Learning Representations of Social Meaning

Jacob Eisenstein
@jacobbeisenstein

Georgia Institute of Technology

August 3, 2017
What does social meaning mean?

Pick up the big red block.
What does social meaning mean?

Pick up the big red block.

► Would you please pick up the big red block?
► That red block needs picked up.
► Hurry up and grab that block.
► Could you, like, pick up the red block?
What is social meaning about?

Idealized model: saying the “same thing” differently, depending on (and conveying information about):

- the speaker/writer
  dialect, gender, race, “latent user attributes”, . . .

- the communicative situation
  register, formality, politeness, accommodation, . . .

- attitude towards the talk and topic
  sentiment, subjectivity, irony, . . .
The department store study

- **Linguistic variable**: (r) in *fourth floor*
- **Social variable**: class (department stores Klein’s, Macy’s, and Sak’s)
- **Situational variable**: feigned misunderstanding

(Labov, 1972)
The department store study

Use of local New York “r-less” variable decreases with
- socioeconomic status
- emphasis.

Figure 13.2: Percentage of all (r - 1) by store for four positions (S = Saks, M = Macy’s, K = Kleins)
The department store study

Use of local New York “r-less” variable decreases with
- socioeconomic status
- emphasis.

Geographical, socioeconomic, and stylistic variation are all linked!
Disentangling social variables

How to assess the impact of multiple social variables?

- **Linguistic variable**: on versus other pronouns in Montreal French
- **Social variables**: age, gender, education, socioeconomic status
- **Model**:

\[
\log \frac{p}{1 - p} = \beta_0 x_0 + \beta_1 x_1 + \ldots + \beta_n x_n \tag{1}
\]
Disentangling social variables

How to assess the impact of multiple social variables?

- **Linguistic variable**: on versus other pronouns in Montreal French
- **Social variables**: age, gender, education, socioeconomic status
- Model:

  \[
  \log \frac{p}{1 - p} = \beta_0 x_0 + \beta_1 x_1 + \ldots + \beta_n x_n \quad (1)
  \]

  *(Sankoff & Labov, 1979)*
Computational models of social meaning

- What is the role of computation in understanding social meaning in language?
- Why do we need representation learning?
- Is social meaning relevant to “core tasks” in NLP?
What is social meaning about?

Idealized model: saying the “same thing” differently, depending on (and conveying information about):

- the speaker/writer
dialect, gender, race, “latent user attributes”, …

- the communicative situation
register, formality, politeness, accommodation, …

- attitude towards the talk and topic
sentiment, subjectivity, irony, …
What is social meaning about?

Idealized model: saying the “same thing” differently, depending on (and conveying information about):

- **the speaker/writer**
  - dialect, gender, race, “latent user attributes”, ...

- **the communicative situation**
  - register, formality, politeness, accommodation, ...

- **attitude towards the talk and topic**
  - sentiment, subjectivity, irony, ...
Social meaning from social media

Large-scale social media data has sparked a renaissance in geographical dialectology.

Philadelphia: 
jawn, ard, 
#phillies, sixers, 
phils, wawa

Chicago: 
#chicago, lbvs, 
chicago, 
blackhawks, 
bears, bogus

New York: 
flatbush, baii, 
brib, bx, staten, 
odee

Philadelphia: 
jawn, ard, 
#phillies, sixers, 
phils, wawa

Chicago: 
#chicago, lbvs, 
chicago, 
blackhawks, 
bears, bogus

New York: 
flatbush, baii, 
brib, bx, staten, 
odee

(Eisenstein et al., 2014; Rahimi et al., 2017) (Grieve et al., 2017)
Representations for geographical dialects

- Language variation generally does not align with geo-political units like cities, states, and nations (Heeringa & Nerbonne, 2001).

- Representation learning for dialect regions:
  - Gaussians (Eisenstein et al., 2010)
  - kd-trees (Roller et al., 2012)
The real picture?
Social variables

- “Demographic” variables: race, gender, age (e.g., Argamon et al., 2007; Eisenstein et al., 2011; Nguyen et al., 2013)

- Eckert (2000): individuals have agency in creating social meanings! Sociolinguistics should focus on locally-defined social categories.
Social variables

- “Demographic” variables: race, gender, age (e.g., Argamon et al., 2007; Eisenstein et al., 2011; Nguyen et al., 2013)

- Eckert (2000): individuals have agency in creating social meanings! Sociolinguistics should focus on locally-defined social categories.

