Human Language Technology Research Institute

Stance Classification in Ideological Debates: Data, Models, Features, and Constraints

Kazi Saidul Hasan and Vincent Ng
Human Language Technology Research Institute
University of Texas at Dallas

Debate Stance Classification

- Given a debate post written for a two-sided debate topic, determine the author's stance (*For* or *Against*)

Debate Stance Classification

- Given a debate post written for a two-sided debate topic, determine the author's stance (*For* or *Against*)

Should abortion be allowed?

Debate Stance Classification

- Given a debate post written for a two-sided debate topic, determine the author's stance (*For* or *Against*)

Should abortion be allowed?

Women should have the ability to choose what they do with their bodies.

Debate Stance Classification

- Given a debate post written for a two-sided debate topic, determine the author's stance (*For* or *Against*)

Should abortion be allowed?

Women should have the ability to choose what they do with their bodies.

For

Debate Stance Classification

- Given a debate post written for a two-sided debate topic, determine the author's stance (*For* or *Against*)

Should abortion be allowed?

Women should have the ability to choose what they do with their bodies.	Technically abortion is murder. They are killing the baby without a justified motive.

For

Debate Stance Classification

- Given a debate post written for a two-sided debate topic, determine the author's stance (*For* or *Against*)

Should abortion be allowed?

Women should have the ability to choose what they do with their bodies.	Technically abortion is murder. They are killing the baby without a justified motive.
For	**Against**

Related Work

- Considered three debate settings
  - US congressional floor debates
    - Thomas et al. (2006), Bansal et al. (2008), Burfoot et al. (2011)..
  - Company-internal debates
    - Murakami and Raymond (2010)
  - Ideological debates
    - Somasundaran and Wiebe (2010), Anand et al. (2011), …

Related Work

- Considered three debate settings
  - US congressional floor debates
    - Thomas et al. (2006), Bansal et al. (2008), Burfoot et al. (2011)..
  - Company-internal debates
    - Murakami and Raymond (2010)
  - Ideological debates
    - Somasundaran and Wiebe (2010), Anand et al. (2011), …
- **More challenging than the other settings**
  - Use of colorful and emotional languages
    - sarcasm, insults, questioning other people's assumptions, …

Goal

- Gain a better understanding of how to improve learning approaches to stance classification in ideological debate

Goal

- Gain a better understanding of how to improve learning approaches to stance classification in ideological debate
  - Examine how the performance of a learning-based stance classification system varies with
    - the **complexity** of the **learning model**

Goal

- Gain a better understanding of how to improve learning approaches to stance classification in ideological debate
  - Examine how the performance of a learning-based stance classification system varies with
    - the **complexity** of the **learning model**
    - the **richness** of the **feature set**

Goal

- Gain a better understanding of how to improve learning approaches to stance classification in ideological debate
  - Examine how the performance of a learning-based stance classification system varies with
    - the **complexity** of the **learning model**
    - the **richness** of the **feature set**
    - **training data** **size** and **quality**

Goal

- Gain a better understanding of how to improve learning approaches to stance classification in ideological debate
  - Examine how the performance of a learning-based stance classification system varies with
    - the **complexity** of the **learning model**
    - the **richness** of the **feature set**
    - **training data** **size** and **quality**

amount

Use automatically labeled data as additional training data

Goal

- Gain a better understanding of how to improve learning approaches to stance classification in ideological debate
  - Examine how the performance of a learning-based stance classification system varies with
    - the **complexity** of the **learning model**
    - the **richness** of the **feature set**
    - **training data size** and **quality**
    - the **application** of extra-linguistic **constraints**

Ensure a stance classifier's outputs are consistent

Plan for the Talk

- Datasets
- Experimental setup for examining how classification performance varies with
  - the complexity of the learning model
  - the richness of the feature set
  - the amount and quality of training data
  - the application of extra-linguistic constraints
- Evaluation

Datasets

- 4 datasets
  - collected from http://www.createdebate.com
  - contain debate posts collected from four debate topics

Topic	Posts	"for" %	Average Sequence Length
Support **Abortion**?	1741	54.9	4.1
Support **Gay Rights**?	1376	63.4	4.0
Support **Obama**?	985	53.9	2.6
Legalize **Marijuana**?	626	69.5	2.5

Datasets

- 4 datasets
  - collected from http://www.createdebate.com
  - contain debate posts collected from four debate topics

Topic	Posts	"for" %	Average Sequence Length
Support **Abortion**?	1741	54.9	4.1
Support **Gay Rights**?	1376	63.4	4.0
Support **Obama**?	985	53.9	2.6
Legalize **Marijuana**?	626	69.5	2.5

Datasets

- 4 datasets
  - collected from http://www.createdebate.com
  - contain debate posts collected from four debate topics

Topic	Posts	"for" %	Average Sequence Length
Support **Abortion**?	1741	54.9	4.1
Support **Gay Rights**?	1376	63.4	4.0
Support **Obama**?	985	53.9	2.6
Legalize **Marijuana**?	626	69.5	2.5

Datasets

- 4 datasets
  - collected from http://www.createdebate.com
  - contain debate posts collected from four debate topics

Topic	Posts	“for” %	Average Sequence Length
Support **Abortion**?	1741	54.9	4.1
Support **Gay Rights**?	1376	63.4	4.0
Support **Obama**?	985	53.9	2.6
Legalize **Marijuana**?	626	69.5	2.5

Plan for the Talk

- Datasets
- Experimental setup for examining how classification performance varies with
  - the complexity of the learning model
  - the richness of the feature set
  - the amount and quality of training data
  - the application of extra-linguistic constraints
- Evaluation

Models

- **Goal**
  - Examine how performance varies with model complexity
    - Examine three types of stance classification models

1st Type: Classification Models

- Binary classifier that assigns a stance label (for/against) to each debate post independently of other posts
  - Each training instance corresponds to a debate post

1st Type: Classification Models

- Binary classifier that assigns a stance label (for/against) to each debate post independently of other posts
  - Each training instance corresponds to a debate post

- To train the binary classifier, we employ
  - a generative model: Naïve Bayes
  - a discriminative model: SVMs

- Can determine which type of models is better for this task

2nd Type: Sequence Models

- Sequence models assume as input a post sequence and output a stance sequence

2^{nd} Type: Sequence Models

- Sequence models assume as input a post sequence and output a stance sequence

- **Motivation**
  - Since a post in a post sequence is a reply to its parent post, its label should be determined in dependent relation to its parent's

2^{nd} Type: Sequence Models

- Sequence models assume as input a post sequence and output a stance sequence

- **Motivation**
  - Since a post in a post sequence is a reply to its parent post, its label should be determined in dependent relation to its parent's

- To train sequence models, we employ
  - a generative model: HMM
  - a discriminative model: CRF

3rd Type: Fine-Grained Models

- Jointly determine the stance label of a debate post and the stance label of each of its sentences

3rd Type: Fine-Grained Models

- Jointly determine the stance label of a debate post and the stance label of each of its sentences

- Why fine-grained models?

