# MeetMux Round 4 - AI–ML Developer Intern

## Project Title: Offline Chat-Reply Recommendation System using Transformers

**Objective:** Build an offline chat-reply recommendation system using Transformers trained on two-person conversation data.

## Approach & Methodology:

1. 1. Preprocessed and tokenized long conversational data using Hugging Face Tokenizer.
2. 2. Utilized GPT-2 Transformer model (preloaded weights) for context-based response generation.
3. 3. Fine-tuned the model on two-person conversation datasets (User A & User B).
4. 4. Implemented training with AdamW optimizer and evaluated with BLEU score.
5. 5. Generated coherent, context-aware replies offline using GPU support if available.

## Model Justification:

GPT-2 was chosen for its strong language modeling and ability to generate contextually relevant responses. It handles long conversational dependencies efficiently while being lightweight enough for offline fine-tuning.

## Evaluation Metrics:

The model was evaluated using BLEU score to measure the similarity between generated replies and actual User A responses. Average BLEU score across test samples indicated good contextual understanding and fluency in responses.

## Deployment Feasibility:

The trained model was saved as Model.joblib for easy offline loading. It can be integrated into chat systems or desktop assistants without internet dependency. The lightweight architecture ensures low inference time and efficient memory usage.

## Tools and Libraries Used:

1. Python 3.10+
2. Transformers, Torch, NumPy, Pandas, scikit-learn, NLTK, Matplotlib
3. Preloaded GPT-2 model from Hugging Face
4. Joblib for model saving and offline deployment

## Results and Output:

The system successfully generated contextually relevant replies based on User B's inputs, showing strong coherence with previous conversation context. BLEU evaluation confirmed satisfactory performance for real-world offline chat recommendation scenarios.

## Conclusion:

This project demonstrates the ability to build an offline chat recommendation system using Transformer models. The fine-tuned GPT-2 model effectively captures conversation context, generates appropriate responses, and operates efficiently in offline environments, meeting the challenge objectives set by MeetMux.

**Submitted By:** Shaurya Srivastava

**Role:** AI–ML Developer Intern Candidate

**Submission:** ChatRec_Model.ipynb | Model.joblib | Report.pdf | ReadMe.txt