

AAAI 2011
San Francisco

**A Microtext Corpus for
Persuasion Detection in Dialog**

Joel Young, Craig Martell, Henry Tucker Gilbert IV
Naval Postgraduate School

Pranav Anand
University of California Santa Cruz

Pedro Ortiz
United States Naval Academy

8 August 2011

Persuasion

- Detecting Persuasion
 - Given a transcribed utterance from a dialog, can we determine if the speaker is attempting persuasion?
- Why Bother?
 - Detecting sexual predators grooming minors
 - Automated training for negotiators
 - Targeted Advertising

Outline

- Related Work and Persuasion Model
- Experimental Design
- Results
- Future Research
- Conclusion

Related Work

- Related Work and Persuasion Model
- Experimental Design
- Results
- Future Research
- Conclusion

Little Published Work

- Pang and Lee, 2004
 - *A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts*
 - Classified sentences in movie reviews as subjective or objective using a combination of Naïve Bayes and SVM
- Palau et al, 2009
 - *Argumentative Mining: The Detection, Classification and Structure of Arguments in Text*
 - Explored structure of argumentation in legal texts

Persuasion:

For persuasion to occur, the persuadee must be resistant to perform the act or hold the attitude in question, and the job of a persuader is to break down that resistance

Revised Cialdini 2007 Persuasion Model

- Reciprocity
- Commitment
- Consistency
- Liking
- Authority
- Social Proof
- Scarcity
- Other
- No Persuasion

Persuasion Model

- *Reciprocity*: People tend to feel indebted to those who do them favors.
- *Commitment*: People are more likely to perform an act if they explicitly commit to it.
- *Consistency*: People are more likely to perform an act they have committed to, when reminded of their previous commitment.
- *Scarcity*: People are more likely to comply if they feel circumstances are fleeting or resources are limited.

Persuasion Model

- *Liking*: People are influenced by those they like, admire, or feel similar to.
- *Authority*: People are influenced by thoughts, words and actions of authority figures. Authority can be embodied in both individuals and organizations.
- *Social Proof*: Expectations and behavior are both influenced by social norms.

NPS Persuasion Corpus

- 18,857 Utterances across 37 Transcripts
 - Rogan: FBI/Police (5249 Utterances)
 - Taylor: Police (6913 Utterances)
 - Waco Texas, Branch Davidian (5871 Utterances)
 - San Diego Police (824 Utterances)
- Converted from audio by source agency
- Punctuation removed, all uppercase, transcriber notes bracketed, names removed
- About 12% persuasion

Annotated Sample Transcript

| Transcript | Line | Type | Speaker | Utterance |
|---------------|------|------------|---------|--|
| Rogan_beta | 221 | Other | ON80 | Yeah but that fiddler isnt gonna cost so much if you walk out easy |
| Rogan_beta | 223 | Other | ON80 | come on <HT01> youre just making it worst on yourself |
| Rogan_charlie | 641 | None | PN1 | Alright [both_hang_up] |
| Rogan_charlie | 691 | Commitment | HT1 | Bring <Wife_First_Name> and Ill come out |

Most Persuasive and Non-Persuasive Terms $P(\text{Persuasion}|\text{Feature})$

| Most Predictive Word | Prob | Least Predictive Word | Prob | Most Predictive Bigram | Prob | Least Predictive Bigram | Prob |
|----------------------|------|-----------------------|------|------------------------|------|-------------------------|------|
| SINCERE | 0.88 | JESUS | 0.04 | YOUR-FRIENDS | 0.94 | YEAH-IM | 0.04 |
| HONORABLE | 0.87 | THANKS | 0.03 | THAT-GUN | 0.93 | ME-IN | 0.04 |
| ANSWERS | 0.86 | SHALL | 0.03 | I-GUARANTEE | 0.93 | HANG-UP | 0.04 |
| CLUBS | 0.86 | HUH | 0.02 | YOUR-FAMILY | 0.93 | NAME-IS | 0.03 |
| LEGITIMATE | 0.85 | SEALS | 0.02 | YOUR-CELLS | 0.92 | I-TRIED | 0.03 |
| ABOARD | 0.83 | HELLO | 0.02 | GET-ALL | 0.92 | MM-HM | 0.03 |
| GUARANTEED | 0.83 | HI | 0.02 | YOUR-SAFETY | 0.92 | MY-WIFE | 0.03 |
| BOUT | 0.83 | CHRIST | 0.01 | GOOD-JOB | 0.92 | OF-GOD | 0.02 |
| TRUSTING | 0.83 | BYE | 0.01 | WHAT-ID | 0.92 | YOU-DOING | 0.02 |
| COOPERATE | 0.82 | HUM | 0.00 | GUN-DOWN | 0.91 | UM-HUM | 0.00 |

Experimental Design

- Related Work and Persuasion Model
- Experimental Design
- Results
- Future Research
- Conclusion

