

# 360° Stance Detection

Sebastian Ruder, John Glover, Afshin Mehrabani, Parsa Ghaffari

Aylien Ltd., Dublin, Ireland

{sebastian, john, afshin, parsa}@aylien.com

## Abstract

The proliferation of fake news and filter bubbles makes it increasingly difficult to form an unbiased, balanced opinion towards a topic. To ameliorate this, we propose 360° Stance Detection, a tool that aggregates news with multiple perspectives on a topic. It presents them on a spectrum ranging from support to opposition, enabling the user to base their opinion on multiple pieces of diverse evidence.

## 1 Introduction

The growing epidemic of fake news in the wake of the election cycle for the 45th President of the United States has revealed the danger of staying within our filter bubbles. In light of this development, research in detecting false claims has received renewed interest (Wang, 2017). However, identifying and flagging false claims may not be the best solution, as putting a strong image, such as a red flag, next to an article may actually entrench deeply held beliefs (Lyons, 2017).

A better alternative would be to provide additional evidence that will allow a user to evaluate multiple viewpoints and decide with which they agree. To this end, we propose 360° Stance Detection, a tool that provides a wide view of a topic from different perspectives to aid with forming a balanced opinion. Given a topic, the tool aggregates relevant news articles from different sources and leverages recent advances in stance detection to lay them out on a spectrum ranging from support to opposition to the topic.

Stance detection is the task of estimating whether the attitude expressed in a text towards a given topic is ‘in favour’, ‘against’, or ‘neutral’. We collected and annotated a novel dataset, which associates news articles with a stance towards a specified topic. We then trained a state-of-the-art

stance detection model (Augenstein et al., 2016) on this dataset.

The stance detection model is integrated into the 360° Stance Detection website as a web service. Given a news search query and a topic, the tool retrieves news articles matching the query and analyzes their stance towards the topic. The demo then visualizes the articles as a 2D scatter plot on a spectrum ranging from ‘against’ to ‘in favour’ weighted by the prominence of the news outlet and provides additional links and article excerpts as context.<sup>1</sup>

The interface allows the user to obtain an overview of the range of opinion that is exhibited towards a topic of interest by various news outlets. The user can quickly collect evidence by skimming articles that fall on different parts of this opinion spectrum using the provided excerpts or peruse any of the original articles by following the available links.

## 2 Related work

Until recently, stance detection had been mostly studied in debates (Walker et al., 2012; Hasan and Ng, 2013) and student essays (Faulkner, 2014). Lately, research in stance detection focused on Twitter (Rajadesingan and Liu, 2014; Mohammad et al., 2016; Augenstein et al., 2016), particularly with regard to identifying rumors (Qazvinian et al., 2011; Lukasik et al., 2015; Zhao et al., 2015). More recently, claims and headlines in news have been considered for stance detection (Ferreira and Vlachos, 2016), which require recognizing entailment relations between claim and article.

<sup>1</sup>The demo can be accessed here: <http://bit.do/aylien-stance-detection-demo>. A screencast of the demo is available here: <https://www.youtube.com/watch?v=WYckOr2NhFM>.

Read the text below paying close attention to detail:

|          |  |
|----------|--|
| TOPIC    | Wind farm  |
| Headline | State House Puts Curbs On Wind Farms Pending Study   |
| Article  | The Tennessee House on Thursday voted to put curbs on wind farms pending a study. Senator Lamar Alexander said, "I am glad the Tennessee General Assembly is taking steps to prohibit the construction of some Tennessee wind farms and instead give the state a chance to study the issue. This will give Tennesseans the opportunity to evaluate whether we want our landscape littered with wind turbines that are over two times as tall as the skyboxes at the University of Tennessee football stadium and produce only a small amount of unreliable electricity." |

What is the article's stance towards the topic (Wind farm)? (required)

- In favour
- Neutral
- Against
- Unrelated

Figure 1: Interface provided to annotators. Annotation instructions are not shown.

### 3 Dataset

#### 3.1 Task definition

The objective of stance detection in our case is to classify the stance of an author’s news article towards a given topic as ‘in favour’, ‘against’, or ‘neutral’. Our setting differs from previous instantiations of stance detection in two ways: a) We focus on excerpts from news articles, which are longer and may be more complex than tweets; and b) we do not aim to classify a news article with regard to its agreement with a claim or headline but with regard to its stance towards a topic.

#### 3.2 Data collection

We collect data using the AYLIEN News API<sup>2</sup>, which provides search capabilities for news articles enriched with extracted entities and other metadata. As most extracted entities have a neutral stance or might not be of interest to users, we take steps to compile a curated list of topics, which we detail in the following.

