

AI-HACKTHON

TEAM MEMBERS

- 1. Devanshu Gupta**
- 2. Rutuja Sanap**
- 3. Yash Sharma**
- 4. Ritika Sharma**
- 5. Anshuman Das**

HACKATHON

GPT-4 POWERED APP CREATION AND EVALS HACKATHON

Overview

Create application from scratch and improve the output with TruLens Evals! Build with the OpenAI models like GPT-4(Vision), GPT-4 Turbo, Assistants API, Text to Speech, Image Generation DALL·E API and more!

Problem STATEMENT



Text to speech

This AI project focuses on creating a Text-to-Speech (TTS) system that transforms written text into natural-sounding speech. Using advanced algorithms and neural networks, the TTS system enhances accessibility, improves interactive applications, and facilitates communication in diverse domains. Its goal is to provide a seamless and human-like audio experience across various platforms and languages.

INITIAL APPROACH


Data Collection and Preprocessing :

Gather a diverse dataset of text samples covering various languages, accents, and speech styles. Preprocess the text data to handle punctuation, formatting, and linguistic nuances effectively.

Feature Extraction: Extract relevant linguistic features from the text data, such as phoneme sequences, stress patterns, and prosodic cues. These features provide essential guidance for synthesizing natural-sounding speech.

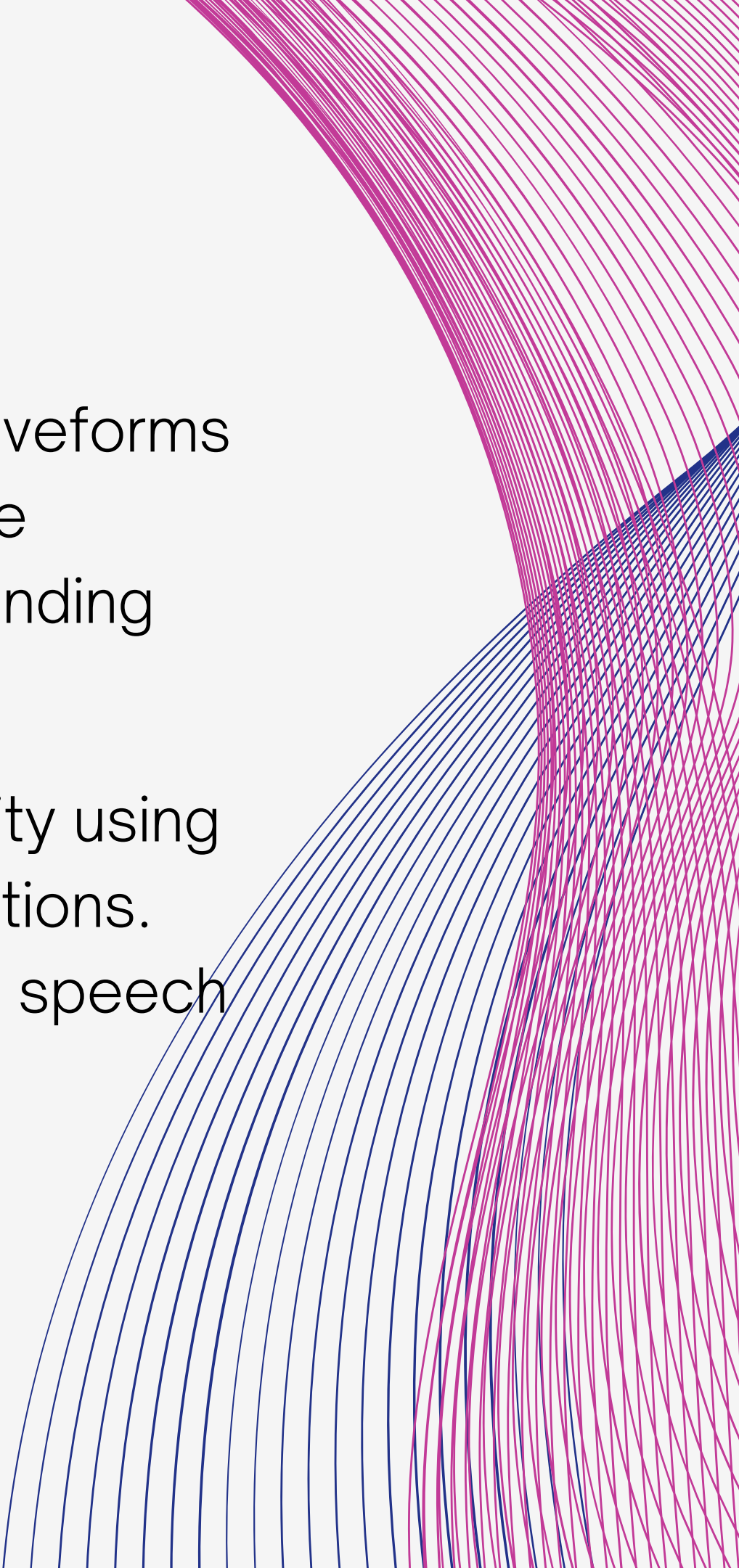
Training the TTS Model: Train the selected model using the preprocessed text data. Optimize the model parameters to minimize the discrepancy between predicted and actual speech outputs. Training should focus on capturing the nuances of intonation, rhythm, and pronunciation.

