Human Language Technology Research Institute

Lightly Supervised Modeling of Argument Persuasiveness

Isaac Persing and **Vincent Ng**
Human Language Technology Research Institute
University of Texas at Dallas

Argumentation Mining

- Traditionally concerned with determining the argumentative structure of a text document
  - identifying its claims and premises and the relationships between them
- Recently expanded to tasks concerning the persuasiveness of arguments
  - **Focus**: how persuasive is your argument?

Example Argument [http://idebate.org]

Motion

This House would ban teachers from interacting with students via social networking websites.

Assertion

Acting as a warning signal for children at risk.

Justification

If a child is aware that private electronic contact between teachers and students is prohibited by law, the child will know the teacher is doing something he is not supposed to if he initiates private electronic contact.

Example Argument [http://idebate.org]

Motion: expresses a stance on the debate's topic

This House would ban teachers from interacting with students via social networking websites.

Assertion

Acting as a warning signal for children at risk.

Justification

If a child is aware that private electronic contact between teachers and students is prohibited by law, the child will know the teacher is doing something he is not supposed to if he initiates private electronic contact.

Example Argument [http://idebate.org]

Motion

This House would ban teachers from interacting with students via social networking websites.

Assertion

Acting as a warning signal for children at risk.

Justification

If a child is aware that private electronic contact between teachers and students is prohibited by law, the child will know the teacher is doing something he is not supposed to if he initiates private electronic contact.

Example Argument [http://idebate.org]

Motion

This House would ban teachers from interacting with students via social networking websites.

Assertion: expresses why author agrees or disagrees with motion

Acting as a warning signal for children at risk.

Justification

If a child is aware that private electronic contact between teachers and students is prohibited by law, the child will know the teacher is doing something he is not supposed to if he initiates private electronic contact.

Example Argument [http://idebate.org]

Motion

This House would ban teachers from interacting with students via social networking websites.

Assertion

Acting as a warning signal for children at risk.

Justification: explains why author believes her assertion

If a child is aware that private electronic contact between teachers and students is prohibited by law, the child will know the teacher is doing something he is not supposed to if he initiates private electronic contact.

Example Argument [http://idebate.org]

Motion

This House would ban teachers from interacting with students via social networking websites.

Assertion

Acting as a warning signal for children at risk.

Justification: explains why author believes her assertion

If a child is aware that private electronic contact between teachers and students is prohibited by law, the child will know the teacher is doing something he is not supposed to if he initiates private electronic contact.

Humans can easily determine this argument is not very persuasive

Scoring Argument Persuasiveness

- Researchers have begun work on automatically scoring an argument's persuasiveness (low score → not persuasive)

- Why bother?
  - could help author understand how persuasive her argument is
    - in persuasive student essays
    - in online debates

Scoring Argument Persuasiveness

- Typical approach: supervised, feature-rich
  - works when labeled training data is abundant
  - Unfortunately, hand-labeling arguments with persuasiveness scores is time-consuming and labor-intensive

- **Our goal**
  - lightly-supervised approach to persuasiveness scoring
    - Significantly reduce reliance on labeled training data

Plan for the Talk

- Corpus and annotation
- Lightly-supervised approach
- Evaluation

Plan for the Talk

- Corpus and annotation
- Lightly-supervised approach
- Evaluation

Corpus and Annotation

- **Corpus**
  - debates from International Debate Education Association website
    - cover a wide range of topics (politics, economics, science, …)
  - 1208 arguments randomly selected from 165 debates

- **Annotation**
  - two native English speakers annotated each argument with its persuasiveness score

Rubric for Scoring Persuasiveness

6: a very persuasive argument

5: a persuasive, or only pretty clear argument

4: a decent, or only fairly clear argument

3: a poor, or only most understandable argument

2: a very unpersuasive or very unclear argument

1: an unclear or missing argument

Distribution over Persuasiveness Scores

1	2	3	4	5	6
3%	12%	20%	21%	20%	24%

Distribution over Persuasiveness Scores

1	2	3	4	5	6
3%	12%	20%	21%	20%	24%

Plan for the Talk

- Corpus and annotation
- Lightly-supervised approach
- Evaluation

Lightly-Supervised Approach

- **Question**: How can we design an approach that can reduce reliance on labeled data?

- **Idea**: use a small number of features (only 5)
  - Each feature encodes a type of error that negatively impacts an argument's persuasiveness
    - more errors → lower persuasiveness score

Five Errors Negatively Impacting Persuasiveness

- motivated by theoretical work on argument persuasiveness

Five Errors Negatively Impacting Persuasiveness

- motivated by theoretical work on argument persuasiveness
- **Grammar Error (GE)**
  - **Motivation**: grammar errors can interrupt the flow of discourse in an argument and reduce its coherence
  - 1 if argument is hard to understand because of grammar errors
  - 0 otherwise

Five Errors Negatively Impacting Persuasiveness

- motivated by theoretical work on argument persuasiveness
- **Grammar Error (GE)**
- **Lack of Objectivity (LO)**
  - **Motivation**: An argument is less persuasive if an author flatly states her personal opinions as evidence for her claim
  - 1 if it displays an inappropriate lack of objectivity
  - 0 otherwise

Five Errors Negatively Impacting Persuasiveness

- motivated by theoretical work on argument persuasiveness
- **Grammar Error (GE)**
- **Lack of Objectivity (LO)**
- **Inadequate Support (IS)**
  - **Motivation**: arguments with more support tend to be more persuasive
  - 2 if support is missing
  - 1 if support is inadequate
  - 0 if support is adequate

