**🧠 Deep Learning for Comment Toxicity Detection with Streamlit**

**📌 Project Overview:**

This project addresses the challenge of toxic comments in online communities by building a deep learning model that detects multiple types of toxicity in user-generated text. The solution includes a trained LSTM-based multi-label classifier and a Streamlit web application for real-time and bulk comment analysis.

**🎯 Problem Statement:**

Online platforms face increasing challenges in moderating toxic content such as harassment, hate speech, and offensive language. Manual moderation is inefficient and inconsistent. This project aims to automate toxicity detection using deep learning, enabling scalable and real-time moderation support.

**💼 Business Use Cases:**

* **Social Media Platforms**: Real-time filtering of toxic comments.
* **Online Forums**: Automated moderation of user posts.
* **Content Moderation Services**: Enhanced moderation pipelines.
* **Brand Safety**: Ensuring safe environments for advertisers.
* **E-learning Platforms**: Protecting students from harmful content.
* **News Websites**: Moderating comment sections on articles.

**🧪 Technical Stack:**

Language – Python

Deep Learning – TensorFlow(Keras)

NLP Preprocessing – Regex, Tokenizer, Padding

Model Architecture – LSTM with Embedding and Dense layers

Evaluation – F1 Score, Threshold Tuning, Confusion Matrix

Visualization – Matplotlib, Seaborn

**🧠 Model Architecture:**

* **Embedding Layer**: Converts tokens to dense vectors.
* **LSTM Layer**: Captures sequential dependencies.
* **Dense Layers**: Outputs multi-label predictions with sigmoid activation.
* **Loss Function**: Custom weighted binary crossentropy to handle class imbalance.
* **Threshold Tuning**: Per-label F1 optimization for decision boundaries.

**📊 Evaluation Metrics:**

* **Accuracy**: Overall prediction correctness.
* **F1 Score**: Per-label harmonic mean of precision and recall.
* **Confusion Matrix**: Visual breakdown of true/false positives/negatives.
* **Threshold vs F1 Curve**: Helps select optimal thresholds for each label.

**🚀 Streamlit App Features:**

* **Single Comment Analysis**: Enter a comment and get toxicity predictions.
* **Bulk Prediction**: Upload a CSV file with comments for batch analysis.
* **Downloadable Results**: Export predictions and scores.
* **Real-time Feedback**: Instant model inference via Streamlit UI.