

Software Engineering and Product Management

Lab 1

**Group Members:**

* Abhishek Anand
* Kaustubh Desale
* Aman Kr. Srivastava
* Himanshu Shekhar Mishra
* Gotam Gorabh

Hate Speech Detection

Hate speech detection is the task of detecting if communication such as text, audio, and so on contains hatred and or encourages violence towards a person or a group of people. This is usually based on prejudice against 'protected characteristics' such as their ethnicity, gender, sexual orientation, religion, age etc.

This is an artificial intelligence-based tool that is designed to analyze hate speeches in real-time audios and texts.

**Abstraction**

It is the simplification of a problem by focusing on only one aspect of the problem while omitting all other aspects. When using the principle of abstraction to understand a complex problem, we focus our attention on only one or two specific aspects of the problem and ignore the rest.

Whenever we omit some details of a problem to construct an abstraction, we construct a model of the problem. In everyday life, we use the principle of abstraction frequently to understand a problem or to assess a situation.

The Abstract of this project are,

1. Hate speech covers many forms of expressions which advocate, incite, promote or justify hatred, violence and discrimination against a person or group of persons for a variety of reasons.
2. As there is , the increasing propagation of hate speech on social media and the urgent need for effective countermeasures have drawn significant investment from governments, companies, and researchers.
3. We aim to develop a method which will automatically detect a hate speech in the

given data. We aim to classify textual content into non-hate or hate speech.

The Key Features are -:

* Analyze hate speech in real-time audios.
* Analyze hate speech in texts data.

**Decomposition:**

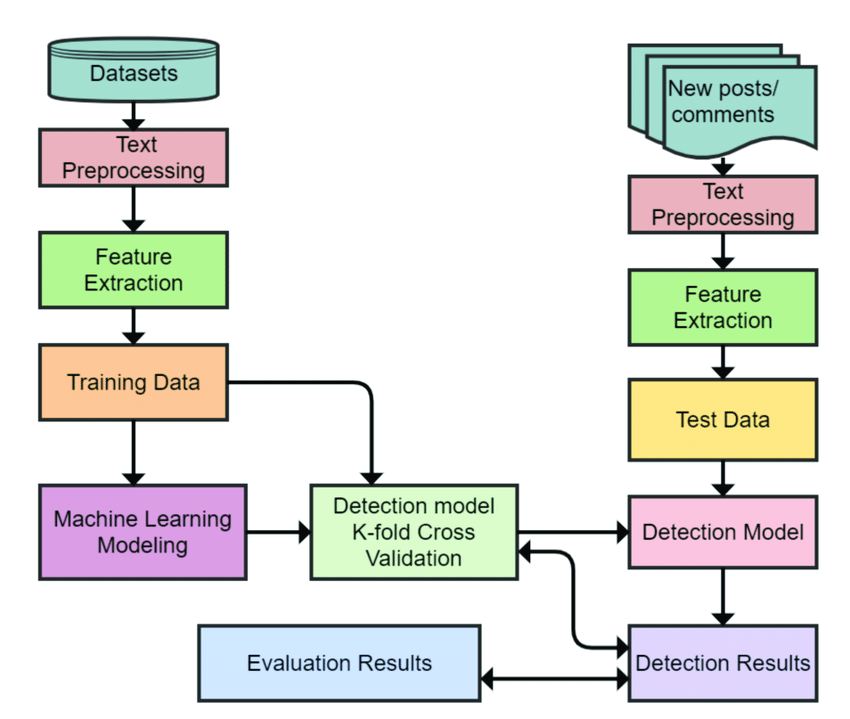
Decomposition is a process of breaking down. It will be breaking down functions into smaller parts. It is another important principle of software engineering to handle problem complexity. This principle is profusely made use by several software engineering techniques to contain the exponential growth of the perceived problem complexity. The decomposition principle is popularly is says the divide and conquer principle.

This hate speech detection system is using a happytransformer framework form Hugging-Face AI community.

As a hate speech classification model, we are using the "Hate-speech-CNERG/dehatebert-mono-english" BERT pre-trained model to do the hate speech analysis, it is an open-source model.

This prototype is combined with knowledgebase model and AI based model, first it checks if a hate keyword is matched in the target text from the knowledgebase file, if it is match (Filtering Process), then it goes to the next step to check with the AI model to understand the meaning behind the sentence, the AI model will return probability to determine that targeted text is a hate speech or not.

Step 01: Filtering - Match hate keywords from the knowledgebase.

Step 02: AI Model - Here, the filtered targeted text will be used to check the probability of hate speech with a AI model.

* **Technologies & Frameworks**
* Python 3.8
* happytransformer
* pyperclip
* pynput
* mysql\_connector
* speech\_recognition