Five Errors Negatively Impacting Persuasiveness

- motivated by theoretical work on argument persuasiveness
- **Grammar Error (GE)**
- **Lack of Objectivity (LO)**
- **Inadequate Support (IS)**
  - **Motivation**: arguments with more support tend to be more persuasive
  - 2 if support is missing
  - 1 if support is inadequate
  - 0 if support is adequate

3 severity levels:
The larger the number, the more severe the error is

Five Errors Negatively Impacting Persuasiveness

- motivated by theoretical work on argument persuasiveness
- **Grammar Error (GE)**
- **Lack of Objectivity (LO)**
- **Inadequate Support (IS)**
- **Unclear Assertion (UA)**
  - **Motivation**: failure to clearly state the assertion makes an argument less persuasive
  - 2 if assertion is incomprehensible w/o reading the justification
  - 1 if unclear how assertion is related to motion w/o justification
  - 0 if assertion is clear

Five Errors Negatively Impacting Persuasiveness

- motivated by theoretical work on argument persuasiveness
- **Grammar Error (GE)**
- **Lack of Objectivity (LO)**
- **Inadequate Support (IS)**
- **Unclear Assertion (UA)**
  - **Motivation**: failure to clearly state the assertion makes an argument less persuasive
  - 2 if assertion is incomprehensible w/o reading the justification
  - 1 if unclear how assertion is related to motion w/o justification
  - 0 if assertion is clear

3 severity levels

Five Errors Negatively Impacting Persuasiveness

- motivated by theoretical work on argument persuasiveness
- **Grammar Error (GE)**
- **Lack of Objectivity (LO)**
- **Inadequate Support (IS)**
- **Unclear Assertion (UA)**
- **Unclear Justification (UJ)**
  - **Motivation**: failure to state an argument's justification for its assertion will make it less persuasive
  - 2 if justification appears unrelated to assertion
  - 1 if justification does not concisely justify the assertion
  - 0 if justification is clear

Five Errors Negatively Impacting Persuasiveness

- motivated by theoretical work on argument persuasiveness
- **Grammar Error (GE)**
- **Lack of Objectivity (LO)**
- **Inadequate Support (IS)**
- **Unclear Assertion (UA)**
- **Unclear Justification (UJ)**
  - **Motivation**: failure to state an argument's justification for its assertion will make it less persuasive
  - 2 if justification appears unrelated to assertion
  - 1 if justification does not concisely justify the assertion
  - 0 if justification is clear

3 severity levels

How to compute each error?

- **Bootstrapping**
  - **Step 1**: for each error, design **heuristics** that can reliably label (a small number of) arguments with error severity values
    - E.g., for Inadequate Support, label an argument with one of its possible values (0, 1, or 2)
  - **Step 2**: Label the remaining arguments by bootstrapping from the seed arguments using EM
    - M-step: estimate the parameters of the generative model
    - E-step: (re)label each argument with the error probabilistically

Generative Model: Naïve Bayes

- 10 features

10 Features

- # grammar errors per sentence in justification
    - Useful for predicting grammar errors

10 Features

- # grammar errors per sentence in justification
- # subjectivity indicators (“morally”, “certain”) in justification
  - Arguments too concerned with the author’s morality or in which the author seems too certain of herself show a lack of objectivity

10 Features

- # grammar errors per sentence in justification
- # subjectivity indicators in justification
- # definite articles in justification
  - An argument with few definite articles is usually less specific and may also be too subjective

10 Features

- # grammar errors per sentence in justification
- # subjectivity indicators in justification
- # definite articles in justification
- # 1st person plural pronouns in justification
  - Justifications that lack objectivity often rely on stories about the author's personal experiences

10 Features

- # grammar errors per sentence in justification
- # subjectivity indicators in justification
- # definite articles in justification
- # 1st person plural pronouns in justification
- # references cited in justification
  - More citations tend to imply more support for claims

10 Features

- # grammar errors per sentence in justification
- # subjectivity indicators in justification
- # definite articles in justification
- # 1st person plural pronouns in justification
- # references cited in justification
- Assertion length
    - short assertions could be unclear
- Justification length
    - short justifications could be unclear

-

10 Features

- # grammar errors per sentence in justification
- # subjectivity indicators in justification
- # definite articles in justification
- # 1st person plural pronouns in justification
- # citations in justification
- Assertion length
- Justification length
- # content lemmas in both assertion and justification
- # subject matches in contingency-cause discourse relation

So far…

- We have labeled each argument with the severity value of each of the five errors
- **Next**: Use these error severity values for scoring argument persuasiveness

Argument Persuasiveness Scoring

- **Assumption**
  - more errors → lower persuasiveness score

Argument Persuasiveness Scoring

- **Assumption**
  - more errors → lower persuasiveness score
  - the persuasiveness score of an argument inversely correlates with the **sum of its five errors' severity values**
    - e.g.,

Argument:
GE=0
LO=0
IS=2
UA=2
UJ=0

→ Sum of severity values = 4

Argument Persuasiveness Scoring

- **Training a Persuasiveness Predictor**
  - Cluster the training arguments by the sum of severity values
  - For each cluster, randomly select *n* arguments and manually label each one with its persuasiveness score
  - Assign to each cluster the average of the *n* scores

Argument Persuasiveness Scoring

- **Training a Persuasiveness Predictor**
  - Cluster the training arguments by the sum of severity values
  - For each cluster, randomly select *n* arguments and manually label each one with its persuasiveness score
  - Assign to each cluster the average of the *n* scores
- **Testing:** For each test argument,
  - compute its sum of severity values
  - assign it to the corresponding cluster
  - predict its persuasiveness score as the cluster's score

Plan for the Talk

- Corpus and annotation
- Lightly-supervised approach
- Evaluation

Evaluation: Goal

- Evaluate ASE, our lightly-supervised approach

Three Scoring Metrics

- **E (Zero-one Loss):**
  - frequency at which a system predicts the wrong score
- **ME (Mean Error):**
  - mean distance between the predicted score and the gold score
- **PC (Pearson's Correlation Coefficient):**
  - Pearson's correlation between the predicted and gold scores

Three Scoring Metrics

- **E (Zero-one Loss):**
  - frequency at which a system predicts the wrong score
- **ME (Mean Error):**
  - mean distance between the predicted score and the gold score
- **PC (Pearson's Correlation Coefficient):**
  - Pearson's correlation between the predicted and gold scores

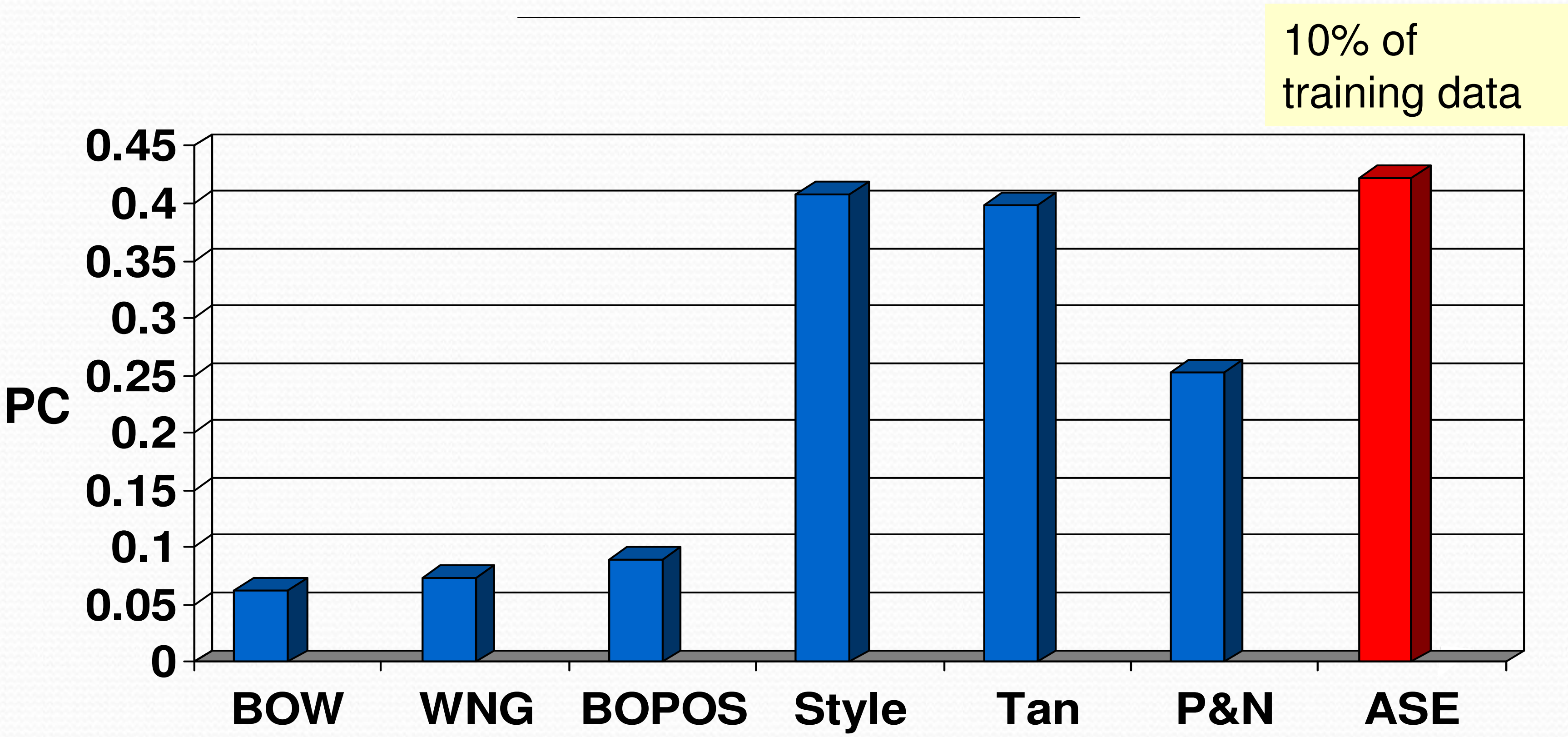
Six Baseline Systems

- Linear SVM regressors trained on different feature sets
- **Bag of words (BOW)**
- **Word n-grams (WNG)**
  - unigrams, bigrams, trigrams
- **Bag of part-of-speech tags (BOPOS)**
- **Style**
  - length, word categories, word complexity, word scores
- **Duplicated Tan et al. (2016)**
  - features for predicting success of persuasion
- **Persing and Ng (2015)**
  - features developed for scoring essay persuasiveness

Evaluation: Setup

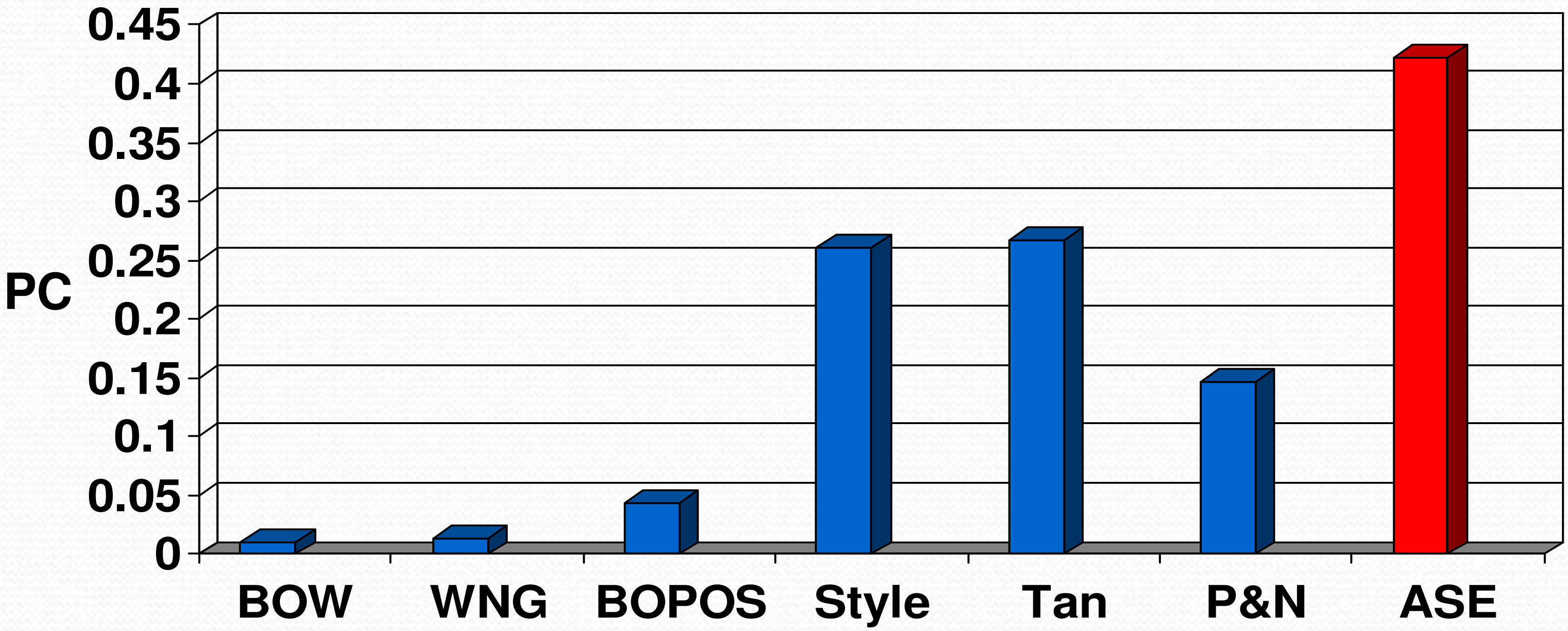
- 5-fold cross validation

Results



10% of training data
PC
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
BOW
WNG
BOPOS
Style
Tan
P&N
ASE

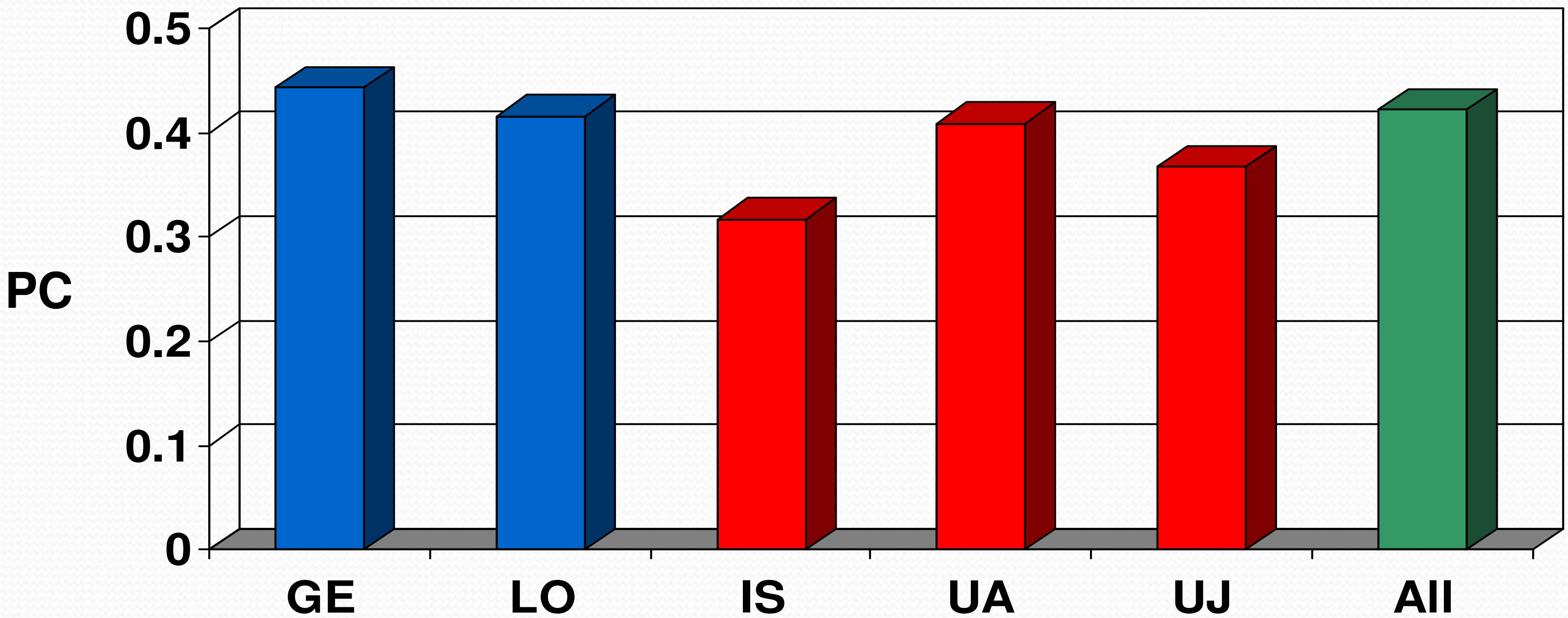
Results with Lightly-Supervised Baselines



