

Predicting and Analyzing Language Specificity in Social Media Posts

Yifan Gao*, Yang Zhong*, Daniel Preotiuc-Pietro and Junyi Jessy Li

UT Austin

UT Austin

Bloomberg LP

UT Austin

* Equal contribution

Motivation

**Specific**

@similar length



I was eating a cookie over the trashcan to catch the crumbs. One of the chocolate chips jumped ship. I considered diving after it. [#Finals](#)

[← Reply](#) [↻ Retweet](#) [★ Favorite](#) [⋮ More](#)

3:33 PM - 13 Nov 18 · Embed this Tweet

**General**

@similar length



How curious that I, looking at you, am thinking exactly what you, looking at me, are thinking and yet... we both cannot understand

[← Reply](#) [↻ Retweet](#) [★ Favorite](#) [⋮ More](#)

4:26 PM - 13 Sep 17 · Embed this Tweet

Specificity

Texts vary in their levels of detail and their involvement of particular concepts, objects and events, i.e., their *specificity*.

The quality of belonging or relating uniquely to a particular subject.
-- Oxford Dictionary

Language specificity

- Impacts reading comprehension
 - (Dixon 1987)
- Connected to several discourse relations
 - (Mann and Thompson 1988)
- Potentially linked to autism
 - (Li et al. 2017)

NLP applications

- Dialogue generation
 - systems favor generic responses such as “I don’t know”
 - (Sordoni et al., 2015; Mou et al., 2016; Li et al 2017)
- Text summarization
 - (Louis and Nenkova, 2011)
- Text quality prediction
 - (Louis and Nenkova, 2013)
- Argumentation mining
 - (Swanson et al., 2015)
- Assessing classroom discussions
 - (Luo and Litman, 2016; Lugini and Litman, 2017)

Specificity prediction: current state

- Training data on news corpora
- Training labels are binary (sentence is either general or specific)

Specificity

A fine grained scale for specificity.



1 = Very General

@general



 Follow

Always too much.



2:48 PM - 27 Jan 2019



2 = General

@general



 Follow

I've been waiting all of August for tomorrow.



2:48 PM - 27 Jan 2019



3 = Specific

@specific



 Follow

So many people at the Turner Center. That will be over by next week.



2:48 PM - 27 Jan 2019



4 = Very specific

@specific



 Follow

The dance this morning on GMA with Alicia Keys made my day :)



2:48 PM - 27 Jan 2019



5 = Extremely Specific

@specific



 Follow

Mascara is the most commonly worn cosmetic, and women will spend an average of \$4,000 on it in their lifetimes.



2:48 PM - 27 Jan 2019

Research Questions

Do personal traits impact one's language specificity?

1

Do temporal factors impact specificity?

2

Can language specificity be predicted accurately?

3

Is specificity prediction useful in a downstream task?

4

Agenda

- Data
- Demographic Analysis
- Temporal Analysis
- Modeling
- Task-based evaluation: mental health (depression) analysis

Data

We introduce a data set of tweets annotated with specificity

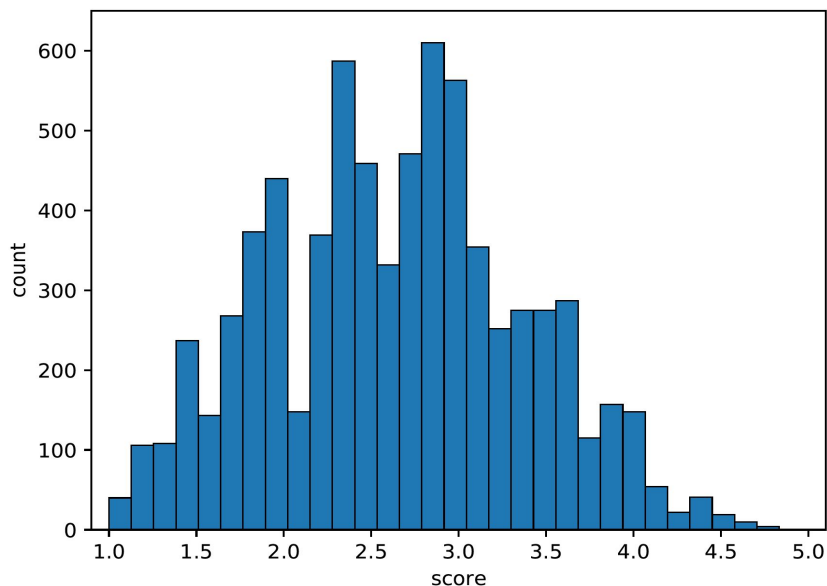
- 7267 tweets
- Sourced from 3665 users with self-reported demographic information
 - (Preotiuc-Pietro et al., 2017)
 - Gender, age, education, income level, faith, political ideology
- Corpus and code available at <https://github.com/cs329yangzhong/specificityTwitter>

