

Who are Tweeting in the 2014 Indonesia's Legislative Elections?

Rischan Mafrur
M. Fiqri Muthohar

Outline

Problem/Intro

Approach

Experiment & Result

Conclusion

Need to know little bit about politics

Additional Information

| Party | Number of Seats in the National Assembly | Leader | Comments |
|--|--|-------------------------------|---|
| Saenuri Party (NFP - New Frontier Party) 새누리당 / 새누리黨 <i>Saenuridang</i> | 156 | Hwang Woo-yea | Conservative; formerly called the Grand National Party. |
| New Politics Alliance for Democracy (NPAD) 새정치민주연합 / 新政治民主聯合 <i>Saejeongchi Minju Yeonhap</i> | 130 | Ahn Cheol-soo Kim Han-gill | Liberal; merged from the Democratic Party and the minor New Political Vision Party . |
| Unified Progressive Party (UPP) 통합진보당 / 統合進步黨 <i>Tonghap Jinbodang</i> | 6 | Lee Jeong-hee | Progressive; merged from the Democratic Labor Party , People's Participation Party , and a faction of the New Progressive Party . |
| Justice Party 정의당 / 正義黨 <i>Jeongeuidang</i> | 5 | Cheon Ho-sun | Progressive; splintered from the Unified Progressive Party . |

Vietnam is a **single-party** state. This means that only one political party, the Communist Party of Vietnam (Đảng Cộng Sản Việt Nam) is legally allowed to hold effective power.

Additional Information

| Party | Leader | Status |
|---|--------------------------|---------------------|
| Democratic Party (Indonesia) (Demokrat) <i>Partai Demokrat</i> | Susilo Bambang Yudhoyono | Governing Coalition |
| Party of the Functional Groups (Golkar) <i>Partai Golongan Karya</i> | Aburizal Bakrie | Governing Coalition |
| Indonesian Democratic Party – Struggle (PDI-P) <i>Partai Demokrasi Indonesia Perjuangan</i> | Megawati Sukarnoputri | Opposition |
| National Awakening Party (PKB) <i>Partai Kebangkitan Bangsa</i> | Muhaimin Iskandar | Governing Coalition |
| United Development Party (PPP) <i>Partai Persatuan Pembangunan</i> | Suryadharma Ali | Governing Coalition |
| Prosperous Justice Party (PKS) <i>Partai Keadilan Sejahtera</i> | Anis Matta | Governing Coalition |
| National Mandate Party (PAN) <i>Partai Amanat Nasional</i> | Hatta Rajasa | Governing Coalition |
| People's Conscience Party (Hanura) <i>Partai Hati Nurani Rakyat</i> | Wiranto | Opposition |
| Great Indonesia Movement Party (Gerindra) <i>Partai Gerakan Indonesia Raya</i> | Suhardi | Opposition |

Project Description

1. Indonesia President Election 2014, now still legislative election.
2. All of political parties **campaign** and tell about the programs
3. The legislative election will held at 9 April 2014
4. Time for legislative **campaign March 16 - April 5** 2014
5. Two types of campaign :
 - Physical campaign
 - Media campaign (TV, Newspaper, Banner, **Online campaign**)



ash Tag

Twitter Campaign

(PKS) Partai Keadilan Sejahtera

(Prosperous Justice Party)

#TolakPartaiPoligami (-)

(Became **Worldwide** trending topic at **9:30 PM - 20 Mar 2014** GMT+7)

#SayaPilihPKS (+)

(Became **Indonesia** trending topic at **8:36 AM - 22 Mar 2014** GMT+7)

Problem

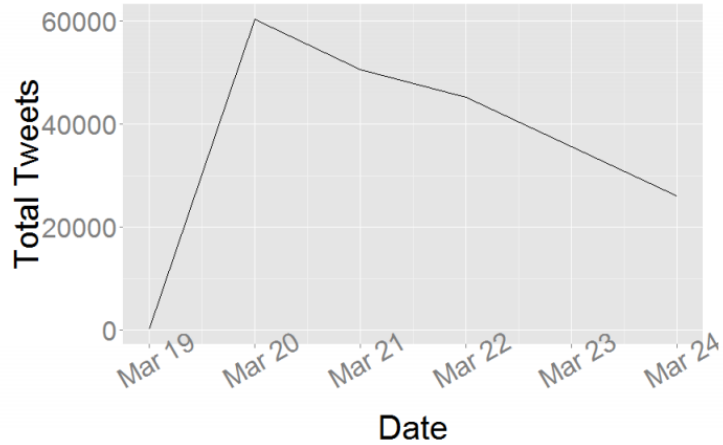
Just to **Know**, who they are, **Who**
are Tweeting it.