3rd Type: Fine-Grained Models

- Jointly determine the stance label of a debate post and the stance label of each of its sentences
- Why fine-grained models?
  - Modeling sentence stances could improve document stance prediction

3rd Type: Fine-Grained Models

- Jointly determine the stance label of a debate post and the stance label of each of its sentences

- Why fine-grained models?
  - Modeling sentence stances could improve document stance prediction
    - Features computed from sentences with a neutral stance should not play any role in determining the document stance

3rd Type: Fine-Grained Models

- Jointly determine the stance label of a debate post and the stance label of each of its sentences
- Why fine-grained models?
  - Modeling sentence stances could improve document stance prediction
    - Features computed from sentences with a neutral stance should not play any role in determining the document stance
- Focus on implementing fine-grained generative models based on NB and HMMs

Generative Story

- To generate a debate post
  - generate its document stance c with P(c)
  - for each sentence in the post
    - generate its sentence stance s with P(s|c)
    - generate each feature f representing the sentence with P(f|s,c)

Generative Story

- To generate a debate post
  - generate its document stance c with P(c)
  - for each sentence in the post
    - generate its sentence stance s with P(s|c)
    - generate each feature f representing the sentence with P(f|s,c)

- Document stance can have one of 2 values: for, against
- Sentence stance can have one of 3 values: for, against, neutral

Generative Story

- To generate a debate post
  - generate its document stance c with P(c)
  - for each sentence in the post
    - generate its sentence stance s with P(s|c)
    - generate each feature f representing the sentence with P(f|s,c)

- Fine-grained NB and fine-grained HMM employ this same story
  - Differ in terms of whether doc stance is generated independently (NB) or in dependent relation to that of the preceding post (HMM)

Generative Story

- To generate a debate post
  - generate its document stance c with P(c)
  - for each sentence in the post
    - generate its sentence stance s with P(s|c)
    - generate each feature f representing the sentence with P(f|s,c)

- **Modeling assumption**
  - P(f | s=neutral, c=for) = P(f | s=neutral, c=against)

Generative Story

- To generate a debate post
  - generate its document stance c with P(c)
  - for each sentence in the post
    - generate its sentence stance s with P(s|c)
    - generate each feature f representing the sentence with P(f|s,c)

- **Modeling assumption**
  - P(f | s=neutral, c=for) = P(f | s=neutral, c=against)
  - If a sentence's stance is neutral, the document's stance has no impact on how likely its feature values are generated

Generative Story

- To generate a debate post
  - generate its document stance c with P(c)
  - for each sentence in the post
    - generate its sentence stance s with P(s|c)
    - generate each feature f representing the sentence with P(f|s,c)

- **Modeling assumption**
  - P(f | s=neutral, c=for) = P(f | s=neutral, c=against)
  - If a sentence's stance is neutral, the document's stance has no impact on how likely its feature values are generated
  - Neutral sentences have no impact on determining doc stance

Generative Story

- To generate a debate post
  - generate its document stance c with $P(c)$
  - for each sentence in the post
    - generate its sentence stance s with $P(s|c)$
    - generate each feature f representing the sentence with $P(f|s,c)$

Generative Story

- To generate a debate post
  - generate its document stance c with P(c)
  - for each sentence in the post
    - generate its sentence stance s with P(s|c)
    - generate each feature f representing the sentence with P(f|s,c)

- The training debate posts are labeled with document stances but not sentence stances

Generative Story

can be estimated in a supervised manner

- To generate a debate post
  - generate its document stance c with $P(c)$
  - for each sentence in the post
    - generate its sentence stance s with $P(s|c)$
    - generate each feature f representing the sentence with $P(f|s,c)$

- The training debate posts are labeled with document stances but not sentence stances

Generative Story

- To generate a debate post
  - generate its document stance c with P(c)
  - for each sentence in the post
    - generate its sentence stance s with P(s|c)
    - generate each feature f representing the sentence with P(f|s,c)

can be estimated in a supervised manner

cannot be estimated in a supervised manner

- The training debate posts are labeled with document stances but not sentence stances

Generative Story

- To generate a debate post
  - generate its document stance c with P(c)
  - for each sentence in the post
    - generate its sentence stance s with P(s|c)
    - generate each feature f representing the sentence with P(f|s,c)

can be estimated in a supervised manner

Treat s as a hidden variable, estimate with EM

- The training debate posts are labeled with document stances but not sentence stances

Plan for the Talk

- Datasets

- Experimental setup for examining how classification performance varies with
  - the complexity of the learning model
  - the richness of the feature set
  - the amount and quality of training data
  - the application of extra-linguistic constraints

- Evaluation

Features

- **Goal**
  - Examine how performance varies with the richness of the feature set
    - Examine three feature sets

Feature Set 1: N-grams

- Unigrams and bigrams collected from the training posts, encoded as binary features indicating their presence/absence

Feature Set 2: Anand et al.'s (2011) Features

- Anand et al. (2011)'s system
  - state-of-the-art system for stance classification in ideological debate

Feature Set 2: Anand et al.'s (2011) Features

- Anand et al. (2011)'s system
  - state-of-the-art system for stance classification in ideological debate
  - Employs 5 types of features
    - **N-grams**
      - Unigrams and bigrams
    - **First N-grams**
      - First unigram, first bigram, first trigram of a debate post
    - **Document statistics**
      - Post length, #words/sentence, % pronouns, % sentiment words,...
    - **Punctuations**
      - Repeated punctuation symbols in a post
    - **Dependency-based features**
      - Argument pairs as features and their generalized form