Data

Annotated for specificity

- 5-point Likert scale
- MTurk
- IAA - Krippendorff's Alpha of 0.507
- ≥ 5 annotations/instance
- QC - experienced US workers; excluded annotators with < 0.3 agreement with majority of others
- Specificity score - average of annotations

Specificity distribution



- Most tweets fall in the middle of the specificity spectrum
- Shows the utility of the fine-grained rating scale

Research Questions

Do personal traits impact one's language specificity?

1

Do temporal factors impact specificity?

2

Can language specificity be predicted accurately?

3

Is specificity prediction useful in a downstream task?

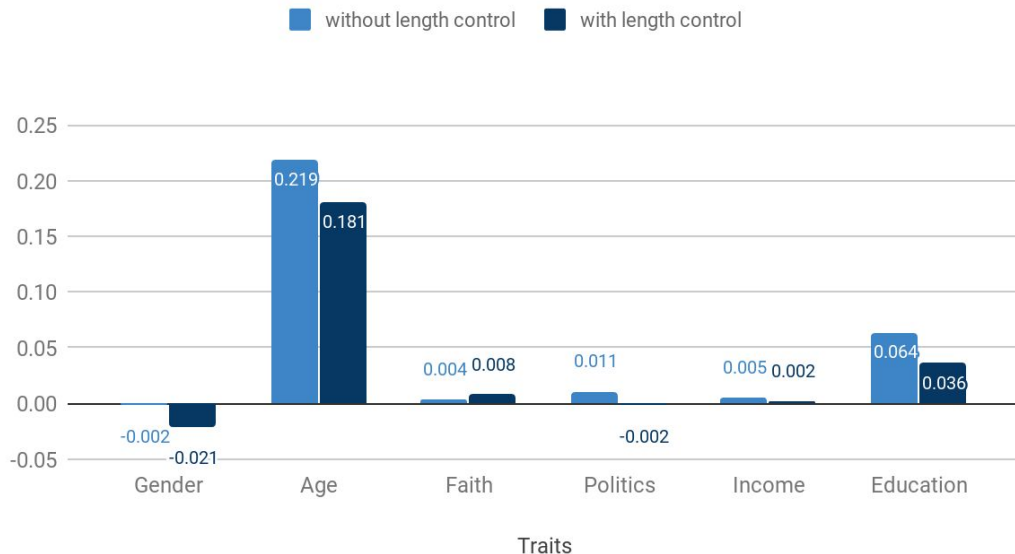
4

Demographic Analysis

Pearson Correlation

- Dependent variable:
average specificity
(per user)
- Control for gender & age
- Control for tweet length
- Bonferroni corrected for
multiple comparisons

Pearson Correlation



Demographic Analysis

Specificity is correlated to:

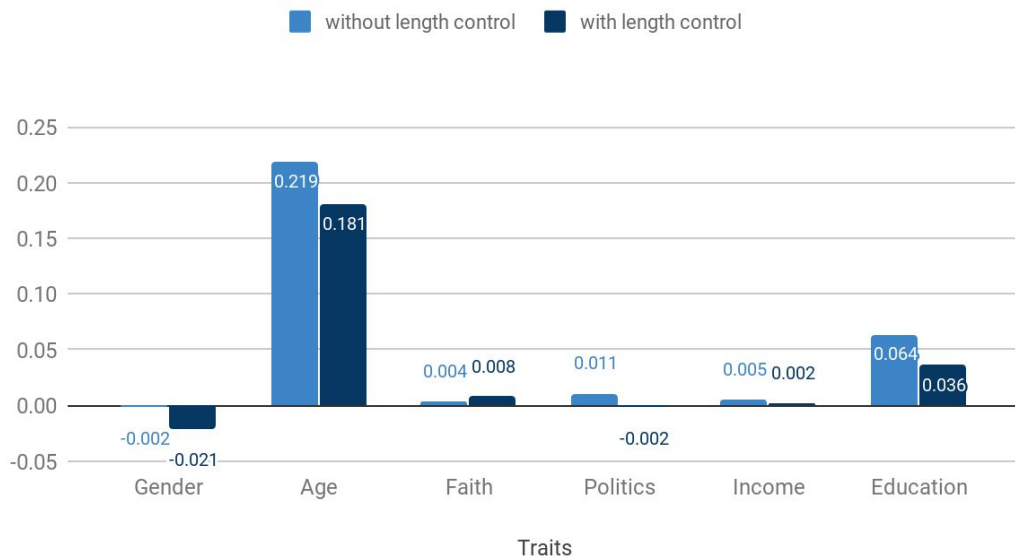
- Age

Correlation remains significant after controlling for tweet length

- Education

Even when controlled for age

Pearson Correlation



Research Questions

Do personal
traits impact
one's
language
specificity?

1

Yes!

Research Questions

Do personal traits impact one's language specificity?

1

Do temporal factors impact specificity?

2

Can language specificity be predicted accurately?

3

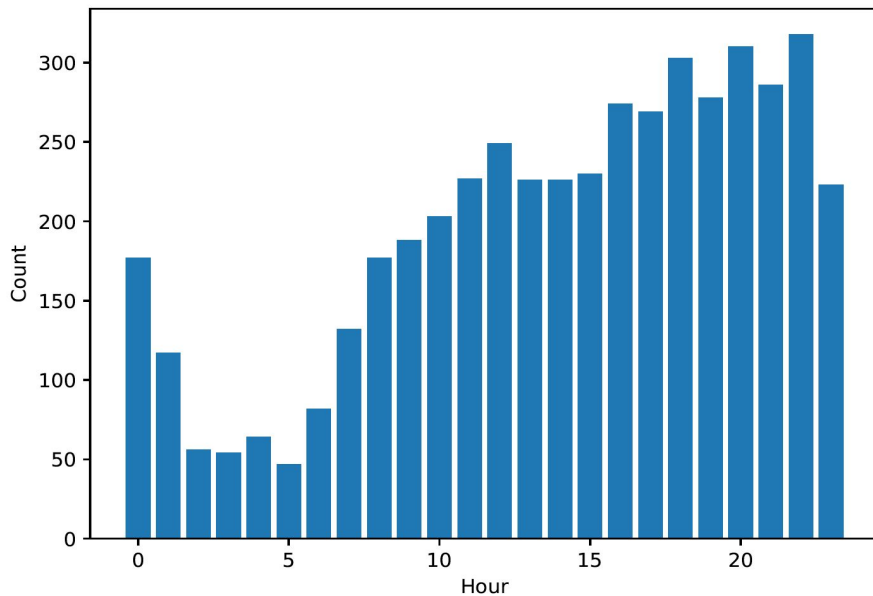
Is specificity prediction useful in a downstream task?

4

Temporal Analysis

- Subset of 4716 tweets can be recovered to their local post time from tweet metadata.
- Example:
 - Original: Mon Dec 21 01:33:05 +0000 2015 -18000
 - Fixed: Mon Dec 20 20:33:05 +0000 2015 (Correct Local Time)

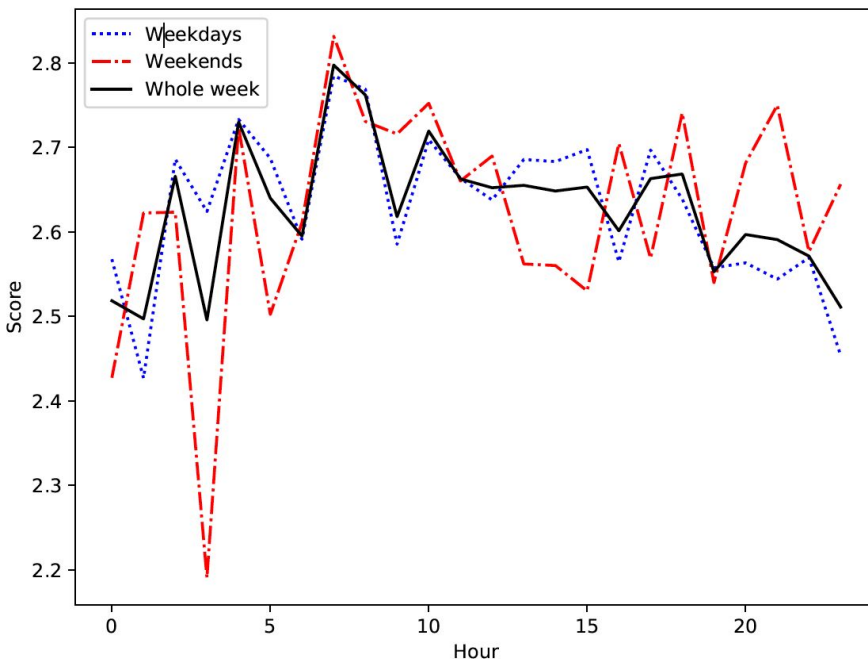
Temporal Analysis



tweets posted across hours of a day.

Temporal Analysis

- Morning tweets are more specific
- Specificity decreases throughout the day
- People tweet less specifically in the evening despite tweeting more frequently



Average tweet specificity throughout the day.

Research Questions

Do temporal factors impact specificity?

2

Yes!

Research Questions

Do personal traits impact one's language specificity?

1

Do temporal factors impact specificity?

2

Can language specificity be predicted accurately?

3

Is specificity prediction useful in a downstream task?

4

Feature Types

- Surface and lexical features
- Tweet content
- Social media content
- Emotion features

Surface and Lexical features

Named Entities

- Indicators of named entities from seven categories:
 - Location, Person, Organization, Money, Percent, Date, and Time

Part-of-speech Tag (POS)

- POS Tag count over eight categories:
 - determiners, nouns, proper nouns, pronouns, adjectives, prepositions, and punctuation.

Concreteness Score

- Average word concreteness score using a lexicon
 - (Brysbaet, Warriner and Kuperman 2014)

Surface and Lexical features

String surface features

- Tweet length (number of tokens)
- Average word length
- Percent of:
 - capital letters
 - numbers
 - non-alphanumeric symbols including punctuation marks

Tweet Content

Word Embeddings

- Average of the embeddings of words in the tweet
 - Pre-trained GloVe word embeddings trained over 2 billion tweets
 - (Pennington, Socher, and Manning 2014)

Brown Clusters

- Count of words in each pre-trained cluster divided by the tweet length.
 - (Turian, Ratinov, and Bengio 2010)

Social Media Content

Captures specific Twitter behaviors:

- URLs
- Contains user mentions
- Is a reply

Emotion Features

Emotional and subjective tweets tend to be less general

(Li and Nenkova 2015, 2016)

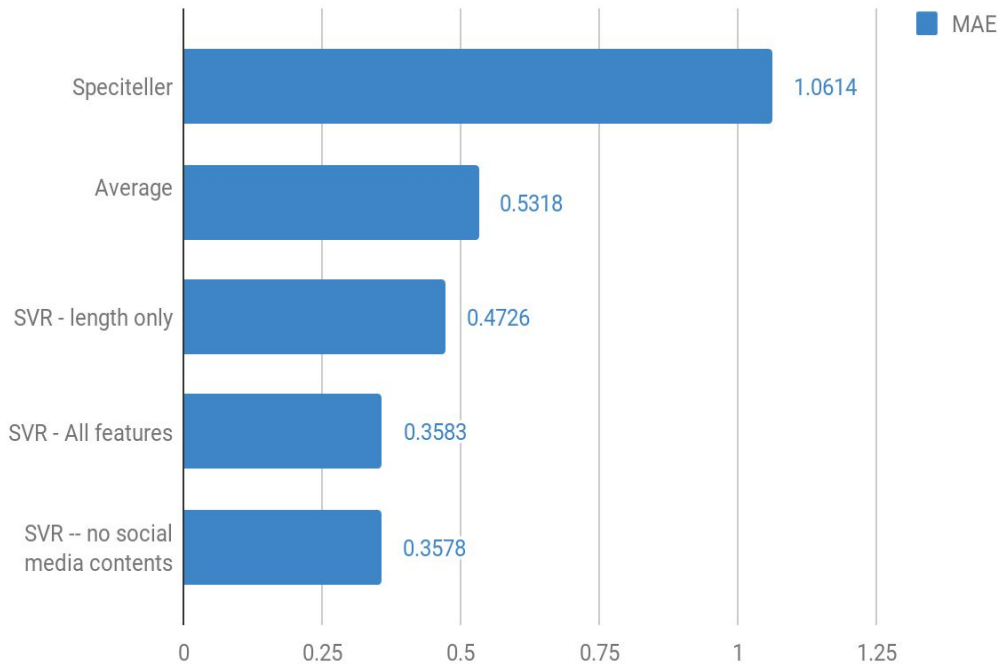
- **Indicators of positive and negative words**
 - Using a sentiment lexicon of 6,800 English words
 - (Hu and Liu, 2004)
- **The percentage of tokens that are emojis**

Predictive Baselines

- **Average**
 - Average specificity from training data
- **Tweet length**
 - Tweet length as the only feature in a SVR model
- **Speciteller**
 - The only available sentence specificity predictor
 - Trained on news corpora (Li and Nenkova 2015)

Modeling Results

- Support Vector Regression
- Data Split:
 - Train: 5767
 - Dev: 500
 - Test: 1000
- Result:
 - Significantly outperforms all baselines



Examples

Tweet	Gold	Spiceller	Tweet Length	Our model
IM HYSTERICALLY CRYING IM SO PROUD OF THEM	1.86	4.91	2.33	2.15
It's November!!!! Crunch Time Bro!!!! Let's Do It!!!	2.43	4.92	2.96	2.37
just reached level 15 on Paradise Island on my Android.	3.63	3.77	2.50	3.59
I'm at Macadams Bar & Grill (5833 SW Macadam Ave, Portland)	4.25	4.96	2.86	3.87

Capital letters and symbols often go with specific info in formal writing

Length alone cannot estimate specificity well

Research Questions

Can language
specificity be
predicted
accurately?

3

Yes!

Research Questions

Do personal traits impact one's language specificity?

1

Do temporal factors impact specificity?

2

Can language specificity be predicted accurately?

3

Is specificity prediction useful in a downstream task?

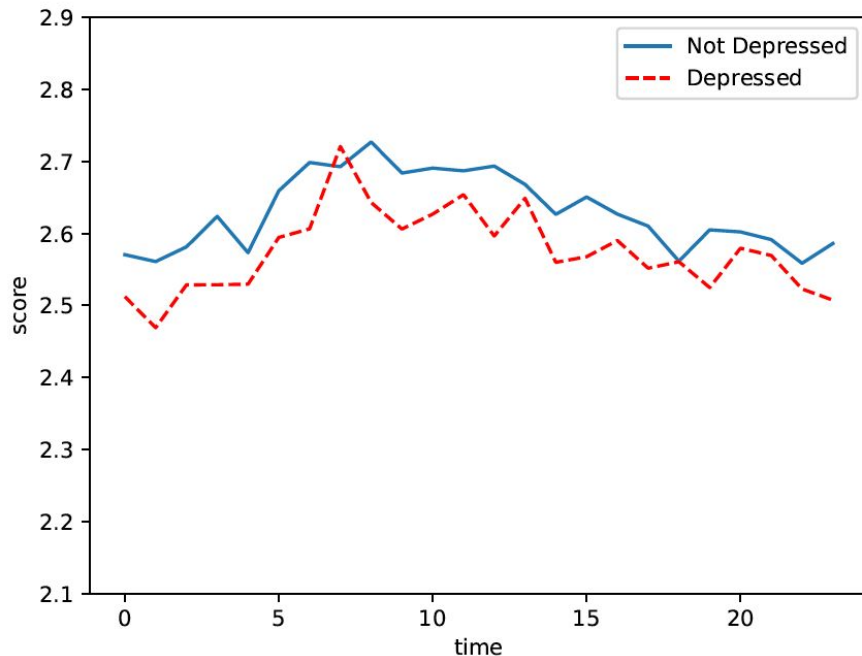
4

Mental Health Analysis

- Does mental health impact one's language specificity?
- A subset of 538 users in our dataset took the Beck Depression Inventory-II (BDI-II) questionnaire.
 - 197 not depressed
 - 95 moderately/severely depressed
- Predict 100 tweets/user with specificity using our best model

Mental Health Analysis

- People with moderate or severe depression tend to write less specific posts
- Regardless of the time of day



Average tweet specificity throughout the day.

Is specificity
prediction
useful in a
downstream
task?

4

Helped with Mental
Health (Depression)
Analysis!

Take Aways

1

New data set and predictive model for fine-grained specificity on social media

2

Social and temporal factors impact language specificity

3

Specificity prediction is useful in downstream analyses

Thank You!



General Thanks

@general



 Follow

Thank you!



2:48 PM - 27 Jan 2019

Contact:

Yifan Gao:

yifan233@utexas.edu

Yang Zhong:

yang.zhong@utexas.edu



Specific Thanks

@specific



 Follow

Thank you for your time and attending our talk on language specificity here at [#AAAI 2019](#) :)



2:48 PM - 27 Jan 2019