🚨 Detection of Cyberbullying Using Machine Learning

Cyberbullying poses a significant challenge in today's digital society. This project leverages Natural Language Processing (NLP) and Machine Learning (ML) techniques to detect various forms of cyberbullying in social media posts and other text-based user content.

📌 Objective

- Automatically detect and classify instances of cyberbullying.
- Distinguish between multiple types of bullying (e.g., gender, religion, ethnicity).
- Develop machine learning models that can generalize well to unseen data.

🗂️ Project Structure

```
cyberbullying-detection/
│
├── dataset/                    # Dataset files used for training and testing
├── models/                     # Trained models (optional, can be generated)
├── visualizations/             # Plots and confusion matrices
├── cyberbullying_model.ipynb   # Main Jupyter Notebook for modeling and evaluation
└── README.md                   # Project overview and documentation
```

📝 Dataset

The dataset used contains labeled text samples from social media, categorized into:

- Age-related bullying
- Ethnicity-related bullying
- Gender-based bullying
- Religion-based bullying
- Other (generic insults or offensive content)
- Not cyberbullying

Each row includes:

- The text content (post)

- The associated label/category

Source: Cyberbullying Classification Dataset – Kaggle

Models Implemented

Multiple machine learning models were evaluated:

• Logistic Regression - Baseline linear model for classification
• Naïve Bayes - Probabilistic classifier well-suited for text
• SVM (Linear Kernel) - High-performance linear classifier
• Random Forest - Ensemble model for improved accuracy

Evaluation Metrics

Models are assessed based on:
- Accuracy
- Precision
- Recall
- F1-Score
- Confusion Matrix

SVM achieved the best balance between precision and recall, making it the top-performing model.

Technologies Used

- Python 3
- Jupyter Notebook
- Scikit-learn
- Pandas
- NumPy
- Matplotlib / Seaborn

Installation & Setup

1. Clone the repository
   git clone https://github.com/your-username/cyberbullying-detection.git
   cd cyberbullying-detection

2. Create a virtual environment (optional)
   python -m venv venv
   source venv/bin/activate # Windows: venv\Scripts\activate

3. Install required packages
   pip install -r requirements.txt

4. Run the notebook
   jupyter notebook cyberbullying_model.ipynb

🚀 Future Improvements

- Integration with real-time social media monitoring APIs
- Deep learning models (e.g., LSTM, BERT)
- Multilingual support
- Deployment as a web application

👨‍💻 Author

Safana Shaik
Safanashaik12@gmail.com
linkedin.com/in/shaik-safana

🙌 Acknowledgements

- Dataset from Kaggle
- Inspiration from real-world moderation systems
- Open-source contributors and ML community