Can representation learning discover locally-defined social variables from text and metadata?
Gender identity and lexical variation in social media (Bamman, Eisenstein, and Schnoebelen, 2014)

- How to model gendered language differences without gender binary?
Literally the simplest possible thing

\(K\)-means clustering on Twitter timelines

hubs blogged recipe fabric
kidd hubs xo =]
wyd #oomf lmbo shyt
n_ggas wyd finna shyt
#nhl #bruins #mlb knicks
Literally the simplest possible thing

*K*-means clustering on Twitter timelines

<table>
<thead>
<tr>
<th>Term</th>
<th>% Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>hubs blogged recipe fabric</td>
<td>90%</td>
</tr>
<tr>
<td>kidd hubs xo =]</td>
<td>80%</td>
</tr>
<tr>
<td>wyd #oomf lmbo shyt</td>
<td>60%</td>
</tr>
<tr>
<td>n_ggas wyd finna shyt</td>
<td>26%</td>
</tr>
<tr>
<td>#nhl #bruins #mlb knicks</td>
<td>11%</td>
</tr>
</tbody>
</table>

- Do women use more standard language?
- Is men’s writing more “informational”?
Literally the simplest possible thing

*K*-means clustering on Twitter timelines

<table>
<thead>
<tr>
<th>Term</th>
<th>% Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>hubs blogged recipe fabric</td>
<td>90%</td>
</tr>
<tr>
<td>kidd hubs xo ]</td>
<td>80%</td>
</tr>
<tr>
<td>wyd #oomf lmbo shyt</td>
<td>60%</td>
</tr>
<tr>
<td>n_ggas wyd finna shyt</td>
<td>26%</td>
</tr>
<tr>
<td>#nhl #bruins #mlb knicks</td>
<td>11%</td>
</tr>
</tbody>
</table>

- Do women use more standard language?
- Is men’s writing more “informational”?
Literally the simplest possible thing

*K*-means clustering on Twitter timelines

<table>
<thead>
<tr>
<th>Slang Terms</th>
<th>% Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>hubs blogged recipe fabric</td>
<td>90%</td>
</tr>
<tr>
<td>kidd hubs xo =]</td>
<td>80%</td>
</tr>
<tr>
<td>wyd #oomf lmbo shyt</td>
<td>60%</td>
</tr>
<tr>
<td>n_ggas wyd finna shyt</td>
<td>26%</td>
</tr>
<tr>
<td>#nhl #bruins #mlb knicks</td>
<td>11%</td>
</tr>
</tbody>
</table>

- Do women use more standard language?
- Is men's writing more "informational"?
Literally the simplest possible thing

*K*-means clustering on Twitter timelines

<table>
<thead>
<tr>
<th>Term</th>
<th>% Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>hubs blogged recipe fabric</td>
<td>90%</td>
</tr>
<tr>
<td>kidd hubs xo =]</td>
<td>80%</td>
</tr>
<tr>
<td>wyd #oomf lmbo shyt</td>
<td>60%</td>
</tr>
<tr>
<td>n_ggas wyd finna shyt</td>
<td>26%</td>
</tr>
<tr>
<td>#nhl #bruins #mlb knicks</td>
<td>11%</td>
</tr>
</tbody>
</table>

- Do women use more standard language?
- Is men’s writing more “informational”?
Literally the simplest possible thing

\textbf{\textit{K}-means clustering on Twitter timelines}

<table>
<thead>
<tr>
<th>% Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>hubs blogged recipe fabric</td>
</tr>
<tr>
<td>kidd hubs xo =]</td>
</tr>
<tr>
<td>wyd #oomf lmbo shyt</td>
</tr>
<tr>
<td>n_ggas wyd finna shyt</td>
</tr>
<tr>
<td>#nhl #bruins #mlb knicks</td>
</tr>
</tbody>
</table>

- Do women use more standard language?
- Is men’s writing more “informational”?
What is social meaning about?

Idealized model: saying the “same thing” differently, depending on (and conveying information about):

- **the speaker/writer**
  - dialect, gender, race, “latent user attributes”, ... 

- **the communicative situation**
  - register, formality, politeness, accommodation, ... 

- **attitude towards the talk and topic**
  - sentiment, subjectivity, irony, ...
What is social meaning about?

Idealized model: saying the “same thing” differently, depending on (and conveying information about):

- the speaker/writer
  
dialect, gender, race, “latent user attributes”, …

- the communicative situation
  
  register, formality, politeness, accommodation, …

- attitude towards the talk and topic
  
sentiment, subjectivity, irony, …
Social-situational variation

“Any code? or do we have to guess at how you did it?”

