ControCurator: Human-Machine Framework For Identifying Controversy

Benjamin Timmermans[1], Kaspar Beelen[2], Lora Aroyo[1], Evangelos Kanoulas[2], Tobias Kuhn[1], Bob van de Velde[2], and Gerben van Eerten[3]

[1]Vrije Universiteit, Amsterdam
[2]University of Amsterdam
[3]CrowdyNews

ABSTRACT

In this paper the ControCurator human-machine framework for identifying controversy in multimodal data is described. The goal of ControCurator is to enable modern information access systems to discover and understand controversial topics and events by bringing together crowds and machines in a joint active learning workflow for the creation of adequate training data. This active learning workflow allows a user to identify and understand controversy in ongoing issues, regardless of whether there is existing knowledge on the topic.

Keywords

ControCurator, Controversy, Debate, Discussion

1. INTRODUCTION

Controversies emerge as divisive public debates. They revolve around issues on which large segments of society harbor opposite views, which are often mediated via news coverage or online social networks. Given the increasing influence of the Web, controversies that occur online can have huge impact on both societal (the outcome of elections or the stock market) as well as individual (the formation of opinions, mental well-being) processes and outcomes. Therefore, inside as well as outside academia, researchers set out to develop methods that effectively detect and monitor controversies. Existing methods focus on using conversation graphs of tweets [3], opinion mining [1] or relied on Wikipedia pages as the yardstick of controversy [2]. The problem with Wikipedia is its reliance on a well structured and existing debate, which may not always be the case. Adding to this is that there is currently no agreed upon definition as to what exactly defines controversy.

2. CONTROCURATOR FRAMEWORK

We define our approach for identifying controversy "ControCurator", because we aim to combine and curate data from crowds and machines in order to identify and better understand controversial issues. As can be seen in Figure 1, the ControCurator architecture consists of two main parts: identifying contrasting viewpoints and identifying controversy. Both parts share the same corpus of discussions, but are distinguished by their purpose.

Detecting controversy in a textual corpus, and blending results requires identifying and quantifying six aspects of the CAPOTE model, which models controversy as a debate containing the following aspects:

1. **Controversy**: Whether discussants describe the debate as controversial.
2. **Actors**: How many actors are involved.
3. **Polarization**: Whether the debate is characterized by polarizing dynamics, i.e. increasing adversarial stances.
4. **Open Debate**: Whether the debate is held in the public sphere and move public opinion.
5. **Time**: Debate has a history, tends to persist over time.
6. **Emotion**: Because the debate impinges on values and beliefs, it often triggers emotional responses.

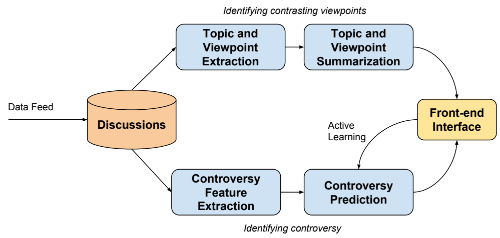

Figure 1: ControCurator Architecture

Given a "conversation" as input, ControCurator extracts features that represent the six aspects of controversy, combines these features using machine learning, and predicts controversy. The granularity of the textual data on which ControCurator operates, therefore, depends on whether we can convert data into a "conversation" at that granularity. In order to achieve this the following granularity were identified:

- **Documents:** ControCurator operates on the level of a "document" content such as a news article. Detecting whether a document is controversial requires ControCurator extracting statements included in the document and viewing a sequence of statements as a conversation across the actors that make these statements.
- **Documents with comments:** ControCurator operates on the level of a "document" such as social media, by accounting for users comments on the document instead of the document content itself.
- **Aggregates:** ControCurator can also operate at higher than document granularity by aggregating documents on the basis of some *anchor*. An anchor can be a news topic, an author of news articles, a hashtag in Twitter etc.

Beyond detecting controversy, ControCurator blends content to present a holistic viewpoint on a controversial topic. Blending is done at a granularity lower than the granularity of controversy detection. For instance, if ControCurator predicts controversy at the level of a hashtag in Twitter, then for the same hashtag it can blend tweets that cover all aspects of the controversy.

