Predicting Responses to Microblog Posts

Yoav Artzi¹, Patrick Pantel² and Michael Gamon²

¹University of Washington, Microsoft Research

Work conducted at Microsoft Research
Tweeting on Twitter

A tweet is 140 characters long

Twitter is a social network news agency

Users respond by replying retweeting
The Problem

• Given a tweet

karlhess karl hess
Facebook has become like a terrible party: i don't know 90% of the people there, there's no *booze*, and i keep checking Twitter.
17 hours ago
The Problem

• Given a tweet
• Predict response
  – Reply
  – Retweet

[Image of tweet: karlhess karl hess
Facebook has become like a terrible party: i don't know 90% of the people there, there's no booze, and i keep checking Twitter.
17 hours ago]

[Diagram showing options: Retweet, Reply]
Microsoft Research @MSFTResearch 26 Jan
Steve Hodges of Microsoft Research in Cambridge, UK, explores how hardware and software come together bit.ly/x4KYGn

Microsoft Research @MSFTResearch 25 Jan
Big, burly challenges: a profile of Microsoft Research's Rane Johnson-Stempson bit.ly/wiUMep

Microsoft Research @MSFTResearch 24 Jan
What can you do to protect your reputation online? Check out some tips from #Msft as part of #DataPrivacyDay bit.ly/MSPriv #msftdpd

Microsoft Research @MSFTResearch 24 Jan
Come meet the funniest man in Microsoft Research bit.ly/Ai4Qgq

Microsoft Research @MSFTResearch 23 Jan
Microsoft Research Cambridge's Chris Bishop speaks about his career, and the "best research environment I know": bit.ly/AtCAIA

Microsoft Research @MSFTResearch 23 Jan
Where many different networks come together: a profile of Jennifer Tour Chayes of Microsoft NERD bit.ly/wHxOqL

Microsoft Research @MSFTResearch 19 Jan
MS Research India hosts TechVista in Kolkata. Researchers, academics, gov't to discuss role of CS in the 21st century.
Microsoft Research @MSFTResearch
Steve Hodges of Microsoft Research in Cambridge, UK, explores how hardware and software come together bit.ly/x4KYGn
Expand

Microsoft Research @MSFTResearch
Big, burly challenges: a profile of Microsoft Research’s Rane Johnson-Stempson bit.ly/wiUMep
Expand

Microsoft Research @MSFTResearch
What can you do to protect your reputation online? Check out some tips from #Msft as part of #DataPrivacyDay bit.ly/MSPriv #msftdpd
Expand

Microsoft Research @MSFTResearch
Come meet the funniest man in Microsoft Research bit.ly/Ai4Qgq
Expand

Microsoft Research @MSFTResearch
Microsoft Research Cambridge’s Chris Bishop speaks about his career, and the “best research environment I know”: bit.ly/AtCAiA
Expand

Microsoft Research @MSFTResearch
Where many different networks come together: a profile of Jennifer Tour Chayes of Microsoft NERD bit.ly/wHxOqL
Expand

Microsoft Research @MSFTResearch
MS Research India hosts TechVista in Kolkata. Researchers, academics, gov’t to discuss role of CS in the 21st century.
Microsoft Research @MSFTResearch 26 Jan
Steve Hodges of Microsoft Research in Cambridge, UK, explores how hardware and software come together bit.ly/x4KYGn
Expand

Microsoft Research @MSFTResearch 25 Jan
Big, burly challenges: a profile of Microsoft Research’s Rane Johnson-Stempson bit.ly/wiUMeP
Expand

Microsoft Research @MSFTResearch 24 Jan
What can you do to protect your reputation online? Check out some tips from #Msft as part of #DataPrivacyDay bit.ly/MSPriv #msftdpc
Expand

Microsoft Research @MSFTResearch 24 Jan
Come meet the funniest man in Microsoft Research bit.ly/Ai4Qgq
Expand

Microsoft Research @MSFTResearch 23 Jan
Microsoft Research Cambridge’s Chris Bishop speaks about his career, and the “best research environment I know”: bit.ly/AtCAiA
Expand

Microsoft Research @MSFTResearch 23 Jan
Where many different networks come together: a profile of Jennifer Tour Chayes of Microsoft NERD bit.ly/wHxOlq
Expand

Microsoft Research @MSFTResearch 19 Jan
MS Research India hosts TechVista in Kolkata. Researchers, academics, gov’t to discuss role of CS in the 21st century.
Motivation

• Good indication of impact
• Increases impact
• So who might care about this?
  – Advertisers
  – Celebrities
  – Media organizations
• Also, a way to rank tweets
Goal

• What triggers a response?
• What features are good for prediction?
• Empirical exploration
Our Approach: Learning

Social Network

Tweets + Response → Extract Features → Learner → Model

Boosted Decision Trees
Maximum Entropy*

*MaxEnt by Chris Quirk, Boosted Decision Trees by Qiang Wu
Our Approach: Testing

Social Network -> Model -> Prediction

Tweet
Experimental Setup

• One week of Twitter data
• Searched for response over two weeks
• Randomly sampled training and testing sets:
  – 750K tweets for training
  – 188K tweets for testing
Results

![Precision vs Recall Graph]

- **Precision**
- **Recall**

---

- **Boosted Decision Trees**
- **MaxEnt**
Results

Hard to predict response, for most tweet, but ...
Results

Hard to predict response, for most tweet, but there exists a large set for which we can predict accurately
Results

![Graph showing the relationship between Precision and Recall. The graph plots a line that decreases as Recall increases.]
Building the Model

• What can we get form the language of the tweet?
• Can we use the social network for prediction?
Features: Sentiment

• How the sentiment of a tweet influences the response behavior?
• Count of negative/positive sentiment words*

@michaelaSYKES_
brother helen.

i love the social side of collge; i hate the lesson side.

