Project Milestone 1

Group members:
- Mads Yar – s193992
- Salik Muneeb – s215133
- Carlos Liger – s243308
- Nacho ?

Topic + Data

Title: *Probabilistic Classification of Political Affiliation in Tweets.*
Data: US Election 2020 Tweets dataset from Kaggle

Short abstract

In this project, we want to explore probabilistic model-based machine learning approaches for classifying political affiliation (Democratic vs. Republican) from Twitter data. Using the US Election 2020 tweets dataset, we want to build generative models or to train a classifier which can capture the textual patterns associated with each political party. The goal is to evaluate whether such models can effectively distinguish between Republican and Democratic tweets while offering uncertainty estimates.