Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts

Jiwei Li¹, Alan Ritter², Claire Cardie³ and Eduard Hovy⁴

Department of Computer Science
¹Stanford University ²Ohio State University
³Cornell University ⁴Carnegie Mellon University

June 22nd, 2014
Life Events

Jiwei Li\textsuperscript{1}, Alan Ritter\textsuperscript{2}, Claire Cardie\textsuperscript{3} and Eduard Hovy\textsuperscript{4}
Life Events on Social Media
Life Events on Social Media

Jessica Jones @jonesalgebra · Sep 27
We're engaged!!!! I could not be more thrilled! We are getting married June 12, 2015!

View photo
Life Events on Social Media

Jessica Jones @jonesalgebra · Sep 27
We’re engaged!!! I could not be more thrilled! We are getting married June 12, 2015!

$00\tilde{z}$ @susiezennario · Dec 17
Haha love school! just got accepted by Harvard
Life Events on Social Media

Jessica Jones @jonesalgebra · Sep 27
We're engaged!!!! I could not be more thrilled! We are getting married June 12, 2015!

$00Z @susiezennario · Dec 17
Haha love school: I just got accepted by Harvard

Paloma Camberos @palomacmbrs · 42m
I got a job offer & idk if I should take it 😞.
Life Events on Social Media

Jiwei Li¹, Alan Ritter², Claire Cardie³ and Eduard Hovy⁴

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts
Life Events on Social Media

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts

accepted to MIT. no words can describe how happy i am. guess hard work really does pay off.

Life Event: University Admission
Event Property (University): MIT
Life Events on Social Media

- Life Event: University Admission
  - Event Property (University): MIT

- Life Event: Engagement
  - Event Property (Engaged to): kylatoast
Life Events on Social Media

- **Life Event:** University Admission
  - **Event Property (University):** MIT

- **Life Event:** Engagement
  - **Event Property (Engaged to):** kyloatoast

- **Life Event:** Receiving Award
  - **Event Property (From):** Norway-America Association
Life Events on Social Media

Why?
Why?

- Better understanding of users
Life Events on Social Media

Why?

- Better understanding of users
- Friend Recommendation
Life Events on Social Media

Why?

- Better understanding of users
- Friend Recommendation
- Online advertising
Outline

- Challenges
- System Overview
- Algorithms
- Experiments
- Conclusion
Challenges
Challenge 1: Major life event is an ambiguous concept!
Challenges

Challenge 1: Major life event is an ambiguous concept!
Challenge 1: Major life event is an ambiguous concept!
**Challenge 1**: Major life event is an ambiguous concept!
Challenge 1: Major life event is an ambiguous concept!
Challenge 1: What are life events?
Challenge 1: What are life events?
Challenges

**Challenge 1**: What are life events?
**Challenge 1:** What are life events?
Challenges

Challenge 2: Noisy Data
Challenges

Challenge 2: Noisy Data

Results for *I get married*

Top / All

Blue Shield of CA @BlueShieldCA · Oct 2
Finding the perfect partner & healthcare plan go hand in hand. Start your life together right
Challenge 2: Noisy Data
Challenge 2: Noisy Data

Retweeted 618 times

Love Quotes @LoveQuotes - 21h
I want to get married once. No divorce & no cheating, just us two till the end.

Random Imagination/ Wish
Challenges

Challenge 2: Noisy Data

Random Imagination/Wish

Some other guys
Challenge 2: Noisy Data

Retweeted 618 times

**Love Quotes @LoveQuotes - 21h**

I want to **get married** once. No divorce & no cheating, just us two till the end.

**Marquita Brown @mbrownNR - 25m**

I'm at the #GSO register of deeds office. Two couples are here to **get married**

**Single Dad @Lonely_Dad - Oct 7**

my dreams died when I **got married**.
**Challenge 3:** Lack of labeled data
Challenges

Challenge 3: Lack of labeled data

- No labeling criteria
Challenge 3: Lack of labeled data

- No labeling criteria
- Life events sparsely distributed
Challenge 3: Lack of labeled data

- No labeling criteria
- Life events sparsely distributed
- Rare events
Challenges

HOW ??

Jiwei Li¹, Alan Ritter², Claire Cardie³ and Eduard Hovy⁴

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts
Challenges

I say

I got accepted by Harvard!!
Challenges

I say

I got accepted by Harvard!!

What you would say?
Challenges

I say

I got accepted by Harvard!!

Congratulations!
Challenges

Congratulations!
great!
Fantastic!
Awesome!
Challenges

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts
Responses based Data Harvesting

Seeds:
congrats, fantastic, cool, ....
Responses based Data Harvesting

Seeds: congrats, fantastic, cool, ....

collect

Conversation Text
Responses based Data Harvesting

Seeds: congrats, fantastic, cool, ....

collect

Conversation Text

LDA

Word Clusters (topics)
Responses based Data Harvesting

Seeds: congrats, fantastic, cool, ....

collect

Conversation Text

LDA

Word Clusters (topics)

manual identification

Meaningful Word Clusters
Responses based Data Harvesting

Seeds: congrats, fantastic, cool, ....

collect

Conversation Text

LDA

Word Clusters (topics)

manual identification

Semi-supervised Data harvesting

Meaningful Word Clusters
Responses based Data Harvesting

Semi-supervised Data harvesting

(Kozareva and Hovy, 2010; Davidov et al, 2007; Iggo and Riloff, 2009)
Responses based Data Harvesting

Semi-supervised Data harvesting

Stream-LDA
(Yao et al, 2009)

More Texts

(Kozareva and Hovy, 2010; Davidov et al, 2007; Igo and Riloff, 2009)
Responses based Data Harvesting

Semi-supervised Data harvesting

(Kozareva and Hovy, 2010; Davidov et al, 2007; Igo and Riloff, 2009)

Stream-LDA
(Yao et al, 2009)

More Texts

collect

More Expression Seeds

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts
Responses based Data Harvesting

Semi-supervised Data harvesting

Stream-LDA
(Yao et al, 2009)

More Texts

collect

More Expression Seeds

(Kozareva and Hovy, 2010; Davidov et al, 2007; Igo and Riloff, 2009)

More Texts
Responses based Data Harvesting

Semi-supervised Data harvesting

(Yao et al, 2009)

(Kozareva and Hovy, 2010; Davidov et al, 2007; Igo and Riloff, 2009)

Stream-LDA

More Texts

collect

More Expression Seeds

LDA

More Texts

Word Clusters

Manual identification

Jiwei Li\textsuperscript{1}, Alan Ritter\textsuperscript{2}, Claire Cardie\textsuperscript{3} and Eduard Hovy\textsuperscript{4}

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts
Responses based Data Harvesting

Semi-supervised Data harvesting

(Yao et al, 2009)

(Kozareva and Hovy, 2010; Davidov et al, 2007; Igo and Riloff, 2009)

More Expression Seeds

collect

More Texts

Word Clusters

LDA
Manual identification

More Texts
Responses based Data Harvesting

![Graph showing the relationship between the number of data retrieved and the number of bootstrapping. The graph has two y-axes: one for 'Num of Data Retrieved' ranging from 0 to 80, and another for 'tweet *10^4' ranging from 0 to 80. The x-axis represents 'Num of Bootstrapping' ranging from 0 to 3. There are four lines: blue for 'replies', red for 'topics', and green for 'tweet *10^4'. Each line shows an upward trend as the number of bootstrappings increases.]

Jiwei Li¹, Alan Ritter², Claire Cardie³, and Eduard Hovy⁴

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts
Responses based Data Harvesting

<table>
<thead>
<tr>
<th>Life Event</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birthday</td>
<td>9.78</td>
</tr>
<tr>
<td>Job</td>
<td>8.39</td>
</tr>
<tr>
<td>Wedding</td>
<td>7.24</td>
</tr>
<tr>
<td>Award</td>
<td>6.20</td>
</tr>
<tr>
<td>Sports</td>
<td>6.08</td>
</tr>
<tr>
<td>Anniversary</td>
<td>5.44</td>
</tr>
<tr>
<td>Give Birth</td>
<td>4.28</td>
</tr>
<tr>
<td>Graduate</td>
<td>3.86</td>
</tr>
<tr>
<td>Death</td>
<td>3.80</td>
</tr>
<tr>
<td>Admission</td>
<td>3.54</td>
</tr>
<tr>
<td>Interview</td>
<td>3.44</td>
</tr>
<tr>
<td>Moving</td>
<td>3.26</td>
</tr>
<tr>
<td>Travel</td>
<td>3.24</td>
</tr>
<tr>
<td>Illness</td>
<td>2.45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Life Event</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vacation</td>
<td>2.24</td>
</tr>
<tr>
<td>Relationship</td>
<td>2.16</td>
</tr>
<tr>
<td>Exams</td>
<td>2.02</td>
</tr>
<tr>
<td>Election</td>
<td>1.85</td>
</tr>
<tr>
<td>New Car</td>
<td>1.65</td>
</tr>
<tr>
<td>Running</td>
<td>1.42</td>
</tr>
<tr>
<td>Surgery</td>
<td>1.20</td>
</tr>
<tr>
<td>Lawsuit</td>
<td>0.64</td>
</tr>
<tr>
<td>Acting</td>
<td>0.50</td>
</tr>
<tr>
<td>Research</td>
<td>0.48</td>
</tr>
<tr>
<td>Essay</td>
<td>0.35</td>
</tr>
<tr>
<td>Lost Weight</td>
<td>0.35</td>
</tr>
<tr>
<td>Publishing</td>
<td>0.28</td>
</tr>
<tr>
<td>Song</td>
<td>0.22</td>
</tr>
</tbody>
</table>

Table 1: List of automatically discovered life event types.
### Responses based Data Harvesting

<table>
<thead>
<tr>
<th>Human Label</th>
<th>Top words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wedding</td>
<td>wedding, love, ring, engagement, engaged, bride, video, marrying</td>
</tr>
<tr>
<td>Graduation</td>
<td>graduation, school, college, graduate, graduating, year, grad</td>
</tr>
<tr>
<td>Relationship</td>
<td>boyfriend, girlfriend, date, check, relationship, see, look</td>
</tr>
<tr>
<td>Anniversary</td>
<td>anniversary, years, year, married, celebrating, wife, celebrate, love</td>
</tr>
<tr>
<td>Admission</td>
<td>admitted, university, admission, accepted, college, offer, school</td>
</tr>
<tr>
<td>Exam</td>
<td>passed, exam, test, school, semester, finished, exams, midterms</td>
</tr>
<tr>
<td>Research</td>
<td>research, presentation, journalism, paper, conference, go, writing</td>
</tr>
<tr>
<td>Job</td>
<td>job, accepted, announce, join, joining, offer, starting, announced, work</td>
</tr>
<tr>
<td>Interview</td>
<td>interview, position, accepted, internship, offered, start, work</td>
</tr>
<tr>
<td>Moving</td>
<td>house, moving, move, city, home, car, place, apartment, town, leaving</td>
</tr>
<tr>
<td>Travel</td>
<td>leave, leaving, flight, home, miss, house, airport, packing, morning</td>
</tr>
<tr>
<td>Vacation</td>
<td>vocation, family, trip, country, go, flying, visited, holiday, Hawaii</td>
</tr>
<tr>
<td>Winning Award</td>
<td>won, award, support, awards, winning, honor, scholarship, prize</td>
</tr>
<tr>
<td>Election</td>
<td>president, elected, run, nominated, named, promotion, cel, selected, business, vote</td>
</tr>
<tr>
<td>Publishing</td>
<td>book, sold, writing, finished, read, copy, review, release, books, cover</td>
</tr>
<tr>
<td>Contract</td>
<td>signed, contract, deal, agreements, agreed, produce, dollar, meeting</td>
</tr>
<tr>
<td>song</td>
<td>video, song, album, check, show, see, making, radio, love</td>
</tr>
<tr>
<td>Acting</td>
<td>play, role, acting, drama, played, series, movie, actor, theater</td>
</tr>
<tr>
<td>Death</td>
<td>dies, passed, cancer, family, hospital, dad, grandma, mom, grandpa</td>
</tr>
<tr>
<td>Give Birth</td>
<td>baby, born, boy, pregnant, girl, lbs, name, son, world, daughter, birth</td>
</tr>
<tr>
<td>Illness</td>
<td>ill, hospital, feeling, sick, cold, flu, getting, fever, doctors, cough</td>
</tr>
<tr>
<td>Surgery</td>
<td>surgery, got, test, emergency, blood, tumor, stomachs, hospital, pain, brain</td>
</tr>
<tr>
<td>Sports</td>
<td>win, game, team, season, fans, played, winning, football, luck</td>
</tr>
<tr>
<td>Running</td>
<td>run, race, finished, race, marathon, ran, miles, running, finish, goal</td>
</tr>
<tr>
<td>New Car</td>
<td>car, buy, bought, cars, get, drive, pick, seat, color, dollar, meet</td>
</tr>
<tr>
<td>Lost Weight</td>
<td>weight, lost, week, pounds, loss, weeks, gym, exercise, running</td>
</tr>
</tbody>
</table>
System Overview

Challenges
Response based Data Harvesting
System Overview
Algorithms
Experiments
Conclusion

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts

System Overview

Challenges
Response based Data Harvesting
System Overview
Algorithms
Experiments
Conclusion

Challenges
Response based Data Harvesting
System Overview
Algorithms
Experiments
Conclusion

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts

System Overview

Challenges
Response based Data Harvesting
System Overview
Algorithms
Experiments
Conclusion

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts

System Overview

Challenges
Response based Data Harvesting
System Overview
Algorithms
Experiments
Conclusion

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts

System Overview

Challenges
Response based Data Harvesting
System Overview
Algorithms
Experiments
Conclusion

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts

System Overview

Challenges
Response based Data Harvesting
System Overview
Algorithms
Experiments
Conclusion

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts

System Overview

Challenges
Response based Data Harvesting
System Overview
Algorithms
Experiments
Conclusion

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts

System Overview

Challenges
Response based Data Harvesting
System Overview
Algorithms
Experiments
Conclusion

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts

System Overview

Challenges
Response based Data Harvesting
System Overview
Algorithms
Experiments
Conclusion

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts

System Overview

Challenges
Response based Data Harvesting
System Overview
Algorithms
Experiments
Conclusion

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts
System Overview

Challenges
Response based Data Harvesting
System Overview
Algorithms
Experiments
Conclusion

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts
System Overview

Challenges Response based Data Harvesting

System Overview

Algorithms

Experiments

Conclusion

Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts
**System Overview**

Input:

- I had beef jerky for lunch
- I got married to Tom
- My friend Chris got married.

Output:

- I got married to Tom
  - Event Category: marriage
- My friend Chris got married
  - Event Category: marriage

**Pipeline 1:** Personal Life Event Identification

**Pipeline 2:** Self-reported Information Identification

**Pipeline 3:** Event Property Extraction

Input:

- I had beef jerky for lunch
- I got married to Tom
- My friend Chris got married.

Output:

- I got married to Tom
  - Event Category: marriage
- My friend Chris got married
  - Event Category: marriage

Throw away:

- I had beef jerky for lunch
Major Life Event Extraction from Twitter based on Congratulations/Condolences Speech Acts
Personal Event Identification

Multi-Class Classifier based on SVM
Personal Event Identification

Multi-Class Classifier based on SVM
Positive Examples for each category: Pre-identified data
Personal Event Identification

Multi-Class Classifier based on SVM
Positive Examples for each category: Pre-identified data
Negative Examples: Random Tweets
Personal Event Identification

Multi-Class Classifier based on SVM
Positive Examples for each category: Pre-identified data
Negative Examples: Random Tweets
Personal Event Identification

Multi-Class Classifier based on SVM

Positive Examples for each category: Pre-identified data
Negative Examples: Random Tweets

- Topic-Tweet probability
Personal Event Identification

**Multi-Class Classifier based on SVM**

- Positive Examples for each category: Pre-identified data
- Negative Examples: Random Tweets
  - Topic-Tweet probability
  - Dictionary
Multi-Class Classifier based on SVM

Positive Examples for each category: Pre-identified data
Negative Examples: Random Tweets

- Topic-Tweet probability
- Dictionary
- Word, NER, POS
- Window Context
Personal Event Identification

Multi-Class Classifier based on SVM:

Split harvested data, training and testing

<table>
<thead>
<tr>
<th>Feature Setting</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word+NER</td>
<td>0.204</td>
<td>0.326</td>
</tr>
<tr>
<td>Word+NER+Dictionary</td>
<td>0.362</td>
<td>0.433</td>
</tr>
<tr>
<td>All</td>
<td>0.382</td>
<td>0.487</td>
</tr>
</tbody>
</table>
Self Information Identification

- **Input**: I had beef jerky for lunch
- **Output**: I got married to Tom
  - Event Category: marriage
  - Married to (property): Tom

- **Input**: My friend Chris got married.
- **Output**: I got married to Tom
  - Event Category: marriage
- **Self-reported Information Identification**
  - **Input**: I got married to Tom
  - **Output**: My friend Chris got married.
Self Information Identification

Negative Examples
Self Information Identification

Negative Examples

- Not self
Self Information Identification

Negative Examples

- Not self
- Random Thought
Self Information Identification

Negative Examples

- Not self
- Random Thought
- Past Tense
Self Information Identification

**Dataset:**
Positive: selected from harvested data
Negative: selected from harvested data
Self Information Identification

**Dataset:**
Positive: selected from harvested data
Negative: selected from harvested data

**Binary SVM Classifier**
Self Information Identification

**Dataset:**
Positive: selected from harvested data  
Negative: selected from harvested data

**Binary SVM Classifier**
- Tense
Self Information Identification

**Dataset:**
Positive: selected from harvested data
Negative: selected from harvested data

**Binary SVM Classifier**
- Tense
- Factuality (could, would, can ... ) (Saurf and Pustejovsky, 2007)
Self Information Identification

**Dataset:**
Positive: selected from harvested data
Negative: selected from harvested data

**Binary SVM Classifier**
- Tense
- Factuality (could, would, can ... ) (Saurf and Pustejovsky, 2007)
- I
Self Information Identification

Dataset:
Positive: selected from harvested data
Negative: selected from harvested data

Binary SVM Classifier
- Tense
- Factuality (could, would, can ... ) \cite{Saurf2007}
- I
- Dependency \cite{Kong2014}
Self Information Identification

**Dataset:**
Positive: selected from harvested data
Negative: selected from harvested data

**Binary SVM Classifier**
- Tense
- Factuality (could, would, can ... ) (Saurf and Pustejovsky, 2007)
- I
- Dependency (Kong et al., 2014)
- Token, NER, POS, window context
# Self Information Identification

<table>
<thead>
<tr>
<th>Feature Setting</th>
<th>Acc</th>
<th>Pre</th>
<th>Rec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bigram+Window</td>
<td>0.76</td>
<td>0.47</td>
<td>0.44</td>
</tr>
<tr>
<td>Bigram+Window + Tense + Factuality</td>
<td>0.77</td>
<td>0.47</td>
<td>0.46</td>
</tr>
<tr>
<td>all</td>
<td>0.82</td>
<td>0.51</td>
<td>0.48</td>
</tr>
</tbody>
</table>
Event Property Identification

### Pipelines

**Pipeline 1: Personal Life Event Identification**
- Input: Tweet content
  - Example: I had beef jerky for lunch.
- Output: Event identified
  - Example: I got married to Tom
  - Category: marriage

**Pipeline 2: Self-reported Information Identification**
- Input: Tweet content
  - Example: My friend Chris got married.
- Output: Event identified
  - Example: I got married to Tom
  - Category: marriage

**Pipeline 3: Event Property Extraction**
- Input: Event of interest
  - Example: Married to (property): Tom
- Output: Event properties
  - Example: My friend Chris got married

*Notes*
- Tweets are filtered and processed through these pipelines to extract and identify major life events.
- The system is designed to handle different types of events and their properties efficiently.
Event Property Identification

Human Labeling

<table>
<thead>
<tr>
<th>Life Event</th>
<th>Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Acceptance, Graduation</td>
<td>Name of University/College</td>
</tr>
<tr>
<td>(b) Wedding, Engagement, Falling love</td>
<td>Name of Spouse/ partner/ bf/ gf</td>
</tr>
<tr>
<td>(c) Getting a job, interview, internship</td>
<td>Name of Enterprise</td>
</tr>
<tr>
<td>(d) Moving to New Places, Trip, Vocation, Leaving</td>
<td>Place, Origin, Destination</td>
</tr>
<tr>
<td>(e) Winning Award</td>
<td>Name of Award, Prize</td>
</tr>
</tbody>
</table>
Event Property Identification

**Sequence Labeling Task, CRF** (Lafferty, et al., 2001)
- Word token, Capitalization, POS, word shape, NER
- A gazetteer of universities and companies
- Context
What benefits brought from Congratulations/Condolences Speech Acts?

- Clean Data
What benefits brought from Congratulations/Condolences Speech Acts?

- Clean Data
- Personal Topic Identification
What benefits brought from Congratulations/Condolences Speech Acts?

- Clean Data
  - Personal Topic Identification
  - Self Report Information
System

What benefits brought from Congratulations/Condolences Speech Acts?

- Clean Data
  - Personal Topic Identification
  - Self Report Information
    User 1: I wish to get married
    User 2: Congratulations !!
Experiments

- End-to-End Experiments
Gold-standard life event dataset
Experiments

Gold-standard life event dataset

- Ask Twitter users to label their own tweets
Gold-standard life event dataset

- Ask Twitter users to label their own tweets
- Ask Turkers to label other people’s tweets.
Experiments

Gold-standard life event dataset

- Ask Twitter users to label their own tweets
- Ask Turkers to label other people’s tweets.
  - 2 Turkers 1 tweet

Inter-rater agreement is 0.54 (Cohen’s kappa)
Authors make final decision

900 positive tweets
60,000 negative tweets
Experiments

Gold-standard life event dataset

- Ask Twitter users to label their own tweets
- Ask Turkers to label other people’s tweets.
  - 2 Turkers 1 tweet
  - Inter-rater agreement is 0.54 (cohen’s kappa)
Experiments

Gold-standard life event dataset

- Ask Twitter users to label their own tweets
- Ask Turkers to label other people’s tweets.
  - 2 Turkers 1 tweet
  - Inter-rater agreement is 0.54 (cohen’s kappa)
  - Authors make final decision
Experiments

Gold-standard life event dataset

- Ask Twitter users to label their own tweets
- Ask Turkers to label other people’s tweets.
  - 2 Turkers 1 tweet
  - Inter-rater agreement is 0.54 (cohen’s kappa)
  - Authors make final decision
- 900 positive tweets
Experiments

Gold-standard life event dataset

- Ask Twitter users to label their own tweets
- Ask Turkers to label other people’s tweets.
  - 2 Turkers 1 tweet
  - Inter-rater agreement is 0.54 (cohen’s kappa)
  - Authors make final decision
- 900 positive tweets
- 60,000 negative tweets
Experiments

Baselines
Experiments

Baselines

- Supervised

Table 3: Performance for different approaches for identifying life events.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supervised</td>
<td>0.13</td>
<td>0.20</td>
</tr>
<tr>
<td>Supervised + Self</td>
<td>0.25</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Jiwei Li$^1$, Alan Ritter$^2$, Claire Cardie$^3$ and Eduard Hovy$^4$
Experiments

Baselines

- Supervised
- Supervised + Self
Experiments

Baselines

- Supervised
- Supervised + Self

<table>
<thead>
<tr>
<th>Approach</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our approach</td>
<td>0.62</td>
<td>0.48</td>
</tr>
<tr>
<td>Supervised</td>
<td>0.13</td>
<td>0.20</td>
</tr>
<tr>
<td>Supervised+Self</td>
<td>0.25</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Table 3: Performance for different approaches for identifying life events.
Experiments

Does bootstrapping help?
Does bootstrapping help?

Table 4: Performance for different steps of bootstrapping for identifying.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>0.65</td>
<td>0.36</td>
</tr>
<tr>
<td>Step 2</td>
<td>0.64</td>
<td>0.43</td>
</tr>
<tr>
<td>Step 3</td>
<td>0.62</td>
<td>0.48</td>
</tr>
</tbody>
</table>
Conclusion

Conclusion
We study the life event extraction problem on Twitter. We propose a framework based on Congratulations/Condolences Speech Acts for data harvesting. We explore different types of features and algorithms for this task.
Conclusion

- We study the life event extraction problem on Twitter
Conclusion

- We study the life event extraction problem on Twitter
- We propose a framework based on Congratulations/Condolences Speech Acts for data harvesting
Conclusion

- We study the life event extraction problem on Twitter
- We propose a framework based on Congratulations/Condolences Speech Acts for data harvesting
- We explore different types features and algorithms for this task
Conclusion

Key idea: solve this problem based on minimum human efforts.
Conclusion

Key idea: solve this problem based on minimum human efforts.

Problems
Conclusion

Key idea: solve this problem based on minimum human efforts.

Problems

- Restricted to event types identified by Congratulations/Condolences Speech Acts.
Conclusion

Key idea: solve this problem based on minimum human efforts.

Problems

- Restricted to event types identified by Congratulations/Condolences Speech Acts.
- No all responses correspond to life events
Conclusion

Key idea: solve this problem based on minimum human efforts.

Problems

- Restricted to event types identified by Congratulations/Condolences Speech Acts.
- No all responses correspond to life events
- Error accumulations.
Thank you!
Thank you!

Questions, Suggestions
Thank you!

Questions, Suggestions

Joint work with

Alan Ritter
Claire Cardie
Eduard Hovy