Datasets:
https://pub.towardsai.net/the-50-best-public-datasets-for-machine-learning-d80e9f030279

1: FAKE NEWS dataset:
This could be nice, but I am still thinking the question. Could it be able to detect fake news from similarity of text? (We could ask that to the teacher).

https://www.kaggle.com/c/fake-news/data
https://www.kaggle.com/datasets/saurabhshahane/fake-news-classification

2: SUICIDE TEXT
Datatext of suicidal text, this is interesting. Only problem no metadata.
https://www.kaggle.com/datasets/nikhileswarkomati/suicide-watch

3: Cyber-bulling dataset:
Detect different types of cyber-bullying like hate speech, aggression, insults and toxicity . :)
https://www.kaggle.com/datasets/saurabhshahane/cyberbullying-dataset
https://ieee-dataport.org/documents/cyber-bullying-types-datasets