Predicting the 2011 Dutch Senate Election Results with Twitter

Erik Tjong Kim Sang & Johan Bos
University of Groningen, The Netherlands
{e.f.tjong.kim.sang,johan.bos}@rug.nl

Dutch Senate Elections

- Held every four years

- Preceded by the Provincial Elections

- Voters choose 566 provincial representatives

- These representatives choose 75 Senate members

- Most recent election: Wednesday 2 March 2011

- Eleven national parties participated and a few local parties

---

Examples of Dutch tweets

- *Maar toch stem ik PVV*
  But still I vote **PVV**

- *Ja ik stem SP en straf zo het kabinet #penw #debat*
  Yes I vote **SP** and punish the cabinet #penw #debate

- *Ja ik zeg natuurlijk: STEM PVDA!!! Maar dat had je waarschijnlijk wel verwacht haha!*
  Yes of course I am saying: VOTE **PVDA**!!! But you would have expected that haha!

Previous work

In 2010, Tumasjan et al. showed that counts of party mentions on Twitter were a good predictor for German election results

Jungherr et al. (2011) show that including other parties and other Twitter data would have dramatically changed the German election predictions

Chung and Mustafaraj (2011), O'Connor et al. (2010) and Gayo-Avello et al. (2011) are also pessimistic about the value of tweet counts for predicting elections

---

Processing steps for tweets

1. harvest Dutch tweets from Twitter

2. select one day of tweets: Wednesday 16 February 2012

3. select tweets containing Dutch party names

4. convert tweet counts to Senate seat numbers

Evaluation

We compared the predictions of our system with the average prediction of two big polling agencies: Politieke Barometer and Maurice de Hond

We did not predict results for the regional elections; only for the Senate elections

We did not predict results for the local parties but assumed they would win one Senate seat

---

Predictions based on tweet counts

Party	tweets	Seats Twitter	Seats PB	Seats MdH	Average polls
PVV	2227	18	12	12	12
VVD	1562	13	14	16	15
CDA	1504	12	9	10	9.5
PvdA	1057	9	13	13	13
SP	839	7	8	7	7.5
GL	748	6	5	3	4
D66	610	5	6	5	5.5
CU	238	2	3	3	3
PvdD	154	1	1	1	1
SGP	139	1	2	2	2
50+	49	0	1	2	1.5
OSF	-	1	1	1	1
offset		21	4	4	-

What went wrong? Negative sentiment:

- *owja ik stem geen pvv voor die dat denken*
  so I'm not voting **pvv** if anyone is thinking that

- *Wel triest van de vvd om de zondagen nu te schrappen wat betreft het shoppen, jammer! Hierbij dus een #fail*
  Sadly, the **vvd** will ban shopping on Sundays, too bad! So here is a #fail

- *Maar ik ben voor kernenergie en tegen de Kunduz missie. Dus waarschijnlijk geen GL stem voor mij.*
  But I am in favor of nuclear power and oppose the Kunduz mission. So probably no **GL** vote for me.

What went wrong? Multiple party names and spam

- *In welke kroeg kan ik vanavond met mensen van **CDA**, **D'66**, **PvdA**, LR, **GL**, **PVV**, **VVD**, **CU** en **SP** het debat kijken in #010?*
  In what pub can I watch tonight's debate with people from **CDA**, **D66**, **PvdA**, LR, **GL**, **PVV**, **VVD**, **CU** and **SP** in Rotterdam?

The top ten users in the one-day selection based on frequency of tweets with political party names had totals of 59, 49, 40, 40, 38, 35, 33, 30, 27 and 26

Additional tweet processing steps

5. remove tweets that mentioned more than one tweet

6. keep only one tweet per user (the first)

7. remove tweets with a negative sentiment

Sentiment analysis

We do not have an automatic sentiment analysis system for Dutch

So two annotaters annotated 1,678 political tweets: either as negative or nonnegative

We agreed on 1,333 tweets (kappa score: 0.59)

We used the sentiments for estimating the probability that a tweet about a party is nonnegative

We assumed that these probabilities remain constant

Results after first normalization steps

Party	One party per tweet	One tweet per user	Both constraints	Sentiment weights	Including sentiment
PVV	22	17	19	0.49	13
VVD	12	13	13	0.68	13
CDA	12	12	12	0.70	12
PvdA	8	8	8	0.69	7
SP	6	8	7	0.90	9
GL	6	7	7	0.81	9
D66	5	5	5	0.94	6
CU	1	2	2	0.67	2
PvdD	1	1	1	1.00	2
SGP	1	1	0	0.86	1
50+	0	0	0	0.93	0
OSF	1	1	1	-	1
offset	29	22	25		23

What went wrong? Demographics:

Netherlands	Age	Dutch Twitter
0.5%	90+	0.4%
3.4%	80-89	0.1%
6.7%	70-79	0.2%
11.1%	60-69	1.0%
13.7%	50-59	4.0%
15.6%	40-49	11.1%
13.2%	30-39	17.8%
12.1%	20-29	28.8%
12.1%	10-19	34.0%
11.6%	0-9	2.6%

Correcting the data for demographic differences

We neither know gender nor age of individual Twitter users

So we cannot correct Twitter demographics for these two variables

The best we can do is to correct the tweet counts per party:
change them to the values derived from the poll seat distribution

Final processing step:

8. multiply tweet counts with weights based on expected seat numbers

---

Results after demographics correction

Party	Population weight	Seats Twitter	Average polls
PVV	0.93	12	12
VVD	1.23	15	15
CDA	0.80	10	9.5
PvdA	1.76	13	13
SP	0.82	8	7.5
GL	0.47	4	4
D66	0.87	5	5.5
CU	1.33	3	3
PvdD	0.49	1	1
SGP	1.84	2	2
50+	2.93	1	1.5
OSF	-	1	1
offset		2	-

Predictions for the elections of 2 March 2011

Party	Result	Seats PB	Seats MdH	Seats Twitter
VVD	16	14	16	14
PvdA	14	12	11	16
CDA	11	9	9	8
PVV	10	11	12	10
SP	8	9	9	6
D66	5	7	5	8
GL	5	4	4	3
CU	2	3	3	3
50+	1	2	2	2
SGP	1	2	2	2
PvdD	1	1	2	2
OSF	1	1	0	1
offset	-	14	14	18

Discussion: Did the tweet counts help?

We looked for the answer to this question by repeating the experiment *without tweets*

Basically, we assumed that the party names were uniformly distributed over the tweets

This required recomputation of the population weights

Results with uniformly distributed party names

Party	Result	Seats Twitter	Population weight
VVD	16	16	2.23
PvdA	14	13	1.93
CDA	11	10	1.41
PVV	10	12	1.78
SP	8	7	1.11
D66	5	5	0.82
GL	5	4	0.59
CU	2	3	0.45
50+	1	1	0.22
SGP	1	2	0.30
PvdD	1	1	0.15
OSF	1	1	-
offset	-	8	

Concluding remarks

- We predicted the Senate election results with reasonable accuracy[1]

- We are not sure if the tweets were really useful for this task

- We have several ideas for improving the system, for example with tweet count normalization based on automatic sentiment analysis and automatic gender and age prediction

[1] all data and software for this experiment are available on http://ifarm.nl/ps2011/ps2011.zip

THE END