2.1 Feature Extraction

This section describes the features ControCurator extracts from a Newswire textual consisting of primarily Guardian articles. The task is to predict whether a Guardian article describes a controversial topic based on the surrounding discussion.The feature space captures the aspects of controversy summarized by the CAPOTE acronym:

1. **Open Debate and Actors (O & A)**: To measure the size of a debate we look, first of all, at number of comments posted below an article, or the scope of actors involved.
2. **Emotion (E)**: This feature type gauges the sentiment profile of a discussions by looking at the amount of clearly positive and clearly negative posts. The sentiment is computed using SentiStrength.
3. **Polarization (P)**: Partyl builds on sentiment features extracted previously, Moreover, we clustering of users and the use of antonyms, to
4. **Indexical aspects (C)**: If something amounts to a controversy, relies partly on whether discussant describe it as such. We apply ”Controversy Lexicons” to score documents as either expressing agreement of disagreement.
5. **Context (persistence of debate (T))**: By computing the textual similarity of a document to the controversial issues listed on Wikipedia, we assess the extent to which it ties in with established and ongoing disputes.

2.2 Learning Algorithms

There are three key points in ControCurator where machine learning algorithms are used: learning the presence of controversy features, learning which data to acquire training data for, and leaning how to cover all aspects of a controversy:

- **Batch Learning:** Having defined the CAPOTE controversy model, and extracted a set of features that can quantify, in some extent, the presence or absence of the six aspects of controversy as defined in the model, Controcurator applies state-of-the-art machine learning, treating the detection of controversy as a classification problem. In order to sufficiently train this algorithm, a large number of labels were acquired through crowdsourcing.
- **Active Learning:** Often, a model trained on some collection (e.g. on comments of Youtube videos) needs to be applied on a different collection (e.g. on comments of Guardian articles). Additional labeled data points may be required to fine-tune the Controcurator controversy detection algorithm on the new collection. Active learning algorithms have been implemented to identify the most informative data points to acquire labels on and train the algorithm. The base training was achieved with the crowdsourced labels.
- **Blending of the results:** Blending results first requires a slicing and dicing of the controversial data into a number of aspects, e.g. subtopics, actors, polarity, time, etc. A blending of data points on the basis of these aspects is then performed so that diversity and coverage over all aspects of a controversial topic is achieved. The algorithms that will be used in order to blend results are all unsupervised, and therefore no label is required.

3. CONTROCURATOR DEMONSTRATOR

In order to demonstrate the corpus, the learned features and the active learning interface to the world, a front-end interface is being developed at http://controcurator.org/browse/. This interface brings together the two parts of the ControCurator architecture: the identification of controversy as a “Barometer”, and the identification of contrasting viewpoints as an “Event Blender” which is developed throughout the next half of this project.

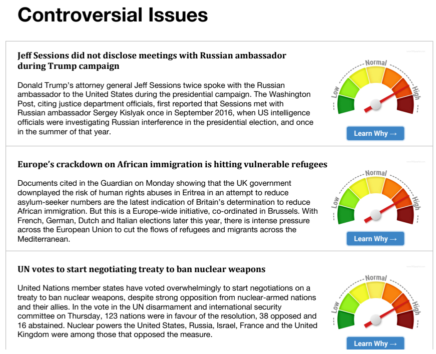

Figure 2: Mockup of ControCurator Interface

The main feature of the controversy barometer is to identify controversial issues, which can be seen on the home screen of the user interface in Figure 2. In this view the user can browse through the listed issues with their controversy score and select one for more in-depth details. A modal view is presented on the side of any page, where the user is asked to go to a document that was selected by the active learning algorithm to be annotated. The user can then annotate that document with a controversy level in the document view.

ACKNOWLEDGEMENTS

This publication was supported by the Dutch national program COMMIT/.

4. REFERENCES

[1] R. Awadallah, M. Ramanath, and G. Weikum. Opinions network for politically controversial topics. In *Proceedings of the first edition workshop on Politics, elections and data*, pages 15–22. ACM, 2012.

[2] S. Dori-Hacohen and J. Allan. Automated controversy detection on the web. In *European Conference on Information Retrieval*, pages 423–434. Springer, 2015.

[3] K. Garimella, G. De Francisci Morales, A. Gionis, and M. Mathioudakis. Quantifying controversy in social media. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pages 33–42. ACM, 2016.