

A Network Tour of Data Science

Professor Pierre Vandergheynst

Project Proposal

ERCOLANI CHIARA
VOROBIEV MIKHAIL
JUILLARD MICHAËL

Our project aims at predicting the outcome of elections based on Twitter data. First of all we are going to scrape data from Tweets regarding various politicians right before an election. Afterwards we are going to perform a sentiment analysis on these Tweets, trying to see if the sentiment changes in correspondence with important events. To do this, we are going to use the sentiment analysis tool CLiPS. This analysis will tell us whether the Twitter population leans towards or against this candidate. Based on this we are going to predict the outcome of the election. After this training phase based on past elections, we are going to try to predict who will win the upcoming elections according to the current twitter sentiment about the candidates. We are going to use simple machine learning algorithms like linear prediction and we are going to train it with about 50-100 analysis from politicians.

Since the tool that we are going to use allows to perform sentiment analysis in different languages, we can focus on predicting a variety of election outcomes from countries all over the world. It will be interesting to see if the sentiment analysis algorithm trained for example in the United States will work as well for politicians in Europe.