

IS POLITICAL AGITATION FLOODING SOCIAL MEDIA?



Wrocław University
of Science and Technology



Mateusz
Baran



Piotr
Kolebski



Mateusz
Wójcik

ARTICLE 105 OF POLISH ELECTION CODE

§ 1. ELECTORAL AGITATION IS PUBLICLY INDUCEMENTING OR ENCOURAGING PEOPLE TO VOTE IN A CERTAIN WAY, INCLUDING, IN PARTICULAR, TO VOTE FOR A CANDIDATE OF A PARTICULAR ELECTION COMMITTEE.

THE AIM

Build a model to detect political agitation and obtain knowledge about polish election campaign.

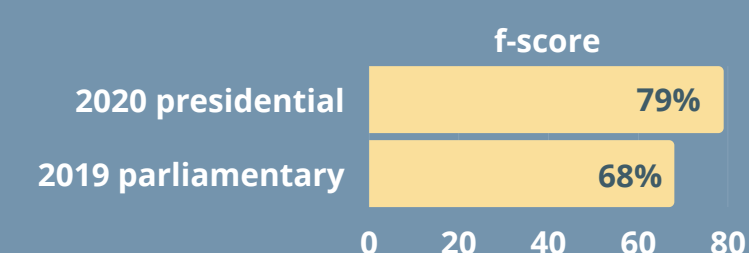
METHOD

We trained a natural language processing model for text classification. The end result of the project is an agitation classifier, a collection of analyses and a web application with an interactive demo.

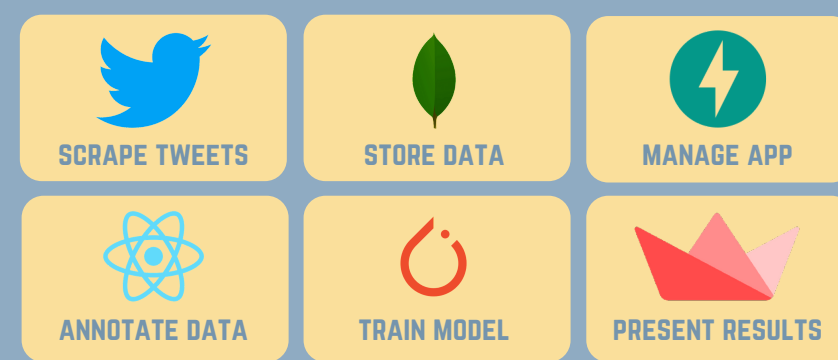
MODEL

We used TF-IDF, FastText and HerBERT to create three neural network classifiers. In the end we assembled all the three models together and built a soft voting ensemble classifier. The result we received was a **79% f-score** in a binary case.

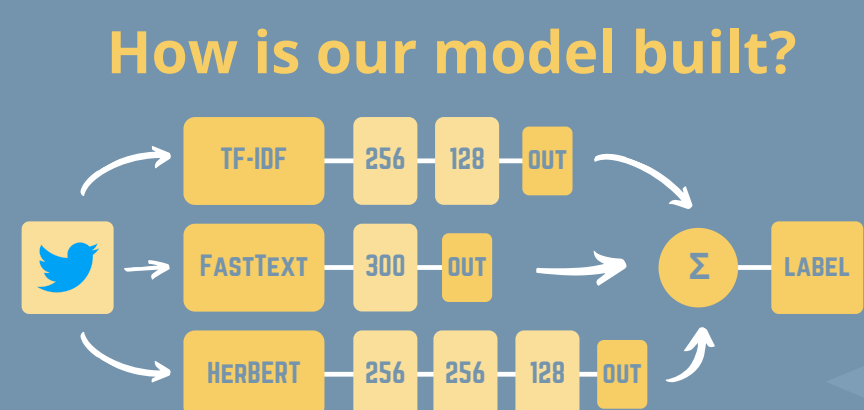
Can we adapt our model to other elections?



TECHNOLOGIES



Tweets had been scrapped and stored in a MongoDB database. For the annotation we created an app based on FastAPI and ReactJS. We used PyTorch to train machine learning models and Streamlit to serve the web application. You can see the results on our website.

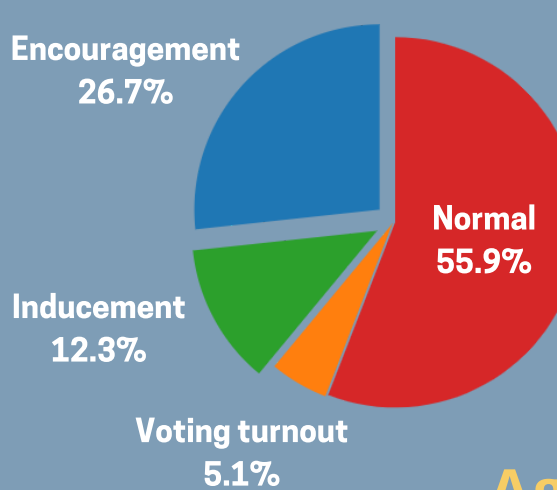


How good is our model?

| | TF-IDF | FASTTEXT | HERBERT | ENSEMBLE |
|----------------|--------|----------|---------|----------|
| NORMAL | 72% | 72% | 80% | 81% |
| VOTING TURNOUT | 50% | 61% | 70% | 70% |
| ENCOURAGEMENT | 48% | 48% | 60% | 60% |
| INDUCEMENT | 46% | 49% | 60% | 61% |
| TOTAL F-SCORE | 54% | 58% | 68% | 69% |
| NON-AGITATION | 75% | 77% | 83% | 83% |
| AGITATION | 62% | 63% | 74% | 75% |
| TOTAL F-SCORE | 69% | 70% | 79% | 79% |

DATASET

We focused on the 2020 Polish presidential election campaign. Tweets were being collected between 05.02.2020 and 12.07.2020. Every tweet in the dataset contains at least one of these hashtags:



Final dataset stats:

- 1 500 000+ total tweets
- 15 790 filtered tweets
- 6 126 annotated tweets
- 0.66 Cohen's Kappa
- 5 annotators

Agitation

Inducement

Idźcie na #WyboryPrezydenckie2020 i zagłosujcie. A jeśli nie zagłosujecie na Andrzeja Dudę, to będzie mi bardzo miło swoją drogą.

Encouragement

Rafał Trzaskowski jest twarzą skompromitowanej partii byłego obozu rządzącego i do tego słańtym prezydentem Warszawy. #wybory2020

Non-agitation

Voting turnout

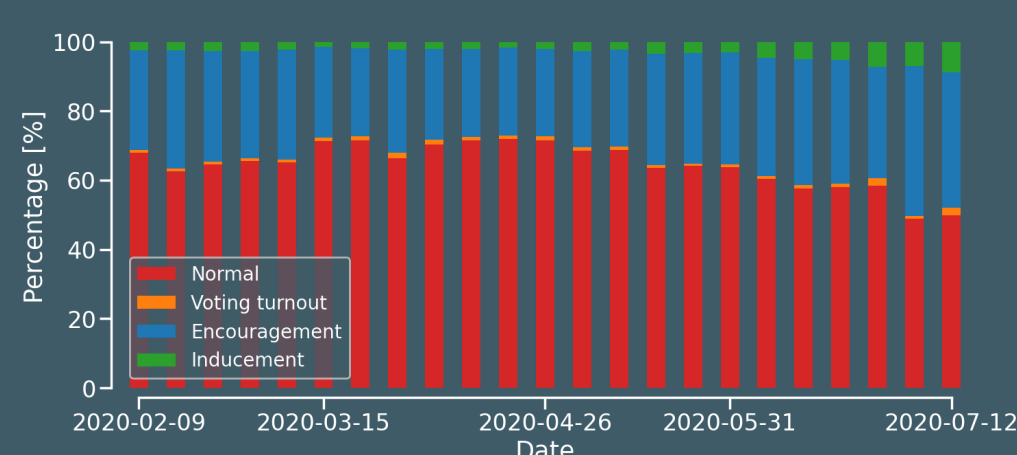
Dzisiaj święto demokracji! II Tura wyborów prezydenckich! Idźcie i głosujecie, każdy głos jest dziś ważny! #wyboryprezydenckie2020

Normal

A może zamiast tych wyborów kopertowych przeprowadzić po prostu ankietę na Twitterze. Tak samo bezpieczne i tajne :D #wybory2020

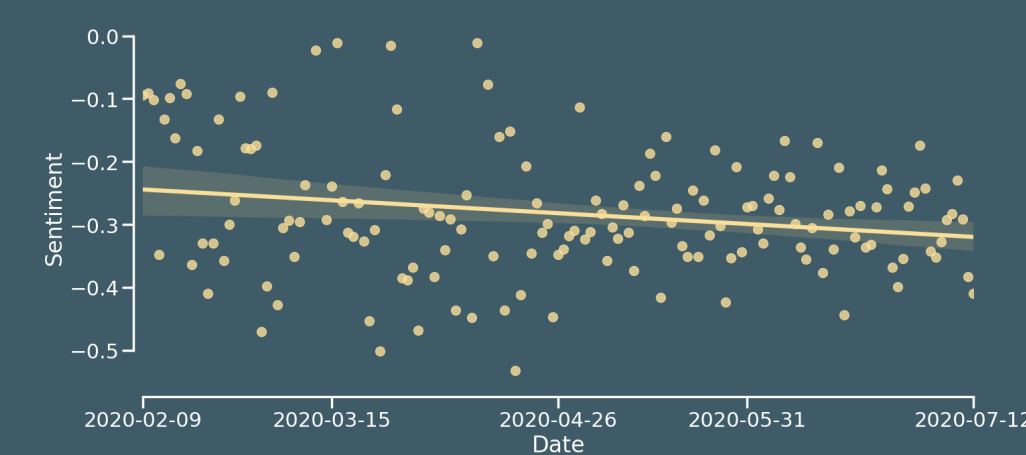
OUTCOMES

How much agitation is there?



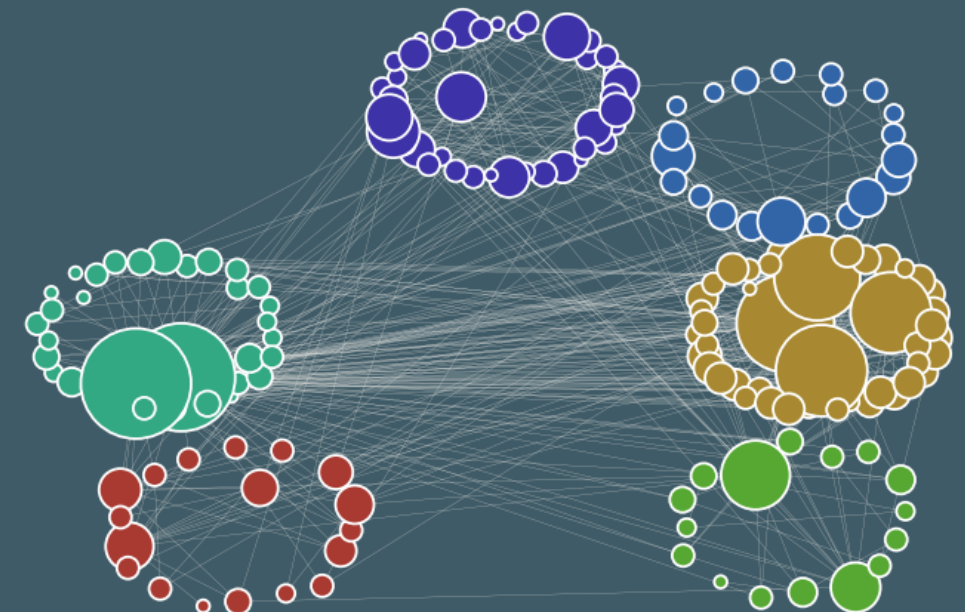
Is agitation flooding Twitter? Between the first to the last weeks of the campaign agitation **grew by 17%** in relation to all of the published tweets. There had been a peak value of political agitation by inducement to vote for particular candidate a week before second round of elections. Twitter is becoming notable agitation spot.

How were the sentiment changing?



Let's take a look at the average daily tweet sentiment - it had been trending downward during the campaign. The closer we got to the election, the **more aggressive** was the agitation becoming. Notice that the sentiment was always negative. Society was becoming more and more politically polarized. Surprised?

Who was in the spotlight?



Most oftenly political opponents were forming one community. Same user group mostly tweeted about **Andrzej Duda** and **Rafał Trzaskowski**. The **rest of the candidates** were mostly focused on themselves.

Are you interested?
Check it out!
smart-wust.pl
mateusz.wojcik@pwr.edu.pl