Approach

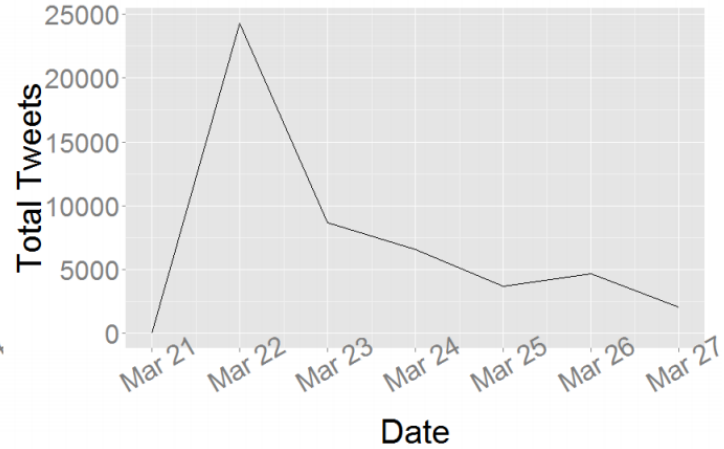
1. Gathered the **dataset** from **twitter**.
2. Picking **username** and then **get all features**.
3. Hand Labeling for **ground truth creation**.
4. Using **Naive Bayes** for classifier
5. **Features Evaluation**
6. Testing to **real dataset**
7. Show the result. (**How many campaign** and **real accounts**)

