Detecting Viewpoint and Bias

Sandeep Avula
asandeep@unc.edu

Questions

- What is the task?
- How is it different from opinion mining?
- What types of features can be used to automatically recognize viewpoint?
- What is needed in order to improve results?

Definitions

- Stance: an overall position held by a person toward an object, idea, or proposition.
- Examples:
  - For vs. against the existence of God
  - Pro-life vs. pro-choice
  - For vs. against universal healthcare
  - For vs. against gay marriage

Detecting Viewpoint

- Recognizing the stance of an author of an argumentative text within a particular debate

Examples

creationism vs. evolution

" .... The belief that God creates and sustains the world is not the same as the belief that God created the world in seven days a few thousand years ago. The former is a theological position which implies the goodness and sovereignty of God, and his continuing involvement with his creation, not a scientific claim. The latter is an empirically testable claim which has been empirically tested and found to be false. We should not allow schools to teach our children things that are just plain wrong, and known to be wrong. This is distinct from allowing the teaching of religion, which is culturally and historically significant and which involves beliefs which are in principle distinct from science."

Examples

for vs. against more gun control

"A handgun ban lowers the availability of guns used in domestic passion murders."

"According to the Supreme Court, [the] states cannot prohibit people's possession of guns. [...] Under the existence of the Second Amendment a total prohibition of people's possession of guns will be illegitimate and unconstitutional."

"[...] The civilian possession of handguns, therefore, increases training that could be useful to the military in the event of a draft. In so far as this is true, any reduction in the ability of citizens to own, use, and train with handguns will reduce the battle readiness of the military."

Examples

for vs. against universal healthcare

“Public insurance is less agile and innovative than private insurers.”

“No matter how much you may like to think otherwise, you could not individually pay for the all the things that make your life the way it is. Your continued insistence that you could survive and build a society similar to the one we have without taxation is incredibly naive.”

“[...] Not to mention, government programs are always a mess and our constitution is designed to limit governmental power for that reason.”

Questions

- How is detecting viewpoint different from opinion mining (e.g., determining whether a movie review expresses a positive or negative opinion)?

Detecting Viewpoint

- S. Somasundaran and J. Wiebe. Recognizing Stances in Ideological on-Line Debates. In *NAACL/HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*. ACL.

Methodology

- Annotate/scrape data from several domains
  - existence of God, universal healthcare, gun rights, gay rights, abortion, and creationism
- Use data from the first two domains to learn about the problem and design different feature representations
- Test different feature representations on the held-out domains
- Evaluate against a distribution-based approach and a unigram classifier

Observations

1. When supporting their side, people do not only express their sentiment, but argue about what is true
   - an argument is a subjective expression about what is true, should be true, or should be done
   - considering positive/negative adjectives (e.g., “good”, “bad”, “terrible”) is not enough

Examples

for vs. against more gun control

"A handgun ban lowers the availability of guns used in domestic passion murders."

"According to the Supreme Court, [the] states cannot prohibit people's possession of guns. [...] Under the existence of the Second Amendment, a total prohibition of people's possession of guns will be illegitimate and unconstitutional."

"[...] The civilian possession of handguns, therefore, increases training that could be useful to the military in the event of a draft. In so far as this is true, any reduction in the ability of citizens to own, use, and train with handguns will reduce the battle readiness of the military."

Observations

2. If a particular positive/negative sentiment is expressed, the target of the sentiment is important in determining the author's stance.

- for and against arguments can contain both positive and negative sentiment words
- the target of the sentiment is important

Examples

for vs. against universal health care

"Yes. Health care is not a privilege, it is a right. In the USA, everything is a commodity, including the precious gift of good health. That's disgusting. I'm so glad I live in Britain."

"We will probably end up with universal healthcare and twenty years from now we will all lament, what happened to the good old days."

"Profits are good for the system."

Observations

3. When supporting their side, people argue about what will, could, and/or should happen.
   - the text surrounding a modal verb may be particularly important
   - modality: likelihood, ability, permission, obligation
   - examples: might, can, may, must, should, etc...

