**Project Proposal:**

Cyberbullying Detector

We take the data from "Mendely Data" on the URL <https://data.mendeley.com/datasets/jf4pzyvnpj/1>.

The data composed from different csv folders that distinguish by the social media platform like twitter, Kaggle, YouTube, and from different types of cyber-bulling, hate speech, racism, sexism, aggression etc.

The data contain almost 500,000 samples, that divide to 0 and 1, 0 represent not bulling text and 1 represent bulling text.

The data contain text the actual comments in text, and the label.  
some of the data (from YouTube for example) have others feature but we are not going to use them because we don’t have them across all the datasets (big part don’t have).

We will try different kind of methods for text representation, like count the words that reuse in the text using TF-IDF under assumption that cyber-bulling message will contain bulling words.

We will try to split the data using some methods like 70% for train and 30% for test, we too will try another split of the data 80% for train and 20% for test, finally we will choose the model that bring us the best result.

As you can see above, we try to solve classification problem, we want to classify texts to two types of messages, bulling messages or not.