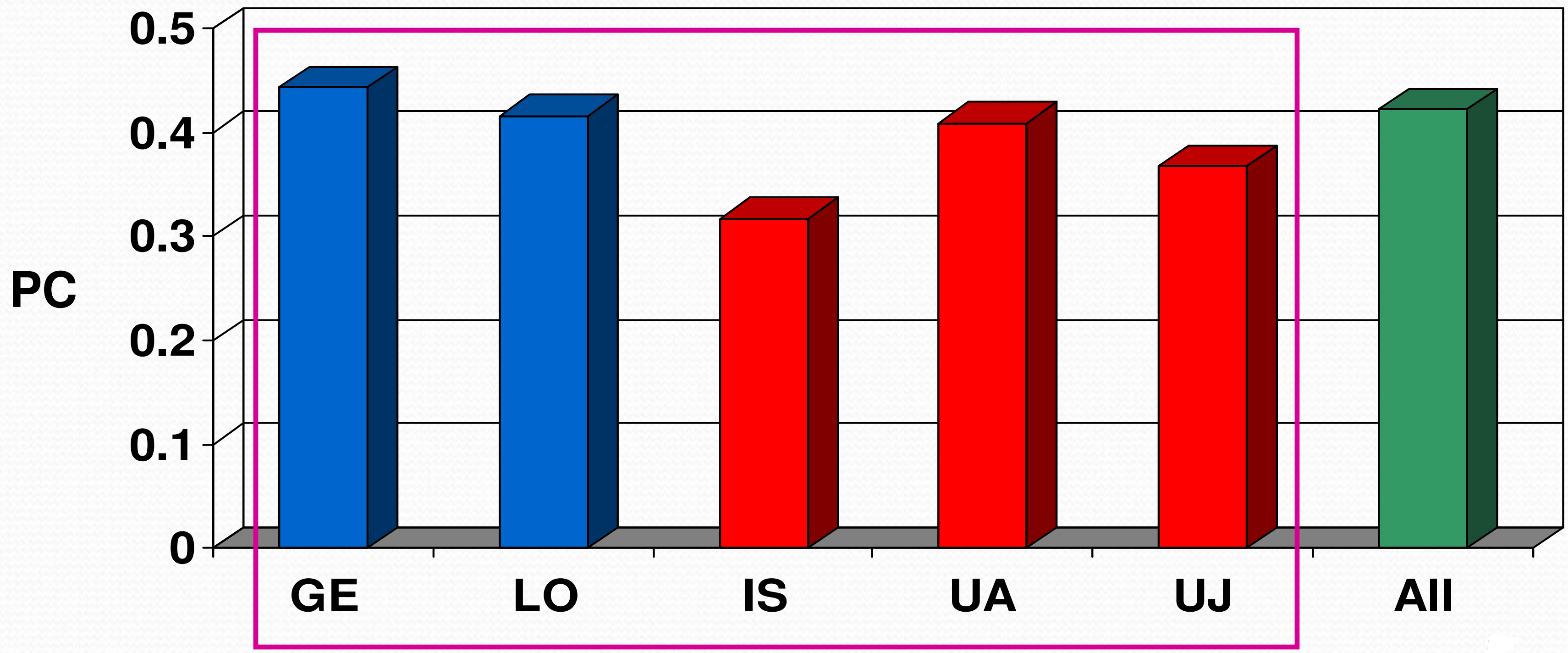
Error Ablation

- Recall that ASE scores persuasiveness by summing the five errors' severity values
- Ablate each of the five errors when scoring persuasiveness