- Ask crowdworkers to annotate the extent of some social variable in text:
  - politeness (Danescu-Niculescu-Mizil et al., 2013)
  - formality (Pavlick & Tetreault, 2016)
- Train a supervised system, examine features.
Construct validity

- Politeness and formality are constructs with strong theoretical support (Brown, 1987; Heylighen & Dewaele, 1999)
- Crowdsourced annotations capture *folk linguistic* intuitions about how these constructs work (and that’s okay)

**But**... are there other forms of social meaning? What are the principle linguistic dimensions of social-situational variation?
You’re Mr. Lebowski, I’m the Dude: Inducing Address Term Formality in Signed Social Networks (Krishnan & Eisenstein, 2015).

Vinodh Krishnan

- How do address terms like dude and Mr. Lebowski convey social meaning?
- What social configurations are likely or unlikely?
We observe address terms $x$ but not relation types $y$.

Inferring $y$ gives a labeling over edges.

Estimating $P(x \mid y)$ gives the distribution over address terms for each edge type.
Complete model specification

\[ P(y, x \mid G; \Theta, \beta, \eta) = P(x \mid y; \Theta)P(y \mid G; \beta, \eta) \]

Bayesian inference answers several questions:

1. What is the social meaning of each dyad?
2. How is social meaning expressed in address terms?
3. Are there structural regularities across networks?
Complete model specification

\[ P(y, x \mid G; \Theta, \beta, \eta) = P(x \mid y; \Theta)P(y \mid G; \beta, \eta) \]

Bayesian inference answers several questions:
1. What is the social meaning of each dyad?
2. How is social meaning expressed in address terms?
3. Are there structural regularities across networks?
Complete model specification

\[ P(y, x \mid G; \Theta, \beta, \eta) = P(x \mid y; \Theta)P(y \mid G; \beta, \eta) \]

Bayesian inference answers several questions:

1. What is the social meaning of each dyad?
2. How is social meaning expressed in address terms?
3. Are there structural regularities across networks?
Complete model specification

\[ P(y, x \mid G; \Theta, \beta, \eta) = P(x \mid y; \Theta)P(y \mid G; \beta, \eta) \]

Bayesian inference answers several questions:
1. What is the social meaning of each dyad?
2. How is social meaning expressed in address terms?
3. Are there structural regularities across networks?
Forms of address

Names such as Barack, Barack Hussein Obama.

Titles such as Ms., Dr., Private, Reverend. Titles can be used for address either by preceding a name (e.g., Colonel Kurtz), or in isolation (e.g., Yes, Colonel).

Placeholder names such as dude, bro, brother, sweetie, cousin, and asshole. These terms can be used for address only in isolation.
Lexicon induction: titles

- Find terms that frequently appear at the beginning of address spans containing the character’s name.
- We manually filter out 17 of 34 candidates, obtaining:

agent, aunt, captain, colonel, commander, cousin, deputy, detective, dr, herr, inspector, judge, lord, master, mayor, miss, mister, miz, monsieur, mr, mrs, ms, professor, queen, reverend, sergeant, uncle
Lexicon induction: placeholder names

- Find terms that are frequently tagged as the unique element in an address span.
- After manually filter out 41 of 96 candidates, we obtain:

  asshole, babe, baby, boss, boy, bro, bud, buddy, cocksucker, convict, cousin, cowboy, cunt, dad, darling, dear, detective, doll, dude, dummy, father, fella, gal, ho, hon, honey, kid, lad, lady, lover, ma, madam, madame, man, mate, mister, mon, moron, motherfucker, pal, papa, partner, peanut, pet, pilgrim, pop, president, punk, shithead, sir, sire, son, sonny, sport, sucker, sugar, sweetheart, sweetie, tiger
Cluster coherence

<table>
<thead>
<tr>
<th>V-cluster</th>
<th>T-cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>sir</td>
<td>FIRSTNAME</td>
</tr>
<tr>
<td>mr+LASTNAME</td>
<td>man</td>
</tr>
<tr>
<td>mr+FIRSTNAME</td>
<td>baby</td>
</tr>
<tr>
<td>mr</td>
<td>honey</td>
</tr>
<tr>
<td>miss+LASTNAME</td>
<td>darling</td>
</tr>
<tr>
<td>son</td>
<td>sweetheart</td>
</tr>
<tr>
<td>mister+FIRSTNAME</td>
<td>buddy</td>
</tr>
<tr>
<td>mrs</td>
<td>sweetie</td>
</tr>
</tbody>
</table>