Feature Set 2: Anand et al.'s (2011) Features

- Anand et al. (2011)'s system
  - state-of-the-art system for stance classification in ideological debate
  - Employs 5 types of features
    - **N-grams**
      - Unigrams and bigrams
    - **First N-grams**
      - First unigram, first bigram, first trigram of a debate post
    - **Document statistics**
      - Post length, #words/sentence, % pronouns, % sentiment words,...
    - **Punctuations**
      - Repeated punctuation symbols in a post
    - **Dependency-based features**
      - Argument pairs as features and their generalized form

composed of statistical and syntactic features

Feature Set 3: Adding Frame-Semantic Features

Feature Set 3: Adding Frame-Semantic Features

- Produce a frame-semantic parse for each sentence in a debate post using SEMAFOR (Das et al., 2010)

Feature Set 3: Adding Frame-Semantic Features

- Produce a frame-semantic parse for each sentence in a debate post using SEMAFOR (Das et al., 2010)
  - Each parse consists of a set of frames and their frame elements
    - **frame**: describes an event mentioned in a sentence
    - **frame element**: person/object participating in the event

Feature Set 3: Adding Frame-Semantic Features

- Produce a frame-semantic parse for each sentence in a debate post using SEMAFOR (Das et al., 2010)
  - Each parse consists of a set of frames and their frame elements
    - **frame**: describes an event mentioned in a sentence
    - **frame element**: person/object participating in the event

- Extract 3 types of features from a frame-semantic parse

Sample Frame-Semantic Parse

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Sample Frame-Semantic Parse

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Sample Frame-Semantic Parse

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Sample Frame-Semantic Parse

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Sample Frame-Semantic Parse

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Sample Frame-Semantic Parse

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Given a parse, extract 3 types of frame-semantic features

Frame-word Interaction Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Frame-word Interaction Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Frame-word Interaction Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Frame-word Interaction Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Frame-word Interaction Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: "woman"
Possession	**Target**: "has"
	Owner: "Every woman"
	Possession: "the right to choose abortion"
Correctness	**Target**: "right"
Choosing	**Target**: "choose"
	Chosen: "abortion"

Possession:every:the

Frame-word Interaction Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Possession:every:the

Frame-word Interaction Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: "woman"
Possession	**Target**: "has"
	Owner: "Every woman"
	Possession: "the right to choose abortion"
Correctness	**Target**: "right"
Choosing	**Target**: "choose"
	Chosen: "abortion"

Possession:every:the
Possession:every:right

Frame-word Interaction Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Possession:every:the
Possession:every:right

Frame-word Interaction Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Possession:every:the
Possession:every:right
Posession:woman:to
…

Frame-Pair Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
	Target: “has”
Possession	**Owner**: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Frame-Pair Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Frame-Pair Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Frame-Pair Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Frame-Pair Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Frame-Pair Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: "woman"
Possession	**Target**: "has"
	Owner: "Every woman"
	Possession: "the right to choose abortion"
Correctness	**Target**: "right"
Choosing	**Target**: "choose"
	Chosen: "abortion"

Possession:People

Frame N-gram Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Frame N-gram Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Frame N-gram Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Frame N-gram Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: "woman"
Possession	**Target**: "has"
	Owner: "Every woman"
	Possession: "the right to choose abortion"
Correctness	**Target**: "right"
Choosing	**Target**: "choose"
	Chosen: "abortion"

People:has

Frame N-gram Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Frame N-gram Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Owner:has

Frame N-gram Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: “woman”
Possession	**Target**: “has”
	Owner: “Every woman”
	Possession: “the right to choose abortion”
Correctness	**Target**: “right”
Choosing	**Target**: “choose”
	Chosen: “abortion”

Frame N-gram Features

Every woman has the right to choose abortion

Frame	Target and frame elements
People	**Target**: "woman"
Possession	**Target**: "has"
	Owner: "Every woman"
	Possession: "the right to choose abortion"
Correctness	**Target**: "right"
Choosing	**Target**: "choose"
	Chosen: "abortion"

People:Possession

How to use frame-semantic features?

- Train two stance classifiers, C_a and C_fs
  - C_a: trained using only Anand et al.'s features
  - C_fs: trained using only the frame-semantic features

How to use frame-semantic features?

- Train two stance classifiers, C_a and C_fs
  - C_a: trained using only Anand et al.'s features
  - C_fs: trained using only the frame-semantic features
- To classify a test post,
  - Linearly combine the output of C_a and C_fs
  - Combination weight tuned to maximize performance on dev set

Plan for the Talk

- Datasets
- Experimental setup for examining how classification performance varies with
  - the complexity of the learning model
  - the richness of the feature set
  - the amount and quality of training data
  - the application of extra-linguistic constraints
- Evaluation

Data

- **Goal**
  - Examine how performance varies with the **amount** and **quality** of training data

Data

- **Goal**
  - Examine how performance varies with the **amount** and **quality** of training data

express all the results as learning curves

Data

- **Goal**
  - Examine how performance varies with the **amount** and **quality** of training data

Use automatically labeled data as **additional** training data

Data

- **Goal**
  - Examine how performance varies with the **amount** and **quality** of training data

Use automatically labeled data as **additional** training data

Determine whether noisily labeled data be used to improve stance classification performance

Data

- **Goal**
  - Examine how performance varies with the **amount** and **quality** of training data

Use automatically labeled data as **additional** training data

Why bother?

Data

- **Goal**
  - Examine how performance varies with the **amount** and **quality** of training data

Use automatically labeled data as **additional** training data

Why bother?

The number of stance-labeled debate posts that can be downloaded from online debate forums is fairly limited

Data

- **Goal**
  - Examine how performance varies with the **amount** and **quality** of training data

Use automatically labeled data as **additional** training data

Goal:
Identify documents where authors express viewpoints on the debate topics of interest and stance-label them heuristically

Data

- **Goal**
  - Examine how performance varies with the **amount** and **quality** of training data

Use automatically labeled data as **additional** training data

Not debate posts
Can be blog posts, news articles, etc.

Goal:
Identify documents where authors express viewpoints on the debate topics of interest and stance-label them heuristically

Data

- **Goal**
  - Examine how performance varies with the **amount** and **quality** of training data

Use automatically labeled data as **additional** training data

Not debate posts
Can be blog posts, news articles, etc.