Examples

"No matter how much you may like to think otherwise, you could not individually pay for the all the things that make your life the way it is.Your continued insistence that you could survive and build a society similar to the one we have without taxation is incredibly naive."

"[...] The civilian possession of handguns, therefore, increases training that could be useful to the military in the event of a draft. In so far as this is true, any reduction in the ability of citizens to own, use, and train with handguns will reduce the battle readiness of the military."

Observations

4. In some debates, a person's stance can be determined simply by the "things" (i.e., nouns) mentioned. Debates often involve prioritization.
   - simple unigrams may help
   - examples from the death penalty debate: God, cost, error, escape, justice
   - examples from the climate change debate?
   - examples from the space exploration debate?

Feature Representation

1. Sentiment-based features: modify the words in the text depending on whether they appear in a sentence with an overall positive or negative sentiment

2. Arguing-lexicon features: modify the words in the text depending on whether they appear in a sentence with a positive or negative argument

3. Modal verb features: subject + modal_verb, modal_verb + object, subject + modal_verb + object

   - "healthcare should", "should [be] available", "healthcare should [be] available"

Sentiment-based Features

1. Segment the text into sentences
2. Predict the sentiment (positive, negative, neutral) of each sentence using a predefined sentiment lexicon (i.e. a set of words associated with positive/negative/neutral sentiment, similar to HW1)
3. Change every content unigram in the sentence to the form **(+/-/=)_unigram**, depending on the sentiment of the sentence
4. Do standard training/testing using this feature set

Sentiment-based Features

"Competition in health care is generally a bad idea. While competition can make sense when it relates to the production of goods, it is not good when it relates to dealing with human life. The main problem is that a primary means of cost-cutting is by skimping on services to people in need."

Sentiment-based Features

sentence segmentation

“Competition in health care is generally a bad idea.”

“While competition can make sense when it relates to the production of goods, it is not good when it relates to dealing with human life.”

“The main problem is that a primary means of cost-cutting is by skimping on services to people in need.”

Sentiment-based Features

predict sentiment of each sentence

"Competition in health care is generally a **bad idea**."

"While competition can **make sense** when it relates to the production of goods, it is **not good** when it relates to dealing with human life."

"The main **problem** is that a primary means of cost-cutting is by skimping on services to people in need."

Sentiment-based Features

predict sentiment of each sentence

“-Competition -health -care”

“=competition =production =goods =human =life.”

“-cost-cutting -services -people.”

Arguing-Lexicon Features

- Similar idea and construction as sentiment-based features
- However, not focused on sentiment
- Instead, focused on positive vs. negative arguing
- Positive and negative arguing lexicon (unigrams, bigrams, and trigrams) derived from human-annotated sentences
- Each n-gram associated with **P(n-gram|positive)** and **P(n-gram|negative)**
- Similar to precision value in LightSIDE
- Sentences classified by comparing sum of probabilities

Arguing-Lexicon Features

Entries indicative of Positive Arguing
be important to, would be better, would need to, be just the, be the true, my opinion, the contrast, show the, prove to be, only if, on the verge, ought to, be most, youve get to, render, manifestation, ironically, once and for, no surprise, overwhelming evidence, its clear, its clear that, it be evident, it be extremely, it be quite, it would therefore

Entries indicative of Negative Arguing
be not simply, simply a, but have not, can not imagine, we dont need, we can not do, threat against, ought not, nor will, never again, far from be, would never, not completely, nothing will, inaccurate and, inaccurate and, find no, no time, deny that

Results

Domain (#posts)	Distribution	Unigram	Sentiment	Arguing	Arg+Sent
Overall (2232)	50	62.50	55.02	62.59	63.93
Guns Rights (306)	50	66.67	58.82	69.28	70.59
Gay Rights (846)	50	61.70	52.84	62.05	63.71
Abortion (550)	50	59.1	54.73	59.46	60.55
Creationism (530)	50	64.91	56.60	62.83	63.96

Discussion

- How is recognizing viewpoint more difficult than detecting positive/negative product reviews?
- How is recognizing viewpoint easier than detecting positive/negative product reviews?
- Can the method from Somasundaran and Wiebe be used for other predictive tasks?