Seminar Report

Topic-Specific Sentiment Analysis for Tweets by German MPs

Department of Statistics
Ludwig-Maximilians-Universität München

By Asmik Nalmpatian and Lisa Wimmer
In Cooperation with Prof. Dr. Paul Thurner (Department of Political Science)
Under the supervision of Matthias Aßenmacher, Ph.D., and Prof. Dr. Christian Heumann (Department of Statistics)
Munich, month day$^{\text{th}}$, 2021

Abstract

Project outline

- **Problem description.** For Twitter data of German MPs, enable automated textual analysis by topic-specific sentiment analysis
- **Key assumptions.**
  - Sentiment analysis is much stronger if topical context is taken into account – while some words, such as "excellent", carry a globally applicable sentiment, others may convey quite different meanings in varying context (see, for example, Thelwall and Buckley (2013)).
  - We will assume that one tweet refers to one (principal) topic and carries a distinct sentiment towards this topic. This is induced by the observation that politicians will mostly tweet ad-hoc in response to a certain event, not as an overall statement of their stance toward some issue. Literature also suggests this is a reasonable assumption (Qiang et al., 2019).
- **Relation to other problems of NLP.** We refer to the problem treated herein as topic-specific sentiment analysis (TSSA). TSSA overlaps with other well-known NLP problems:
  - **Sentiment analysis (SA).** TSSA may be interpreted as a methodological variant, taking into account topical context that is lost on standard SA.
  - **Topic modeling (TM).** TM retrieves topics without sentiment classification and is thus subsumed under TSSA.
  - **Aspect-based sentiment analysis (ABSA).** TSSA is closely related to ABSA. In particular, what we call topics, exhibiting several aspects, could be identified with entities from ABSA. However, we observe that problems and solutions framed in ABSA often take a slightly different focus, and sometimes employ approaches that deviate from the idea we pursue here. Therefore, we deliberately introduce this distinction to emphasize the difference in intuitions we perceive as follows:
    - ABSA is frequently discussed in the context of product reviews and thus mostly concerned with a specific type of document: the topic treated is quite clear (reviews are typically posted on a product detail page or similar dedicated location), and the document always carries sentiment (sometimes, even with an explicit rating attached).
    - Whether or not such a prior sentiment information is available, ABSA is interested in the particular aspects that constitute overall sentiment (which all convey polarity, possibly of conflicting nature). These aspects are typically assembled in higher-level entities.
    - A major task in ABSA is therefore the decomposition of the documents into single-aspect fragments, which may be interpreted as a reshaping of the corpus with a modified definition of the smallest-element level.

- Once this is achieved, SA is performed on each single-aspect fragment.
- Our setting, on the other hand, deals with a different kind of document. As tweets are publicly posted without explicit connection to a specific target, we will usually not know what a document is about, much less its polarity – many documents will not even convey sentiment.
- Also, as mentioned above, tweets can be expected to have a narrow focus on a single topical issue.
- We identify three potential sub-tasks of TSSA that relate to ABSA in varying degree:
  - **Stand-alone TM.** We might be only interested in extracting topics from a set of tweets, which is equivalent to the ABSA-subtask of aspect extraction (i.e., finding entity-aspect pairs), albeit on document level only. If the goal is to label tweets according to pre-specified topics, we are more in the realm of multi-label classification.
  - **Stand-alone SA.** We might also be purely interested in SA without relation to single aspects (e.g., because the data were collected in such a fashion that they are related to a very specific topic from the beginning, or because we wish to analyze general polarity of tweets by a certain source). Here, we emphasize in particular the topic orientation of TSSA: we still want to leverage latent (sub-)topics because we believe it will aid classification by providing contextual information. Rather than being an explicit target of sentiments, topics and aspects are indicators of contextual clusters then.
  - **TSSA.** This is indeed a form of ABSA with the specific assumption that one document carries exactly one aspect and one associated sentiment (topic is then synonym to entity). We underline that topics and aspects do not serve as mere document separators, but are viewed as indicative for different contexts that need to be accounted for. By this, we differentiate with respect to approaches that only reshape corpora according to aspects and then perform topic-agnostic SA[1].
- Summing up the above, our approach incorporates all general ideas present in ABSA, but does so with a slightly different focus, and special emphasis on topical context.

