Stance Detection: A Review

Assoc. Prof. Dilek Küçük
Chief Researcher
TÜBİTAK Marmara Research Center - Energy Institute

This talk is based on the following publication:

Outline

- Introduction
- Stance Detection and Related Problems
- A Generic System Architecture
- A Historical Perspective
- Approaches to Stance Detection
- Datasets and Evaluation Metrics
- Software and Tools
- Stance Detection Experiments on Turkish Tweets
- Application Areas
- Outstanding Issues
- Conclusions
Introduction
Definition 1.1 (Stance Detection). For an input in the form of a piece of text and a target pair, stance detection is a classification problem where the stance of the author of the text is sought in the form of a category label from this set: \{Favor, Against, Neither\}. Occasionally, the category label of Neutral is also added to the set of stance categories [Mohammad et al. 2016b] and the target may or may not be explicitly mentioned in the text [Augenstein et al. 2016a; Mohammad et al. 2016b].

Definition 1.2 (Multi-target Stance Detection). For an input in the form of a piece of text and a set of related targets, multi-target stance detection is a classification problem where the stance of the text author is sought as a category label from this set: \{Favor, Against, Neither\} for each target and each stance classification (for each target) might have an effect on the classifications for the remaining targets [Sobhani 2017].

Definition 1.3 (Cross-target Stance Detection). Cross-target stance detection is a classification problem where the stance of the text author is sought for a specific target as a category label from this set: \{Favor, Against, Neither\}, in a settings where stance annotations are available for (though related but) different targets, i.e., there is not enough stance-annotated training data for the target under consideration [Augenstein et al. 2016a; Xu et al. 2018].
Definition 1.4 (Rumour Stance Classification). For an input in the form of a piece of text and a rumour pair, rumour stance classification is a problem where the position of the text author towards the veracity of the rumour is sought for, in the form of a category label from this set: \{Supporting, Denying, Querying, Commenting\}. As the set of possible category labels, a subset of this set such as \{Supporting, Denying\} is occasionally employed [Zubiaga et al. 2018].

Definition 1.5 (Fake News Stance Detection). For an input in the form of news headline and a news body pair (where the headline and body parts may belong to different news articles), this is a classification problem where the stance of the body towards the claim of the headline is sought for, in the form of a category label from this set: \{Agrees, Disagrees, Discusses (the same topic), Unrelated\}. This problem is defined in order to facilitate the task of fake news detection [PNC 2017].
### Table 1. Sample Tweets from SemEval 2016 Stance Dataset [Mohammad et al. 2016b].

<table>
<thead>
<tr>
<th>Tweet</th>
<th>Stance Target</th>
<th>Stance</th>
<th>Sentiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT @TheCLF: Thanks to everyone in Maine who contacted their legislators in support of #energyefficiency funding! #MEpoli #SemST</td>
<td>Climate Change is a Real Concern</td>
<td>Favor</td>
<td>Positive</td>
</tr>
<tr>
<td>We live in a sad world when wanting equality makes you a troll... #SemST</td>
<td>Feminist Movement</td>
<td>Favor</td>
<td>Negative</td>
</tr>
<tr>
<td>I don't believe in the hereafter. I believe in the here and now. #SemST</td>
<td>Atheism</td>
<td>Favor</td>
<td>Neither</td>
</tr>
<tr>
<td>@violencehurts @WomenCanSee The unborn also have rights #defendthe8th #SemST</td>
<td>Legalization of Abortion</td>
<td>Against</td>
<td>Positive</td>
</tr>
<tr>
<td>I'm conservative but I must admit I'd rather see @SenSanders as president than Mrs. Clinton. #stillVotingCOP #politics #SemST</td>
<td>Hillary Clinton</td>
<td>Against</td>
<td>Negative</td>
</tr>
<tr>
<td>I have my work and my faith... If that's boring to some people, I can't tell you how much I don't care. ~Madonna Ciccone #SemST</td>
<td>Atheism</td>
<td>Against</td>
<td>Neither</td>
</tr>
<tr>
<td>@BadgerGeno @kreichert27 @jjackbahlman Too busy protesting :) #LoveForAll #BackdoorBadgers #SemST</td>
<td>Hillary Clinton</td>
<td>Neither</td>
<td>Positive</td>
</tr>
<tr>
<td>@ShowTruth You're truly unwelcome here. Please leave. #ygk #SemST</td>
<td>Legalization of Abortion</td>
<td>Neither</td>
<td>Negative</td>
</tr>
<tr>
<td>@Maisie_Williams everyone feels that way at times. Not just women #SemST</td>
<td>Atheism</td>
<td>Neither</td>
<td>Neither</td>
</tr>
</tbody>
</table>
Stance Detection and Related Problems
A Generic System Architecture

(a) Training phase.
A Generic System Architecture

(b) Testing phase.
A Historical Perspective

Earlier Work on Stance Detection [2006 - 2015]

