Twitter & Polls: Analyzing and estimating political orientation of Twitter users in India General #Elections2014

Abhishek Bhola
Advisor: Dr. Ponnurangam Kumaraguru

M.Tech Thesis Defense
06-June-2014
Thesis Committee

- Dr. Ayesha Choudhary, JNU
- Dr. Vinayak Naik, IIIT-Delhi
- Dr. PK (Chair), IIIT-Delhi
Presentation Outline

• Research Motivation
• Research Aim
• Related Work
• Research Gap
• Data Analysis
• Classification of Political Orientation
• System Design
• Conclusion
• Limitations and Future Work
Presentation Outline

• Research Motivation
• Research Aim
• Related Work
• Research Gap
• Data Analysis
• Classification of Political Orientation
• System Design
• Conclusion
• Limitations and Future Work
India General Elections 2014

• Elections for the 16th Lok Sabha
• Total Seats: 543
• Number of Registered Voters: 813 million
• Newly registered voters: 100 million
• Money at stake: $5 billion
• Number of parties registered with the Election Commission: 1616
• National Parties: 6
• State Parties: 47
• Number of candidates: 8000
India General Elections 2014
Elections and Social Media

• Major parties battling it out:
  – Aam Aadmi Party (AAP)
  – Bhartiya Janta Party (BJP)
  – Indian National Congress (INC)

• Their PM Candidates:
  – Arvind Kejriwal
  – Narendra Modi
  – Rahul Gandhi

• Internet Users in India: 243 million
  (by June 2014)

• Facebook Users: 114.8 million
• Twitter Users: 33 million
Elections and Social Media

- Google+ Hangout: Interaction with party workers
- What’sApp: To send bulk messages
- Facebook: Televised Interviews and ad campaigns
- Instagram: Pictures of party rallies uploaded
- YouTube: Videos of rallies uploaded
- Google’s Election Hub
Research Motivation

- Social Media could sway 3-4% of urban votes: IAMAI
- Increase in elections related data- 600% from 2009
- Almost all leaders and parties are on Twitter
- Extensively used for communicating and interaction
Presentation Outline

• Research Motivation
• Research Aim
• Related Work
• Research Gap
• Data Analysis
• Classification of Political Orientation
• System Design
• Conclusion
• Limitations and Future Work
Research Aim

• To analyze and draw meaningful inferences from the collection of tweets collected over the entire duration of elections
• To check the feasibility of development of a classification model to identify the political orientation of the twitter users based on the tweet content and other user based features.
• To develop a system to analyze and monitor the election related tweets on daily basis.
Presentation Outline

• Research Motivation
• Research Aim
• Related Work
• Research Gap
• Data Analysis
• Classification of Political Orientation
• System Design
• Conclusion
• Limitations and Future Work
Tumasjan et. al. *Predicting elections with twitter: What 140 characters reveal about political sentiment*, 2010

Jungherr et. al. *Why the pirate party won the german election of 2009 or the trouble with predictions*, 2012


Conover et. al. *Predicting the political alignment of twitter users*, 2011
Related Work

Cohen et. al.  
*Classifying political orientation on twitter: Its not easy!, 2013*

Simplify360: Calculation of SSI

NExT Centre, NUS Weekly Infographics

Twitritis + Wright State University

Political Figures

Politically Active

Politically modest
Presentation Outline

• Research Motivation
• Research Aim
• Related Work
• Research Gap
• Data Analysis
• Classification of Political Orientation
• System Design
• Conclusion
• Limitations and Future Work
Research Gap

• No previous attempts to classify the political orientation of users in the Indian scenario
• No previous work explored both ‘Pro’ as well as ‘Anti’ views
Presentation Outline

• Research Motivation
• Research Aim
• Related Work
• Research Gap
• Data Analysis
• Classification of Political Orientation
• System Design
• Conclusion
• Limitations and Future Work
Data Collection