*Sentiment lexicon provided by Livia Polanyi
Building the Model

![Graph showing precision and recall for sentiment analysis]
Features: Posting

• Tweeter posting trends are influenced by time and day of the week
• Does it influence response behavior?
• Included features:
  – Local time of posting
  – Day of the week
Building the Model

Precision

Recall

+posting

sentiment
Features: Content

• 45 simple features over the content of the tweet
• Manually developed by observing large number of tweets

# stop words
# user references
# hash tags
% non English*
# tokens

RT @yoavartzi: What's "minimally supervised"? How do you prove supervision to be minimal? << good point. lightly sup is better #emnlp

28 Jul via TweetDeck  ⭐ Favorite  ✭ Retweet  ✅ Reply

*English lexicon provided by Lucy Vanderwend
Building the Model

![Graph showing precision and recall for different features: +content, +posting, and sentiment.](image-url)
Features: Lexical Ratio Buckets

• Detect lexical items indicating towards certain response behavior
  – 14M bigrams
  – 400K hashtags
  – Collected from 186M tweets
• Use as flags on each tweet that has them
Features: Lexical Ratio Buckets

• Detect lexical items indicating towards certain response behavior
  – 14M bigrams
  – 400K hashtags
  – Collected from 186M tweets

• Use as flags on each tweet that has them

• Issues:
  – Scalability of learning
  – Sparsity
Features: Lexical Ratio Buckets
Collapsing

• For every lexical item $l$: 

\[
\begin{align*}
\{ & \text{tweets containing } l \text{ that} \\
& \text{received no response} \\
\{ & \text{tweets containing } l \text{ that} \\
& \text{received a response}
\end{align*}
\]
Features: Lexical Ratio Buckets

Collapsing

• For every lexical item \( l \):

\[
\begin{align*}
\text{tweets containing } l \text{ that received no response} \\
\text{tweets containing } l \text{ that received a response}
\end{align*}
\]
Features: Lexical Ratio Buckets

Collapsing

• For every lexical item \( l \):

\[
\frac{\text{tweets containing } l \text{ that received no response}}{\text{tweets containing } l \text{ that received a response}} \rightarrow n
\]
Features: Lexical Ratio Buckets

Collapsing

• For every lexical item $l$:

  - Define each such $n$ as a feature
  - Trigger feature $n$ for each sample that contains $l$
Building the Model
Features: Social

• What are the characteristics of the user’s network?
• Simple social statistics
  – Number of followers
  – Number of followings
Building the Model

Precision

Recall

+socialNet
+lexical
+content
+posting
sentiment
Features: User History

• Aggregate historical response to user
• 3 months of Twitter data
  – Over 2 billion tweets
• Compute statistics
  – For example: ratio of tweets retweeted
Building the Model

![Precision vs Recall with different features]

- **+history**
- **+socialNet**
- **+lexical**
- **+content**
- **+posting**
- **sentiment**
No Local Content Features

![Graph showing precision and recall for different features.](image-url)
No Aggregate Features

![Graph showing precision vs recall for 'All features' and 'No aggregate features']

- **Precision**
- **Recall**

Legend:
- Blue line: All features
- Red line: No aggregate features
Examples

@BoingBoing
Boing Boing

Help find the stolen scripts for GAME OF THRONES goo.gl/d5V14

@ImSoCelebrity
Jeremy Drummond

#IfAliensAttack I hope they kill all people 16 and pregnant.

Response
Examples

@BoingBoing
Boing Boing

Help find the stolen scripts for GAME OF THRONES goo.gl/d5Vl4

@ImSoCelebrity
Jeremy Drummond

#IfAliensAttack I hope they kill all people 16 and pregnant.
Examples

@VidaOfficial
VIDA

Just discovered 'Jamie's Italian'...food is incredible!! Got pure foodbaby now thanks mr oliver! XxCatxX @VidaOfficial

@emilieautumn
Emilie Autumn

On another, more pleasant note (because there always is one, and it's usually a B flat), I ate six apples on camera this weekend.
Examples

@VidaOfficial
VIDA

Just discovered 'Jamie's Italian'...food is incredible!! Got pure foodbaby now
thanks mr oliver! XxCatxX @VidaOfficial

@emilieautumn
Emilie Autumn

On another, more pleasant note (because there always is one, and it's usually a B flat), I ate six apples on camera this weekend.
Examples

@VidaOfficial
VIDA

Just discovered 'Jamie's Italian'...food is incredible!! Got pure foodbaby now thanks mr oliver! XxCatxX @VidaOfficial

Response

@emilieautumn
Emilie Autumn

On another, more pleasant note (because there always is one, and it's usually a B flat), I ate six apples on camera this weekend.

Response
Conclusions

• Local content matters less
  – Or harder to capture
• Despite chronological trends on Twitter, posting time matters less
• Historical behavior is a good indicator
• Twitter is largely a social game
• People are sensitive to certain phrases
Future Work

• New features, such as:
  – Clique specific language features
  – Denseness of user’s social network
  – Mentions of named entities
  – Tweet topic

• Predicting more:
  – Distinguishing between replies and retweets
  – Numerical predictions
  – Predicting length of conversation thread
Thank you for listening
[fin]