- Earlier work are carried out on
  - Congressional-floor debates
  - Company internal discussions
  - Online social, political, and ideological debates (in public forums)
  - Online debates about products
  - Spontaneous speech (a single study by Levow et al. (2014))
  - Student essays
  - Tweets (few studies)

- Approaches in earlier work
  - Few rule-based methods
  - Supervised learning methods (SVM, decision tree, random forest, HMM, CRF, ILP, …)
A Historical Perspective

Stance Detection Competitions [2016 - 2017]

- **SemEval-2016** shared task on stance detection in English tweets
  (Mohammad et al., 2016)
  - **Targets:** Atheism, Climate change is a real concern, Feminist movement, Hillary Clinton, Legalization of abortion, Donald Trump

- **NLPCC-ICCPOL-2016** shared task on stance detection in Chinese microblogs
  (Xu et al., 2016)
  - **Targets:** iPhone SE, Set off firecrackers in the Spring Festival, Russia's anti terrorist operations in Syria, Two child policy, Prohibition of motorcycles and restrictions on electric vehicles in Shenzhen, Genetically modified food, Nuclear test in DPRK

- **IberEval-2017** shared task on stance detection in Spanish and Catalan tweets
  (Taulé et al., 2017)
  - **Target:** Independence of Catalonia
Approaches to Stance Detection

Table 5. Temporal Distribution of Published Papers on Stance Detection

<table>
<thead>
<tr>
<th>Publication Year</th>
<th>Number of Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006 – 2010</td>
<td>5</td>
</tr>
<tr>
<td>2011 – 2014</td>
<td>8</td>
</tr>
<tr>
<td>2015 – 2016</td>
<td>38</td>
</tr>
<tr>
<td>2017 – 2019</td>
<td>78</td>
</tr>
</tbody>
</table>
Fig. 4. A word cloud of the algorithms used for stance detection problem in the published papers included in this survey paper.
Approaches to Stance Detection

- **Feature-based machine learning approaches**
  - SVM, Logistic Regression, Naïve Bayes, Decision Trees, ANN, ILP, kNN, ...

- **Deep learning approaches**
  - LSTM (RNN), RNN, GRU (RNN), CNN

- **Ensemble learning approaches**
  - Random Forest, Majority Voting, Proprietary ensemble learners, Boosting, Bagging, ...
<table>
<thead>
<tr>
<th>Authors</th>
<th>Domain</th>
<th>Annotation Classes</th>
<th>Target(s)</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Mohammad et al. 2016a]</td>
<td>Tweets (English)</td>
<td>Favor, Against, Neither</td>
<td>Atheism, Climate change is a real concern, Feminist movement, Hillary Clinton, Legalization of abortion, Donald Trump</td>
<td>4,870 tweets</td>
</tr>
<tr>
<td>[Mohammad et al. 2017]</td>
<td>Tweets (English)</td>
<td>Favor, Against, Neither for stance; Positive, Negative, and Neither for sentiment</td>
<td>Atheism, Climate change is a real concern, Feminist movement, Hillary Clinton, Legalization of abortion, Donald Trump</td>
<td>4,870 tweets</td>
</tr>
<tr>
<td>[Xu et al. 2016b]</td>
<td>Microblogs (Chinese)</td>
<td>Favor, Against, None</td>
<td>iPhone SE, Set off firecrackers in the Spring Festival, Russia’s anti terrorist operations in Syria, Two child policy, Prohibition of motorcycles and restrictions on electric vehicles in Shenzhen, Genetically modified food, Nuclear test in DPRK</td>
<td>4,000 annotated and 2,400 unannotated tweets</td>
</tr>
<tr>
<td>[Taulé et al. 2017]</td>
<td>Tweets (Catalan &amp; Spanish)</td>
<td>Favor, Against, None</td>
<td>Independence of Catalonia</td>
<td>5,400 tweets in Spanish and 5,400 tweets in Catalan</td>
</tr>
</tbody>
</table>
### Datasets

<table>
<thead>
<tr>
<th>Authors</th>
<th>Domain</th>
<th>Annotation Classes</th>
<th>Target(s)</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Sobhani et al. 2017]</td>
<td>Tweets</td>
<td>Favor, Against, Neither</td>
<td>{Clinton-Sanders}, {Clinton-Trump}, {Cruz-Trump}</td>
<td>4,455 tweets</td>
</tr>
<tr>
<td>[Küçük 2017b]</td>
<td>Tweets</td>
<td>Favor, Against</td>
<td>Galatasaray, Fenerbahçe</td>
<td>700 tweets</td>
</tr>
<tr>
<td>[Küçük and Can 2018]</td>
<td>Tweets</td>
<td>Favor, Against</td>
<td>Galatasaray, Fenerbahçe</td>
<td>1,065 tweets</td>
</tr>
<tr>
<td>[Murakami and Raymond 2010]</td>
<td>Online debates</td>
<td>Support, Oppose</td>
<td>Selected five ideas</td>
<td>481 comments about five ideas</td>
</tr>
<tr>
<td>[Darwish et al. 2017]</td>
<td>Tweets</td>
<td>Favor (Positive), Against (Negative)</td>
<td>Transfer of two islands from Egypt to Saudi Arabia</td>
<td>33,024 tweets</td>
</tr>
<tr>
<td>[Hercig et al. 2017]</td>
<td>News comments</td>
<td>In Favor, Against, Neither</td>
<td>Miloš Zeman, Smoking ban in restaurants</td>
<td>5,423 news comments</td>
</tr>
</tbody>
</table>
Evaluation Metrics