Keywords for Election related tweets

Elections related tweets
Data Collection

Screen names of Twitter Profiles

130 Twitter Profiles
<table>
<thead>
<tr>
<th>Twitter Data</th>
<th>What we had with us:</th>
</tr>
</thead>
<tbody>
<tr>
<td>As per reports, what Twitter had:</td>
<td>• #Tweets, Sept 25 – May 16: 18.21 million</td>
</tr>
<tr>
<td></td>
<td>• #Tweets, Jan 1- May 12: 13.09 million</td>
</tr>
<tr>
<td></td>
<td>• 600% rise in #Tweets from 2009 elections</td>
</tr>
<tr>
<td></td>
<td>• 2009 elections only 1 politician had an account with 6K followers</td>
</tr>
<tr>
<td></td>
<td>• 23.37 % of Tweets</td>
</tr>
<tr>
<td></td>
<td>• Difference in the keywords used</td>
</tr>
</tbody>
</table>


† Twitter’s official blog: https://blog.twitter.com/2014/indias-2014-twitterelection
Methodology for Analysis

• Tweets were picked from the database
• Different fields were exploited for different analysis
• Python, Matplotlib and Excel were used to plot graphs
20th Jan: Arvind Kejriwal protests against Delhi Police
27th Jan: Rahul Gandhi’s interview with Arnab
05th Mar: Election dates declared
14th Feb: Kejriwal resigned as Delhi CM
07th Apr: 1st phase of elections and BJP’s manifesto
Hourly Frequency Analysis

- Per hour tweets frequency
- Mean: 4073.08 Tweets
- Std deviation: 3072.77
- UCL: 10218.63

Timeline (Hours of each day)

Number of Tweets per hour

27th Jan 2100hrs, Rahul Gandhi’s Interview

05th Mar, 1900hrs, ECI declares election dates

23
Hour v/s Day of the Week

- Max #Tweets during weekdays
- Max Tweets on Tuesdays and Wednesdays
- Tweeting activity goes higher during the second half of the day
- Most of the relevant events were on weekdays
Who tweets how much

- Inverse Law Proportion
- Graph for the tweets of month of all the months
- April has the highest number of unique users
- April is the month with a single user tweeting $>10^4$ times
- **815,425** total unique users
## Top 5 active tweeters

<table>
<thead>
<tr>
<th>Username</th>
<th>Followers</th>
</tr>
</thead>
<tbody>
<tr>
<td>BJP4India</td>
<td>82,280</td>
</tr>
<tr>
<td>Bond_2014-32</td>
<td>32,927</td>
</tr>
<tr>
<td>Tips4DayTrader</td>
<td>24,117</td>
</tr>
<tr>
<td>MINDMONEY</td>
<td>23,283</td>
</tr>
<tr>
<td>AshikonFire</td>
<td>19,611</td>
</tr>
</tbody>
</table>

Clearly supporting BJP and anti AAP or Congress.
Which party received maximum mentions

![Graph showing #Tweets over weeks for BJP, AAP, and Cong]
Comparing electoral results with tweet share

General Elections 2014

Delhi Assembly Elections 2013

AAP  BJP  Cong

Percentage of vote share  Percentage of tweet share

AAP  BJP  Cong

Percentage of vote share  Percentage of tweet share
No Hashtags ‘AntiBJP’ was found in the top 5 list over all the weeks.
Analyzing the popularity of Modi & Kejriwal: #Followers

Date

#Followers
0 500000 1000000 1500000 2000000 2500000 3000000 3500000 4000000

Kejriwal  Modi
Analyzing the popularity of Modi & Kejriwal: #Followers

Percentage Change in #Followers

Date

Analyzing the popularity of Modi & Kejriwal: Retweet Frequency on Tweets

Kejriwal

Avg = 887

Modi

Avg = 612
Tradeoff b/w #Followers & Retweet Frequency

• Klout score uses a lot of factors viz.,
  – #Followers
  – #Friends
  – #Retweets on each tweet
• Pearson’s correlation between
  – #Followers and Klout score
  – Avg. #retweets on tweets and Klout score

Narendra Modi’s popularity was more than Arvind Kejriwal based on the number of followers
Presentation Outline

• Research Motivation
• Research Aim
• Related Work
• Research Gap
• Data Analysis
• Classification of Political Orientation
• System Design
• Conclusion
• Limitations and Future Work
1000 random twitter user profiles tweeting about India Elections were selected

10 Data Annotators decided their political orientation on the basis of tweets

Pro
• AAP
• BJP
• Cong
• Can’t Say

Anti
• AAP
• BJP
• Cong
• Can’t Say

Other information about these profiles such as #Followers, #Friends, Tweets between Mar 20 – Apr 10 was also collected
Agreement between the annotators

- Confusion Matrix for the 1st set of 250 instances (Pro)

<table>
<thead>
<tr>
<th></th>
<th>Annotator 1</th>
<th>Annotator 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AAP</td>
<td>BJP</td>
</tr>
<tr>
<td>AAP</td>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>BJP</td>
<td>6</td>
<td>76</td>
</tr>
<tr>
<td>CONG</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>CAN’T SAY</td>
<td>11</td>
<td>11</td>
</tr>
</tbody>
</table>

Observed agreement: 183/250 = 0.732

Hypothetical
Chance agreement

Cohen’s Kappa coefficient,
\[ \kappa = \frac{Pr(a) - Pr(e)}{1 - Pr(e)} \]

- \( Pr(e) = 0.375 \)
- \( \kappa = 0.571 \)
# Annotation Results

<table>
<thead>
<tr>
<th>Party</th>
<th>Pro</th>
<th>Anti</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>133</td>
<td>205</td>
</tr>
<tr>
<td>BJP</td>
<td>447</td>
<td>85</td>
</tr>
<tr>
<td>CONG</td>
<td>33</td>
<td>135</td>
</tr>
<tr>
<td>CAN’T SAY</td>
<td>387</td>
<td>575</td>
</tr>
</tbody>
</table>

### PRO

- AAP: 39%
- BJP: 45%
- CAN’T SAY: 3%
- CONG: 13%

### ANTI

- AAP: 20%
- BJP: 8%
- CAN’T SAY: 58%
- CONG: 14%
Text Based Classification

- Text of 200 tweets collected
- Stop words, URLs, Hashtags and user mentions removed
- Words like ‘RT’, ‘&amp’, ‘ka’, ‘ke’, ‘ki’ etc. were also removed
- Vector based on TF-IDF of every term and each user was formed

\[
TF_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}}
\]

Number of times the term ‘i’ is used by user ‘j’

\[
IDF_i = \log \left( \frac{|U|}{1 + |U_i|} \right)
\]

Total users

Number of terms used by the user ‘j’ in ‘k’ tweets

Users using the term ‘i’
## Text Based Classification: Results

### Results for all 613 ‘Pro’ Instances

<table>
<thead>
<tr>
<th>Party</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>0.381</td>
<td>0.061</td>
<td>0.105</td>
</tr>
<tr>
<td>BJP</td>
<td>0.736</td>
<td>0.975</td>
<td>0.839</td>
</tr>
<tr>
<td>CONG</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Instances:** 613  
**Attributes:** 9312  
**Classifier:** Random Forest, 10 folds cross-validation  
**Efficiency:** 72.36%

### Results for equal ‘Pro’ Instances

<table>
<thead>
<tr>
<th>Party</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>0.424</td>
<td>0.758</td>
<td>0.543</td>
</tr>
<tr>
<td>BJP</td>
<td>0.481</td>
<td>0.394</td>
<td>0.433</td>
</tr>
<tr>
<td>CONG</td>
<td>0.308</td>
<td>0.121</td>
<td>0.174</td>
</tr>
</tbody>
</table>

**Instances:** 99  
**Attributes:** 2442  
**Classifier:** Random Forest, 10 folds cross-validation  
**Efficiency:** 42.42%

- **Good Efficiency**
- **But 0 Precision and Recall for Congress**

- We tried 2-class classification with all 3 possible pairs and got 65.15% efficiency for AAP-BJP

- **Efficiency falls**
- **Precision and Recall for Congress not 0**
## Text Based Classification: Results

### Results for all 425 ‘Anti’ Instances

<table>
<thead>
<tr>
<th>Party</th>
<th>Pre</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>0.489</td>
<td>0.863</td>
<td>0.624</td>
</tr>
<tr>
<td>BJP</td>
<td>0.313</td>
<td>0.059</td>
<td>0.099</td>
</tr>
<tr>
<td>CONG</td>
<td>0.447</td>
<td>0.157</td>
<td>0.232</td>
</tr>
</tbody>
</table>

**Classifier:** Random Forest

**Efficiency:** 47.75%

### Results for equal ‘Anti’ Instances

<table>
<thead>
<tr>
<th>Party</th>
<th>Pre</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>0.321</td>
<td>0.529</td>
<td>0.4</td>
</tr>
<tr>
<td>BJP</td>
<td>0.47</td>
<td>0.365</td>
<td>0.411</td>
</tr>
<tr>
<td>CONG</td>
<td>0.388</td>
<td>0.224</td>
<td>0.284</td>
</tr>
</tbody>
</table>

**Classifier:** Random Forest

**Efficiency:** 37.25%

---

Less instances of BJP result in low Precision and recall values.
Hashtags Based Classification

• Hashtags represent the topic of the tweet
• Picked up all the hashtags in the last 200 tweets of the user
• Computed the user vector in same manner as in text based classification
• Terms in this case were the hashtags instead of words used in the tweets
Hashtags Based Classification: Results

Results for all 613 ‘Pro’ Instances

<table>
<thead>
<tr>
<th>Instances: 613</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attributes: 1398</td>
</tr>
<tr>
<td>Classifier: Random Forest</td>
</tr>
<tr>
<td>Efficiency: 75.49%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Party</th>
<th>Pre</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>0.759</td>
<td>0.167</td>
<td>0.273</td>
</tr>
<tr>
<td>BJP</td>
<td>0.756</td>
<td>0.983</td>
<td>0.856</td>
</tr>
<tr>
<td>CONG</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Results for all 425 ‘Anti’ Instances

<table>
<thead>
<tr>
<th>Instances: 425</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attributes: 1182</td>
</tr>
<tr>
<td>Classifier: Random Forest</td>
</tr>
<tr>
<td>Efficiency: 50.35%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Party</th>
<th>Pre</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>0.5</td>
<td>0.946</td>
<td>0.654</td>
</tr>
<tr>
<td>BJP</td>
<td>0.875</td>
<td>0.165</td>
<td>0.277</td>
</tr>
<tr>
<td>CONG</td>
<td>0.286</td>
<td>0.045</td>
<td>0.077</td>
</tr>
</tbody>
</table>

- Efficiency improved by 2-3%, even with equal number of instances
- Precision and recall values remain 0 for Congress with all ‘Pro’ instances
User Features Based Classification

1. #Friends
2. #Followers
3. Following AAP?
4. Following BJP?
5. Following Congress?
6. #AAP related words
7. #BJP related words
8. #Congress related words
9. #AAP related hashtags
10. #BJP related hashtags
11. #Congress related hashtags
User Features Based Classification

1. #Friends
2. #Followers

Scatter Plot b/w #Followers and #Friends for users ‘Pro’ to each party

AAP  BJP  cong

Number of Friends

Number of Followers
User Features Based Classification: Results

**Results for all 613 ‘Pro’ Instances**

<table>
<thead>
<tr>
<th>Party</th>
<th>Pre</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>0.455</td>
<td>0.376</td>
<td>0.412</td>
</tr>
<tr>
<td>BJP</td>
<td>0.781</td>
<td>0.855</td>
<td>0.816</td>
</tr>
<tr>
<td>CONG</td>
<td>0.429</td>
<td>0.182</td>
<td>0.255</td>
</tr>
</tbody>
</table>

**Results for equal ‘Pro’ Instances**

<table>
<thead>
<tr>
<th>Party</th>
<th>Pre</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>0.483</td>
<td>0.412</td>
<td>0.44</td>
</tr>
<tr>
<td>BJP</td>
<td>0.439</td>
<td>0.486</td>
<td>0.462</td>
</tr>
<tr>
<td>CONG</td>
<td>0.588</td>
<td>0.606</td>
<td>0.597</td>
</tr>
</tbody>
</table>

Good Efficiency

Precision and Recall not 0

Improvement in efficiency by 6.4%
### User Features Based Classification: Results

#### Equal instances of ‘Pro’ AAP-BJP

<table>
<thead>
<tr>
<th>Party</th>
<th>Pre</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>0.606</td>
<td>0.62</td>
<td>0.615</td>
</tr>
<tr>
<td>BJP</td>
<td>0.781</td>
<td>0.59</td>
<td>0.603</td>
</tr>
</tbody>
</table>

#### Equal instances of ‘Pro’ AAP-Cong

<table>
<thead>
<tr>
<th>Party</th>
<th>Pre</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>0.706</td>
<td>0.727</td>
<td>0.716</td>
</tr>
<tr>
<td>CONG</td>
<td>0.719</td>
<td>0.697</td>
<td>0.708</td>
</tr>
</tbody>
</table>

- For BJP-Cong also the efficiency was > 60%
- This method works well for 2-class classification
- For ‘Anti’ category, there was a 6% improvement with this method
Network Based Classification

- Network formed on the basis of
  - Retweets
  - User mentions
- Undirected, without weights graph
- Users formed the nodes
- Used Gephi 0.8.2
- Community detection algorithm
  - Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre, Fast unfolding of communities in large networks
  - Fastest for large networks
Network Based Classification: Results

- With all 613 ‘Pro’ instances
- #Communities: 11
- 3 major communities
- Rest had 0.05% of nodes
- Modularity Score: 0.402
Network Based Classification: Results

- With all 613 ‘Pro’ instances

<table>
<thead>
<tr>
<th>Party</th>
<th>Pre</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>0.709</td>
<td>0.672</td>
<td>0.690</td>
</tr>
<tr>
<td>BJP</td>
<td>0.939</td>
<td>0.850</td>
<td>0.897</td>
</tr>
<tr>
<td>CONG</td>
<td>0.326</td>
<td>0.576</td>
<td>0.451</td>
</tr>
</tbody>
</table>

#Nodes: 6022
#Edges: 13693
Modularity Score: 0.402
Efficiency: 78.31%
Network Based Classification: Results

- With equal instances of ‘Pro’ category
- #Communities: 8
- 3 major and rest with 0.05% of nodes

<table>
<thead>
<tr>
<th>Party</th>
<th>Pre</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>0.856</td>
<td>0.733</td>
<td>0.794</td>
</tr>
<tr>
<td>BJP</td>
<td>0.769</td>
<td>0.952</td>
<td>0.860</td>
</tr>
<tr>
<td>CONG</td>
<td>0.818</td>
<td>0.897</td>
<td>0.857</td>
</tr>
</tbody>
</table>

| #Nodes: 1193 |
| #Edges: 1489 |
| Modularity Score: 0.582 |
| Efficiency: 80.00% |
Presentation Outline

• Research Motivation
• Research Aim
• Related Work
• Research Gap
• Data Analysis
• Classification of Political Orientation
• System Design
• Conclusion
• Limitations and Future Work
System Design

- Requirement: To be able to see and analyze the election related tweets on daily basis
- Used PHP and Javascript to develop the portal [http://bheem.iiitd.edu.in/IndiaElections](http://bheem.iiitd.edu.in/IndiaElections)
- Refreshes at a 24 hour interval, but displays the tweets at an interval of 5000 ms
Realtime Tweets

India General Elections 2014

18 new Tweets

KJS Arora@KanwaljiSingA
Wen Bjp n sangh parivar put Twitter n fb to good use in LS polls it becums Jurassic park take on Star Wars #144inAmethi

Gurmeet S. Randhir@gorgandhir
@M_Lokhi BJP must thank Cong' 4 their Mindless Election Camp'n as they wouldn't hav got such Overwhelming Support & Sure Victory in d end.

Naresh Dixit@nareshdixit2
RT @ANI_news: Sushil Modi: More than 50 JDU MLAs helping BJP win elections, they know JDU has slipped to 3rd position don't want RJD to rega?

