NLU project exercise lab: 11

Mateo Rodriguez (239076)

University of Trento

md.rodriguezromero@studenti.unitn.it

For the first part of the lab, there were two tasks at hand, subjectivity and polarity detection. For the first task BERT was used along side some fine-tuning in order to classify objective and subjective sentences, the second model was a simpler SVM that classified if a document is positive or negative. The idea behind the lab is to test the advantageous results from this paper [1] where objective sentence removal improves polarity classification; and indeed there was an increase in performance when objective sentences were removed in the tests performed. For the second part a model was designed to predict the positions of the aspects in a sentence, the model performance is capped by the aspect prediction performance, and the standard metrics for performance are likely misrepresenting the actual performance of the model.

1. Introduction

- For the first part of the lab, there are two main models, one that uses BERT for subjectivity classification, and for polarity an SVM was used
- For the second part the final model used is based on RoBERTa and predicts both aspects and their polarity

To get a robust estimate of a models' performance the testing was executed using stratified k-folds, with 10 folds. Training for subjectivity was done using the subjectivity dataset and for polarity the movies dataset, both available from NLTK. For the second part the model was trained and tested for 3 runs of 150 epochs each on the laptop dataset from semeval 2014 task 4.

2. Implementation details

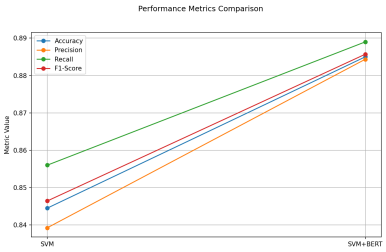
Starting with the BERT based model, it uses 2 extra linear layers that process BERT's CLS embeddings and thanks to sigmoid, outputs a value between 0 and 1, where 0 means subjective and 1 objective. Training was done with BCELoss and AdamW. For the polarity task, an SVM with a sigmoid kernel proved to be the best performing. The data was tokenized using TF-IDF and the labels were of course also converted into a binary representation. Now for the combined processing the main problem was the size of the sentences in the documents, because of this it was decided to remove the stopwords reducing the size significantly and not affecting performance.

For the second part of the lab, multiple models were designed and tested in search of the best way to tackle the aspect extraction problem. The original idea was to treat each aspect as a class label, making the task a classification task, but there were two problems with this approach, first the total number of classes was almost as big as the available sentences, meaning the data to train was not big enough and the second problem was that the model had to also predict the aspects in the test and validation so there were classes without train samples. Because of these problem I opted for another strategy, where the idea was to make the model predict only the polarity of the aspects but this had to be done in the correct position of the input string. This also didn't seem to be working as testing with various different models led to unsatisfactory results. Another approach was also experimented with, involving the model predicting a mask that indicates the positions of aspects within the input sentences. This alternative also did not yield successful results.

The final approach to the problem was to change the task for the aspect extraction from classification into regression, by making the model predict a tensor of shape [10,2] for the aspects and [10] for the polarity, where the task would be to predict the start and end position of the aspects in the sentences, and a maximum of 10 aspects per sentence can be predicted (the maximum number of aspects per sentence in all of the data was 9). The model designed for this uses Roberta as the base model, where the aspects are processed form the last hidden state using linear layers and a convolutional layer that halves the tensor size from [batch,20,4] to the desired output of size [batch,10,2]. For the polarity prediction the model uses a single linear layer processing the hidden state of the CLS token. For this configuration cross-entropy was used for the polarity and MSE for the aspect prediction, and the optimizer was AdamW.

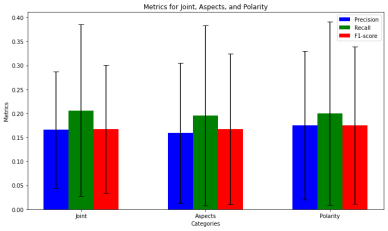
3. Results

The figure below shows the results of the first part of the lab



The model for subjectivity prediction achieved 0.9917 accuracy and 0.9883 F1-score

For the second part of the lab, results were the following:



The results from the first figure confirm what was reported in [1], and regarding the second part of the lab, the model evaluation is the most problematic point since the predicted position of an aspect can be partially correct by selecting only part of

the ground truth or more words surrounding the correct aspect, leading to equal penalization as a completely wrong position. So it's likely the performance of the model is better than what's portrayed by the metrics alone.

4. References

[1] B. Pang and L. Lee, "A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts." Ithaca, NY: Cornell University. [Online]. Available: https://aclanthology.org/P04-1035.pdf