On an intrusion task

- Raters found the intruder in 73% of cases for the full model (M4).
- ... versus 52% in the text-only model (M1).
Star Wars
Network features

\begin{align*}
\text{+3.73} & \quad \text{t} \quad \text{t} \quad \text{t} \\
-6.48 & \quad \text{t} \quad \text{v} \\
-1.05 & \quad \text{t} \quad \text{v} \\
+1.23 & \quad \text{v} \quad \text{v} \\
\end{align*}
What is social meaning about?

Idealized model: saying the “same thing” differently, depending on (and conveying information about):

- the speaker/writer
  - dialect, gender, race, “latent user attributes”, . . .
- the communicative situation
  - register, formality, politeness, accommodation, . . .
- attitude towards the talk and topic
  - sentiment, subjectivity, irony, . . .
What interactional and linguistic patterns distinguish online communities?

How do automatically-identified linguistic styles relate to annotated constructs such as formality and politeness?
The **stance triangle** is proposed as a representation for interactional meaning (Du Bois, 2007)

- Alignment (or not) between speakers
- Evaluation of (and positioning w.r.t.) the object of discussion and the talk itself
The **stance triangle** is proposed as a representation for interactional meaning (Du Bois, 2007)

- Alignment (or not) between speakers
- Evaluation of (and positioning w.r.t.) the object of discussion and the talk itself

Hard to annotate!  
(Fitzpatrick et al 2014)
Hypotheses:

- Online communities differ in their characteristic stances
- Stance is revealed (in part) by a lexicon of **stance markers**
**STANCE LEXICON: SEED LIST**

- List of markers from Biber and Finegan (1989)  
  - certainty adverbs *actually, of course, in fact*
  - affect markers *amazing, thankful, sadly*
  - hedges *kind of, maybe, something like*

- Dialog act markers from Jurafsky et al. (1998)  
  - *oh yeah, nah, wow*
**STANCE LEXICON: EXPANSION**

Similar words from word embeddings trained on Reddit comments

- **Seed stance markers**
  - **Seed term** | **Expanded terms**
  - significantly | considerably, substantially, dramatically
  - certainly | surely, frankly, definitely
  - incredibly | extremely, unbelievably, exceptionally

- **228 additional markers**

- **Seed dialog acts**
  - **Seed term** | **Expanded terms**
  - nope | nah, yup, nevermind
  - great | fantastic, terrific, excellent

- **112 additional markers**
### A Multidimensional Lexicon for Interpersonal Stancetaking

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Negative Phrases</th>
<th>Positive Phrases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dim-1</td>
<td>beautifully, pleased, thanks, spectacular, delightful</td>
<td>just, even, all, no, so</td>
</tr>
<tr>
<td>Dim-2</td>
<td>suggests that, demonstrates, conclude, demonstrated, demonstrate</td>
<td>lovely, awww, hehe, aww, haha</td>
</tr>
<tr>
<td>Dim-3</td>
<td>funnier, hilarious, disturbing, creepy, funny</td>
<td>thanks, ideally, calculate, estimate, calculation</td>
</tr>
<tr>
<td>Dim-4</td>
<td>phenomenal, bummed, enjoyed, fantastic, disappointing</td>
<td>hello, thx, hehe, aww, hi</td>
</tr>
</tbody>
</table>
**A MULTIDIMENSIONAL LEXICON FOR INTERPERSONAL STANCETAKING**

<table>
<thead>
<tr>
<th>Dim-1</th>
<th>Beautifully, pleased, thanks, spectacular, delightful</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Just, even, all, no, so</td>
</tr>
<tr>
<td>Dim-2</td>
<td>Suggests that, demonstrates, conclude, demonstrated, demonstrate</td>
</tr>
<tr>
<td></td>
<td>Lovely, awww, hehe, aww, haha</td>
</tr>
<tr>
<td>Dim-3</td>
<td>Funnier, hilarious, disturbing, creepy, funny</td>
</tr>
<tr>
<td></td>
<td>Thanks, ideally, calculate, estimate, calculation</td>
</tr>
<tr>
<td>Dim-4</td>
<td>Phenomenal, bummed, enjoyed, fantastic, disappointing</td>
</tr>
<tr>
<td></td>
<td>Hello, thx, hehe, aww, hi</td>
</tr>
</tbody>
</table>
A MULTIDIMENSIONAL LEXICON FOR INTERPERSONAL STANCETAKING