- **Solutions.** We offer two different approaches to tackle this problem:

  1. Methods of standard machine learning, more precisely, a combination of lexicon-based techniques and off-the-shelf classifiers
  → Accessible to researchers without strong data science background and/or large computational resources

[1] E.g., Sivakumar and Reddy (2017), Zhao et al. (2016), Zhao et al. (2017) first assign aspects, but then apply standard classifiers / global dictionaries.

2. Methods of deep learning, more precisely, a variant of BERT
   → Powerful tool for more advanced researchers

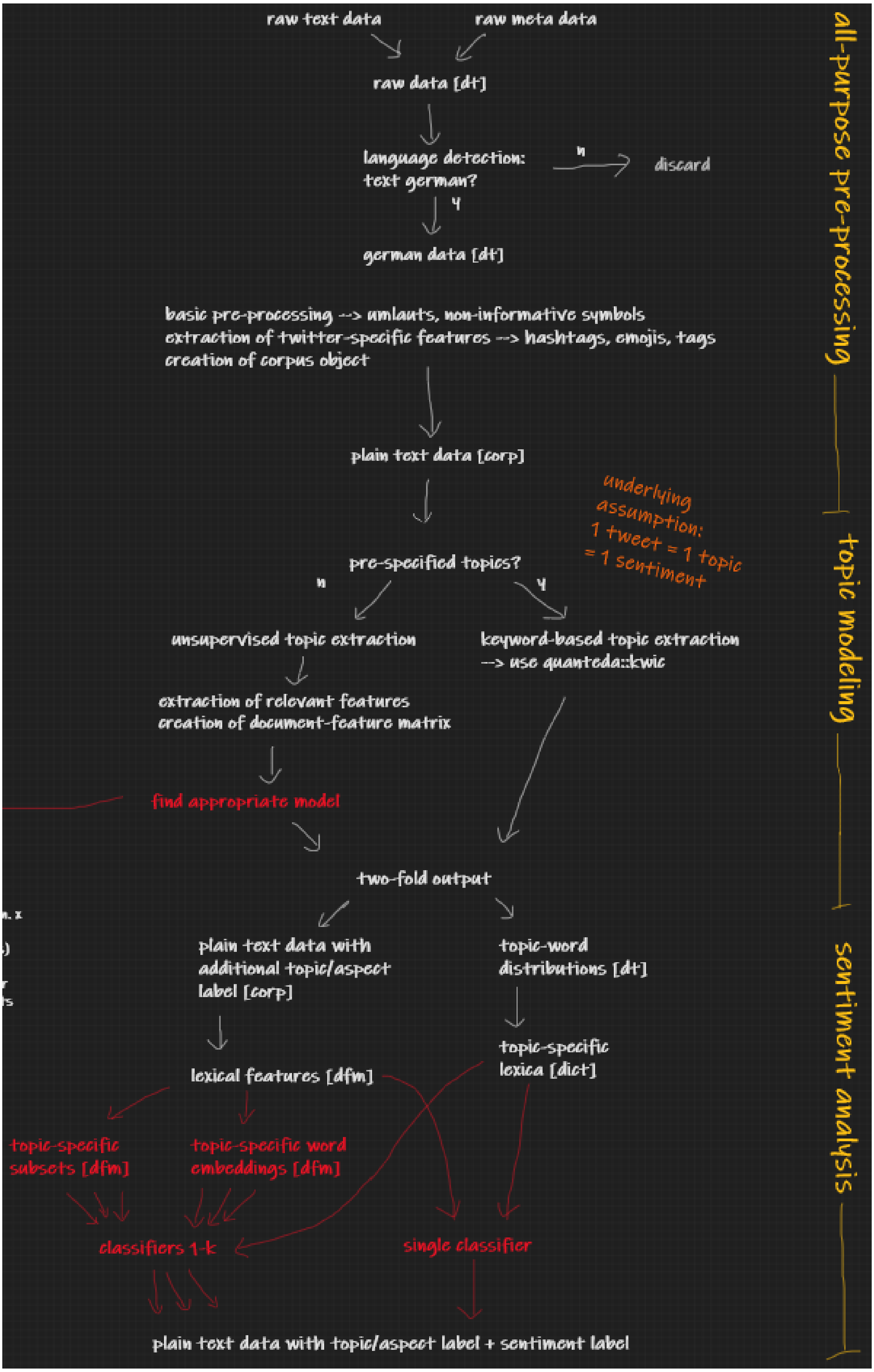
Part I: proposed pipeline

- Deep learning methods, BERT in particular, are capable of actually learning the intrinsic structures constituting language (grammatical patterns, inter-term relations), and can draw from this knowledge to perform downstream tasks such as sentiment analysis.
- By contrast, lexicon-based and machine learning approaches are confined to learning from features derived from the bag-of-words assumption[2]. This makes the handling of phenomena like negations much harder.
- Nonetheless, both are widely applied in SA tasks.
- Lexicon-based approaches identify sentiment-bearing words and assign them polarity scores, an aggregation of which on document level then yields document polarity. In this, they are able to work fully unsupervised.
- ML approaches make use of a variety of features (e.g., unigrams), and learn their association to the target from labeled training instances. Typically, a combination of three types of features is used:
  - Document-level lexical features (unigrams, number of exclamation marks, ...)
  - Lexicon-based features (number of positive-polarity words, ...)
  - Word embeddings (learned from large, possibly corpus-external feature-co-occurrence matrices)
- For the incorporation of topical context, we choose to take a sequential approach, where topics are assigned first, and then SA is performed.
