

# 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 and finetuned on SST-2 [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 improve 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.

In the next step we train our model. We prevent train-test leakage with the known train-validation-test split method. Furthermore we exclude all Twitter domain names from the web archive because Sentiment140 is based on statements from Twitter. The ratios of this split depends on the amount of samples we actually get in the end. We then compare our model and the baseline model with regards to the performance on Sentiment140 dataset.

This leads us to the following two research questions:

- Does finetuning on uncertain statements improves sentiment classification?
- 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, visualizing 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 (inclusive datasheet)
- Visualizations and Analysis of said dataset
- A DistilBERT model trained on said dataset (inclusive model card)

- 
- [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)