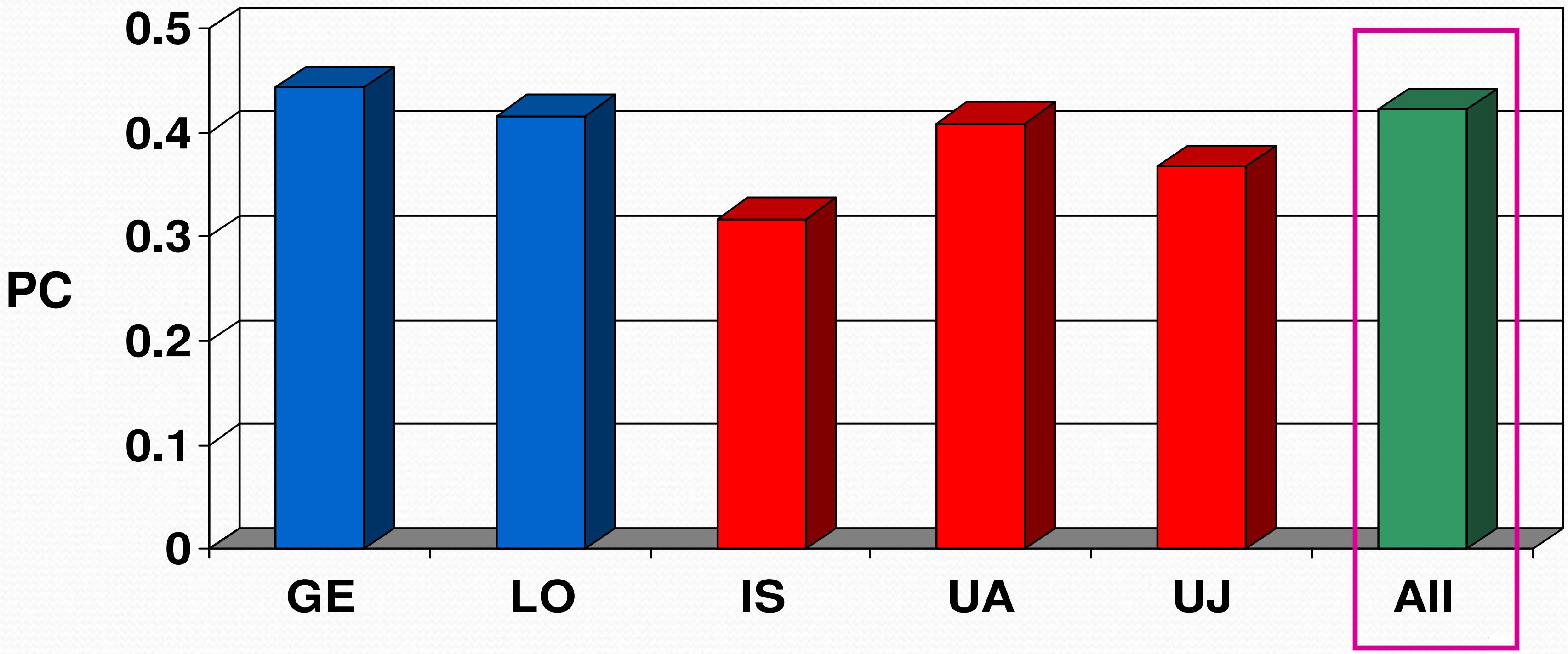
Results: Error Ablation



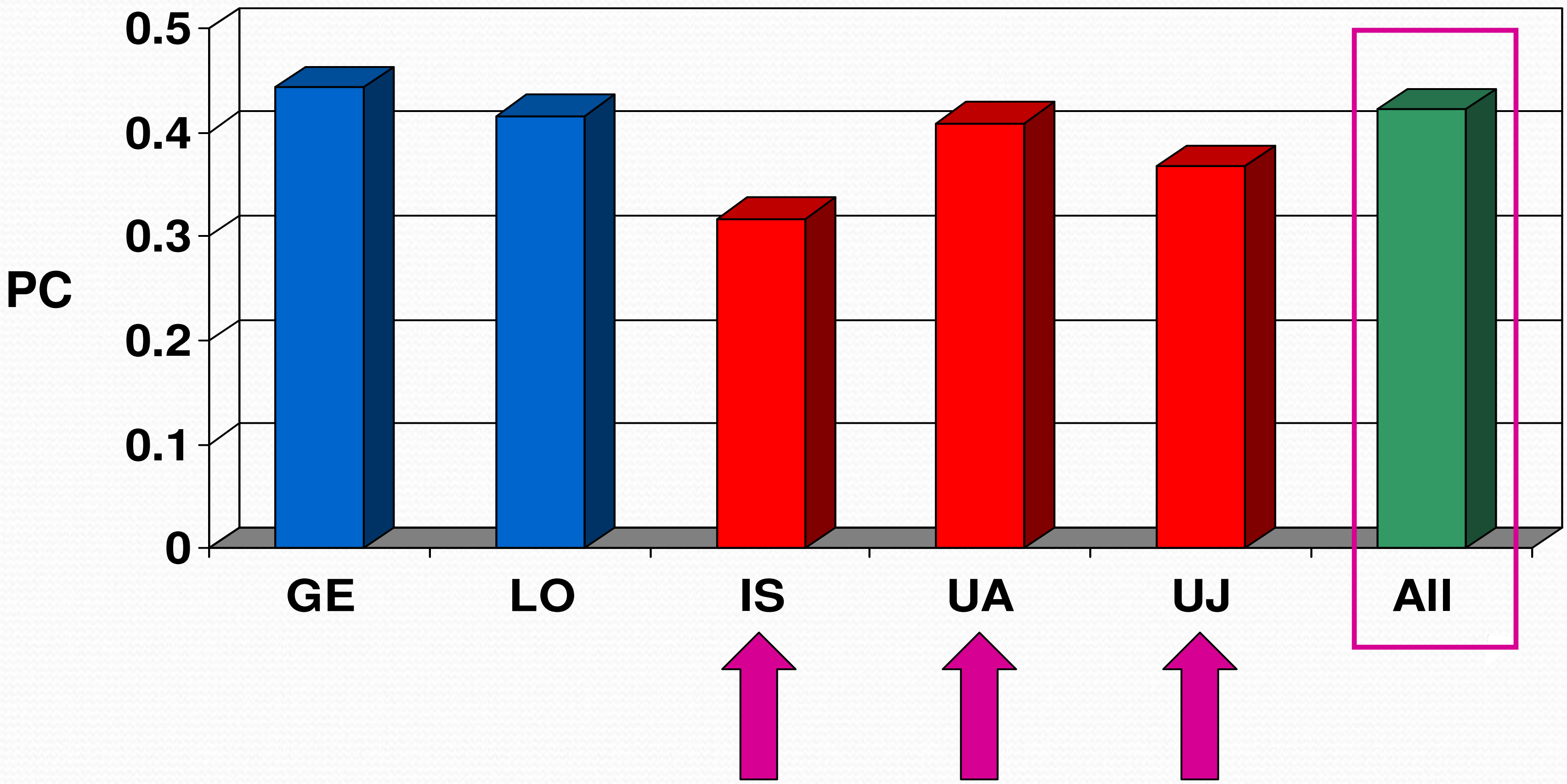
Results: Error Ablation



Results: Error Ablation



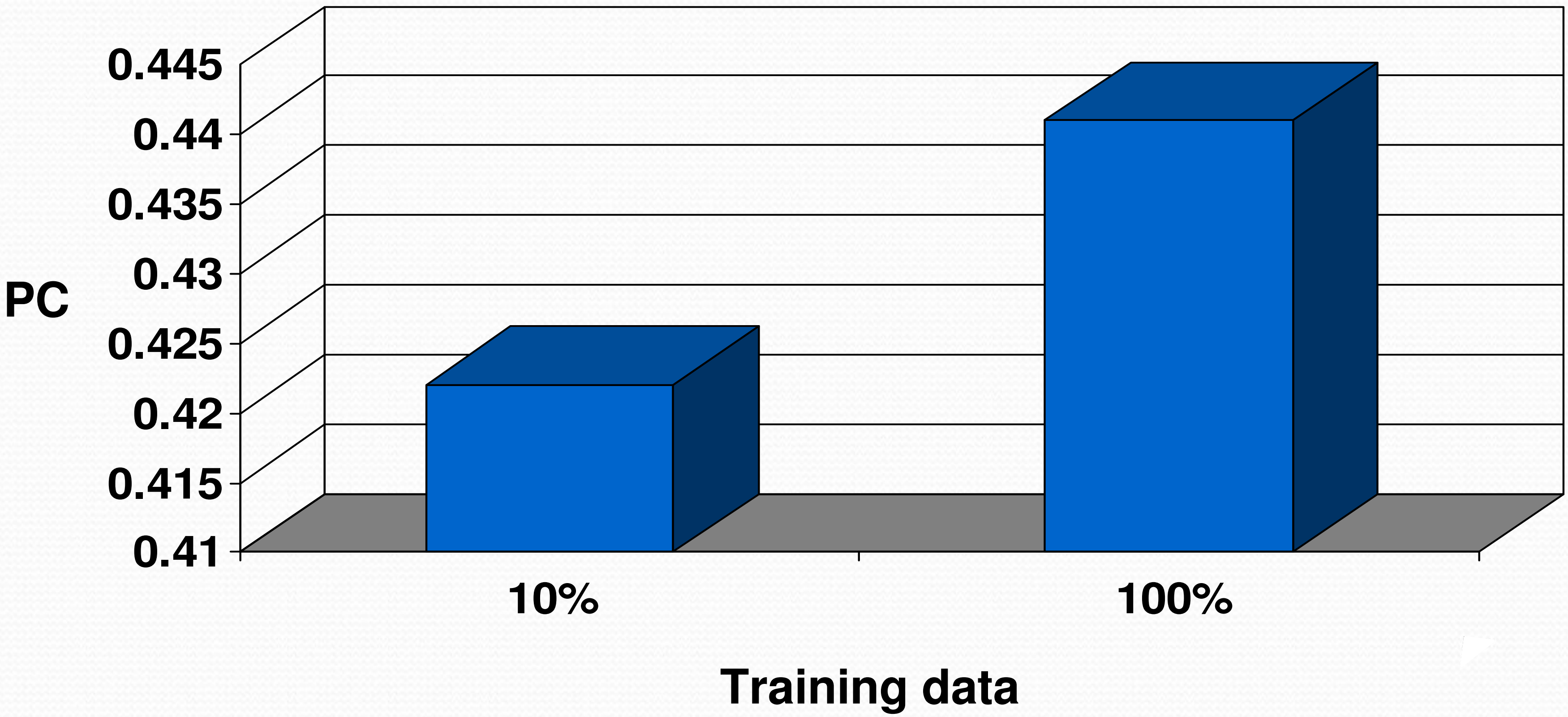
