**Project Name:** Bharatlaw Text-to-Speech Task

**Problem Statement:**

Convert a given text summary into an audio summary using python.

**Objective:**

1. Research different TTS (Text-to-Speech) technologies.
2. Implement all researched technologies for given summary.

**Approach:**

* At first, I read and understood the objective and task of the given project. Then I searched about available text-to-speech technologies using Bard/Gemini, ChatGPT, google search and listed down them for implementation.
* Then I researched about each model and learned how to implement them.
* After that, I implemented the models individually and tested their result.
* Lastly, I choose the best models based on their results and performances.

**Technology used:**

Google Collab, Kaggle Notebook

**Result:**

* **Models I Found out:**

1. Suno/bark
2. Facebook/mms-tts-eng
3. Microsoft/speecht5\_tts

* **Bark performance**:
* It gave very good and pretty accurate result.
* It has many preset voices available and that too for different gender and languages.
* The input text can be improvised too add more human like behaviours like laugh, cry etc.
* Sample audio of the given summary:

<https://drive.google.com/file/d/1kYlI5b926Ja7NnSvO_uVP-rA_WkrFo4E/view?usp=drive_link>

<https://drive.google.com/file/d/1s9hcz7EMl6T1pFg4AJlPdNxmsfHKl6Zi/view?usp=drive_link>

* Bark will take time to generate audio.
* **Facebook mms\_tts performance**:
* It also gave very good and pretty decent result.
* It is easy to use for long texts and gives result pretty fast.
* It supports over 1000 languages.
* Sample audio of given summary:

<https://drive.google.com/file/d/1w27depj7c-J9Sst9boYWvV03MCKaPIIT/view?usp=drive_link>

<https://drive.google.com/file/d/1r5Faqdf_w_2q4wCNGSaAfBmiBEAzuGjM/view?usp=drive_link>

* **Microsoft speecht5\_tts performance:**
* It also gives good results.
* Sample audio output for given summary:

<https://drive.google.com/file/d/1HAc-0GE1ACXR9h_Mzn3ShgN9B7h_4842/view?usp=drive_link>

<https://drive.google.com/file/d/1yQu8gVNdYCMGAwPfR-PvslGs_dgtU_bD/view?usp=drive_link>

**Future Scope:**

There are many other open-source TTS models available like :

* metavoiceio/metavoice-1B-v0.1
* coqui/XTTS-v2
* Amphion
* Tortoise tts etc.

The voices also can be cloned to user voice.

**References and resources used:**

I followed hugging face documentations, YouTube tutorials, official website documentations, GitHub repos as resource and references. Here are few: https://huggingface.co/tasks/text-to-speech, https://huggingface.co/facebook/mms-tts-eng, https://huggingface.co/microsoft/speecht5\_tts, https://github.com/open-mmlab/Amphion, https://github.com/neonbjb/tortoise-tts, https://github.com/coqui-ai/TTS, https://www.youtube.com/watch?v=lPitjhhodaw&t=21s etc.