Experiment & Result

Tweet Distribution



(a) #TolakPartaiPoligami hashtag



(b) #SayaPilihPKS hashtag

Fig. 1: Tweets distribution

> head(tw_tolak, 2)

```
[1] "{\"created_at\":\"Mon Mar 31 06:52:03 +0000 2014\", \"id\":\"450525915072569344\", \"id_str\":\"450525915072569344\", \"text\": \"Sapi sapi pada kepedean menang pemilu, ini Indonesia, bukan timur tengah #TolakPartaiPoligami\", \"source\":\"\\u003ca href= \"http://\\u003eubersocial.com\" rel=\\\"nofollow\\\" \\u003eubersocial for Android\\u003c/a\\u003e\", \"truncated\":false, \"in_reply_to_status_id\":null, \"in_reply_to_status_id_str\":null, \"in_reply_to_user_id\":null, \"in_reply_to_user_id_str\":null, \"in_reply_to_screen_name\":null, \"user\":{\"id\":\"99937376\", \"id_str\":\"99937376\", \"name\":\"Dedi Eka Putra\", \"screen_name\":\"dedi012\", \"location\":\"jakarta\", \"url\":\"http://\\u003ededikaputra.com\", \"description\":\"just ordinary\", \"protected\":false, \"followers_count\":67, \"friends_count\":493, \"listed_count\":0, \"created_at\":\"Mon Dec 28 12:30:35 +0000 2009\", \"favourites_count\":22, \"utc_offset\":25200, \"time_zone\":\"jakarta\", \"geo_enabled\":true, \"verified\":false, \"statuses_count\":1138, \"lang\":\"en\", \"contributors_enabled\":false, \"is_translator\":false, \"is_translation_enabled\":false, \"profile_background_color\":\"1A1B1F\", \"profile_background_image_url\":\"http://\\u003eps.twimg.com/\\u003eprofile_background_images/\\u003e244270118/\\u003earsenal-fc-logo.jpg\", \"profile_background_image_url_https\":\"https://\\u003eps.twimg.com/\\u003eprofile_background_images/\\u003e244270118/\\u003earsenal-fc-logo.jpg\", \"profile_background_tile\":false, \"profile_image_url\":\"http://\\u003eps.twimg.com/\\u003eprofile_images/\\u003e447676801423142912/\\u003exk1ppnq3_normal.jpeg\", \"profile_image_url_https\":\"https://\\u003eps.twimg.com/\\u003eprofile_images/\\u003e447676801423142912/\\u003exk1ppnq3_normal.jpeg\", \"profile_link_color\":\"2FC2EF\", \"profile_sidebar_border_color\":\"181A1E\", \"profile_sidebar_fill_color\":\"252429\", \"profile_text_color\":\"666666\", \"profile_use_background_image\":true, \"default_profile\":false, \"default_profile_image\":false, \"following\":null, \"follow_request_sent\":null, \"notifications\":null}, \"geo\":null, \"coordinates\":null, \"place\":null, \"contributors\":null, \"retweet_count\":0, \"favorite_count\":0, \"entities\":{\"hashtags\":[{\"text\":\"TolakPartaiPoligami\", \"indices\":[74, 94]}], \"symbols\":[], \"urls\":[], \"user_mentions\":[], \"favorite_id\":false, \"retweeted\":false, \"filter_level\":\"medium\", \"lang\":\"in\"}}
```

```
[2] "{\"created_at\":\"Mon Mar 31 08:19:17 +0000 2014\", \"id\":\"450547869494177792\", \"id_str\":\"450547869494177792\", \"text\": \"Pepohonan di jalan TB simatupang banyak \\\"dikotori\\\" oleh foto caleg Sapi #TolakPartaiPoligami\", \"source\":\"\\u003ca href= \"http://\\u003eubersocial.com\" rel=\\\"nofollow\\\" \\u003eubersocial for Android\\u003c/a\\u003e\", \"truncated\":false, \"in_reply_to_status_id\":null, \"in_reply_to_status_id_str\":null, \"in_reply_to_user_id\":null, \"in_reply_to_user_id_str\":null, \"in_reply_to_screen_name\":null, \"user\":{\"id\":\"99937376\", \"id_str\":\"99937376\", \"name\":\"Dedi Eka Putra\", \"screen_name\":\"dedi012\", \"location\":\"jakarta\", \"url\":\"http://\\u003ededikaputra.com\", \"description\":\"just ordinary\", \"protected\":false, \"followers_count\":67, \"friends_count\":493, \"listed_count\":0, \"created_at\":\"Mon Dec 28 12:30:35 +0000 2009\", \"favourites_count\":22, \"utc_offset\":25200, \"time_zone\":\"jakarta\", \"geo_enabled\":true, \"verified\":false, \"statuses_count\":1140, \"lang\":\"en\", \"contributors_enabled\":false, \"is_translator\":false, \"is_translation_enabled\":false, \"profile_background_color\":\"1A1B1F\", \"profile_background_image_url\":\"http://\\u003eps.twimg.com/\\u003eprofile_background_images/\\u003e244270118/\\u003earsenal-fc-logo.jpg\", \"profile_background_image_url_https\":\"https://\\u003eps.twimg.com/\\u003eprofile_background_images/\\u003e244270118/\\u003earsenal-fc-logo.jpg\", \"profile_background_tile\":false, \"profile_image_url\":\"http://\\u003eps.twimg.com/\\u003eprofile_images/\\u003e447676801423142912/\\u003exk1ppnq3_normal.jpeg\", \"profile_image_url_https\":\"https://\\u003eps.twimg.com/\\u003eprofile_images/\\u003e447676801423142912/\\u003exk1ppnq3_normal.jpeg\", \"profile_link_color\":\"2FC2EF\", \"profile_sidebar_border_color\":\"181A1E\", \"profile_sidebar_fill_color\":\"252429\", \"profile_text_color\":\"666666\", \"profile_use_background_image\":true, \"default_profile\":false, \"default_profile_image\":false, \"following\":null, \"follow_request_sent\":null, \"notifications\":null}, \"geo\":null, \"coordinates\":null, \"place\":null, \"contributors\":null, \"retweet_count\":0, \"favorite_count\":0, \"entities\":{\"hashtags\":[{\"text\":\"TolakPartaiPoligami\", \"indices\":[72, 92]}], \"symbols\":[], \"urls\":[], \"user_mentions\":[], \"favorite_id\":false, \"retweeted\":false, \"filter_level\":\"medium\", \"lang\":\"in\"}}
```

> |

Picking username from dataset

Why we need to doing this?

We want to know **how many accounts** (in total)
from **the all of tweets**.

