Statistical Consulting

Topic-Specific Sentiment Analysis for Tweets by German MPs

Department of Statistics
Ludwig-Maximilians-Universität München

Munich, month day^th, 2021

Authors Asmik Nalmpatian
Lisa Wimmer

Project Partner Prof. Dr. Paul Thurner
Department of Political Science

Supervisors Matthias Aßenmacher, Ph.D.
Prof. Dr. Christian Heumann
Department of Statistics

Abstract

Contents

List of Figures

List of Tables

1 Introduction

The advance of social media has sparked a fundamental change to public debate. Political coverage and discourse has spilled from traditional outlets over to online locations whose accessibility has lowered entrance barriers to welcome a wide audience (Bode, 2017). Social media exhibit certain properties that make them attractive to politicians seeking to broadcast their message. From a supply-side perspective it is easy to publish content: posting on social media is cheap, does not require approval of any authority, and allows for full control over the intended presentation. In an environment that spins information at enormous speed the resulting ability to react to events in real-time offers a distinct advantage over the inertia of traditional channels (Stier et al., 2018). On the receiving end, social media grant unprecedented access to target audiences. The industry's oligopolistic structure sees people from heterogeneous backgrounds convene on few global platforms to exchange their views. While not all of them show political interest, research suggests that the low cost of online engagement prompts a large number of users to participate to some extent (Jost et al., 2018). A contrary but equally important aspect is the evolution of echo chambers. Users view content according to their perceived preferences and thus often end up in community niches populated by like-minded people. This clustering process creates groups disproportionately receptive to certain messages and has played an important role in large-scale propaganda for events like US presidential elections. Echo chambers are particularly suited for the direct communication among users social media offer: chat and commenting functions enable a dialogue between politicians and their electorate that is hard to achieve via traditional channels, and allow to deploy the power of emotion in shaping opinion (Hasell and Weeks, 2016). These opportunities have propelled internet platforms to a level at least equal to the former hegemons of political debate. Twitter, in particular, has emerged as a medium for political information. In what might have been inconceivable a few years ago politicians actively convey messages to the public via tweets (van Vliet et al., 2020). The impact of this disruption to the political environment is complex and most certainly has positive as well as worrisome aspects. Yet, from a purely scientific point of view, activity on social media creates on its way a vast amount of publicly accessible data that benefits the research community with a constant source of analytical potential.

A question frequently posed in political analysis is the assessment of public opinion toward a particular matter. However, the textual data, gathered from social media, that might hold the answer command the use of specific tools subsumed under the field of *natural language processing (NLP)*. NLP has gained much traction with the rise of deep learning methods and become virtually ubiquitous, techniques ranging from simple approaches to gigantic neural networks powering search engines and the like (see, for example, Torfi et al. (2020)). Statistically speaking, the above problem translates into the classification of texts into instances of certain sentiments (typically, *positive* and *negative*). Building upon the assumption that sentiment may be expressed differently in varying contexts, such *sentiment analysis* is often combined with some form of *topic modeling*.

It is the goal of this project to make analysis of social media texts in a political context more easily accessible to researchers. We focus on the analysis of public sentiment in a topic-aware manner for texts collected from Twitter posts by German MPs.
Our contribution is two-fold:

1. We explore how topic-specific sentiment analysis can be implemented, considering (1) standard machine learning techniques and (2) more complex deep learning models.
2. We provide extensive teaching material on both approaches, composed as a coherent online course, to educate researchers on addressing NLP problems in their own work.

The remainder of this report is organized as follows. First, we state the problem and proposed solution of this project in more detail in section 2. Section 3 provides some theoretical context for topic modeling and sentiment analysis from which we derive what we call *topic-specific sentiment analysis (TSSA)*. We proceed in section 4 by sketching our data collection and cleaning process and then present our proposal for conducting TSSA, laying out for both approaches the underlying methodology and discussing the results of applying them to the data at hand. Section 5 outlines how the findings from this analysis translate to the proposed teaching material. Afterwards, in section 6, we critically assess the findings and limitations of the project, and conclude with a brief summary in section 7.

2 Project Outline

2.1 Problem Description

Access to data on social media platforms provides the social sciences with a source of cheap information readily available for empirical analysis. Exploiting these data requires certain skills that may exceed basic statistical knowledge. The first obstacle to overcome is data collection, i.e., retrieving online information in an automated and resource-efficient manner that results in a suitable data structure. Analysis of the thus acquired data is complicated by the special nature of textual data. Texts are not arbitrary sequences of interchangeable characters but complex constructs that carry meaning on multiple linguistic levels. Analyzing text entails handling the sheer size of languages' vocabularies, grammatical rules and irregularities, individual preferences in expression, and delicate phenomena like colloquialisms or irony, to name only a few (section 4.1.2 will address the challenges arising from Twitter data in more detail).
Sentiment analysis, a task frequently encountered in social media analysis, is therefore not straightforward despite posing a standard classification problem.
A task frequently encountered in social media analysis is the assessment of public sentiment. Although sentiment analysis poses a standard classification problem, the textual form of data input makes modeling the relation between features and targets far from straightforward.

2.2 Project Goal

- **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-sub-task 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.

◦ 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.

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

- 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)

3 General Theoretical Context

3.1 Topic Modeling

General context: topic modeling, sentiment analysis, why both combined
TSSA as classification task that can be solved with ml oder dl

- 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).

3.2 Sentiment Analysis

3.3 Topic-Specific Sentiment Analysis (TSSA)

- 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.
  - 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)

4 Analytical Proposal

4.1 Data

4.1.1 Data Collection

4.1.2 Data Pre-Processing

4.2 Standard Machine Learning Solution

4.2.1 Methodology

Automated machine learning pipeline

foo

Methodological concepts

foo

4.2.2 Results

4.3 Deep Learning Solution

4.3.1 Methodology

Deep transfer learning

foo

BERT

foo

4.3.2 Results

5 Knowledge Transfer

5.1 Static Material

5.2 Live Teaching

6 Discussion

6.1 Analytical Proposal

6.2 Knowledge Transfer

7 Conclusion

foo

A Appendix

B Electronic Appendix

Data, code and figures are provided in electronic form.

References

Bode, L. (2017). Gateway political behaviors: The frequency and consequences of low-cost political engagement on social media, *Social Media + Society* **3**.

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**.

Hasell, A. and Weeks, B. E. (2016). Partisan provocation: The role of partisan news use and emotional responses in political information sharing in social media, *Human Communication Research* **42**: 641–661.

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.

Jost, J. T., Barbera, P., Bonneau, R., Langer, M., Metzger, M., Nagler, J., Sterling, J. and Tucker, J. A. (2018). How social media facilitates political protest: Information, motivation, and social networks, *Advances in Political Psychology* **39**(1).

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)*.

Stier, S., Bleier, A., Lietz, H. and Strohmaier, M. (2018). Election campaigning on social media: Politicians, audiences, and the mediation of political communication on facebook and twitter, *Political Communication* **35**(1): 50–74.

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.

Torfi, A., Shirvani, R. A., Keneshloo, Y., Tavaf, N. and Fox, E. A. (2020). Natural language processing advancements by deep learning: A survey, *CoRR* **abs/2003.01200**.
URL: *https://arxiv.org/abs/2003.01200*

van Vliet, L., Törnberg, P. and Uitermark, J. (2020). The twitter parliamentarian database: Analyzing twitter politics across 26 countries, *PLoS ONE* **15**(9).

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 report 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 reports 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.