Week 7: Project Deliverables

Project Topic: Hate Speech Detection using Transformers (Deep Learning)

Group Name: Data Defenders

Batch Code: LISUM34

Team Members' Details:

Name	Malika Hafiza Pasha	Biswadip Bhattacharyya	Oluseun Omobulejo
Email	malikahafizap@gmail.com	bbhattac@syr.edu	Seun.omobulejo@outlook.com
Country	United States	United States	United Kingdom
College/Company	California State University – Dominguez Hills	Syracuse University	University of Westminster
Specialization	Data Science		

Problem Description:

Any verbal, written, or behavioral communication that targets or uses derogatory or discriminatory language against an individual or group on the basis of who they are—that is, based on their religion, ethnicity, nationality, race, color, ancestry, sex, or another identity factor—is referred to as hate speech. We will walk you through a Python and machine learning hate speech detection model in this problem.

Sentiment categorization is often the process involved in hate speech detection. Therefore, training a model on data that is often used to classify attitudes can result in a model that can identify hate speech from a given text passage. Therefore, in order to complete the objective of developing a hate speech recognition model, we will use Twitter to find tweets that include hate speech.

Business Understanding:

The rise of social media and online communication platforms has increased the dissemination of hate speech. It is crucial for businesses, social media platforms, and communities to identify and address hate speech to maintain a safe and inclusive environment. Effective detection and moderation can enhance user experience, comply with regulations, and protect brand reputation.

Project Lifecycle Along with Deadline:

1. Project Initiation (20th July - 22nd July)	• Defining the project scope and objectives. • Forming the project team and assign roles. • Gathering initial data and resources.
2. Data Collection and Preparation (23rd July - 29th July)	• Collecting a diverse dataset of text samples, including examples of hate speech and non-hate speech. • Cleaning and preprocessing the data (e.g., remove stop words, tokenize text).
3. Model Development (30th July - 8th August)	• Choosing and implementing suitable Transformer models (e.g., BERT, RoBERTa). • Training the models on the prepared dataset. • Fine-tuning the models for optimal performance.
4. Evaluation and Optimization (9th August - 13th August)	• Evaluating the models using appropriate metrics (e.g., accuracy, precision, recall, F1-score). • Optimizing the models based on evaluation results.
5. Deployment and Testing (14th August - 16th August)	• Deploying the final model to a suitable environment. • Conducting thorough testing to ensure reliability and performance.
6. Documentation and Reporting (17th August - 20th August)	• Preparing the final report and documentation. • Submitting the project deliverables.

Data Intake Report:

Dataset Details:	
1. Dataset Name	train_E6oV3lV
2. Dataset storage location	Twitter hate speech (kaggle.com)
3. Base format of the file	csv
4. Size of the data	2.95 MB
5. Total number of observations	31962
6. Total number of files	1
7. Total number of features	3

8. Proposed Approach	There are no missing vales in this dataset

GitHub Repository Link:

https://github.com/malikahafizap/Hate_Speech_Detection_using_Transformers