



Data Science Intern at Data Glacier

Project: Hate Speech Detection Using Transformers (Deep Learning)

Week 7: Project Deliverable 1

Team Name: Fibo

Team Members: Mahyar Arani

Email: arani.mahyar@gmail.com

Batch code: LISUM19

Submission date: Apr 19, 2023

Submitted to: Data Glacier

Problem Description:

The problem at hand is detecting hate speech in Twitter tweets using a machine learning model. Hate speech is a type of communication that attacks or uses derogatory or discriminatory language against a person or group based on their religion, ethnicity, nationality, race, color, ancestry, sex, or other identity factors. The goal is to develop a model that can accurately classify whether a tweet contains hate speech or not.

Business Understanding:

Detecting hate speech in social media is an important task for businesses, governments, and society as a whole. Hate speech can have a detrimental impact on the mental health and wellbeing of individuals who are targeted by it. It can also lead to the spread of misinformation, the polarization of society, and even violence. By developing a hate speech detection model, businesses and governments can proactively identify and address hate speech on social media, leading to a safer and more inclusive online environment. Furthermore, businesses can use this model to understand their customers' sentiments and improve their products and services accordingly.

Project Lifecycle

Weeks	Date	plan
Weeks 07	Apr 19, 2022	Problem Statement, Data Collection, Data Report
Weeks 08	Apr 26, 2022	Data Preprocessing
Weeks 09	May 2, 2022	Feature Extraction
Weeks 10	May 9, 2022	Building the Model
Weeks 11	May 16, 2022	Model Evaluation
Weeks 12	May 23, 2022	Flask Development + Heroku
Weeks 13	May 30, 2022	Final Submission (Report + Code + Presentation)