

Measuring political confidence using Twitter sentiment analysis: a Belgian example

Olivier Van Poppel
Marco Chi Chung Fong

Table of Contents

- Introduction
- Literature review
- Methodology
- Results
- Future Work

Introduction and literature review

What we analyzed and why we were interested in the topic.

Introduction

- Why?
 - Politics is our passion
 - Twitter → perfect place for discussion
 - Does Twitter reflect reality?
 - Opinion polls are flawed
- What?
 - Compare opinion polls with Twitter Sentiment analysis results
 - Find confidence in government
 - Accuracy for Flemish and federal government

Literature review: a recap

- Focus on Politics & Sentiment Analysis(SA) on Twitter
 - SA: very well documented
 - Politics: Importance of microblogs
- No research on using Twitter data to analyse government confidence
 - Other research: focus on election prediction
 - Importance of sarcasm, etc
- No related research about Belgian Politics and the Dutch Language

Methodology

How did we obtain our results?

Methodology: RobBERT

- A state-of-the-art Dutch BERT model by KU Leuven researchers for text classification
- Pre-trained model for Dutch language natural language processing including sentiment analysis
- We utilize a modified version by Statistiek Vlaanderen for this project in Python



RobBERT

A Dutch RoBERTa-based Language Model

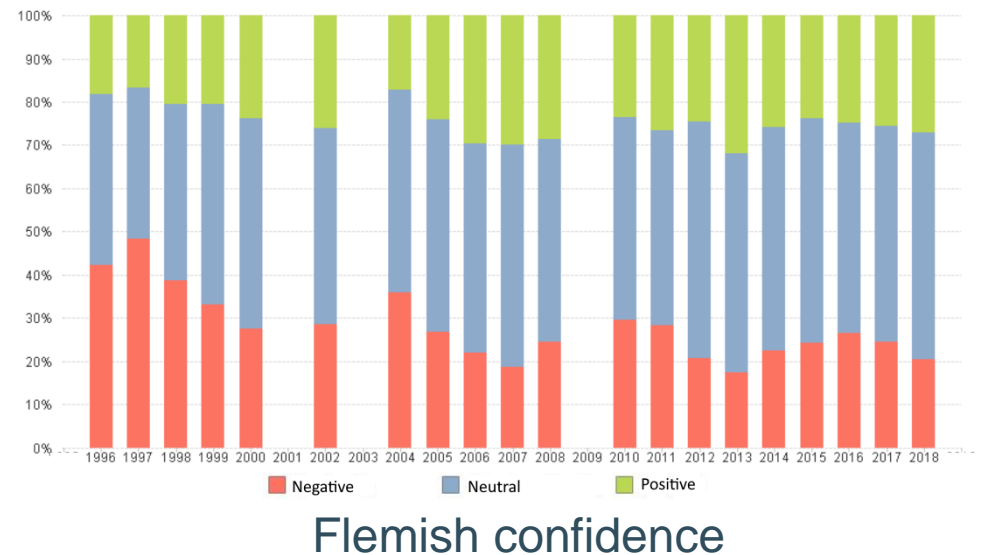
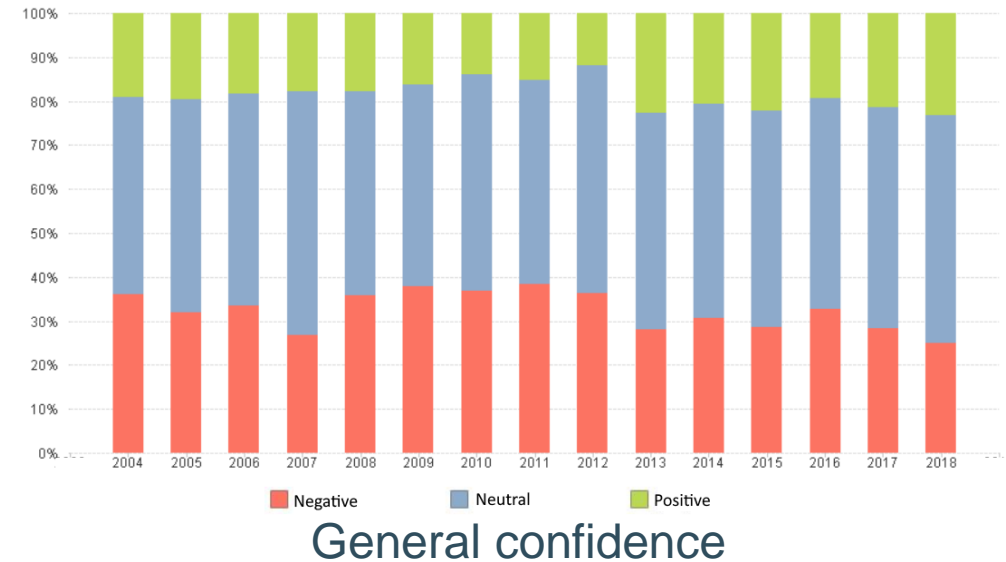
Methodology: Dataset

```
{
  "mentions": [
    "CarlFoulon",
    "crevits"
  ],
  "hashtags": [
    "ikmagdatzeggen"
  ],
  "urls": null,
  "emojis": [
    "😏"
  ],
  "text": "@CarlFoulon @crevits Hahaha. Is wel ons minister van onderwijs he ... enne... Ik heb nog voor haar ministerie gewerkt #ikmagdatzeggen",
  "created_at": "2018-11-09T10:40:43Z",
  "author_id": 294119752,
  "public_metrics": {
    "retweet_count": 0,
    "reply_count": 1,
    "like_count": 5,
    "quote_count": 0,
    "impression_count": 0
  },
  "id": 1060844625974517800,
  "sentiment": "NEGATIVE",
  "keyword": "ministerie",
  "gov_type": "flemish"
}
```

- NoSQL based
- Large Datasets involved (600k tweets) from Twitter API
- Categorized Attributes for sentiment analysis

Methodology: SCV-survey

- Compare SCV-survey \leftrightarrow Twitter data
- Compare 3 different levels:
 - General confidence
 - Flemish government
 - Federal government
- High degree of neutral sentiment
- Strong negative sentiment.



Methodology: Keywords

- Very difficult
 - Very little prior research
 - Keywords often neglected
 - Contacted researchers
- Key research: “Let them tweet cake”
 - Comprehensive list
- Key findings:
 - Names of politicians
 - Balance generalizability vs specificity



QR code to Github
containing our
keyword list

Methodology: Bias

- Model bias vs Twitter data bias
 - Model bias: out of scope
- Twitter bias: 3 types
 - Media bias
 - Government bias
 - Intensive user bias

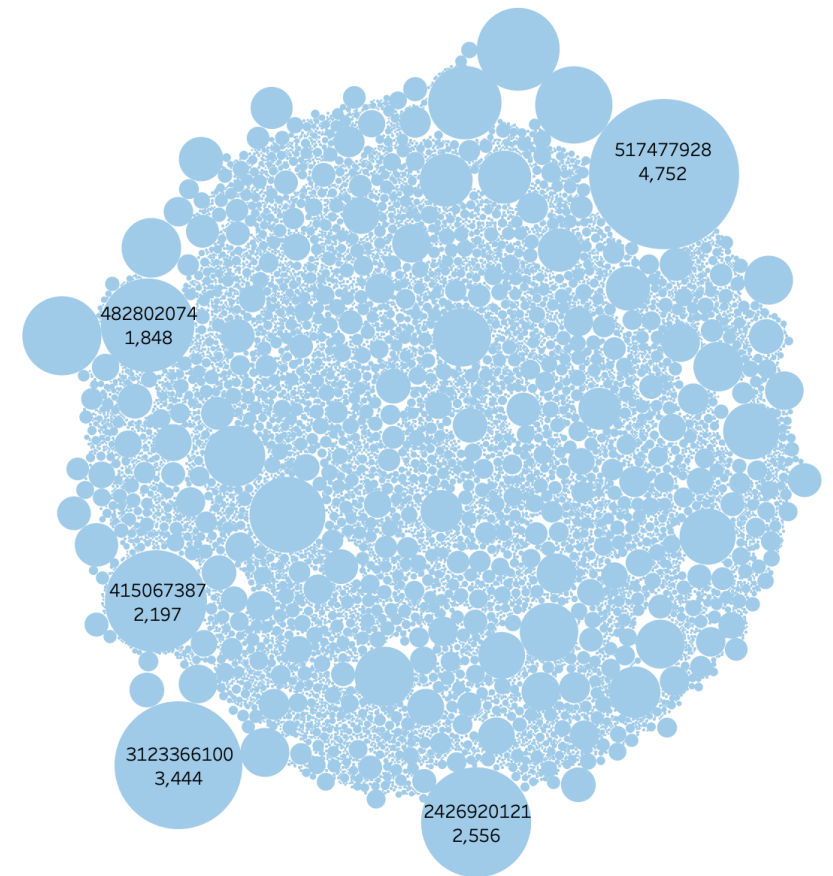
Methodology: Bias

- Media bias:
 - Media accounts are not voters
 - Overwhelmingly neutral
 - Removal less impact than expected
- Government bias:
 - Accounts owned by government
 - Not voters!
 - Impact small but expected

Methodology: Bias

- Intensive user bias:
 - Users with multiple appearances in dataset
 - Some users appear +4000 times
 - More tweets → more influence
- Everyone needs equal vote
- Solution: user weighting

Packed Bubble Chart of Users' Tweeting Frequencies



Methodology: Bias

- User weighting
 - Little research
 - Needed to be made from scratch
- Final formula:

$$user\ weight_{year}^{sentiment} = \frac{\#user\ tweets\ with\ the\ sentiment\ level\ for\ year}{\#total\ user\ tweets\ for\ year}$$

Results

What were our final results?

Results: Number of sentiments analyzed

Year	General conf	Flemish conf	Federal conf
2010	250	140	107
2011	4056	2472	2452
2012	9119	6279	5847
2013	12907	8555	8389
2014	23409	15759	14499
2015	32904	23686	17958
2016	44225	31670	24238
2017	57933	43583	28697
2018	91048	69836	44249
TOTAL SUM	298049	215320	148536

GRAND TOTAL = 661905

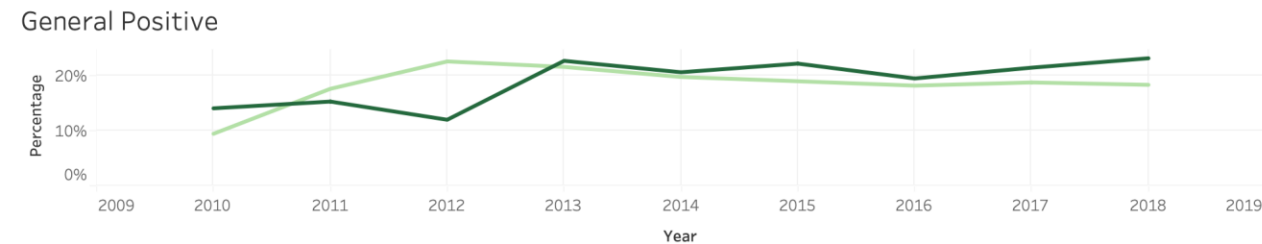
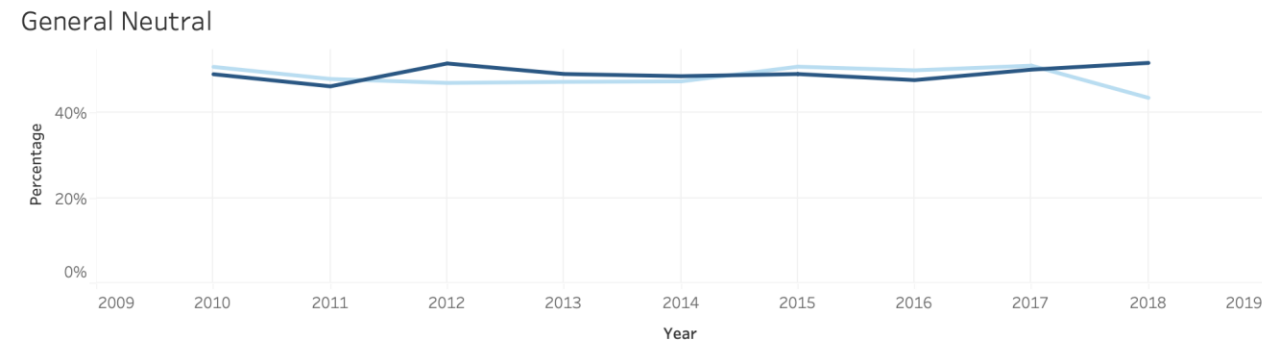
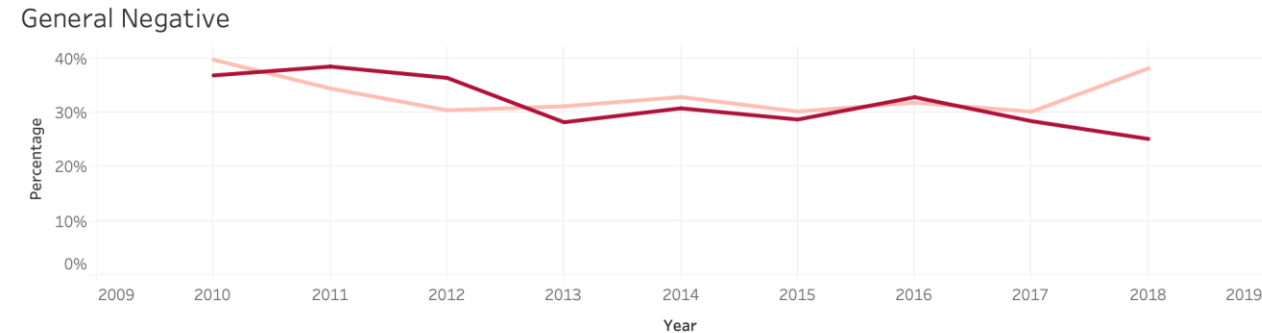
Results: comparison metrics

- Goals:
 - Validate use case of Twitter sentiment analysis
 - Measure weighting formula effectiveness
- Structure:
 - Compare results before/after weighting
 - Use symmetric mean absolute percentage error (SMAPE) to determine accuracy
 - Check for large differences with Autorank

$$\text{SMAPE} = \frac{100}{n} \sum_{t=1}^n \frac{|F_t - A_t|}{|A_t| + |F_t|}$$

Results: General confidence

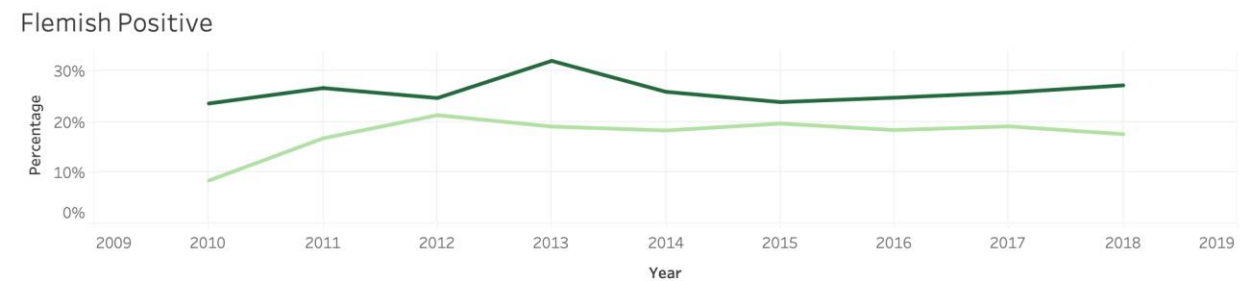
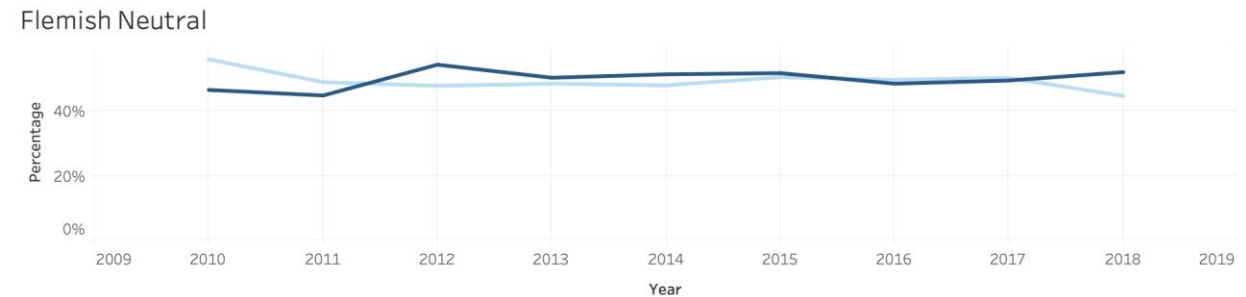
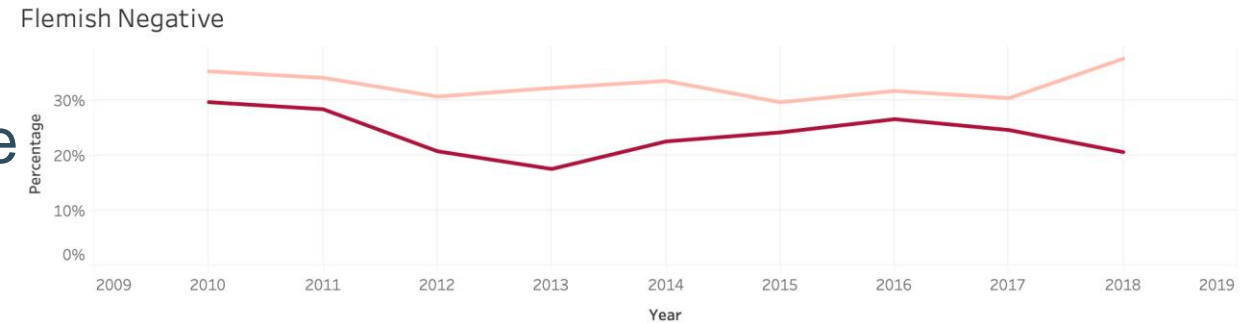
- Strong neutral presence
 - Both before/after weighting
- Weighting → significant improvements
 - Much less significant differences
 - Neutral sentiment almost spot on
- SMAPE and Autorank improved



light color: analysis result; dark color: SCV-survey result

Results: Flemish confidence

- Neutral is more accurate than Negative and Positive
- The model undermines the negative sentiment and overestimates the positive sentiment

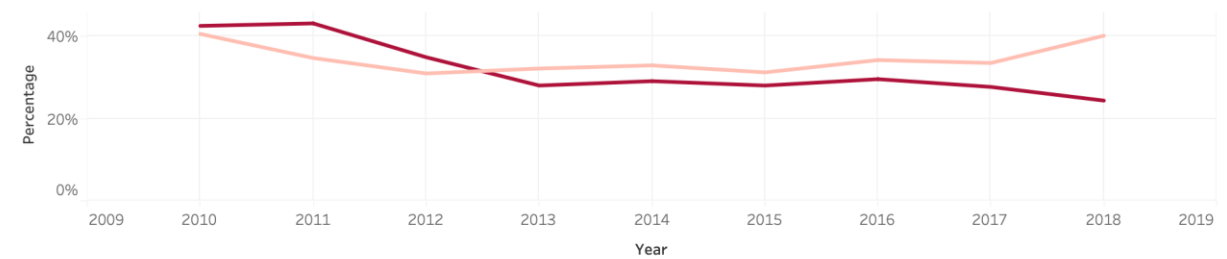


light color: analysis result; dark color: SCV-survey result

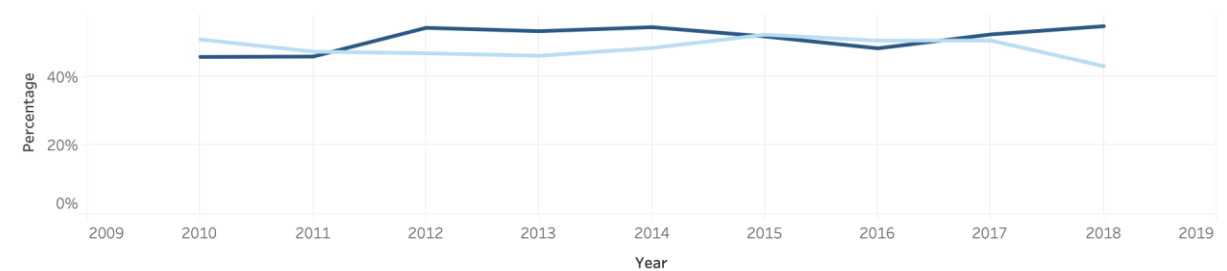
Results: Federal confidence

- Negative and Neutral are better than Positive
- The model does not perform well in during 2010 - 2013

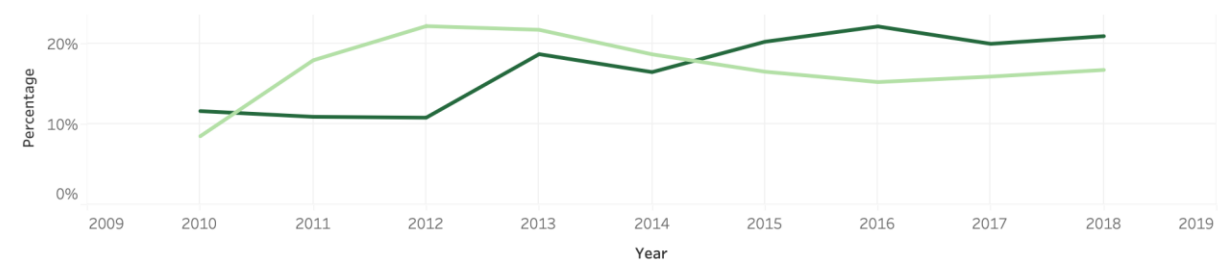
Federal Negative



Federal Neutral



Federal Positive



light color: analysis result; dark color: SCV-survey result

Key Results Takeaways

- We can utilize this Twitter Sentiment Analysis method to predict government confidence in Belgium
- The accuracy of the model performs much better in the General government level (avg. 6% SMAPE) rather than Federal (9%) and Flemish (13%) government level
- We have discovered that the model deviate the most in positive sentiment prediction

Future Work

Suggestions for future researchers on this topic

Suggested Future Work

- **Researching on User Weighting**
 - prevent bias from a small number of users dominating the results.
- **Standardized Keyword Creation**
 - robust framework focusing on government names and minimizing out-of-scope topics.
- **Addressing Bot Accounts**
 - introduce methods to identify and filter out tweets from bot accounts to avoid biased data.
- **Including Twitter Usage in Opinion Polls**
 - include a question about Twitter usage in government confidence polls
 - establish a benchmark for evaluating Twitter sentiment analysis and examine its impact on sentiment.

Thank you!