Goal:
Identify documents where authors express viewpoints on the debate topics of interest and stance-label them heuristically

How to collect and heuristically stance-label such documents?

How to incorporate such documents into the training process?

How to collect documents noisily labeled with stance information?

How to collect documents noisily labeled with stance information?

- 2 steps
  1. Create using commonsense knowledge a list of phrases that are reliable indicators of both stances for each debate topic

How to collect documents noisily labeled with stance information?

- 2 steps
  1. Create using commonsense knowledge a list of phrases that are reliable indicators of both stances for each debate topic

Abortion	
For	**Against**
I think abortion should be legal.	I think abortion should not be legal.
I support abortion.	I do not support abortion.
I think abortion should be allowed.	I think abortion should not be allowed.

How to collect documents noisily labeled with stance information?

- 2 steps
  1. Create using commonsense knowledge a list of phrases that are reliable indicators of both stances for each debate topic

Abortion	
For	**Against**
I think abortion should be legal.	I think abortion should not be legal.
I support abortion.	I do not support abortion.
I think abortion should be allowed.	I think abortion should not be allowed.

  2. Use each phrase as an exact search query to retrieve documents from the Web
     - Heuristically label each retrieved document using the stance associated with each phrase

How to incorporate these noisily labeled documents into the training process?

- How to use noisily labeled documents in combination with the (cleanly labeled) debate posts in the training process?

How to incorporate these noisily labeled documents into the training process?

- How to use noisily labeled documents in combination with the (cleanly labeled) debate posts in the training process?

- Train two stance classifiers
  - C_c: trained on only the debate posts
  - C_c+n: trained on debate posts and noisily labeled documents

How to incorporate these noisily labeled documents into the training process?

- How to use noisily labeled documents in combination with the (cleanly labeled) debate posts in the training process?

- Train two stance classifiers
  - C_c: trained on only the debate posts
  - C_c+n: trained on debate posts and noisily labeled documents

- To classify a test post,
  - Linearly combine the output of these two classifiers
  - Combination weight tuned to maximize performance on dev set

Plan for the Talk

- Datasets
- Experimental setup for examining how classification performance varies with
  - the complexity of the learning model
  - the richness of the feature set
  - the amount and quality of training data
  - the application of extra-linguistic constraints
- Evaluation

Constraints

- **Goal**
  - Examine how author constraints (ACs) impact stance classification performance

Constraints

- **Goal**
  - Examine how author constraints (ACs) impact stance classification performance

ACs are inter-post constraints that specify that two posts written by the same author for the same debate topic should have the same stance

How to enforce ACs?

- Use ACs to postprocess the output of a stance classifier

How to enforce ACs?

- Use ACs to postprocess the output of a stance classifier

P1: 0.7 (for), 0.3 (against)
P2: 0.2 (for), 0.8 (against)
P3: 0.7 (for), 0.3 (against)
P4: 0.3 (for), 0.3 (against)
P5: 0.2 (for), 0.8 (against)
P6: 0.9 (for), 0.1 (against)
P7: 0.6 (for),0.4 (against)
P8: 0.1 (for), 0.9 (against)
P9: 0.4 (for), 0.6 (against)

How to enforce ACs?

- Use ACs to postprocess the output of a stance classifier

P1: 0.7 (for), 0.3 (against)
P2: 0.2 (for), 0.8 (against)
P3: 0.7 (for), 0.3 (against)
P4: 0.3 (for), 0.3 (against)
P5: 0.2 (for), 0.8 (against)
P6: 0.9 (for), 0.1 (against)
P7: 0.6 (for),0.4 (against)
P8: 0.1 (for), 0.9 (against)
P9: 0.4 (for), 0.6 (against)

Find the posts written by the same author

How to enforce ACs?

- Use ACs to postprocess the output of a stance classifier

P1: 0.7 (for), 0.3 (against)
P2: 0.2 (for), 0.8 (against)
P3: 0.7 (for), 0.3 (against)
P4: 0.3 (for), 0.3 (against)
P5: 0.2 (for), 0.8 (against)
P6: 0.9 (for), 0.1 (against)
P7: 0.6 (for),0.4 (against)
P8: 0.1 (for), 0.9 (against)
P9: 0.4 (for), 0.6 (against)

Find the posts written by the same author

How to enforce ACs?

- Use ACs to postprocess the output of a stance classifier

P1: 0.7 (for), 0.3 (against)
P2: 0.2 (for), 0.8 (against)
P3: 0.7 (for), 0.3 (against)
P4: 0.3 (for), 0.3 (against)
P5: 0.2 (for), 0.8 (against)
P6: 0.9 (for), 0.1 (against)
P7: 0.6 (for),0.4 (against)
P8: 0.1 (for), 0.9 (against)
P9: 0.4 (for), 0.6 (against)

Sum up the probabilistic votes cast by these posts

How to enforce ACs?

- Use ACs to postprocess the output of a stance classifier

P1: 0.7 (for), 0.3 (against)
P2: 0.2 (for), 0.8 (against)
P3: 0.7 (for), 0.3 (against)
P4: 0.3 (for), 0.3 (against)
P5: 0.2 (for), 0.8 (against)
P6: 0.9 (for), 0.1 (against)
P7: 0.6 (for),0.4 (against)
P8: 0.1 (for), 0.9 (against)
P9: 0.4 (for), 0.6 (against)

Sum up the probabilistic votes cast by these posts

1.4 (for), 2.6 (against)

How to enforce ACs?

- Use ACs to postprocess the output of a stance classifier

P1: 0.7 (for), 0.3 (against)
P2: 0.2 (for), 0.8 (against)
P3: 0.7 (for), 0.3 (against)
P4: 0.3 (for), 0.3 (against)
P5: 0.2 (for), 0.8 (against)
P6: 0.9 (for), 0.1 (against)
P7: 0.6 (for),0.4 (against)
P8: 0.1 (for), 0.9 (against)
P9: 0.4 (for), 0.6 (against)

Assign to each of them the stance that receives more votes

1.4 (for), 2.6 (against)

How to enforce ACs?

