Approach:

Concept:

The Emotion-Enhanced AI Voiceover Model will leverage advanced natural language processing (NLP) and speech synthesis techniques to generate voiceovers that reflect the intended emotions of the input text. The model will analyze the sentiment of the text and modulate the generated audio accordingly.

Sentiment Analysis:

Utilize a pre-trained NLP model RoBERTa to classify the emotion of the input text into predefined categories such as happiness, exclamation, anger, sadness/distress, and politeness.

Emotion-Driven Speech Synthesis:

Integrating an AI-based text-to-speech (TTS) engine capable of modulating pitch, tone, and pace to express different emotions.

Developing a mapping mechanism to translate the classified emotion into specific modulation parameters for the TTS engine.

Interactive User Interface:

Creating a web-based interface where users can input text and receive the corresponding emotionally expressive audio output in real-time.

Ensuring the interface is intuitive and supports quick input and output processing.

Technical Stack:

NLP Model:

TTS Engine:

Web Interface: