Inducing Domain-Specific Sentiment Lexicons from Unlabeled Corpora

William L. Hamilton, Kevin Clark, Jure Leskovec, Dan Jurafsky

Speaker: Yupei Du

Outline

- *Motivations*
- Framework
- Recreating known lexicons
- Inducing lexicons

Motivations

BEFORE

- Domain-specific sentiment lexicons are crucial to CSS
- Lexical sentiment is hugely influenced by context
  (Domain: "soft" & History: "Terrific")
- Hand-made lexicons are expensive & time-consuming
- Web-scale lexicons induce biases in specific domain

THIS WORK

- Small sets of seed words universal across domains
- Modest-sized corpora
- Confidence score

Outline

- Motivations
- *Framework*
- Recreating known lexicons
- Inducing lexicons

Framework

- 3 Steps
  1. Embeddings
  2. Constructing Graphs
  3. Propagating
- Variants
- Rubostness

Embeddings

- Construct Co-occurrence Matrix M^{PPMI}

  Each Element:

$$M_{i,j}^{PPMI} = \max\{\log(\frac{\hat{p}(w_i, w_j)}{\hat{p}(w_i)\hat{p}(w_j)}), 0\}$$

  Where $\hat{p}$ is smoothed window-based co-occur *(c=0.75,r=4)*

- Perform a SVD composition to build embeddings:

$$M^{PPMI} = U\Sigma V^T \quad w_i^{SVD} = (U)_i$$

- Outperform Word2Vec & GloVe

Constructing Graphs

- Connecting nearest k neighbors according to cosine-similiarty

$$E_{i,j} = \textbf{arccos}\,(-\frac{w_i^T w_j}{||w_i||\ ||w_j||})$$

As w_i and w_j get closer, value of edges increase from $\frac{\pi}{2}$ to π

Propagating

- Transition Matrix

$$T = D^{\frac{1}{2}} E D^{\frac{1}{2}}$$

$$\text{thus } T_{i,j} = (\sum_{k=1}^{N} w_{i,k} \sum_{k=1}^{N} w_{k,j}) w_{i,j}$$

- Update Until Numerical Convergence

  $p \in R^{|V|}$: vector of word sentiment

  $s \in R^{|S|}$: vector of seeds

$$p^{(t+1)} = \beta T p^{(t)} + (1 - \beta) s$$

- Obtain both positive and negative scores

Propageting

- Final Score

$$\overline{P}^{P}(w_i) = \frac{P^{P}(w_i)}{P^{P}(w_i) + P^{P}(w_i)}$$

- seed words

Domain	Positive seed words	Negative seed words
Standard English	good, lovely, excellent, fortunate, pleasant, delightful, perfect, loved, love, happy	bad, horrible, poor, unfortunate, unpleasant, disgusting, evil, hated, hate, unhappy
Finance	successful, excellent, profit, beneficial, improving, improved, success, gains, positive	negligent, loss, volatile, wrong, losses, damages, bad, litigation, failure, down, negative
Twitter	love, loved, loves, awesome, nice, amazing, best, fantastic, correct, happy	hate, hated, hates, terrible, nasty, awful, worst, horrible, wrong, sad

manually selected to be context insensitive

Variants & Rubostness

Variants

- Transition Matrix
- label propagation

Robustness

- Running over B random equally-sized seed sets
- B = 50 & size_of_subsets = 7
- Confidence

Outline

- Motivations
- Framework
- *Recreating known lexicons*
- Inducing lexicons

Recreating known lexicons

Domains

Domain	Standard English	Finance	Twitter
Lexicon	General Inquirer Lexicon with Continuous valence	Handcrafted	SemEval task 10E
Embeddings	Off-the-shelf Word2Vec on Google News(10^{11})	-	Rothe et al. (2016)
Corpus	COHA(10^7)	financial 8K documents(10^7)	-

Recreating known lexicons

Baseline

1. PMI
   Computing the pointwise mutual information between the seeds and the targets without using propagation.
2. Velikovich et al.(2010)
   Similar to this work with alternative propagations and vectors

State-of-the-art

1. DENSIFIER
   Learning orthogonal transformations
2. WordNet-based
   Performing LP over WordNet-derived graph (Standard English)

Recreating known lexicons

Evaluation Setup

- Binary Classification

  Positive vs. Negative
- ternary classification

  Positive, Neutral and Negative with a prior distribution
- Kendall τ correlation

  With continuous human-annotated polarity scores

Findings

- SENTPROP achieves state-of-the-art level on large corpus
- Maintain high accuracy on modest-sized corpus
- High quality word vectors have a drastic impact on performance