Total Number of Tweets

[1] #TolakPartaiPoligami : 222,444 tweets

[2] #SayaPilihPKS : 48,135 tweets

Dataset I : 7,000 from [1] + 3,000 from [2]

Dataset II : All tweets - Dataset I

Ground Truth Creation

From 10,000 tweets (dataset I) are only came from 1,680 twitter accounts, And the result of hand labeling =>

| real | campaign | total |
|-------|----------|-------|
| 1,479 | 201 | 1,680 |

TABLE I: Labeled dataset I overview

Classifier Performance for dataset I

| | | Predicted | |
|------|----------|-----------|----------|
| | | real | campaign |
| True | real | 1467 | 12 |
| | campaign | 15 | 186 |

(a) Confusion Matrix

| | real | campaign |
|-----------|------|----------|
| precision | 0.99 | 0.92 |
| recall | 0.98 | 0.94 |
| accuracy | 0.98 | 0.98 |

(b) Classifier Performance

TABLE III: Confusion matrix and classifier performance

Features Evaluation

| Value | Rank | Value | Rank | Value | Rank |
|--------|---------------|-------|---------------|-------|---------------|
| 1538.0 | age | 0.45 | age | 0.50 | age |
| 1433.8 | followings | 0.41 | followings | 0.25 | mention ratio |
| 1321.3 | followers | 0.40 | num tweets | 0.23 | num tweets |
| 1238.5 | mention ratio | 0.38 | mention ratio | 0.13 | API ratio |
| 1214.8 | num tweets | 0.37 | followers | 0.12 | hashtag ratio |
| 1206.5 | hashtag ratio | 0.35 | hashtag ratio | 0.11 | tweet rate |
| 1150.8 | tweetrate | 0.33 | tweet rate | 0.09 | URL ratio |
| 481.7 | reputation | 0.18 | reputation | 0.08 | followings |
| 84.6 | API ratio | 0.06 | API ratio | 0.01 | reputation |
| 0 | URL ratio | 0 | URL ratio | 0.008 | followers |

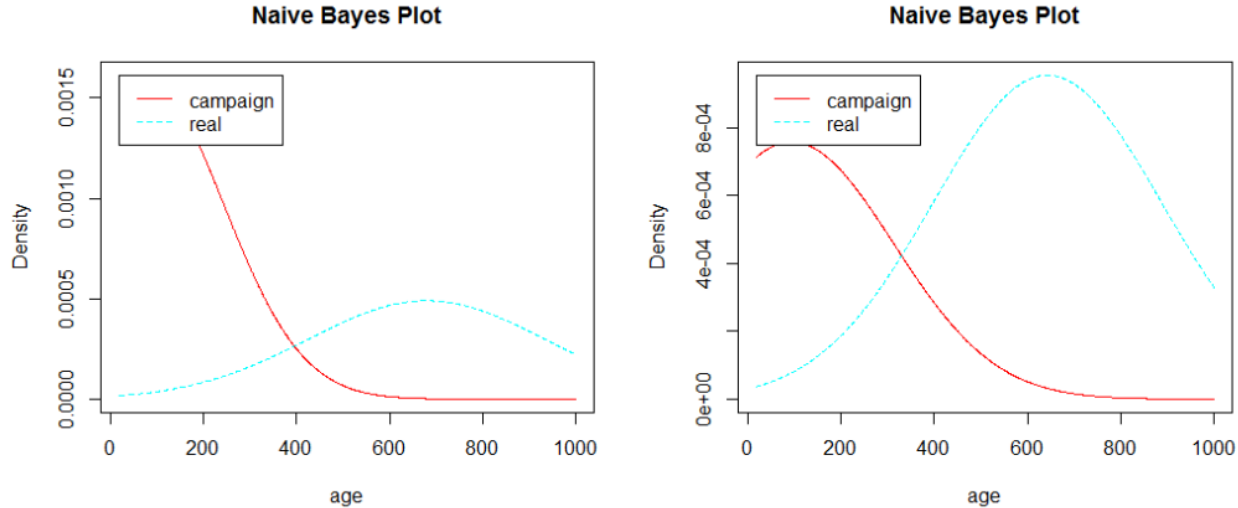
(a) Chi Squared

(b) Information Gain

(c) ReliefF

TABLE II: Features Evaluation

Features Plot



(a) Age of twitter accounts from #TolakPartaiPoligami hashtag
(b) Age of twitter accounts from #SayaPilihPKS hashtag

