GUVI

DATA SCIENCE JOB FAIR-2024

DATA SCIENTIST

AI ML

Hate Speech

Objective:

The objective of this project is to analyze and mitigate hate speech in online platforms through a comprehensive approach that encompasses contextual analysis, temporal analysis, user profiling, geospatial analysis, false positive/negative analysis, and hate speech detection. By addressing these aspects, the project aims to enhance our understanding of hate speech dynamics, identify patterns and trends over time, profile users prone to engaging in hate speech, analyze geographic distributions, evaluate model performance, and ultimately develop effective strategies to detect and mitigate hate speech online.

Data Understanding:

The dataset comprises the following features:

Dataset: Link

Features:

File_id: Unique identifier for each file containing textual context.

U**ser_id:** Identifier for the user posting the context.

S**ubforum_id:** Identifier for the subforum where the context was posted.

N**um_contexts:** Number of contexts associated with each file.

L**abel:** Label indicating whether the context contains hate speech or not.

Approach:

Contextual Analysis:

- Analyze the textual context provided in the dataset to understand the linguistic patterns, sentiment, and thematic content associated with hate speech.
- Utilize natural language processing (NLP) techniques to extract meaningful insights from the text data.
- Explore word embeddings and topic modelling to identify key themes and topics within hate speech contexts.

Temporal Analysis:

- Investigate temporal patterns and trends in hate speech occurrences over time.
- Analyze whether there are spikes or fluctuations in hate speech activity during specific periods or events.
- Utilize time series analysis techniques to model and forecast hate speech trends.

User Profiling:

- Profile users based on their propensity to engage in hate speech.
- Identify demographic characteristics or behavioural patterns associated with users who are more likely to produce hate speech content.

Geospatial Analysis:

- Explore the geographic distribution of hate speech occurrences.
- Analyze whether there are regional or spatial variations in hate speech prevalence.
- Visualize hate speech hotspots using geospatial mapping techniques.

False Positive/Negative Analysis:

- Evaluate the performance of hate speech detection models by analyzing false positive and false negative instances.
- Examine instances where the model misclassifies non-hate speech as hate speech (false positives) or vice versa (false negatives).
- Investigate the content and target of misclassified instances to identify common patterns or challenges.

Hate Speech Detection:

- Develop and train Deep Learning models to detect hate speech in online contexts.
- Experiment with various classification algorithms.
- Evaluate model performance using appropriate metrics.

Skills Outcome:

- Proficiency in data preprocessing techniques for textual data.
- Understanding of natural language processing (NLP) techniques for text analysis.
- Ability to perform exploratory data analysis (EDA) to derive meaningful insights from datasets.
- Experience in building and training Deep Learning models for classification tasks.
- Competence in evaluating model performance and interpreting results.
- Skills in data visualization and storytelling to communicate findings effectively.

Submission:

- Provide a well-commented Python file (*.ipynb / *.py) containing the complete code for the project, organized into sections for data preprocessing and Analysis.
- Upload the same into github with proper Readme file.
- Presentation on the entire project, including Problem Statement, Tools Used, Approaches and Insights Found.

Evaluation metrics:

- Project evaluation will be done in the live session and have to showcase the approaches done to complete the project
- You are supposed to write a code in a modular fashion (in functional blocks)
- Maintainable: It can be maintained, even as your codebase grows.
- Portable: It works the same in every environment (operating system)
- You have to maintain your code on GitHub.(Mandatory)
- You have to keep your GitHub repo public so that anyone can check yourcode.(Mandatory)
- Proper readme file you have to maintain for any project development(Mandatory)
- Follow the coding standards: https://www.python.org/dev/peps/pep-0008/
- You should include basic workflow and execution of the entire project in the readme file on GitHub

GitHub Repo:

The attached reference document will help you use GitHub effectively. - **Link**