

Week 7 Deliverables

Orkun

19 June 2024

Team Member's Details

- **Group Name:** Orkun
- **Name:** Orkun Kınay
- **Email:** orkunkinay@sabanciuniv.edu
- **Country:** Turkey
- **College/Company:** Sabancı University
- **Specialization:** NLP

Problem Description

The goal of this project is to develop a hate speech detection model using Twitter data. The dataset consists of tweets labeled as hate speech or non-hate speech, which will be used to train and evaluate machine learning models. The primary objective is to accurately classify tweets and mitigate the spread of hate speech on social media platforms.

Business Understanding

Hate speech on social media is a growing concern, leading to negative impacts on individuals and communities. Detecting and filtering hate speech is crucial for maintaining a safe and inclusive online environment. This project aims to provide a reliable solution for automated hate speech detection, which can be integrated into social media platforms for real-time monitoring and intervention.

Project Lifecycle and Deadline

- **Week 7:** Project initiation, problem description, and data intake.
- **Week 8:** Data understanding and preliminary data analysis.
- **Week 9:** Data cleaning and transformation.
- **Week 10:** Exploratory Data Analysis (EDA).
- **Week 11:** EDA presentation and proposed modeling technique.
- **Week 12:** Model selection, building, and evaluation.

Data Intake Report

- **Name:** Twitter Hate Speech Detection
- **Report Date:** 19 June 2024
- **Internship Batch:** LISUM33
- **Version:** 1.0
- **Data Intake by:** Orkun Kınay
- **Data Intake Reviewer:** Data Glacier
- **Data Storage Location:** <https://www.kaggle.com/datasets/vkrahul/twitter-hate-speech>

Tabular Data Details

- **Total Number of Observations:** 31962
- **Total Number of Files:** 2
- **Total Number of Features:** 3
- **Base Format of the File:** .csv
- **Size of the Data:** 4.7 MB

Proposed Approach

- No missing values

GitHub Repo Link

<https://github.com/orkunkinay/Hate-Speech-Detection/tree/main>