Final Project of NLU

Michele Presutto 229208
university of Trento
michele.presutto@studenti.unitn.it

1. Introduction

Sentiment Analysis (SA), is a field of Natural Language Processing that uses computational techniques to identify and extract subjective information from source materials. This includes determining the overall sentiment of a piece of text (e.g. positive, negative, neutral) as well as identifying specific subjective phrases or sentences.

Sentiment analysis is useful in a variety of real-world applications, such as customer service, market research, and social media monitoring. It can be used to understand the public opinion about a particular topic or brand, and can help businesses make informed decisions. This is possible due to the amount of text data, that has been collected in the years thanks to internet and the social applications, used to train and test the different models and algorithm for Sentiment Analysis.

Sentiment analysis is not a trivial task and it's composed by different steps, starting from analysing the raw text, preprocess the text and in the end train a classifier and testing how well it does prediction. The purpose of this project is to build a sentiment analysis extractor by using a dataset of movie review and, as pre-processing steps, removing the objective part of each review. The motivation behind this approach is given by the general fact that objective sentences, which are those that convey factual information rather than opinions, do not contribute to sentiment evaluation and can even be detrimental to the performance of the model.

2. Task Formalisation

Sentiment Analysis can be defined as a supervised learning task and so each sentence in the dataset is coupled with a label which will tell us the polarity of the relative sentence. The focus of this project is to test how the polarity classifier will behave when the data fed into it is processed by removing the objective sentences and compare the performances with the non-processed one.
we trained 3 different classifier:

- **Polarity baseline Classifier**: is the polarity classifier trained on the raw data, without the pre-processing steps involving the object removal part.
- **Subjective Classifier**: is the model trained on the subjective dataset. this model is used to remove the objective phrases in the pre-processing steps.
- **Polarity Classifier**: is the classifier trained on the data pre-processed using the subjectivity classifier.

doing so we are testing if the performances of the model improves by removing the objective part.
The process of testing began with training the Polarity Baseline Classifier without selecting subjective sentences. This approach was chosen in order to have a clear understanding of the impact of considering only the subjective sentences in the classification process. It's worth mentioning that the classification being used is binary, which means that the prediction made for a single review can only be positive or negative and does not take into consideration the neutrality of a sentence.

The concept of using another classifier that is trained only on subjectivity to remove the objective context from the reviews is an interesting one, as it highlights the importance of the subjective sentences in determining the difference between positive and negative. For example, if we were to take a movie review, we would see that many sentences in the review would refer to objective facts such as who the actor or director is, or the plot of the movie. However, other sentences would refer to the viewer's personal experience and therefore would contain subjective evaluations based on their feelings and opinions.

It is due to the fact that objective sentences can act as noise in the data and so the removal of these sentences can lead to an improvement in the accuracy of the classification. This is because subjective sentences contain valuable information that can accurately distinguish between positive and negative sentiment, while objective sentences may not provide the same level of insight. By considering only the subjective sentences, the classifier can be trained to focus on the most relevant information and make more accurate predictions.

3. Data description and Analysis

For this project, two separate datasets were utilized for training and evaluating models. These datasets were obtained from the Natural Language Toolkit (NLTK) [1] library and both were introduced by Pang and Lee in June 2004 [2]

3.1. Subjectivity Dataset

The subjectivity dataset used in this project consists of 10,000 sentences. Half of these sentences (5,000) are objective and the other half (5,000) are subjective. The dataset has a total of approximately 240,000 words and a vocabulary size of 23,906. The minimum sentence length is 10 words, the maximum is 120, and the average is 24 words. The most frequent word in the dataset, excluding punctuation, is "the".

The subjectivity dataset is represented as a list of sentences, each of which is a list of strings. The target labels were not included in the original dataset, so they were added manually. The dataset is split into a training set and a test set and is used to train and evaluate the binary classifier for subjectivity.

3.2. Polarity Dataset

The Movie Reviews dataset is used for polarity classification. It contains 2,000 reviews, half of which are positive and half negative. The dataset has a larger number of words than the previous dataset, with a total of 1,583,820 words and a vocabulary size of 39,768. Like the previous dataset, the most common word without pre-processing is ”the.”

The Movie Reviews dataset is structured differently than the previous dataset, consisting of a list of reviews, each of which is a list of sentences (which are themselves lists of strings). The maximum length of a review is 115 sentences, the minimum is 1 sentence, and the average is 24 sentences per review.

The ground truth targets for each review are not included in the dataset, but instead each review is paired with a binary label of ”positive” or ”negative.” The dataset is split into a train and test set, which is used to train the polarity classifier to predict whether a movie review is positive or negative.

statistics for movie review dataset		
statistics	negative	positive
number of review	1000	1000
number of words	751256	832564

Table 1: *table showing difference between the positive review and negative ones, we can clearly see that the positive review have in general more words.*

4. Models

In this section i describe the models used as baseline and the final model used to improve performances in comparison with the baseline.

