Yashraj Parmar

Data Glacier

19 May 2025

Week 7: Deliverables

1) **Information:**
    a) Group Name: NLP Sentinels
    b) Name: Yashraj Parmar
    c) Email: yparmar2024@gmail.com
    d) Country: United States of America
    e) College: Stevens Institute of Technology
    f) Specialization: Natural Language Processing
    g) Internship Batch: LISUM44
2) **Problem Description:**
    a) 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 such as religion, ethnicity, nationality, race, colour, ancestry, sex or identity factors. Hate speech can be seen on various social media platforms which goes against most of social media platform's ideologies of keeping their website safe for all users. Social media platforms develop models to determine if a piece of text is considered hate speech or not.
3) **Business Understanding:**
    a) Most social media platforms struggle to detect and moderate hate speech effectively due to the complexity of language, context, and slang. Automating

hate speech detection improves content moderation, user safety, brand reputation and legal compliance. A well-trained model can assist in flagging toxic content quickly and at scale.

4) Project Life Cycle along with Deadline:

a) Hate Speech Detection Timeline:

Dates	**Plan**
Week 7: 5/12 - 5/19	Define the problem and business context, outline the project plan, and set up the GitHub repository.
Week 8: 5/19 – 5/26	Analyze the dataset to understand its structure, identify issues like missing values or outliers, and plan your data handling strategies.
Week 9: 5/26 – 6/2	Perform data cleaning and transformation using multiple techniques, applying and reviewing different methods.
Week 10: 6/2 – 6/9	Conduct exploratory data analysis (EDA) to uncover insights, visualize trends, and finalize the approach for modeling.
Week 11: 6/9 – 6/16	Create and present an EDA presentation for business users, ending with model recommendations for technical users.
Week 12: 6/16 – 6/23	Build and evaluate various models from different families, select one that aligns with business needs, and (if applicable) create a dashboard.
Week 13: 6/23 – 6/30	Finalize and present the project with a complete report, working code, select and justify the best solution.