



Data Science Intern at Data Glacier

Project: Hate Speech Detection using Transformers (Deep Learning)

Week 7: Deliverables

Name: Pranav Boyapati

University: University of Texas at Dallas

Email: pranav13b@gmail.com

Country: United States

Specialization: Data Science

Batch Code: LISUM34

Date: July 19, 2024

Submitted to: Data Glacier

Table of Contents:

1. Project Plan	3
2. Problem Statement	3

1. Project Plan

Weeks	Date	plan
Weeks 07	July 19, 2024	Problem Statement, Data Collection, Data Report
Weeks 08	July 26, 2024	Data Preprocessing
Weeks 09	August 2, 2024	Feature Extraction
Weeks 10	August 9, 2024	Building the Model
Weeks 11	August 16, 2024	Model Result Evaluation
Weeks 12	August 23, 2024	Flask Development + Heroku
Weeks 13	August 30, 2024	Final Submission (Report + Code + Presentation)

2. Problem Statement

The term hate speech is understood as any type of verbal, written or behavioural communication that attacks or uses derogatory or discriminatory language against a person or group based on what they are, in other words, based on their religion, ethnicity, nationality, race, colour, ancestry, sex or another identity factor. In this problem, We will take you through a hate speech detection model with Machine Learning and Python.

Hate Speech Detection is generally a task of sentiment classification. So for training, a model that can classify hate speech from a certain piece of text can be achieved by training it on a data that is generally used to classify sentiments. So for the task of hate speech detection model, We will use the Twitter tweets to identify tweets containing Hate speech.