

Explicit Sentiment Analysis with Language Patterns about Uncertainty

Our goal with this project is divided into two parts. First of all we extract a dataset about semantic uncertainty from the web archive data [2]. This dataset is then compared to Sentiment140 [1] to clarify how well our dataset conforms with this popular sentiment classification dataset. In the second part a sentiment classifier based on DistilBERT [3] is trained on our dataset using transfer learning. This model is benchmarked on the previous mentioned dataset to see if our dataset is suited to train sentiment classification.

We will use specific language patterns about uncertainty to extract samples from web archive data. An overview of the patterns can be found here [4, p. 43]. To classify these samples into positive/negative sentiments we will mainly use GPT-3. Since we cannot fully trust results based on GPT-3 we will verify some of the labels manually. We plan to analyse our dataset based on what topics the internet is most uncertain about and how those topics changed over time using circular packing charts. After the labeling process we want to compare strong positive/negative topics of our data with those of the dataset mentioned above. So the deliverables from this step are the dataset and visualizations of the comparison.

In the next step we train our model. We prevent train-test leakage with the known train-validation-test split method. The ratios of this split depends on the amount of samples we actually get in the end. We then compare the performance of our model to other models (Quelle mit Ergebnissen von anderen modellen zu diesem datensatz) based on the dataset from above. With these evaluation results we can determine if our data can be used to train sentiment classification.

This leads us to the following two research questions:

- Can a dataset consisting only of uncertain statements be used to extract the sentiment of other statements?
- What are the topics the internet is most uncertain about and have those topics changed over time?

The project will be split in the following three work packages:

- Dataset extraction from the web archive data
- Labeling, analyzing, visualising and the dataset and comparing it to Sentiment140
- Train a model and evaluate it on said dataset

Deliverables of the project are:

- A Dataset consisting of only uncertain statements
- Visualizations and Analysis of said dataset
- A DistilBERT model trained on said dataset

-
- [1] GO, Alec ; BHAYANI, Richa ; HUANG, Lei: Twitter sentiment classification using distant supervision. In: *Processing* 150 (2009), 01
- [2] KIESEL, Johannes ; KNEIST, Florian ; ALSHOMARY, Milad ; STEIN, Benno ; HAGEN, Matthias ; POTTHAST, Martin: Reproducible Web Corpora. In: *Journal of Data and Information Quality* 10 (2018), dec, Nr. 4, S. 1–25

- [3] SANH, Victor ; DEBUT, Lysandre ; CHAUMOND, Julien ; WOLF, Thomas: Distil-BERT, a distilled version of BERT: smaller, faster, cheaper and lighter. In: *CoRR* abs/1910.01108 (2019). – URL <http://arxiv.org/abs/1910.01108>
- [4] VINCZE, Veronika: Uncertainty detection in natural language texts. In: *PhD, University of Szeged* 141 (2014)

- x What is your research goal/the research question you are trying to answer?
 - x Is there any previous/related work?
 - x What are the necessary steps (research plan)? How will those be distributed among team members?
 - x What data will you use? What preprocessing steps are necessary? Do you need to annotate the data? How will this be done? Manually? With the assistance of e.g. GPT-3?
 - x What will be the train-validation-test split? How do you prevent train-test leakage?
 - x Are there baseline datasets or baseline methods for your approach that you will compare your results to?
 - x Will you train a new model? Will you use an existing model architecture/a pre-trained model? Finetuning, head retraining?
 - What experiments will you conduct? What will be the evaluation methods/quality criteria/metrics?
 - x How will you be able to answer the research question based on those experiments?
-
- x The cleaned up/refined dataset (if you release data, provide a datasheet <https://arxiv.org/abs/1803.03324>)
 - (x) The trained model (if you release a model, provide a model card <https://arxiv.org/abs/1810.03993>)
 - (x) For shared tasks: A task description and evaluation scripts
 - x Evaluation results and plots
 - x Visualizations and/or interactive demonstrations (if applicable - specify what and in which suitable format)