- Use ACs to postprocess the output of a stance classifier

P1: 0.7 (for), 0.3 (against)
P2: 0.2 (for), 0.8 (against)
P3: 0.7 (for), 0.3 (against)
P4: 0.3 (for), 0.3 (against)
P5: 0.2 (for), 0.8 (against)
P6: 0.9 (for), 0.1 (against)
P7: 0.6 (for),0.4 (against)
P8: 0.1 (for), 0.9 (against)
P9: 0.4 (for), 0.6 (against)

Assign to each of them the stance that receives more votes

P3: against
P5: against
P8: against
P9: against

How to enforce ACs?

- Use ACs to postprocess the output of a stance classifier

P1: 0.7 (for), 0.3 (against)
P2: 0.2 (for), 0.8 (against)
P3: 0.7 (for), 0.3 (against)
P4: 0.3 (for), 0.3 (against)
P5: 0.2 (for), 0.8 (against)
P6: 0.9 (for), 0.1 (against)
P7: 0.6 (for),0.4 (against)
P8: 0.1 (for), 0.9 (against)
P9: 0.4 (for), 0.6 (against)

Assign to each of them the stance that receives more votes

P3: against
P5: against
P8: against
P9: against

- **Goal**: examine how ACs impact stance classification

Plan for the Talk

- Datasets
- Experimental setup for examining how classification performance varies with
  - the complexity of the learning model
  - the richness of the feature set
  - the amount and quality of training data
  - the application of extra-linguistic constraints
- Evaluation

Evaluation: Goal

- Examine how stance classification performance varies with the four factors concerning data, features, models and constraints

Evaluation: Setup

- 5-fold cross validation
- Evaluation metric: accuracy

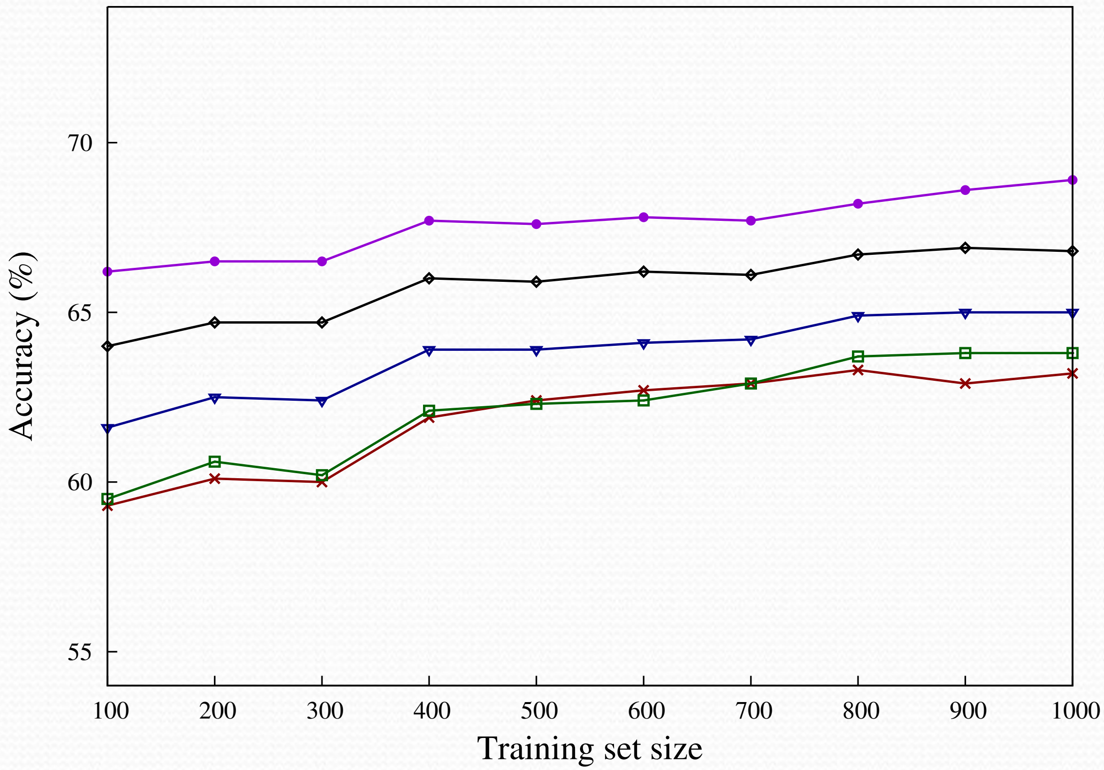
Recap

- We have 4 evaluation datasets
  - Abortion, Gay Rights, Obama, Marijuana
- We have 6 learning models
  - Naïve Bayes (NB), SVM, HMM, CRF, NB-f, HMM-f

