion X will be referred to as an instance. Note that the absolute scores have no inherent meaning -- they are used only as a means to convey that the instances with higher scores correspond to a greater degree of emotion X than instances with lower scores.

Input: Documents (raw tweets)

Output: An Emotion Label and an Intensity Score

Objective:

- Identify the **best** suited emotion label for corresponding tweets
- Predict an intensity score for the corresponding emotion label

Approach: Solve the above multi-task problem through the following approaches and compare their performances.

- single Bi-GRU layer (128 units)
- single Bi-GRU layer (128 units) with word-attention (attention-dimension can be taken as 100 (max_sent_length))
- two Bi-GRU layers on top of one another (64 units each in each layer)
- two Bi-GRU layers on top of one another (64 units each in each layer) with word-attention (attention-dimension can be taken as 100 (max_sent_length))

Dataset: Perform the experiments on WASAA dataset (*emoint.csv*) comprising of 4-emotion classes (Anger, Disgust, Fear, Sadness) with 1500 instances in each class.

Structure of the data: 4 columns (review, emotion, labels, intensity).

In labels: 0 means Anger, 1 means Fear, 2 means Joy and 3 means Sadness

Use GloVe 100D pre-trained embeddings shared earlier with you all.

Experiment and evaluation: For each approach mentioned above, fit your model for 5 epochs and report the following after fitting is complete:

- a. Total Loss
- b. Individual Losses (per task)
- c. Accuracy
- d. Pearson Score
- e. Precision, Recall and F1-score
- f. Classification Report
- g. Confusion Matrix

Hint: Use `model.evaluate()` function to calculate the first four values.

Note: *Refer the Emotion Detection Implementation tutorial shared with you all.*