

PROJECT PROPOSAL

Hate Speech Detection and Analysis in Indian Social Media(Hinglish)

Problem Statement:

Social media in India sees a lot of hate speech, especially during political events, and most detection tools only work in English. This project focuses on identifying hate content in Indian languages like Hindi, Tamil, Bengali, and Telugu, and visualizing when and where it spikes to help create a real-world impact.

Data Source:

Source 1: Web Scraping Twitter or X

Source 2: Web Scraping YouTube comments on Political or News videos

Source 3: Finding a similar dataset from Kaggle

to Do

1)Preprocessing + Language Detection

LangDetect, **fastText**, or **CLD3** to detect language

2)Text Classification

Naive Bayes / SVM (for TF-IDF)

BERT / MuRIL (Google's multilingual BERT for Indian languages)

3) Visualization

Create **Geo heatmaps**

Deliverables:

Classification Model for hate speech in multiple Indian languages.

Geo Heatmap to identify geographical hotspots of hate content(If location is available like IP)

Final Report outlining methodology, insights, model performance, and policy relevance.