Fear is falling, people tweeting having fun

Applying SOTA emotion classifier to analyze public option regarding covid vaccination on polish twitter

Kemal Erdem, Piotr Mazurek, Piotr Rarus

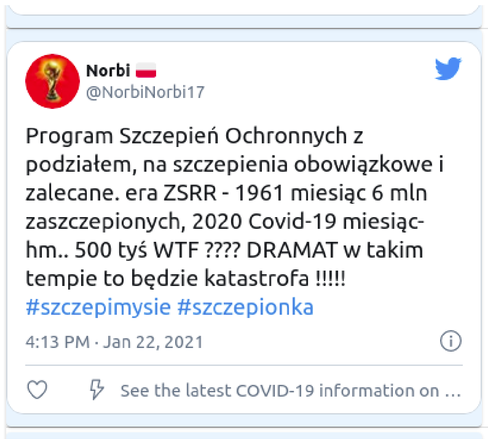
Department of Computer Science and Management, Wrocław University of Science and Technology
Wybrzeże Stanisława Wyspiańskiego 27, 50-370 Wrocław

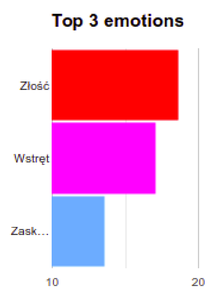
smapwr.github.io/project-polish-emotions-classifier/

1. Introduction

The project's goal was to **measure the change of emotions, expressed in tweets posted with hashtag *#szczepionka***, on polish twitter from 1st December 2020 until 21 January 2021.

We trained a SOTA transformer-based polish language model that is capable of identifying one of 9 emotions in a given text message.





We measure the change in the distribution of each emotion over time and show how changing trends can be linked to specific events.

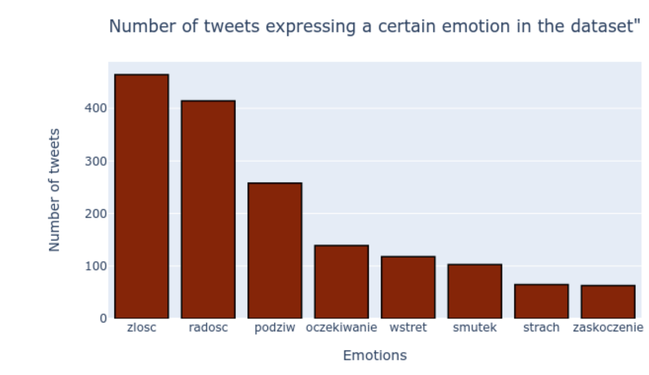
We also present the tool that you can use to identify the most common emotions expressed in your Twitter feed.

The model achieves acceptable accuracy rates in the task of identifying emotions and in our opinion can be successfully used for monitoring social media.

2. Data sets

The data set, which we used to train our models, was acquired through scrapping posts on diverse topics from Twitter. 31 annotators labeled **4212 tweets** with one of **nine emotions**. To ensure the validity of the annotation, each tweet was labeled by 3 independent annotators. We calculated Cohen's kappa coefficient and determined its value as *0.51*. Most of the tweets were annotated as *neutral* (over 50% of the dataset).

To train our models, we calculated their sentence and sequence embeddings using pretrained Herbert[1]. We improved their quality by removing any URLs and replacing emotes with actual words, as Herbert is unable to tokenize them properly.



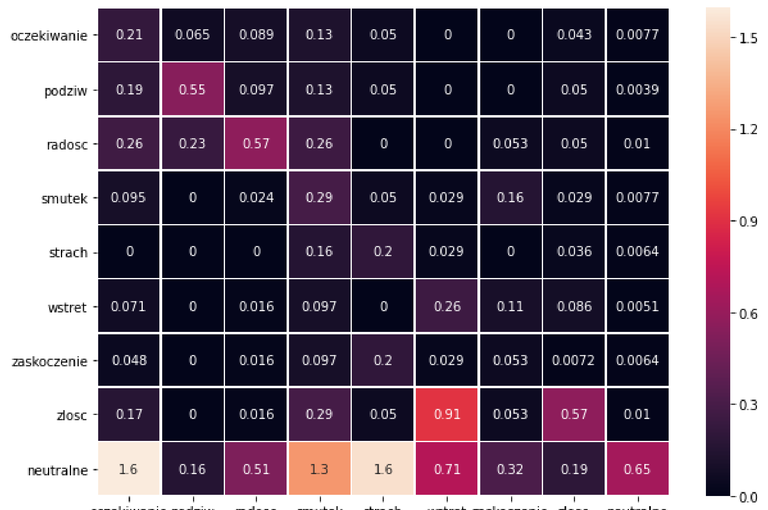
Additionally, another data set, without labels, has been made available to us. This data set comprises over million tweets and was preprocessed the same way.