Results: Error Ablation



Lightly vs. Fully-Supervised ASE

- Train ASE with 100% of the training data

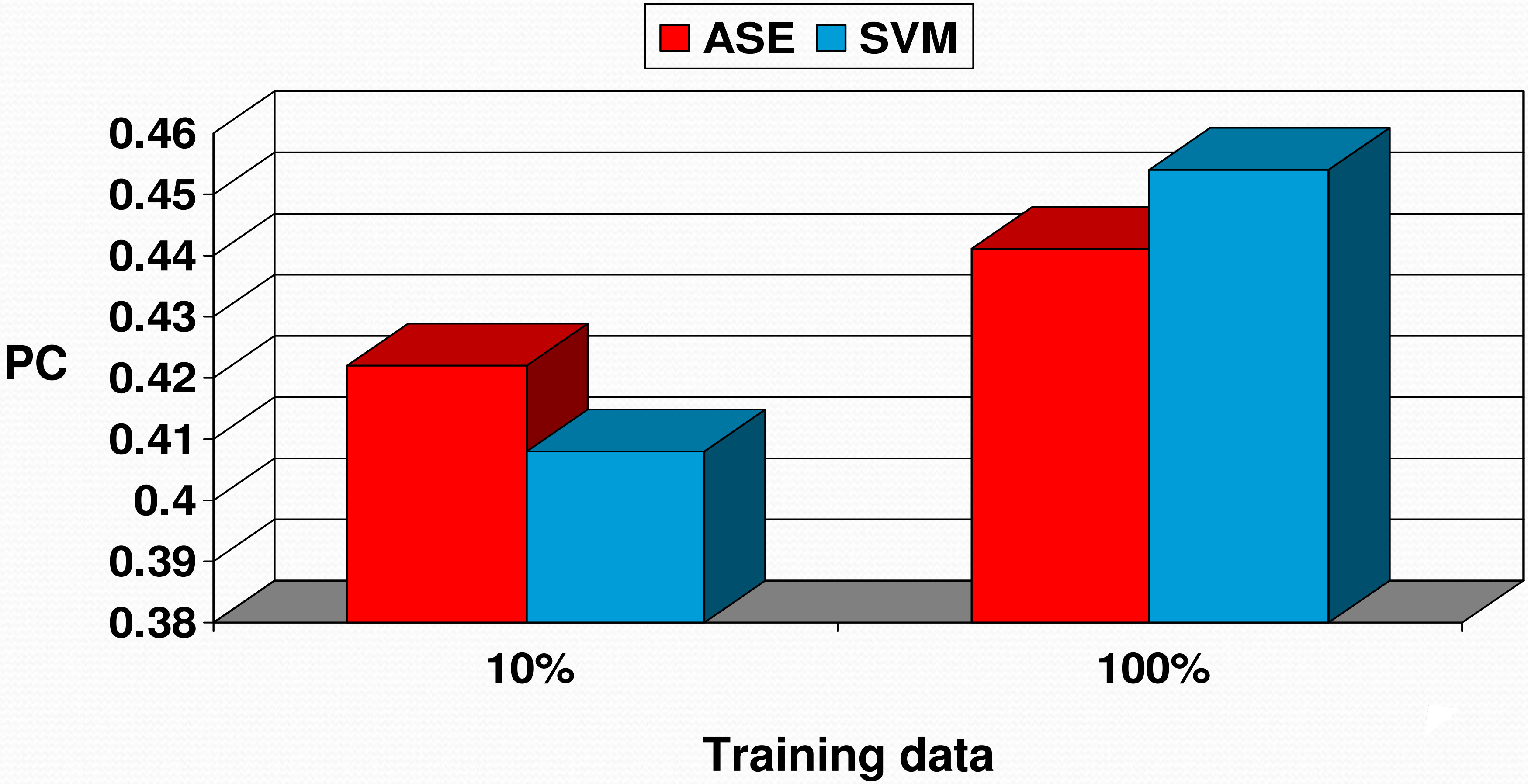
Results: Lightly vs. Fully Supervised ASE