4.1. Naive Bayes Classifier

The Naive Bayes classifier is used as the first model, following the approach used by Bo Pang and Lillian Lee, to establish a baseline for comparison with subsequent results.

4.2. Transformers

For the second model, I utilized a straightforward fully connected layer to categorize the embeddings obtained from a pre-trained model. Specifically, the pre-trained model used is the all-mpnet-base-v2, that is a pre-trained language model developed by the Hugging Face Transformers library [3]. It is based on the Multilingual Pretrained Model Network (MPNet) [4] architecture, which is a variant of the Transformer architecture. The base version of this model has been fine-tuned on a large multilingual corpus and has been trained to perform a variety of natural language processing tasks, including text classification, sequence labeling, and question answering.
The model is designed to handle a wide range of text inputs and to be highly transferable, for this reason i decided to used this model without finetunig.
I used this model to extract embedding both for the subjectivity dataset and for the polarity one. The embedding dimension obtained after passing the sentence to the model is of length 768.

5. Testing

The initial step taken in the testing process was to calculate the accuracy of the Baseline (Naive-Bayes) classifier, taking into account the removal of objective sentences and without removing them.
To compare the models performances it’s been used the accuracy metrics formally described here:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (1)$$

The results are displayed in Table 2 (Naive Bayes classifier).
In the second step, a fully-connected network was trained to perform the classification of the embeddings. This network was trained multiple times to evaluate its performance under various types of preprocessing for the dataset or network parameters. These tests were carried out using GPUs provided by Google Colaboratory and trained for 70 epochs each, with a division of the dataset into training and testing portions of 77%/33%, respectively.

The first test using this network was performed without any pre-processing of the dataset, resulting in the lowest accuracy. The second test involved removing punctuation to see how much the accuracy could improve, with a resulting increase in accuracy of 6.17%. The third test involved first training a classifier for subjectivity(Table 3) to allow for the removal of objective sentences from the dataset, followed by training the polarity model, resulting in an increase in accuracy of 2.98%.

The final test consisted of first filtering each review to remove objective sentences, followed by the removal of punctuation. This last test did bring small improvement to the accuracy, resulting with an accuracy around 86%.

Polarity	
models	accuracy
Naive Bayes	81.4%
Naive Bayes+Objective Revomal(OR)	84.4%
all-mpnet	78.5%
all-mpnet+punct removal	83%
all-mpnet+OR	85.5%
all-mpnet+objectivity+punct removal	86%

Table 2: *table showing the accuracy of the different models the best performance is obtained by using all-mpnet for the encoding and a FC neural network for classification and applying Objective Removal on each sentence of the data*

Subjectivity	
models	accuracy
Naive Bayes	91.1%
all-mpnet	94%

Table 3: *table showing the accuracy of the models used for classify the subjectivity used in the preprocessing of the text. The best result are obtained by using all-mpnet*

6. Conclusions

In conclusion, the focus of this paper was to investigate the impact of removing objective sentences from a dataset and evaluate the performance of sentiment analysis. Three classifiers were trained and tested on two different datasets. The first was the Polarity Baseline Classifier, trained on the raw data without any pre-processing. The second was the Subjective Classifier, trained on the subjective dataset to remove objective phrases in the pre-processing steps. The third was the Polarity Classifier, trained on the pre-processed data using the Subjective Classifier. The results showed that removing objective sentences from the data improved the accuracy of the sentiment analysis. This highlights the importance of subjective sentences in determining the sentiment of a review and the fact that objective sentences can act as noise in the data. The use of a pre-trained language model in combination with a fully connected layer showed promising results, outperforming the Naive Bayes classifier. The findings of this paper suggest that removing objective sentences from the data and focusing on subjective sentences can lead to improved results in sentiment analysis.

7. References

[1] E. Loper and S. Bird, "Nltk: The natural language toolkit," *arXiv preprint cs/0205028*, 2002.

[2] M. Hu, Y. Peng, Z. Huang, D. Li, and Y. Lv, "Open-domain targeted sentiment analysis via span-based extraction and classification," *arXiv preprint arXiv:1906.03820*, 2019.

[3] "sentence-transformers library huggingface," https://huggingface.co/sentence-transformers.

[4] K. Song, X. Tan, T. Qin, J. Lu, and T.-Y. Liu, "Mpnet: Masked and permuted pre-training for language understanding," *Advances in Neural Information Processing Systems*, vol. 33, pp. 16 857–16 867, 2020.