Project Proposal: Aggressive language classifier
Natural Language Processing

Students
Manoj Kolpe Lingappa
Alan Gomez

Introduction
Nowadays, the use of social media is part of our everyday life. There are side effects of extensive use of the internet such as verbal aggression and cyberbullying. A lot of studies have been carried out to study the semantic field of aggression and it's aspects in multicultural communities [1]. The main area of focus in this field is to develop a machine learning model that can detect toxicity in online conversations. Toxicity is anything disrespectful, rude, or making someone leave a conversation [2].

Objective
The objective of this small project is to train a classifier that enables to detect the aggressive language in a word document of speech or text.

Approach
Jigsaw toxic comment data [2] is a set of data containing toxic comments and binary indicators to represent whether the data is toxic or non-toxic. For this project, a small part of Jigsaw toxic comment data will be used to train, test, and validate a supervised learning classifier in order to identify aggressive words. Data preprocessing is done to remove unnecessary items from the dataset. After data preprocessing, data is split into train, test, and validate dataset.
N-gram features from the data are taken and weight them according to the TFIDF value. Then, a logistics regression model is used for the classification. Finally, the hyperparameters of the machine learning model are tuned to get the best result [3].

Reference

1. Gordeev, D. (2016, August). Detecting state of aggression in sentences using CNN. In International Conference on Speech and Computer (pp. 240-245). Springer, Cham.
2. https://www.kaggle.com/c/jigsaw-multilingual-toxic-comment-classification/data
3. Detecting Hate Speech and Offensive Language on Twitter using Machine Learning: An N-gram and TFIDF based Approach Aditya Gaydhani, Vikrant Doma, Shrikant Kendre and Laxmi Bhagwat.Department of Computer Engineering, Maharashtra Institute of Technology, Pune.