Is ASE's persuasiveness scoring method too simplistic?

- What if we train an **SVM regressor** using the **five errors as features**?

Results: ASE vs. SVM for Scoring

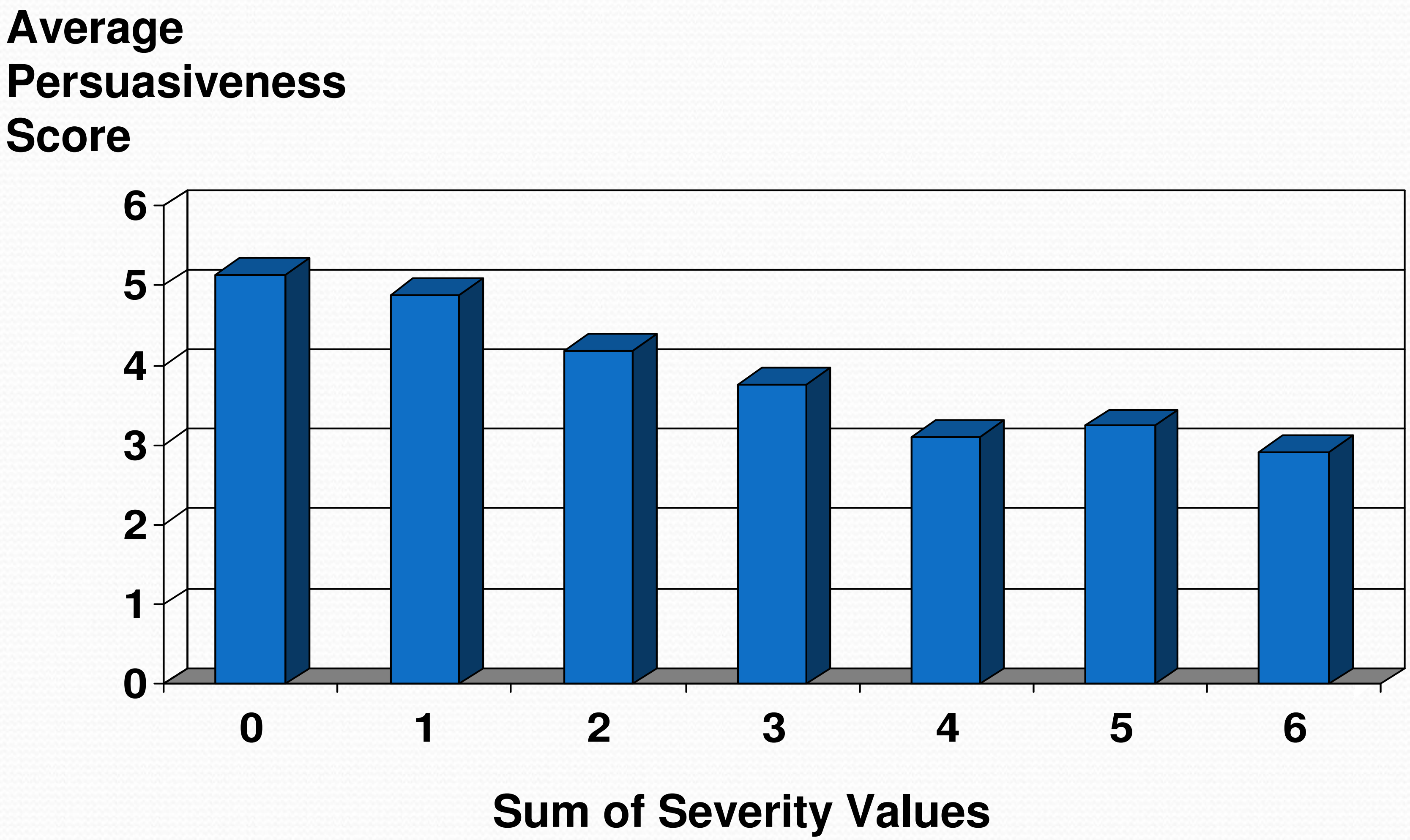


Is ASE's assumption correct?

- ASE assumes that an argument with more errors is less persuasive

- How can we validate this assumption?
  1. cluster arguments by the sum of severity values
  2. average the gold persuasiveness scores of the arguments in each cluster

Results: Clustering



Summary

- Proposed a lightly-supervised approach to persuasiveness scoring that outperformed competing baselines
- Made our annotated corpus publicly available