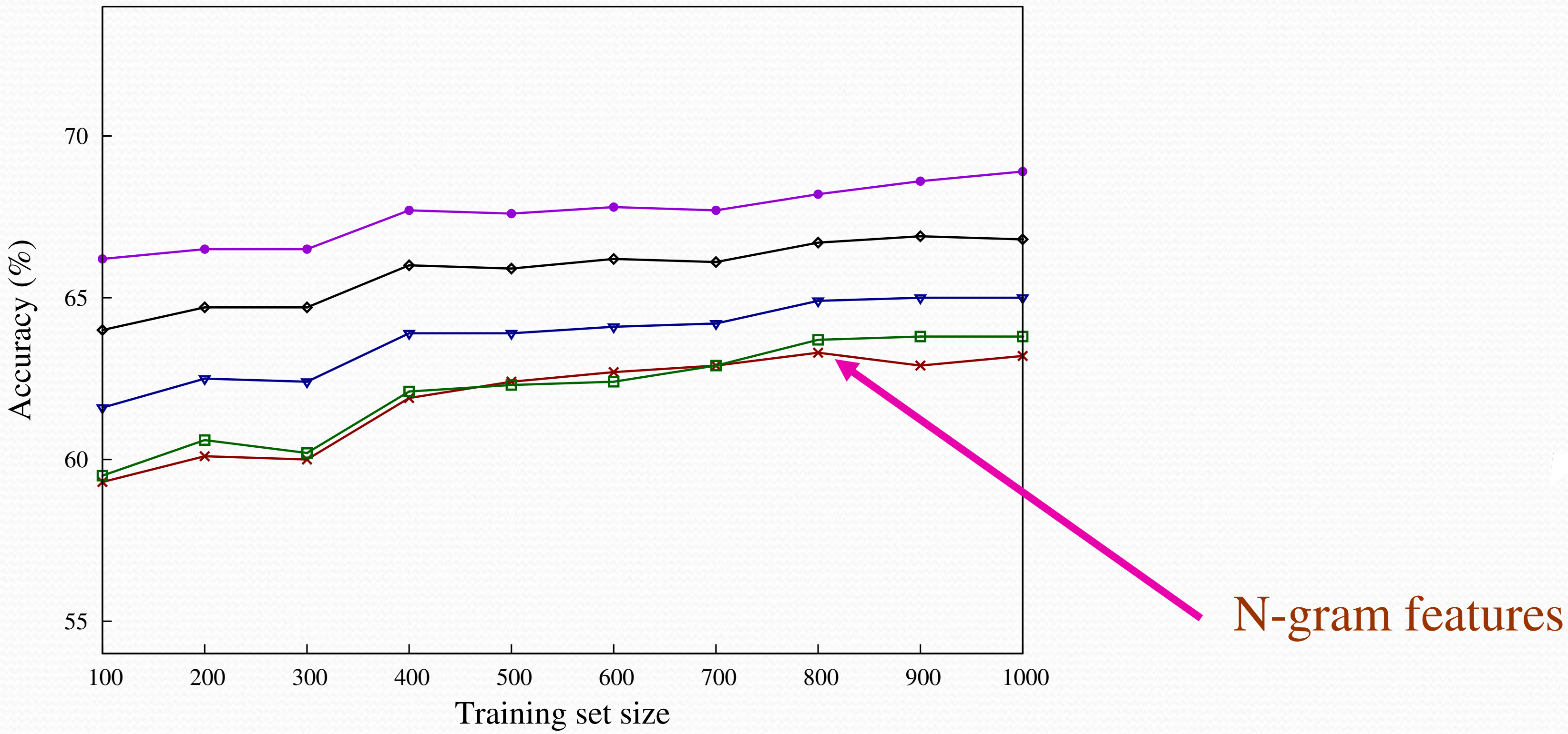
Recap

- We have 4 evaluation datasets
  - Abortion, Gay Rights, Obama, Marijuana
- We have 6 learning models
  - Naïve Bayes (NB), SVM, HMM, CRF, NB-f, HMM-f
- There are 4x6=24 model-dataset combinations
- For each combination, we plot a graph
- Each graph has 5 learning curves