3. Models

We developed several families of models to classify emotions: Dense model, RNN model, semi-supervised pseudo-labeling and semi-supervised VAE M2. Pseudo-labeling is most straightforward semi-supervised approach, it involves two models Teacher and Student. First Teacher model is trained with supervised data, then it is used to predict labels of unsupervised data set. We use those pseudo-labeled data to train the Student model. Semi-supervised VAE M2 follows the ideas presented in [2].

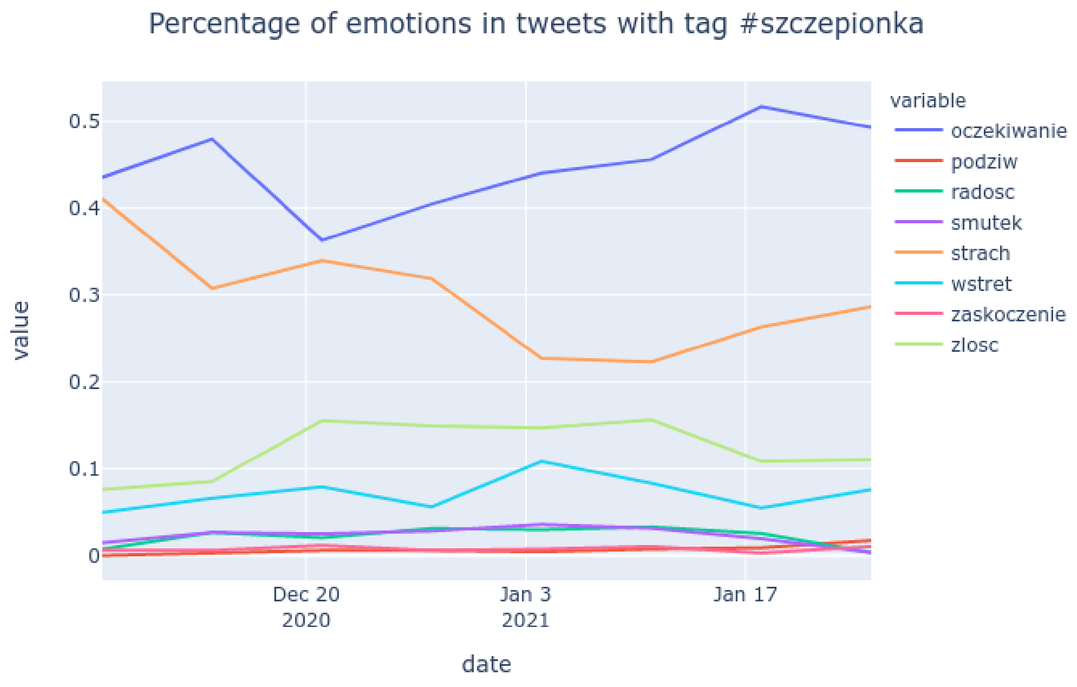
Model	Accuracy	Top 2 acc.	F1 score
Dense	0.56	0.71	0.30
RNN	**0.68**	**0.79**	**0.35**
TS	0.65	0.72	0.15
VAE M2	0.60	0.76	0.31

Despite using significantly smaller dataset the simple (GRU-based[3]) **RNN** model outperforms weakly supervised models and achieves the **best performance**.



4. Results

Below we show the changes in the percentage of concrete emotions in tweets over time. We measure only a top emotion returned by our best model (RNN).



The **anger** (green) is on a **steady rise** since the beginning of the study which may be connected to the problems with the vaccine availability.
On the other hand the **fear** (orange) **drops** significantly from almost **40%** to only **25%** of the tweets.
The later we measure **the higher percentage** of tweets express "**anticipation**" (blue), and during the whole period it was our most common emotion.

5. Conclusions

1. Emotions about vaccination change over time
2. We can observe the rise of anger, the rise of expressed anticipation and the fall of fear
3. It is possible to detect emotions in a text using relatively small data sets

6. References

[1] Rybak et al. KLEJ: Comprehensive benchmark for polish language understanding. Association for Computational Linguistics, 2018.

[2] Shakir Mohamed Max Welling Diederik P. Kingma, Danilo J. Rezende. Semi-supervised learning with deep generative models. *Neural Information Processing Systems (NIPS)*, 2014.

[3] Chung et al. Empirical evaluation of gated recurrent neural networks on sequence modeling. 2014.