**Topics** We define a topic to include named entities, but also more abstract, controversial keywords such as ‘gun control’ and ‘abortion’. We compile a diverse list of topics that people are likely to be interested in from several sources: a) We retrieve the top 10 entities with the most mentions in each month from November 2015 to June 2017 and filter out entities that are not locations, persons, or organizations and those that are generally perceived as neutral; b) we manually curate a list of current important political figures; and c) we use DBpedia to retrieve a list of controversial topics. Specifically, we included all of the topics mentioned in the Wikipedia

<sup>2</sup><https://newsapi.aylien.com/>

| Topic type    | # topics | Examples             |
|---------------|----------|----------------------|
| Popular       | 44       | Arsenal F.C., Russia |
| Controversial | 300      | Abortion, Polygamy   |
| Political     | 22       | Ted Cruz, Xi Jinping |
| Total         | 366      |                      |

Table 1: Types and numbers of retrieved topics.

list of controversial issues<sup>3</sup> and converted them to DBpedia resource URIs (e.g. <http://en.wikipedia.org/wiki/Abortion> → <http://dbpedia.org/resource/Abortion>) in order to facilitate linking between topics and DBpedia metadata. We then used DBpedia types (Auer et al., 2007) to filter out all entities of type Place, Person and Organisation. Finally, we ranked the remaining topics based on their number of unique outbound edges within the DBpedia graph as a measure of prominence, and picked the top 300. We show the final composition of topics in Table 1. For each topic, we retrieve the most relevant articles using the News API from November 2015 to July 2017.

**Annotation** For annotation, we need to trade-off the complexity and cost of annotation with the agreement between annotators. Annotating entire news articles places a large cognitive load on the annotator, which leads to fatigue and inaccurate annotations. For this reason, we choose to annotate excerpts from news articles. In internal studies, we found that providing a context window of 2-3 sentences around the mention of the entity together with the headline provides suffi-

<sup>3</sup>[https://en.wikipedia.org/wiki/Wikipedia:List\\_of\\_controversial\\_issues](https://en.wikipedia.org/wiki/Wikipedia:List_of_controversial_issues)

## Stance Detection

In order to evaluate a series of articles, enter a keyword to fetch the results from AYLIEN News API and another keyword to fetch the results from Stance Detection model. Please note that these two keywords could be different.



Figure 2: 360° Stance Detection interface. News articles about a query, i.e. ‘Ireland AND brexit’ are visualized based on their stance towards a specified topic, i.e. ‘ireland’ and the prominence of the source. Additional information is provided in a table on the right, which allows to skim article excerpts or follow a link to the source.

cient context to produce a reliable annotation. If the entity is not mentioned explicitly, we provide the first paragraph of the article and the headline as context. We annotate the collected data using CrowdFlower with 3 annotators per example using the interface in Figure 1. We retain all examples where at least 2 annotators agree, which amounts to 70.5% of all examples.

**Final dataset** The final dataset consists of 32,227 pairs of news articles and topics annotated with their stance. In particular, 47.67% examples have been annotated with ‘neutral’, 21.9% with ‘against’, 19.05% with ‘in favour’, and 11.38% with ‘unrelated’. We use 70% of examples for training, 20% for validation, and 10% for testing according to a stratified split. As we expect to encounter novel and unknown entities in the wild, we ensure that entities do not overlap across splits and that we only test on unseen entities.

## 4 Model

We train a Bidirectional Encoding model (Augenstein et al., 2016), which has achieved state-of-the-art results for Twitter stance detection on our dataset. The model encodes the entity using a bidi-

rectional LSTM (BiLSTM)<sup>4</sup>, which is then used to initialize a BiLSTM that encodes the article and produces a prediction. To reduce the sequence length, we use the same context window that was presented to annotators for training the LSTM. We use pretrained GloVe embeddings (Pennington et al., 2014) and tune hyperparameters on a validation set. The best model achieves a test accuracy of 61.7 and a macro-averaged test F1 score of 56.9.<sup>5</sup> It significantly outperforms baselines such as a bag-of-n-grams (accuracy: 46.3; F1: 44.2).

## 5 360° Stance Detection Demo

The interactive demo interface of 360° Stance Detection, which can be seen in Figure 2, takes two inputs: a news search query, which is used to retrieve news articles using News API, and a stance target topic, which is used as the target of the stance detection model. For good results, the stance target should also be included as a keyword in the news search query. Multiple keywords can be provided as the query by connecting them with

<sup>4</sup>We tried other encoding strategies, such as averaging pretrained embeddings, but this performed best.

<sup>5</sup>These scores are comparable to those achieved in (Augenstein et al., 2016). Compared to tweets, stance in news is often more subtle and thus more challenging to detect, while our dataset contains more diverse entities than previous ones.



Figure 3: 360° Stance Detection visualizations for example queries and topics.

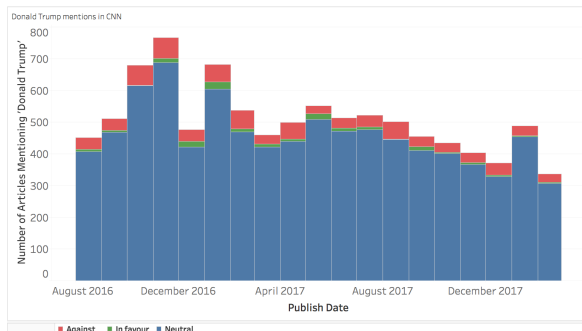


Figure 4: Visualization distribution of stance towards Donald Trump and number of CNN news articles mentioning Donald Trump from August 2016 to January 2018.

