Data Science Project: Cyberbullying Classification

1. **Project Overview**

With rise of social media coupled with the Covid-19 pandemic, cyberbullying has reached all-time highs. We can combat this by creating models to automatically flag potentially harmful tweets as well as break down the patterns of hatred. In this project, by using the provided data, you are expected to train (at least) three classification models for cyberbullying classification, report and discuss the evaluation results. You are encouraged to explore both the traditional text classification methods and the deep learning methods.

2. **Dataset Description**

As social media usage becomes increasingly prevalent in every age group, a vast majority of citizens rely on this essential medium for day-to-day communication. Social media's ubiquity means that cyberbullying can effectively impact anyone at any time or anywhere, and the relative anonymity of the internet makes such personal attacks more difficult to stop than traditional bullying.

On April 15th, 2020, UNICEF issued a warning in response to the increased risk of cyberbullying during the COVID-19 pandemic due to widespread school closures, increased screen time, and decreased face-to-face social interaction. The statistics of cyberbullying are outright alarming: 36.5% of middle and high school students have felt cyberbullied and 87% have observed cyberbullying, with effects ranging from decreased academic performance to depression to suicidal thoughts.

In light of all of this, this dataset (i.e., cyberbullying_tweets.csv) contains more than 47,000 tweets labelled according to the class of cyberbullying:

- Age;
- Ethnicity;
- Gender;
- Religion;
- Other type of cyberbullying;
- Not cyberbullying

The dataset has been balanced in order to contain ~8000 tweets in each class.
Warning: These tweets either describe a bullying event or are the offense themselves, therefore explore it to the point where you feel comfortable.

3. **Additional Resources**

The dataset is provided in Paper [1]. You may also refer to some useful preprocessing and visualization codes at [2].

[1]. J. Wang, K. Fu, C.T. Lu, "SOSNet: A Graph Convolutional Network Approach to Fine-Grained Cyberbullying Detection," Proceedings of the 2020 IEEE International Conference on Big Data (IEEE BigData 2020), December 10-13, 2020.
[2]. https://www.kaggle.com/lizakonopelko/cyberbullying-on-twitter-visualization