\[
F = \frac{F_{\text{Favor}} + F_{\text{Against}}}{2}
\]

\[
F_{\text{Favor}} = \frac{2 \times P_{\text{Favor}} \times R_{\text{Favor}}}{P_{\text{Favor}} + R_{\text{Favor}}}
\]

\[
F_{\text{Against}} = \frac{2 \times P_{\text{Against}} \times R_{\text{Against}}}{P_{\text{Against}} + R_{\text{Against}}}
\]

\[
P_{\text{Favor}} = \frac{\text{Correct}_{\text{Favor}}}{\text{Correct}_{\text{Favor}} + \text{Spurious}_{\text{Favor}}}
\]

\[
P_{\text{Against}} = \frac{\text{Correct}_{\text{Against}}}{\text{Correct}_{\text{Against}} + \text{Spurious}_{\text{Against}}}
\]

\[
R_{\text{Favor}} = \frac{\text{Correct}_{\text{Favor}}}{\text{Correct}_{\text{Favor}} + \text{Missing}_{\text{Favor}}}
\]

\[
R_{\text{Against}} = \frac{\text{Correct}_{\text{Against}}}{\text{Correct}_{\text{Against}} + \text{Missing}_{\text{Against}}}
\]

\[
F = \frac{F_{\text{Favor}} + F_{\text{Against}} + F_{\text{Neither}}}{3}
\]

\[
\text{Accuracy} = \frac{\text{Correct classifications}}{\text{All classifications}}
\]
Software and Tools

- Few papers present visualization systems/tools for stance detection.

- Many papers use the following machine learning tools, libraries in their stance detection experiments:
  - Weka
  - Scikit-learn package
  - Keras
  - Theano
  - Gensim
  - SVMlight
  - FastText
  - Brainy
Stance Detection Experiments on Turkish Tweets

Table 1. A Summary of the Three Versions of the Stance-Annotated Tweet Data Set

<table>
<thead>
<tr>
<th>Stance Data Set</th>
<th># of Annotators</th>
<th># of Tweets Annotated</th>
<th>For Target-1</th>
<th>For Target-2</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Favor</td>
<td>Against</td>
<td>Favor</td>
</tr>
<tr>
<td>Version-1</td>
<td>1</td>
<td></td>
<td>175</td>
<td>175</td>
<td>175</td>
</tr>
<tr>
<td>Version-2</td>
<td>2</td>
<td></td>
<td>173</td>
<td>173</td>
<td>173</td>
</tr>
<tr>
<td>Version-3</td>
<td>2</td>
<td></td>
<td>269</td>
<td>268</td>
<td>269</td>
</tr>
</tbody>
</table>

Target-1: Galatasaray    Target-2: Fenerbahçe

https://github.com/dkucuk/Stance-Detection-Turkish-V1
https://github.com/dkucuk/Stance-Detection-Turkish-V2
https://github.com/dkucuk/Stance-Detection-Turkish-V3
Stance Detection Experiments on Turkish Tweets


Table 6. Evaluation Results of the SVM Classifiers Utilizing Unigrams+Hashtag Use+Named Entities as Features, with Named Entities Extracted by the NER Tool.

<table>
<thead>
<tr>
<th>Target</th>
<th>Class</th>
<th>Stance Data Set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Version-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P (%)</td>
</tr>
<tr>
<td>Target-1</td>
<td>Favor</td>
<td>75.6</td>
</tr>
<tr>
<td></td>
<td>Against</td>
<td>87.9</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>81.8</td>
</tr>
<tr>
<td>Target-2</td>
<td>Favor</td>
<td>71.8</td>
</tr>
<tr>
<td></td>
<td>Against</td>
<td>81.3</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>76.5</td>
</tr>
</tbody>
</table>

Application Areas

- Opinion surveys/polling
- Trend and market analysis/forecast
- Recommendation systems
- Public health surveillance
- Information retrieval
- Stance summarization
- Rumour classification
- Fake news detection
- Automatic fact checking
Outstanding Issues

- Cross-lingual and multilingual stance detection
- Stance detection in other media content and robots
- Stance detection for decision making
- Stance detection in data streams
Stance detection is usually defined as the automatic determination of the position of a post owner (as in favor of or against) towards a specific target, based on the content of the post.

In general, stance detection is performed on the following text genres:
- microblogs (tweets, mostly),
- posts published in online debate forums,
- news articles and comments.

Along with a number of related problems such as sentiment analysis, controversy detection, and argument mining, it is a crucial process to elicit useful information from the underlying content, most of the time, regarding controversial issues or elections/referendums.
Thank You