DNAPopularNews@DNAPopularNews
Congress facing its worst time, says Narendra Modi. This is the worst ever time for Congress, BJP prime ministr... http://t.co/UDjOifsaoJB
# Trending Politicians

## India General Elections 2014

<table>
<thead>
<tr>
<th>Image</th>
<th>Name</th>
<th>Screen Name</th>
<th>Followers</th>
<th>Klout Score</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td>Narendra Modi</td>
<td>narendramodi</td>
<td>3892386</td>
<td>88.9544376106022</td>
</tr>
<tr>
<td><img src="image2.png" alt="Image" /></td>
<td>Shashi Tharoor</td>
<td>ShashiTharoor</td>
<td>2158589</td>
<td>63.8670452261145</td>
</tr>
<tr>
<td><img src="image3.png" alt="Image" /></td>
<td>Omar Abdullah</td>
<td>abdullah_omar</td>
<td>405903</td>
<td>82.2925779719157</td>
</tr>
<tr>
<td><img src="image4.png" alt="Image" /></td>
<td>Rajeev Chandrasekhar</td>
<td>rajeev_mp</td>
<td>127523</td>
<td>80.4229036163932</td>
</tr>
<tr>
<td><img src="image5.png" alt="Image" /></td>
<td>Dr Manmohan Singh</td>
<td>PMOIndia</td>
<td>1202471</td>
<td>80.2100714321393</td>
</tr>
<tr>
<td><img src="image6.png" alt="Image" /></td>
<td>Milind Deora</td>
<td>miinddeora</td>
<td>131956</td>
<td>79.6566835279628</td>
</tr>
</tbody>
</table>
India General Elections 2014

Location
What’s Trending

India General Elections 2014

Realtime Tweets

Gopal Kumar @gopal2315
Monday, 05 May 2014 at 11:00:00
RT @EPROPolitics Huge cheer among crowd as #NarendraModi arrives in #Amethi #NaMoINAmethi

Aditya Aggerwal @adi_207
Monday, 05 May 2014 at 11:00:01
As soon as #NaMo arrives in Amethi, it started to rain! ;} #AcheDinaAaneWaleHai #NaMoINAmethi

Saurav Dixit @sauravdixit1976
Monday, 05 May 2014 at 11:00:21
RT @NtiCentral: #NaMoINAmethi Narendra Modi releases manifesto for Amethi http://t.co/BxZ8rFWV0

Right Now I/O Feed @rightnowio_feed
Monday, 05 May 2014 at 11:00:26
RT @chinavkrd: Saffron surge in Amethi as people rush for Modi’s rally “NaMo http://t.co/ZX2EuR1Namensamethiin

Sentiment

varanasi bjp campaign irani rally priyanka Gandhi people Modi says watch ec pm nitya india ganga aap arvind
Sentiment

India General Elections 2014

Per hour sentiment score for BJP

Per hour sentiment score for AAP

Per hour sentiment score for Cong

Hour

Score
Presentation Outline

• Research Motivation
• Research Aim
• Related Work
• Research Gap
• Data Analysis
• Classification of Political Orientation
• System Design
• Conclusion
• Limitations and Future Work
Conclusion

• Twitter activity is directly proportional to the real time activities
• Data is particularly higher on weekdays
• BJP emerged as the leader in tweet share as well as seat share in both Assembly and General Elections
• Predicting the political orientation with content based methods is not easy
• The transliteration and sarcasm used in the text can be possible causes for poor performance of content based methods
Conclusion

• The user features based classification can improve the efficiency, but not for the ‘Anti’ category
• Prediction of ‘Anti’ political orientation is even more difficult
• Network based methods worked best for the Indian users
• A 2-class classification gives better results in all the methods as compared to 3-class classification
• A system to monitor and analyze the recent tweets was also developed
Presentation Outline

• Research Motivation
• Research Aim
• Related Work
• Research Gap
• Data Analysis
• Classification of Political Orientation
• System Design
• Conclusion
• Limitations and Future Work
Limitations & Future Work

• API rate limit puts a restriction on the data collected
• Political views of people is difficult to judge only on the basis of tweets
• Too much of content by BJP made the results biased towards BJP
• Future work can include to see the change in sentiments post elections
• To look at if the parties that lost the elections were still active and trending
Acknowledgement

- Dr. PK, IIIT-D
- Anupama Aggarwal, PhD Scholar, IIIT-D
- Data Annotators
- CERC@IIIT-D
- Precogs
- Family and Friends
References


References


Feedback/ Suggestions:

pk@iiitd.ac.in
abhishek1231@iiitd.ac.in