<table>
<thead>
<tr>
<th>Dim-1</th>
<th>Informational vs. Involved Style</th>
</tr>
</thead>
<tbody>
<tr>
<td>- beautifully, pleased, thanks, spectacular, delightful</td>
<td></td>
</tr>
<tr>
<td>+ just, even, all, no, so</td>
<td></td>
</tr>
<tr>
<td>- suggests that, demonstrates, conclude, demonstrated, demonstrate</td>
<td></td>
</tr>
<tr>
<td>+ lovely, awww, hehe, aww, haha</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dim-3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- funnier, hilarious, disturbing, creepy, funny</td>
<td></td>
</tr>
<tr>
<td>+ thanks, ideally, calculate, estimate, calculation</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dim-4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- phenomenal, bummer, enjoyed, fantastic, disappointing</td>
<td></td>
</tr>
<tr>
<td>+ hello, thx, hehe, aww, hi</td>
<td></td>
</tr>
</tbody>
</table>
Involved Style
Informational Style
Formal Style
Informal Style
MODEL VALIDATION: INTRINSIC EVALUATIONS

Word intrusion task

<table>
<thead>
<tr>
<th>suggests that</th>
<th>demonstrates</th>
<th>awww</th>
<th>conclude</th>
<th>demonstrated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Word intrusion task

<table>
<thead>
<tr>
<th>suggests that</th>
<th>demonstrates</th>
<th>conclude</th>
<th>demonstrated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| suggests that | demonstrates | awww     | conclude     | demonstrated |
|---------------|--------------|----------|--------------|
|               |              |          |              |

intruder term
## MODEL VALIDATION: INTRINSIC EVALUATIONS

### Word intrusion task

<table>
<thead>
<tr>
<th>suggests that</th>
<th>demonstrates</th>
<th>conclude</th>
<th>demonstrated</th>
</tr>
</thead>
<tbody>
<tr>
<td>awww</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Inter-rater reliability:** Krippendorf’s $\alpha = 0.73$
- **Model precision:** 0.82 (chance precision = 0.20)
MODEL VALIDATION: EXTRINSIC EVALUATIONS

- Dimension 2:
  - "suggests that, demonstrates"
  - vs. "lovely, awww, hehe"

- Dimension 5:
  - "lovely, stunning, wonderful"
  - vs. "nvm, cmon, smh"
MODEL VALIDATION: EXTRINSIC EVALUATIONS

Predicting cross-posting in multiple subreddits
Predicting cross-posting in multiple subreddits

Classification task

Predict whether a pair of subreddits has higher or lower membership overlap

Dataset

Paired each of the top 100 popular subreddits with five other subreddits with highest and lowest membership overlap

<table>
<thead>
<tr>
<th>High-Overlap Pairs</th>
<th>Low-Overlap Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>r/blog, r/announcements</td>
<td>r/soccer, r/nosleep</td>
</tr>
<tr>
<td>r/politics, r/technology</td>
<td>r/Minecraft, r/personalfinance</td>
</tr>
<tr>
<td>r/pokemon, r/wheredidthesodago</td>
<td>r/programming, r/gonewild</td>
</tr>
</tbody>
</table>
MODEL VALIDATION: EXTRINSIC EVALUATIONS

Predicting cross-posting in multiple subreddits

Models & Results

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOW-COSINE</td>
<td></td>
</tr>
<tr>
<td>STANCE-COSINE</td>
<td></td>
</tr>
<tr>
<td>BOW-SVD</td>
<td></td>
</tr>
<tr>
<td>STANCE-SVD</td>
<td></td>
</tr>
</tbody>
</table>
Predicting cross-posting in multiple subreddits

Models & Results

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOW-COSINE</td>
<td>66.13%</td>
</tr>
<tr>
<td>STANCE-COSINE</td>
<td>64.31%</td>
</tr>
<tr>
<td>BOW-SVD</td>
<td>77.48%</td>
</tr>
<tr>
<td>STANCE-SVD</td>
<td>84.93%</td>
</tr>
</tbody>
</table>

Interactional style predicts cross-posting behavior
Computational models of social meaning

- What is the role of computation in understanding social meaning in language?
- Why do we need representation learning?
- Is social meaning relevant to “core tasks” in NLP?
Overcoming language variation in sentiment analysis (Yang & Eisenstein, 2017)

- Can we exploit social network homophily to make document classification more robust to social variation?
Language variation: a challenge for NLP

“I would like to believe he’s sick rather than just mean and evil.”