‘AND’ or ‘OR’ as in Figure 2.

When these two inputs are provided, the application retrieves a predefined number of news articles (up to 50) that match the first input, and analyzes their stance towards the target (the second input) using the stance detection model. The stance detection model is exposed as a web service and returns for each article-target entity pair a stance label (i.e. one of ‘in favour’, ‘against’ or ‘neutral’) along with a probability.<sup>6</sup>

The demo then visualizes the collected news articles as a 2D scatter plot with each (x,y) coordinate representing a single news article from a particular outlet that matched the user query. The x-axis shows the stance of the article in the range  $[-1, 1]$ . The y-axis displays the prominence of the news outlet that published the article in the range  $[1, 1000000]$ , measured by its Alexa ranking<sup>7</sup>. A table displays the provided information in a complementary format, listing the news outlets of the articles, the stance labels, confidence scores, and

<sup>6</sup>We leave confidence calibration (Guo et al., 2017) for future work.

<sup>7</sup><https://www.alexa.com/>

prominence rankings. Excerpts of the articles can be scanned by hovering over the news outlets in the table and the original articles can be read by clicking on the source.

360° Stance Detection is particularly useful to gain an overview of complex or controversial topics and to highlight differences in their perception across different outlets. We show visualizations for example queries and three controversial topics in Figure 3. By extending the tool to enable retrieval of a larger number of news articles and more fine-grained filtering, we can employ it for general news analysis. For instance, we can highlight the volume and distribution of the stance of news articles from a single news outlet such as CNN towards a specified topic as in Figure 4.

## 6 Conclusion

We have introduced 360° Stance Detection, a tool that aims to provide evidence and context in order to assist the user with forming a balanced opinion towards a controversial topic. It aggregates news with multiple perspectives on a topic, annotates them with their stance, and visualizes them on a spectrum ranging from support to opposition, allowing the user to skim excerpts of the articles or read the original source. We hope that this tool will demonstrate how NLP can be used to help combat filter bubbles and fake news and to aid users in obtaining evidence on which they can base their opinions.

## Acknowledgments

Sebastian Ruder is supported by the Irish Research Council Grant Number EBPPG/2014/30 and Science Foundation Ireland Grant Number SFI/12/RC/2289.

## References

- Sören Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. 2007. Dbpedia: A nucleus for a web of open data. In *The semantic web*, Springer, pages 722–735.
- Isabelle Augenstein, Tim Rocktäschel, Andreas Vlachos, and Kalina Bontcheva. 2016. *Stance Detection with Bidirectional Conditional Encoding*. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. <http://arxiv.org/abs/1606.05464>.
- Adam Faulkner. 2014. Automated classification of stance in student essays: An approach using stance target information and the wikipedia link-based measure. *Science* 376(12):86.
- William Ferreira and Andreas Vlachos. 2016. Emergent: a novel data-set for stance classification. In *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: Human language technologies*. pages 1163–1168.
- Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q. Weinberger. 2017. *On Calibration of Modern Neural Networks*. *Proceedings of ICML 2017* <http://arxiv.org/abs/1706.04599>.
- Kazi Saidul Hasan and Vincent Ng. 2013. Stance classification of ideological debates: Data, models, features, and constraints. In *Proceedings of the Sixth International Joint Conference on Natural Language Processing*. pages 1348–1356.
- Michal Lukasik, Trevor Cohn, and Kalina Bontcheva. 2015. Classifying tweet level judgements of rumours in social media. *arXiv preprint arXiv:1506.00468*.
- Tessa Lyons. 2017. *News Feed FYI: Replacing Disputed Flags with Related Articles*. <https://bit.ly/2BTvu0x>.
- Saif Mohammad, Svetlana Kiritchenko, Parinaz Sobhani, Xiaodan Zhu, and Colin Cherry. 2016. Semeval-2016 task 6: Detecting stance in tweets. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*. pages 31–41.
- Jeffrey Pennington, Richard Socher, and Christopher D. Manning. 2014. *Glove: Global Vectors for Word Representation*. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing* pages 1532–1543. <https://doi.org/10.3115/v1/D14-1162>.
- Vahed Qazvinian, Emily Rosengren, Dragomir R Radev, and Qiaozhu Mei. 2011. Rumor has it: Identifying misinformation in microblogs. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, pages 1589–1599.
- Ashwin Rajadesingan and Huan Liu. 2014. Identifying users with opposing opinions in twitter debates. In *International conference on social computing, behavioral-cultural modeling, and prediction*. Springer, pages 153–160.
- Marilyn A Walker, Pranav Anand, Robert Abbott, and Ricky Grant. 2012. Stance classification using dialogic properties of persuasion. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, pages 592–596.
- William Yang Wang. 2017. "Liar, Liar Pants on Fire": A New Benchmark Dataset for Fake News Detection. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL 2017)*.
- Zhe Zhao, Paul Resnick, and Qiaozhu Mei. 2015. Enquiring minds: Early detection of rumors in social media from enquiry posts. In *Proceedings of the 24th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, pages 1395–1405.