Data Collection:

- Collect large text datasets from social media platforms like Twitter and online forums like Wikipedia.
- Ensure the datasets contain instances of cyberbullying, such as hate speech tweets and comments characterized by personal attacks.

Preprocessing:

- Clean the text data by removing unnecessary characters, normalizing the text to lowercase, tokenizing it into words, and removing stop words and punctuation.
- Optionally, perform stemming or lemmatization to reduce words to their root forms and handle contractions and abbreviations.

Folder Redirection and Login:

- Redirect to the project folder location using MySQL.
- Access the login page to authenticate and gain access to the system.

Feature Extraction:

- Extract relevant features from the preprocessed text data, such as bag-of-words representations, TF-IDF scores, or word embeddings.
- These features will be used as input for the machine learning models.

Dataset Upload:

- Navigate to the dataset page after successful login.
- Upload the dataset containing text data from sources like Twitter and Wikipedia.

Model Training:

- Train machine learning models, such as logistic regression, Naive Bayes, and decision tree classifiers, on the preprocessed and feature-extracted data.
- Use labeled data to train the models to distinguish between offensive and non-offensive text.

Text Input and Prediction:

- Train the machine learning models on the uploaded dataset.
- Enter new text data for prediction, such as social media posts or comments.

- Utilize the trained models to predict whether the text contains instances of cyberbullying (offensive or non-offensive).

Model Evaluation:

- Evaluate the trained models' performance using metrics like accuracy, precision, recall, and F1-score.
- Use techniques like cross-validation to ensure the models generalize well to new data.

Prediction:

- Once trained and evaluated, deploy the models to predict whether new text data contains instances of cyberbullying.
- Use the trained models to classify text as offensive or non-offensive, helping to identify and address cyberbullying behavior.

Result Analysis:

- Evaluate the prediction results using performance metrics like accuracy, precision, recall, and confusion matrix.
- Visualize the analysis through charts or graphs to understand the effectiveness of the cyberbullying detection system.

Continuous Improvement:

- Continuously monitor and update the models as new data becomes available and as the nature of cyberbullying evolves.
- Incorporate feedback and insights to improve the models' accuracy and effectiveness in detecting cyberbullying.