Result

Large corpus

- Google News & Twitter

Method	AUC	Ternary F1	τ
SENTPROP	90.6	58.6	0.44
DENSIFIER	**93.3**	**62.1**	**0.50**
WordNet	89.5	58.7	0.34
Majority	–	24.8	–

Method	AUC	Ternary F1	τ
SENTPROP	86.0	**60.1**	0.50
DENSIFIER	**90.1**	59.4	**0.57**
Sentiment140	86.2	57.7	0.51
Majority	–	24.9	–

Modest corpus

- Finance & COHA

Method	AUC	Ternary F1
SENTPROP	**91.6**	**63.1**
DENSIFIER	80.2	50.3
PMI	86.1	49.8
Velikovich et al. (2010)	81.6	51.1
Majority	–	23.6

Method	AUC	Ternary F1	τ
SENTPROP	**83.8**	**53.0**	**0.28**
DENSIFIER	77.4	46.6	0.19
PMI	70.6	41.9	0.16
Velikovich et al. (2010)	52.7	32.9	0.01
Majority	–	24.3	–

Outline

- Motivations
- Framework
- Recreating known lexicons
- *Inducing lexicons*

Community-specific lexicons

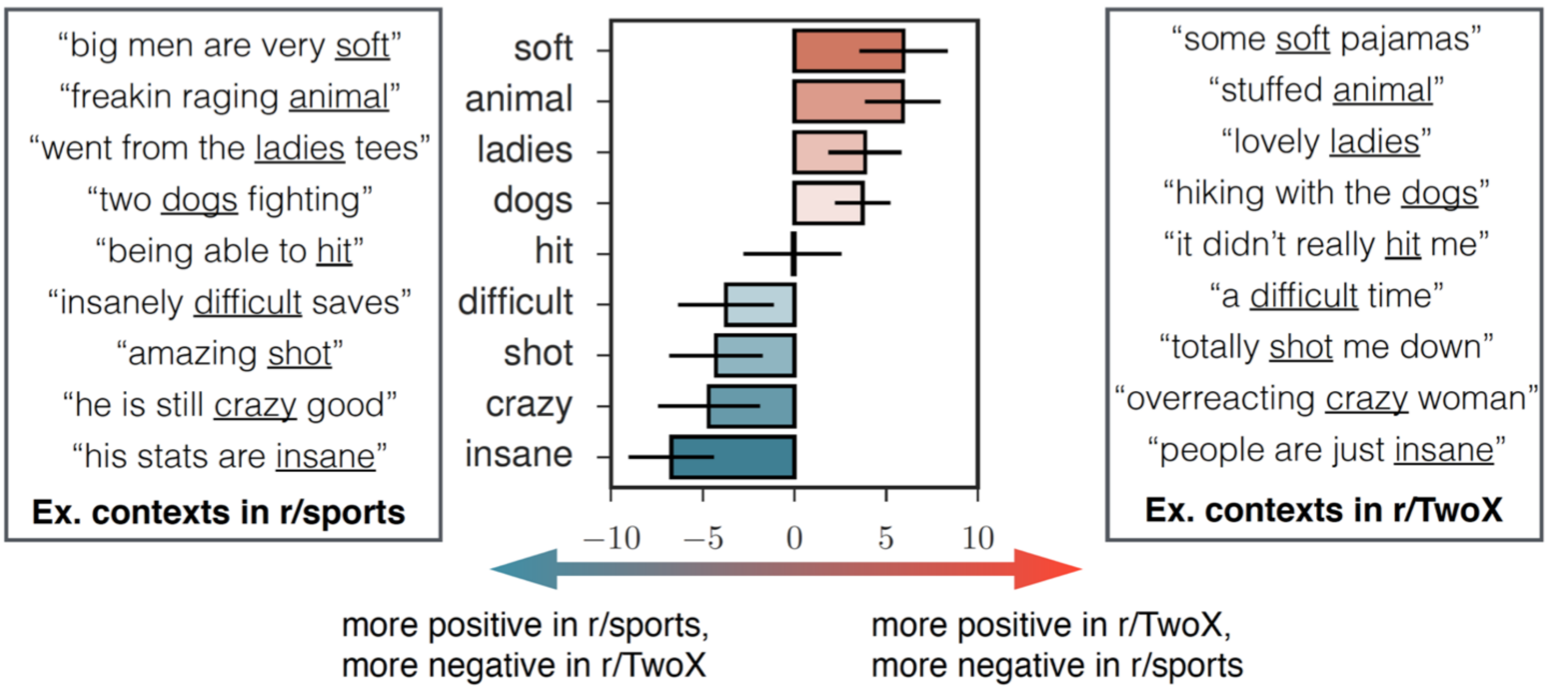
Setup

- Twitter seeds
- Comments of top 250 sub-Reddits 2014
  Top 5000 non-stop words
- Random set of 1000 community pairs
  Overlap more than half top 5000 words. τ is computed on top 25% shared words in case of perventing noise and neutral words

