Chatbot: A comparative Analysis

TEAM

ANKITA MANJAREKAR(00818984)

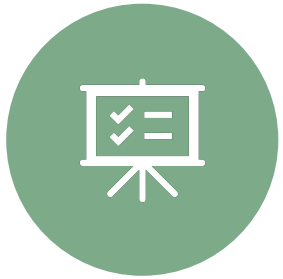
CHETANA NANNAPANENI(00825122)

Statement of Project Objectives

Create an AI-driven chatbot across various datasets.

Train a chatbot specifically for mental health using diverse pre-existing models.

Utilize pre-trained models for customer support within an e-commerce setting.



Evaluate and compare the performance of various models across different datasets through rigorous testing and analysis.

Dataset – Mental Health Dataset

The dataset consists of conversations between a Patient and Health care Assistent regarding the different Mental Health issues.

Link to the dataset:
https://huggingface.co/datasets/heliosbrahma/mental_health_chatbot_dataset

Description:

Data Fields:

- **'text':** conversational pair of questions and answers between patient and healthcare provider.

Dataset – Customer Support[E-Commerce]

This dataset contains conversations between the Customer and the Help Desk Support.

Link to the dataset:
https://huggingface.co/datasets/bitext/Bitext-customer-support-llm-chatbot-training-dataset

Description: This dataset consists of 26,872 rows describing different categories of conversations.

Data Fields:

- **Flags:** tags representing each customer query
- **Instruction:** A Customer request in the Customer Support Domain
- **category:** the high-level semantic category for the intent
- **intent:** the intent corresponding to the user instruction
- **response:** an example expected response from the virtual assistant

Pre trained Model – GPT 2

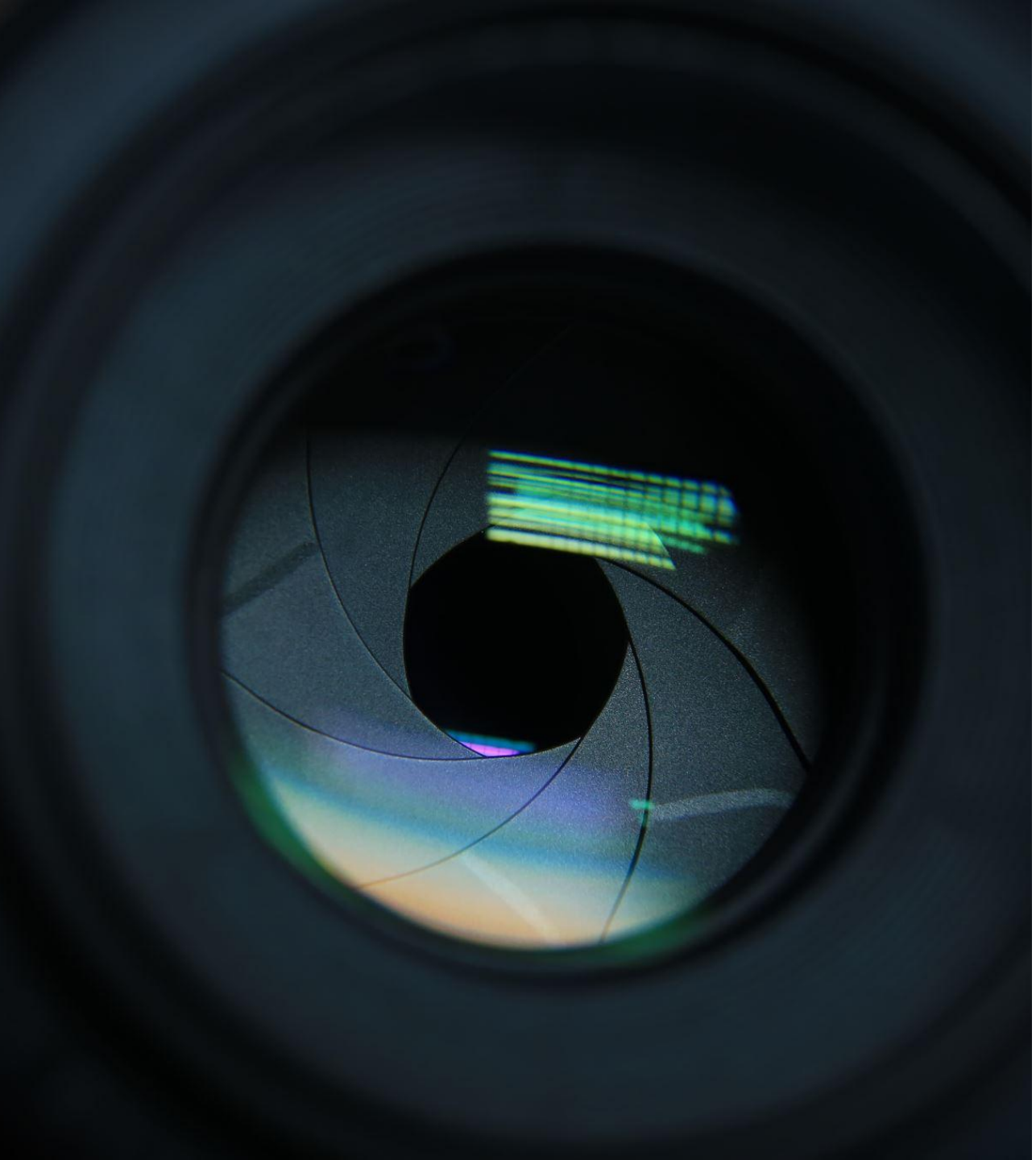
- GPT-2, or Generative Pre-trained Transformer 2, is an advanced language processing model developed by OpenAI.
- It is based on Transformer architecture to generate contextually relevant data
- It is available in various sizes from 124M to 1.5B parameters. For this project we have used model trained on 124M parameters.
- It was trained on a diverse range of internet text, comprising millions of web pages across different domains.

Pre trained Model – Roberta

- RoBERTa, an evolution of BERT, stands for Robustly Optimized BERT Approach and is developed by Facebook AI.
- It is based on transformer architecture and consists of multiple layers od self- attention mechanisms to give better performance.
- It improves upon BERT by optimizing training techniques and hyperparameters, resulting in enhanced performance.
- It was trained on a vast corpus of text from various sources, enabling it to capture a broader understanding of language.

Pre trained Model – BERT

- BERT (Bidirectional Encoder Representations from Transformers) is a Bidirectional Architecture that helps in processing text to the left and right of the context word.
- It is based on Transformer architecture which utilizes self-attention mechanisms for efficient text processing
- The model we used is vast text corpora through unsupervised learning which is generally used to predict missing words
- We have fine tuned this model to give better performance specific tasks by training on task-specific datasets



Demo

Comparative Analysis

Accuracy	GPT 2	Roberta	BERT
Mental Health Dataset	40.56	57.68	38.46
Customer Support Dataset	41.67	60	63.64

Challenges

- Fine tuning large models such as BERT would require a much higher performance hardware like GPUs and TPUs.
- more extensive experimentation is required to select the optimal hyperparameters required to give the best results, which was difficult with limited resources and time.

Thank you