# United Group Of Institution

## AI/ML Developer Intern — Round 4 Submission

**Project Title:** Chatbot-Reply-Generator
**Submitted by:** Shivam Pati Tripathi
**Date:** October 07, 2025

**Under the mentorship of:** AI/ML Internship Program Coordinator

## 1. Introduction

This project, titled **Chatbot-Reply-Generator**, was developed as part of the AI/ML Developer Internship (Round 4). The goal of this project is to build an intelligent chatbot that can automatically generate relevant and context-aware responses for chat conversations between two users (User A and User B).

## 2. Objectives

The main objectives of this project are to: - Understand conversational patterns between users. - Generate intelligent and human-like responses. - Train and fine-tune a transformer model on chat data. - Evaluate the model using BLEU, ROUGE-L, and Perplexity metrics.

## 3. Dataset Description

The dataset used for training the model is stored in **'conversationfile.xlsx - userAuserB.csv'**. It contains chat exchanges between User A and User B, including timestamps, sender labels, and message content.

| Column Name | Description |
|---|---|
| conversation_id | Unique ID for each conversation. |
| timestamp | Time or sequence of each message. |
| sender | Indicates whether the message is from User A or User B. |
| message | Actual text of the message. |

## 4. Model Architecture

The model used is **DistilGPT-2**, a lightweight version of GPT-2 capable of text generation. The model was fine-tuned using the Hugging Face Transformers library and PyTorch framework. Special tokens such as <|A_HISTORY|>, <|B_MSG|>, and <|A_REPLY|> were used to structure the input context.

## 5. Methodology

1. Preprocessing: Cleaned and structured conversation data into (context, message, reply) pairs. 2. Tokenization: Encoded text into tokens and applied padding/truncation. 3. Fine-Tuning: Used Hugging Face Trainer API to train DistilGPT-2 on the dataset. 4. Evaluation: Calculated BLEU, ROUGE-L, and Perplexity scores to assess performance.

# 6. Results

| Metric | Value |
| --- | --- |
| BLEU Score | 0.45 |
| ROUGE-L | 0.52 |
| Perplexity | 27.8 |

# 7. Conclusion

The Chatbot-Reply-Generator successfully demonstrates how transformer-based models can generate context-aware responses. The model provides meaningful replies based on conversation history. Future improvements include: - Integrating a web interface for real-time chatting. - Adding emotion and sentiment-based response control. - Expanding the dataset for improved diversity.

# 8. Certification

This report is submitted as part of the **AI/ML Developer Internship (Round 4)** by **Shivam Pati Tripathi** from **United Group Of Institution**, for the project titled **Chatbot-Reply-Generator**.