INITIAL APPROACH



Speech Synthesis: Create algorithms to generate speech waveforms from text predictions using waveform synthesis, concatenative synthesis, or parametric synthesis for high-quality, natural-sounding speech.

Evaluation: Assess synthesized speech quality and intelligibility using metrics like Mean Opinion Score and subjective human evaluations. Refine models and algorithms based on feedback to enhance speech synthesis performance.

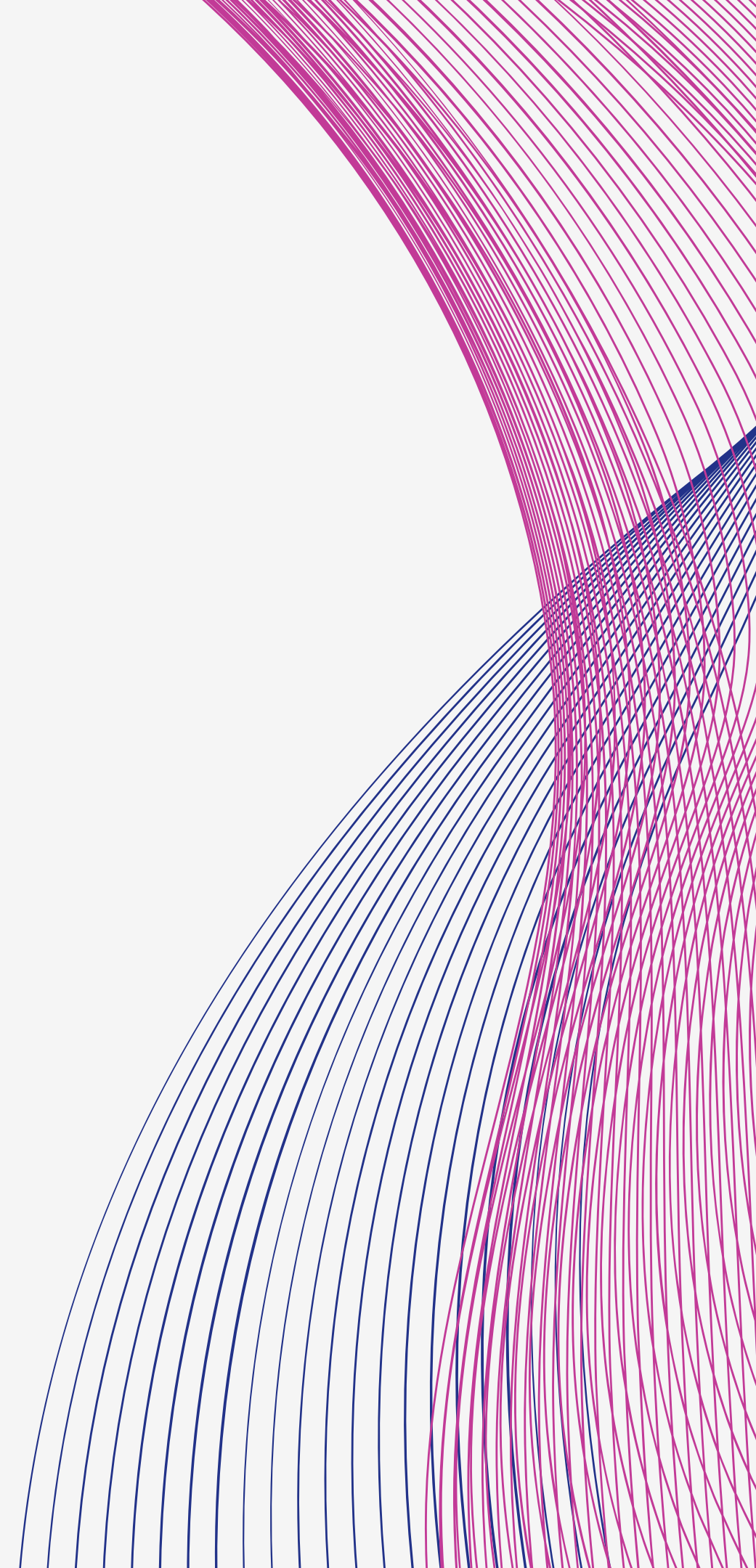


Application

- Narrate a written blog post
- Produce spoken audio in multiple languages
- Give real time audio output using streaming



Thank you!





Any Questions?

1180393933