Features

- Unigram
- Bigram
- Orthogonal Sparse Bigram (OSB)
- Gappy Bigrams

START_the, START_purple, START_dog,
the_purple, the_dog, the_END, purple_dog, purple_END, dog_END

Gappy Bigrams

START_0_the, START_1_purple, START_1_dog,
the_0_purple, the_1_dog, the_2_END, purple_0_dog,
purple_1_END, dog_0_END

Orthogonal Sparse Bigrams

The Purple Dog

- **Validation**
 - 6-fold cross validation with 36 transcripts
 - Train on three, test on 1 (four combinations)
- **Experiments**
 - Naive Bayes w/ Laplace Smoothing
 - Maximum Entropy
 - Radial Basis Kernel Support Vector Machine (SVM)
 - All parameters tuned via grid search

Results

- Related Work and Persuasion Model
- Experimental Design
- Results
- Future Research
- Conclusion

6-fold Cross Validation Results

| Naïve Bayes | | | | | |
|-------------|---------------|---------------|---------------|------------------|---------------------|
| Features | Precision | Recall | F-score | Baseline F-score | Percent Improvement |
| Unigrams | 0.4996 | 0.4052 | 0.4450 | 0.2018 | 120.5 |
| Bigrams | 0.4572 | 0.4172 | 0.4334 | 0.2018 | 114.8 |
| Gappy | 0.5072 | 0.4572 | 0.4772 | 0.2018 | 136.5 |
| OSBs | 0.5402 | 0.3712 | 0.4358 | 0.2018 | 116.0 |

- OSBs are the least performant
- SVM and Max Entropy performed slightly worse but are more precise

Train on Three – Test on One

Naval Postgraduate School
Department of Computer Science
Analyzing Microtext 8 August 2011

Trained with all, except SDPolice (18033 utterances)

| Features | Precision | Recall | F-score | Baseline F-score | Percent Improvement |
|----------|-----------|--------|---------|------------------|---------------------|
| Unigrams | 0.731 | 0.560 | 0.635 | 0.297 | 113.8 |
| Bigrams | 0.773 | 0.482 | 0.594 | 0.297 | 100.0 |
| GBGs | 0.737 | 0.596 | 0.659 | 0.297 | 121.9 |
| OSBs | 0.851 | 0.447 | 0.586 | 0.297 | 97.3 |

Trained with all, except Waco (12986 post)

| Features | Precision | Recall | F-score | Baseline F-score | Percent Improvement |
|----------|-----------|--------|---------|------------------|---------------------|
| Unigrams | 0.270 | 0.444 | 0.335 | 0.119 | 181.5 |
| Bigrams | 0.242 | 0.530 | 0.332 | 0.119 | 179.0 |
| GBGs | 0.244 | 0.532 | 0.334 | 0.119 | 180.7 |
| OSBs | 0.258 | 0.481 | 0.336 | 0.119 | 182.4 |

Persuasion Detection is Very Hard

- Training set size is important especially with short utterances and low exemplar count
- Much better F-Scores vs. guessing persuasion

But Doable

Future Research

- Related Work and Persuasion Model
- Experimental Design
- Results
- **Future Research**
- Conclusion

Future Work

- More annotation of more transcripts in the NPS Persuasion Corpus
- Add contextual features
- Add part-of-speech and dialog act tagging
- Ensemble Learning Techniques (Boosting/Bagging)

Detecting Persuasion in Dialog Possible but Difficult

- Developed annotation scheme and code book
- Built annotated persuasion corpus
- Tested machine learning and feature extraction techniques

Questions?

Detecting Persuasion in Dialog Possible but Difficult

Want Our Corpus?

Email me: jdyoung@nps.edu

BACKUP

Naval Postgraduate School
Department of Computer Science
Analyzing Microtext 8 August 2011

BACKUP SLIDES

- Look for instances where one party tries to make the other feel indebted to them
 1. The favor done should not be part of a previously established agreement, otherwise reminding the person of the agreement would fall under consistency
 2. In negotiations, common examples include instances where the negotiator conveys to the hostage taker how hard he is working for him but needs something in return
 3. This is not a *commitment* (see below) of the form “If you do this, I’ll do that”
- Look for cases where the negotiator asks for a favor, which is then rejected, but then follows with a request for a smaller favor. In cases like this one, do not tag the original request with reciprocity, but only the subsequent smaller requests.

Commitment

- Look for any kind of deal making (“If you do this, I’ll do that...”)
- May have to be taken in context: (TSU 1993)

| | | |
|----|-------|---|
| 1 | JIM | As soon as we get these kids hooked up, I’m going to go back and talk to these commanders – |
| 2 | STEVE | All right. Okay. |
| 3 | JIM | –about that perimeter motion, okay? |
| 4 | STEVE | Right. |
| 5 | JIM | And, and some of these issues we’ve discussed. |
| 6 | STEVE | Right, okay. |
| 7 | JIM | I’ve got to round them up. There’s a bunch of them in here. |
| 8 | STEVE | Okay. |
| 9 | JIM | And – |
| 10 | STEVE | And call me back then? |

Steve wants a commitment from Jim to call him back after he speaks with his commanders. Utterance 10 should be tagged with *commitment*, even though there is no specific “if-then” phrase.

