

WEEK 7 REPORT

Group Name:

	Member 1	Member 2
Name	Hyejoon	Ky Dang
Email		Keith.dang1610@gmail.com
Country	UK	Vietnam
College/ Company		
Specialization	NLP	NLP

Problem description:

Hate speech is a form of communication, whether verbal, written, or behavioral, that attacks or discriminates against an individual or group based on their inherent characteristics, such as religion, ethnicity, nationality, race, gender, or other identity factors. The emergence of hate speech on social media platforms like Twitter poses significant challenges, including creating a toxic environment for users and impacting the platform's reputation.

To address this issue, the goal is to develop a machine learning-based hate speech detection model. This model will classify tweets as hate speech or not, using sentiment classification techniques. By leveraging a dataset of tweets commonly used for sentiment analysis, the model will learn to identify patterns associated with hate speech. This project combines data preprocessing, feature extraction, and machine learning to build an effective classifier capable of addressing the problem of online hate speech in a scalable and automated way.

Business Context

The widespread use of social media platforms like Twitter has enabled users to share opinions, ideas, and engage in conversations globally. However, this openness also facilitates the propagation of hate speech, which includes derogatory or discriminatory expressions targeting individuals or groups based on religion, ethnicity, nationality, race, gender, or other identity factors. Detecting and addressing hate speech is a pressing concern to maintain a healthy online ecosystem, protect users from harm, and ensure regulatory compliance.

Business understanding:

-Developing a hate speech detection model isn't just a technical challenge but also an effort to address real-world problems.

-The challenge lies in handling the nuanced nature of language, such as sarcasm, context, or cultural references, while minimizing false positives (overflagging neutral text) and false negatives (missing hate speech)

Project lifecycle

Weeks	Due date	Plan
Week 7	11/19/2024	Business Understanding
Week 8	11/26/2024	Data understanding
Week 9	12/02/2024	Data cleansing and transformation done on the data.
Week 10	12/09/2024	EDA performed on the data
Week 11	12/16/2024	EDA Presentation and proposed modeling technique
Week 12	12/23/2024	Model Selection and Model Building/Dashboard
Week 13	12/30/2024	Document the challenge

Github Repo link:

- Individual GitHub links:
 - o KyDang:
 - o Hyejoon: <https://github.com/candy0543/HateSpeech.git>
- Group GitHub Link: <https://github.com/candy0543/HateSpeech.git>