- While there exist successful implementations of joint topic-sentiment modeling (and theoretical arguments are strong, particularly so if inference is relevant), we decide against a simultaneous analysis for the following reasons:
  - This first part addresses researchers from other disciplines, so methods should be fairly easy to implement and explain. Joint models, however, are usually rather complicated, hierarchical Bayesian architectures.
  - As outlined above, we would like to be able to tackle sub-tasks individually.
  - A likely scenario entails TSSA for topics that are pre-specified based on a specific research interest (given in form of keywords, for instance). It is unclear how such a setting would be handled with joint models.
- However, it is not clear whether we should use a single classifier after topic labeling or one classifier per topic.

[2] Well, you could probably hand-craft grammatical features, but this would become incredibly complicated.

- Single classifier
  - Pro: lean architecture; topic-agnostic features need to be learned only once; single-best classifier is easily found and tuned
  - Con: topic-specific features, such as word embeddings, are hard to incorporate
  - Examples: single global lexicon with polarity as topic-prevalence-weighted average across sentiment-bearing words (Naskar et al., 2016); aspect as single categorical variable in otherwise global feature set (baseline of Sem-Eval 2016 ABSA task, Pontiki et al. (2016))
- Per-topic classifier
  - Pro: easy handling of topic-specific features
  - Con: many models in parallel; topic-agnostic features are learned from smaller subsets; unclear which classifier is best (single best, best for each sub-task?)
  - Examples: domain-specific lexica and purely lexicon-based (Thelwall and Buckley, 2013); domain-specific lexica and per-domain ML classifiers (Choi et al., 2009); standard features and per-topic ML classifiers (Ficamos and Liu, 2016); per-topic word embeddings and LSTM (Jebbara and Cimiano, 2016)