- Commitment can also be one party emphasizing an agreement has been made.

Consistency

- When one party (party A) makes a reference to a previous commitment by Party B in order to persuade Party B into some action or request.
- When one party (Party A) makes reference to a want or need by Party B in order to influence some kind of belief or action.

- Any utterance which implies explicitly or implicitly that time is a factor
 1. A hostage taker setting a deadline for demands to be met
 2. A negotiator claiming that a situation may get worse in the future unless the hostage taker acts now
- Any time “secret” information is used to influence another party’s decision. For example a negotiator says

let you all in on a little secret okay? These guards out here, they’re pushing me to get something done and I am trying to hold them back because I know you all are up to your word

during a prison riot situation in one of the Taylor transcriptions.

- Any kind of compliment in order to influence decisions. Most times, annotators agree that influential compliments are given from the negotiator to the hostage taker, however there can be cases where the hostage taker uses *liking* to influence the negotiator.
 1. Can be brown-nosing and insincere
 2. Words and phrases like “buddy” and “friend”
- Any reference to similar life experiences
- Any expression of affection towards one party; phrases like “I like you” and “it’s important to me that you make it out of this” are examples.

- Appealing or making reference to a higher authority or expert in order to persuade or influence one party's beliefs or actions
 1. A negotiator making reference to his boss's needs in order to influence the hostage taker
 2. A negotiator making reference to an authority figure in the hostage taker's life like a parent or older sibling
- Any request for action or belief to a hostage taker from an authority figure. If an utterance like "Please put the gun down and come outside" comes from the negotiator, then that utterance is not persuasive. However, if the same utterance were said by the hostage taker's mother, then the utterance would be persuasive, because the mother is an authority figure. This requires the annotators to understand the context.

Social Proof

- Any reference to what is normal or customary in situations (a social norm). The negotiator might make reference to what a judge would normally do in order to influence the hostage taker.
- Any appeal to what a group thinks the person should do. A negotiator might make reference to the hostage taker's friends or family, claiming they all think he or she should give up. In the following example from the Rogan transcript, the negotiator (PN60) is trying to convince the hostage taker (HT01) to give up.

| | | |
|---|------|---|
| 1 | PN60 | Suppose you got a healthy body and a healthy mind, right? |
| 2 | HT01 | [Laughs] I wouldn't bet on that. [Laughs] |
| 3 | PN60 | Well hell that's quite a bit just that one right there. What? Well I don't know what to tell you know, you got all '- [?] |
| 4 | HT01 | Huh? |
| 5 | PN60 | – your friends talkin' to ya and tryin' to give you advice and these people who know you and like you. |

In Utterance 5, the negotiator uses peer pressure by making reference to the hostage taker's friends and their opinion that he should give up.

- The “other” category is a catchall for any utterance that annotators view as persuasive but does not fit the above. Here are some examples agreed on by the annotators:
 1. An appeal to the hostage taker to think about their children
 2. An emphatic plea by the negotiator using words like “guarantee,” “absolutely,” etc., in order to gain trust
 3. Reasons why certain actions should be performed (justifications)

6-fold Cross Validation Results

Naval Postgraduate School
Department of Computer Science
Analyzing Microtext 8 August 2011

| Naive Bayes | | | | | |
|------------------------|---------------|---------------|---------------|------------------|---------------------|
| Features | Precision | Recall | F-score | Baseline F-score | Percent Improvement |
| Unigrams | 0.4996 | 0.4052 | 0.4450 | 0.2018 | 120.5 |
| Bigrams | 0.4572 | 0.4172 | 0.4334 | 0.2018 | 114.8 |
| Gappy | 0.5072 | 0.4572 | 0.4772 | 0.2018 | 136.5 |
| OSBs | 0.5402 | 0.3712 | 0.4358 | 0.2018 | 116.0 |
| Maximum Entropy | | | | | |
| Unigrams | 0.5430 | 0.3700 | 0.4376 | 0.2018 | 116.8 |
| Bigrams | 0.5950 | 0.3012 | 0.3960 | 0.2018 | 96.2 |
| Gappy | 0.5280 | 0.3126 | 0.3902 | 0.2018 | 93.4 |
| OSBs | 0.6042 | 0.2564 | 0.3562 | 0.2018 | 76.5 |
| Support Vector Machine | | | | | |
| Unigrams | 0.4968 | 0.3582 | 0.4134 | 0.2018 | 104.9 |
| Bigrams | 0.5188 | 0.3406 | 0.4080 | 0.2018 | 102.2 |
| Gappy | 0.5516 | 0.3142 | 0.3966 | 0.2018 | 96.5 |
| OSBs | 0.5498 | 0.2920 | 0.3770 | 0.2018 | 86.8 |