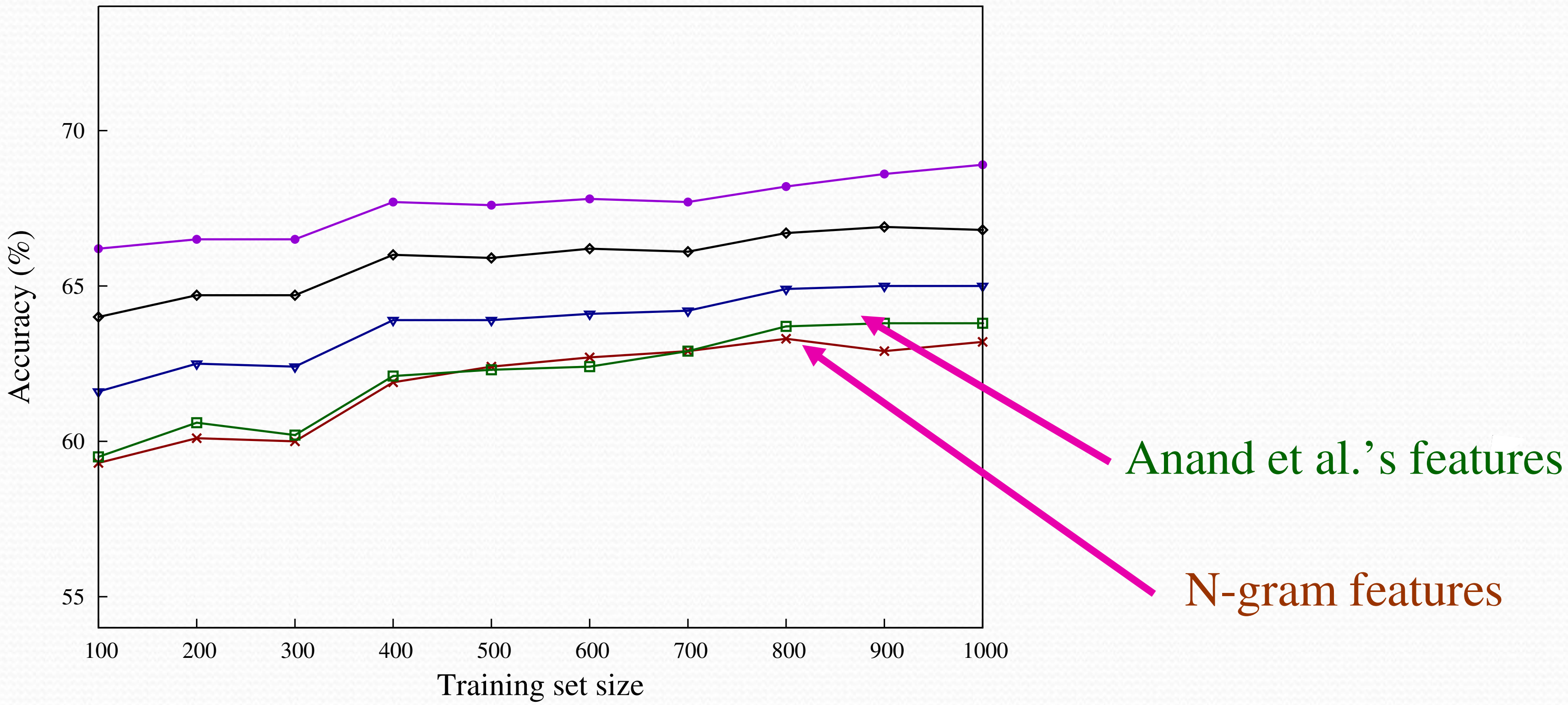
Graph for Gay Rights-HMM



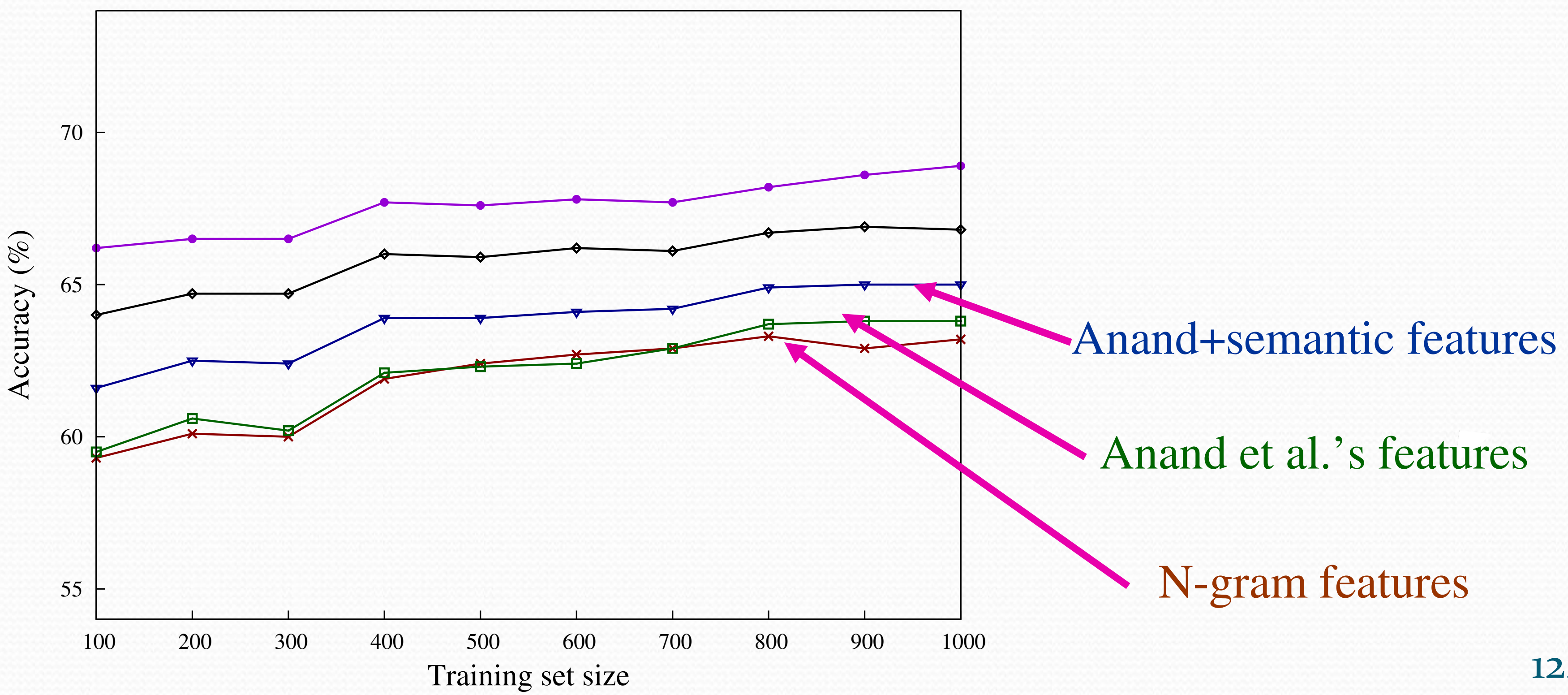
Graph for Gay Rights-HMM



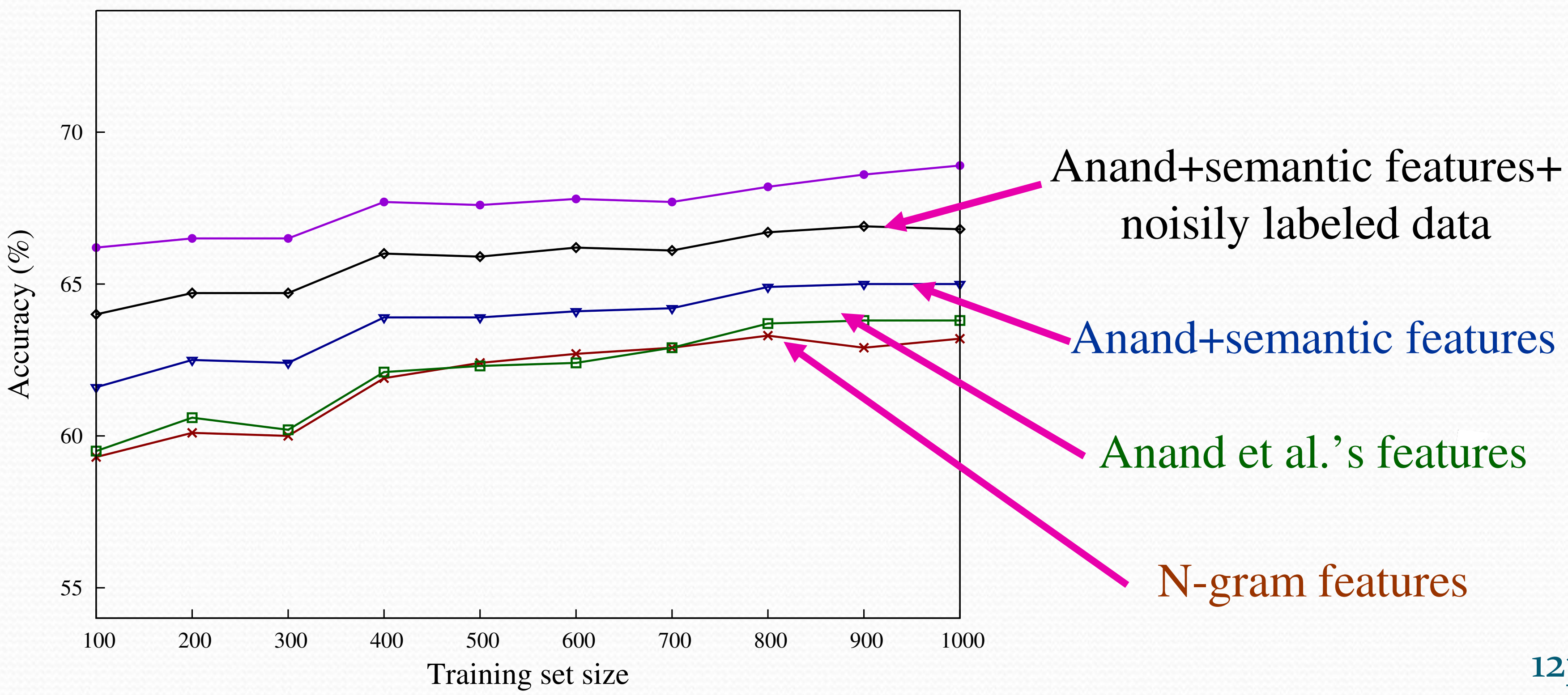
Graph for Gay Rights-HMM



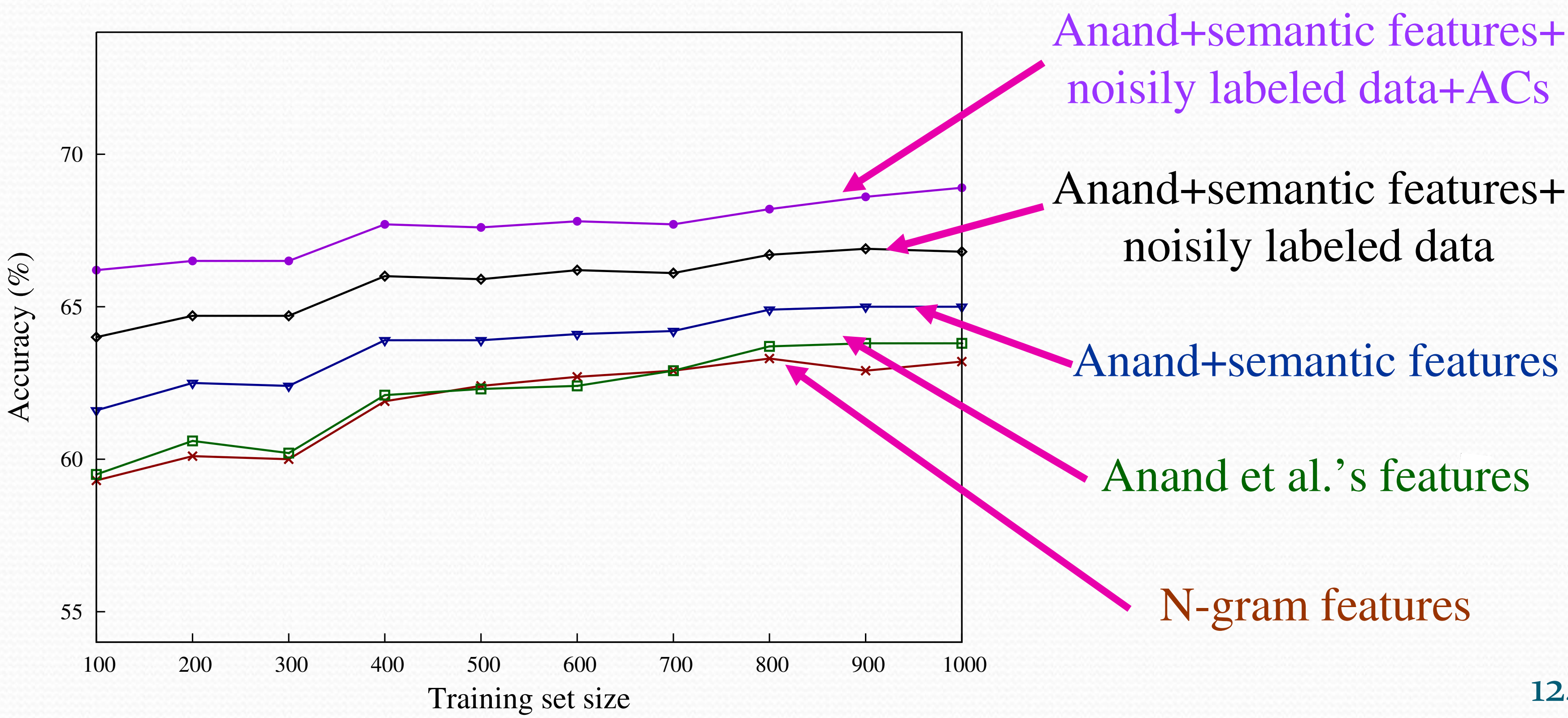
Graph for Gay Rights-HMM



Graph for Gay Rights-HMM



Graph for Gay Rights-HMM



Goal

- Given the 24 graphs corresponding to the 24 model-dataset combinations, we analyze stance classification performance

Richness of Feature Set

- Is Anand's feature set stronger than the N-gram feature set?

Richness of Feature Set

- Is Anand's feature set stronger than the N-gram feature set?

- Not always.
  - In some cases Anand's feature set yields better performance
  - In other cases it's the other way round

Richness of Feature Set

- Are frame-semantic features useful?

Richness of Feature Set

- Are frame-semantic features useful?
- Yes. Apart from a few cases in Abortion, adding semantic features to Anand’s feature set yields significant improvements

Amount of Training Data

- Can we improve performance simply by training on a larger amount of (cleanly labeled) debate posts?

Amount of Training Data

- Can we improve performance simply by training on a larger amount of (cleanly labeled) debate posts?

- Yes. As the number of training posts increases, we see significant improvements on all debate topics
  - 1.5 (Abortion); 2.4 (Gay Rights), 2.0 (Obama), 3.1 (Marijuana)

Quality of Training Data

- Does using noisily labeled documents help improve performance?

Quality of Training Data

- Does using noisily labeled documents help improve performance?
- Yes. Adding noisily labeled documents improves performance significantly regardless of the learning model.

Usefulness of Author Constraints

- Are ACs useful?

Usefulness of Author Constraints

- Are ACs useful?

- Yes. Adding ACs consistently yields significant improvements on all debate topics
  - 7% (Abortion); 3% (Gay Rights); 4% (Obama); 1% (Marijuana)

Models

- Which model is better, NB or SVM?

Models

- Which model is better, NB or SVM?

- No clear winner
  - SVM beats NB in 17% of the cases
  - NB beats SVM in 27% of the cases
  - they are statistically indistinguishable in the remaining cases
  - Neither generative nor discriminative models seems better

Model Complexity

- Are the sequence models better than their non-sequence counterparts?