(Yang & Eisenstein, 2017)
Language variation: a challenge for NLP

“I would like to believe he’s sick rather than just mean and evil.”

“You could’ve been getting down to this sick beat.”

(Yang & Eisenstein, 2017)
Personalization by ensemble

- **Goal:** personalized conditional likelihood, $P(y \mid x, a)$, where $a$ is the author.
- **Problem:** We have labeled examples for only a few authors.
Personalization by ensemble

- **Goal:** personalized conditional likelihood, \( P(y \mid x, a) \), where \( a \) is the author.
- **Problem:** We have labeled examples for only a few authors.
- **Personalization ensemble**

\[
P(y \mid x, a) = \sum_k P_k(y \mid x) \pi_a(k)
\]

- \( P_k(y \mid x) \) is a basis model
- \( \pi_a(\cdot) \) are the ensemble weights for author \( a \)
Homophily to the rescue?

Are language styles **assortative** on the social network?
Evidence for linguistic homophily

Pilot study: is classifier accuracy assortative on the Twitter social network?

\[
\text{assort}(G) = \frac{1}{\#|G|} \sum_{(i,j) \in G} \delta(y_i = \hat{y}_i)\delta(y_j = \hat{y}_j) \\
+ \delta(y_i \neq \hat{y}_i)\delta(y_j \neq \hat{y}_j)
\]
Evidence for linguistic homophily

Pilot study: is classifier accuracy \textbf{ assortative } on the Twitter social network?

\[
\text{assort}(G) = \frac{1}{\#|G|} \sum_{(i,j) \in G} \delta(y_i = \hat{y}_i) \delta(y_j = \hat{y}_j) \\
+ \delta(y_i \neq \hat{y}_i) \delta(y_j \neq \hat{y}_j)
\]

![Graphs showing assortativity over rewiring epochs for different interactions: follow, mention, retweet. The red dashed line represents the original network, and the blue line represents random rewiring.](image)
Network-driven personalization

- For each author, estimate a **node embedding** \( e_a \) (Tang et al., 2015).
- Nodes who share neighbors get similar embeddings.

\[
\pi_a = \text{SoftMax}(f(e_a))
\]

\[
P(y \mid x, a) = \sum_{k=1}^{K} P_k(y \mid x) \pi_a(k)
\]
Results

Improvements over ConvNet baseline:
- +2.8% on Twitter Sentiment Analysis
- +2.7% on Ciao Product Reviews

NLSE is prior state-of-the-art (Astudillo et al., 2015).
<table>
<thead>
<tr>
<th>More positive</th>
<th>More negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>1    banging  loss  fever  broken</td>
<td>dear  like  god  yeah  wow</td>
</tr>
<tr>
<td>fucking</td>
<td></td>
</tr>
<tr>
<td>2    chilling  cold  ill  sick  suck</td>
<td>satisfy  trust  wealth  strong  lmao</td>
</tr>
<tr>
<td>3    ass  damn  piss  bitch  shit</td>
<td>talent  honestly  voting  win  clever</td>
</tr>
<tr>
<td>4    insane  bawling  fever  weird  cry</td>
<td>lmao  super  lol  haha  hahaha</td>
</tr>
<tr>
<td>5    ruin  silly  bad  boring  dreadful</td>
<td>lovatics  wish  beliebers  ariana-tors  kendall</td>
</tr>
</tbody>
</table>
What’s the right representation for social meaning?

Word senses are only ever defined relative to a set of interests. The set of senses defined by a dictionary may or may not match the set that is relevant for an NLP application (Kilgarriff, 1997).
What’s the right representation for social meaning?

Word senses are only ever defined relative to a set of interests. The set of senses defined by a dictionary may or may not match the set that is relevant for an NLP application (Kilgarriff, 1997).

- Annotating the full social meaning of an utterance is not possible.
- The way forward: representation learning plus linguistics for model design, lexicon and metadata selection.
- Single-membership models are only the first step!
Acknowledgments

- **Collaborators**: David Bamman, Ming-Wei Chang, Jim Fitzpatrick, Scott Kiesling, Vinodh Krishnan, Umashanthi Pavalanathan, Tyler Schneebelen, Yi Yang

- **Sponsors**: National Science Foundation, Air Force Office of Scientific Research, National Institutes for Health.