Findings

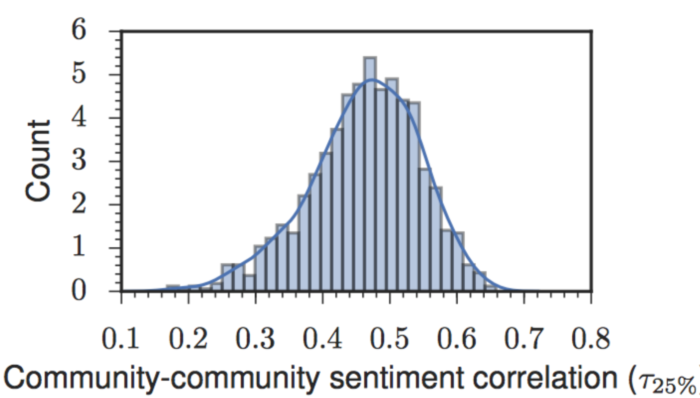
- Sentiment can differ across communities
- Conflicting communities may have similar sentiment

Results



r/TwoX & r/Sports : $\tau_{25} = 0.41$

r/TwoX & r/TheRedPill : $\tau_{25} = 0.58$



Historical lexicons

Setup

- All adijectives with counts above 100 in a given decade and top 5000 non-stop words each year
- Ternary classification
- Phenomena: amelioration / pejoration

Result

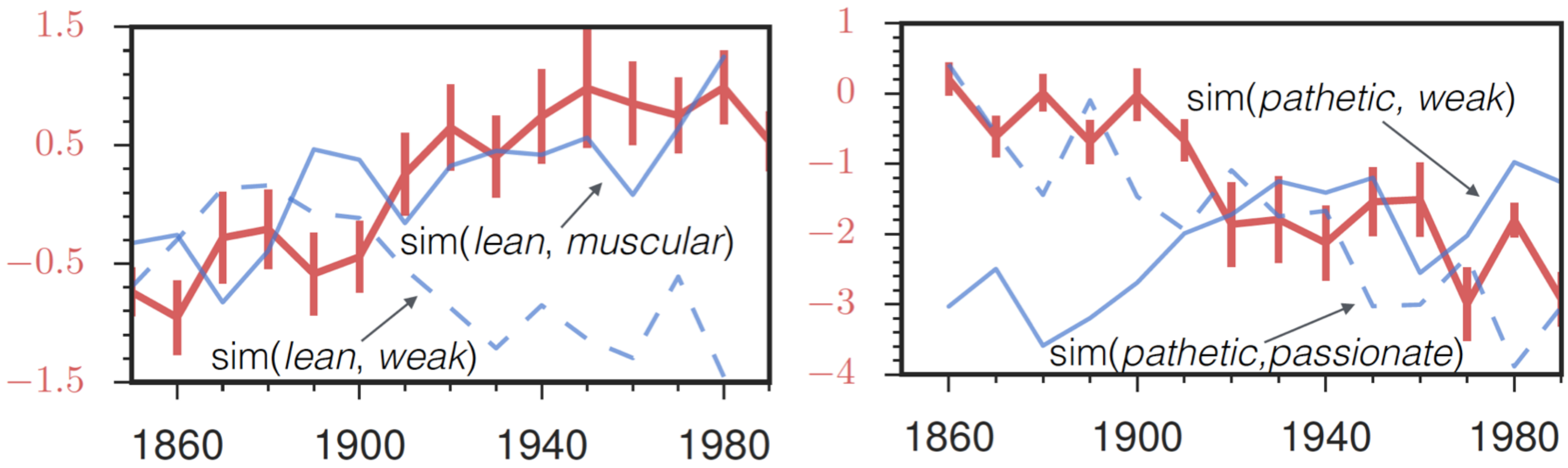
- Over 5% words completely changed in 150 years
- Over 25% words changed in 150 years

Examples

- *lean*, *pathetic* and *sorry*

Results

- Amelioration VS. Pejoration



Thank You!!!