- Proposition: use single classifier with lexical features and embeddings derived from topic-agnostic features, plus polarity counts from topic-specific lexica (topic-agnostic meaning those that are selected into the topic lexica, or at least those that are selected with different polarities)


raw text data
raw meta data
raw data [dt]
language detection:
text german?
n
discard
y
german data [dt]
basic pre-processing --> umlauts, non-informative symbols
extraction of twitter-specific features --> hashtags, emojis, tags
creation of corpus object
plain text data [corp]
underlying
assumption:
1 tweet = 1 topic
= 1 sentiment
pre-specified topics?
n
y
unsupervised topic extraction
keyword-based topic extraction
--> use quanteda::kwic
extraction of relevant features
creation of document-feature matrix
find appropriate model
two-fold output
plain text data with
additional topic/aspect
label [corp]
topic-word
distributions [dt]
topic-specific
lexica [dict]
lexical features [dfm]
topic-specific
subsets [dfm]
topic-specific word
embeddings [dfm]
classifiers 1-k
single classifier
plain text data with topic/aspect label + sentiment label
all-purpose pre-processing
topic modeling
sentiment analysis

Contents

List of Figures

List of Tables

1 Introduction

foo

2 Project Outline

2.1 Problem Description

2.2 Proposed Solution

3 Theoretical Background

foo

4 Data

4.1 Problem Description

4.2 Proposed Solution

5 Part I: Topic-Specific Sentiment Analysis in `R`

foo

6 Part II: Topic-Specific Sentiment Analysis with `BERT`

foo

7 Discussion

foo

8 Conclusion

foo

A Appendix

B Electronic Appendix

Data, code and figures are provided in electronic form.

References

Choi, Y., Kim, Y. and Myaeng, S.-H. (2009). Domain-specific sentiment analysis using contextual feature generation, *Proceedings of the 1st International CIKM Workshop on Topic-Sentiment Analysis for Mass Opinion*.

Ficamos, P. and Liu, Y. (2016). A topic based approach for sentiment analysis on twitter data, *International Journal of Advanced Computer Science and Applications* **7**.

Jebbara, S. and Cimiano, P. (2016). Aspect-based sentiment analysis using a two-step neural network architecture, *Proceedings of the Third SemWebEval Challenge*, pp. 153–167.

Naskar, D., Mokaddem, S., Rebollo, M. and Onaindia, E. (2016). Sentiment analysis in social networks through topic modeling, *Proceedings of the Tenth Language Resources and Evaluation Conference*.

Pontiki, M., Galanis, D. and Papageorgiou, H. (2016). Semeval-2016 task 5: Aspect based sentiment analysis, *Proceedings of the Tenth International Workshop on Semantic Evaluation*.

Qiang, J., Qian, Z., Li, Y., Yuan, Y. and Wu, X. (2019). Short text topic modeling techniques, applications, and performance: A survey, *Journal of Latex Class Files* **14**(8).

Sivakumar, M. and Reddy, U. S. (2017). Aspect based sentiment analysis of students opinion using machine learning techniques, *Proceedings of the International Conference on Inventive Computing and Informatics (ICICI)*.

Thelwall, M. and Buckley, K. (2013). Topic-based sentiment analysis for the social web: the role of mood and issue-related words, *Journal of the American Society for Information Science and Technology* **64**(8): 1608–1617.

Zhao, Y., Qin, B., Liu, T. and Tang, D. (2016). Social sentiment sensor: a visualization system for topic detection and topic sentiment analysis on microblog, *Multimedia Tools and Applications* **75**: 8843–8860.

Zhao, Y., Qin, B., Liu, T. and Tang, D. (2017). Social sentiment sensor: a visualization system for topic detection and topic sentiment analysis on microblog, *Sustainability* **9**(10).

Declaration of Authorship

We hereby declare that the report submitted is our own unaided work. All direct or indirect sources used are acknowledged as references. We are aware that the Thesis in digital form can be examined for the use of unauthorized aid and in order to determine whether the report as a whole or parts incorporated in it may be deemed as plagiarism. For the comparison of our work with existing sources we agree that it shall be entered in a database where it shall also remain after examination, to enable comparison with future Theses submitted. Further rights of reproduction and usage, however, are not granted here. This paper was not previously presented to another examination board and has not been published.