

# CYBER-BULLYING DETECTION USING MACHINE LEARNING

## How to set up and run our project:

“It is a Machine Learning model built using classification algorithm like Naïve Bayes, SVM, Random Forest and Decision Tree Algorithms. The Model will take text as an input and detects whether there is any kind of cyberbullying detection is present in the text and classifies the text to which kind of cyberbullying it is.”

## Steps to run the ML Model on local Computer:

- 1. Install Python:** If you don't already have Python installed on your computer, download and install the latest version from the official Python website.
- 2. Install the Python extension for VS Code:** Open VS Code and click on the Extensions icon on the left-hand side of the window. Search for "Python" and click on the "Install" button to install the extension.
- 3. Create a new Python environment:** In VS Code, open a new terminal window by clicking on Terminal in the top menu bar and selecting "New Terminal". Then, create a new Python environment by running the following command in the terminal:  
**`"python -m venv myenv"`**  
This will create a new virtual environment named "myenv" in the current directory.
- 4. Activate the Python environment:** In the terminal, activate the Python environment by running the following command:  
**`.\myenv\Scripts\activate`**  
This will activate the environment and prepare it for use.
- 5. Install the required packages:** If the cyberbullying detection model requires any Python packages that are not already installed in the virtual environment, install them using pip by running the following command: **`pip install package_name`**  
Replace "package\_name" with the name of the required package.
- 6. Open the Notebook file:** In VS Code, click on File in the top menu bar and select "Open Folder". Navigate to the directory where you saved the cyberbullying detection model Jupyter Notebook file and select it to open it.

7. **Run the Notebook:** Once you have opened the Notebook file, you can run each cell of code by clicking on it and pressing Shift + Enter. This will execute the code in the cell and display any output.

In order to run this, we need some requirements for system:

- Windows/Linux/Mac operating system.
- 4 GB RAM

### **Challenges Faced During the Development:**

1. **Ambiguity of language:** Cyberbullying can take many forms and can be expressed using ambiguous language that can be difficult to detect. For example, sarcasm or teasing might be used in a message, which can be interpreted differently by different people. Thus, the model must be trained to identify such nuances in language and interpret them in the context of the message.
2. **Data availability and quality:** Building a good detection model requires a large amount of high-quality training data that includes different types of cyberbullying. However, such data can be difficult to obtain, especially for less common types of cyberbullying.
3. **Diversity of platforms:** Cyberbullying can take place across a variety of digital platforms, each with its own unique language and features. The model must be able to generalize across different platforms and adapt to new forms of cyberbullying.
4. **Balancing false positives and false negatives:** A model that is too sensitive can result in too many false positives, which can be frustrating for users. On the other hand, a model that is too specific can result in too many false negatives, which can lead to harmful messages going undetected.
5. **Ethical considerations:** The development of a cyberbullying detection model raises important ethical questions, such as the potential for the model to infringe on individual privacy or be used for harmful purposes. The model must be designed and implemented with ethical considerations in mind to prevent unintended consequences.