Fig. 2: One of the features : the age of twitter accounts

Final Result

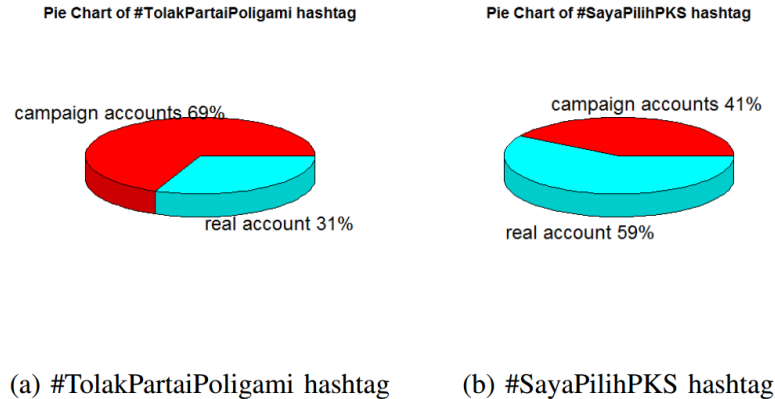


Fig. 3: Percentage of campaign and real accounts

From dataset II, **#TolakPartaiPoligami** total tweets are 215,444 only came from 9,651 twitter accounts.

#SayaPilihPKS, total tweets are 45,135 came from 5,639 twitter accounts.

#TolakPartaiPoligami hashtag has been classified to 6,621 (**69%**) campaign accounts and 3,030 (**31%**) real accounts.

#SayaPilihPKS hashtag has 2,334 (**41%**) campaign accounts and 3,305 (**59%**) real accounts.

Conclusion

In this research we **did not consider** about the number of retweets.

From the result we know that #TolakPartaiPoligami hashtag has almost 70% campaign accounts, and the second #SayaPilihPKS has more than 40% campaign accounts.

When there are many newspapers or online media were reported that there are many people who hate or love with one of party because many people tweeting about it, **It could not be used as a basis of truth.**

REFERENCES

- [1] B. J. Jansen, M. Zhang, K. Sobel, and A. Chowdury, "Twitter power: Tweets as electronic word of mouth," *Journal of the American Society for Information Science and Technology*, vol. 62(11), pp. 2169–2188, 2009.
- [2] M. Thelwall, K. Buckley, and G. Paltoglou, "Sentiment in twitter events," *Journal of the American Society for Information Science and Technology*, vol. 62(2), pp. 406–418, 2011.
- [3] H. Becker, M. Naaman, and L. Gravano, "Beyond trending topics: Real-world event identification on twitter," in *ICWSM*, Barcelona, July 17-21 2011.
- [4] T. A. Small, "What the hashtag? a content analysis of canadian politics on twitter," *Journal Information, Communication and Society*, vol. 14(6), pp. 872–895, 2011.
- [5] R. D. L. Wigand, "Tweets and retweets: Twitter takes wing in government," *Journal Information Polity*, vol. 16(3), pp. 215–224, August 2011.
- [6] S. E. Cho and H. W. Park, "Government organizations' innovative use of the internet: The case of the twitter activity of south korea's ministry for food, agriculture, forestry and fisheries," *Journal Scientometrics*, vol. 90(1), pp. 9–23, January 2012.
- [7] H. Kwak, G. Lee, H. Park, and S. Moon, "What is twitter, a social network or a news media?" in *Proceedings of the 19th international conference on World wide web*, 2010, pp. 591–600.
- [8] S. Yardi, D. Romero, G. Schoenebeck, and danah boyd, "Detecting spam in a twitter network," *First Monday*, vol. 15, pp. 1–4, January 2010.
- [9] A. H. Wang, "Dont follow me: Spam detection in twitter," in *International Conference on Security and Cryptography (SECURITY)*, July 26-28 2010.
- [10] Z. Chu, S. Glanvecchio, H. Wang, and S. Jajodia, "Detecting automation of twitter accounts: Are you a human, bot, or cyborg?" *Journal IEEE Transactions on Dependable and Secure Computing*, vol. 9(6), pp. 811–824, November 2012.
- [11] J. Song, S. Lee, and J. Kim, "Dont follow me: Spam detection in twitter," in *Proceedings of the 14th international conference on Recent Advances in Intrusion Detection*, September 20-21 2011.
- [12] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, "Detecting spammers on twitter," in *Collaboration, Electronic messaging, AntiAbuse and Spam Conference (CEAS)*, July 2010.
- [13] C. Yang, R. C. Harkreader, and G. Gui, "Die free or live hard?empirical evaluation and new design for fighting evolving twitter spammers," in *Proceedings of the 14th international conference on Recent Advances in Intrusion Detection*, 2011.