

Desenvolvimento de uma I.A capaz de gerar voz humana a partir de textos.

Tema: Síntese de voz e Geração de Áudio.

Objetivo:

Desenvolver um modelo de inteligência artificial capaz de gerar áudio realista a partir de textos e outras entradas. O foco será na síntese da fala humana buscando o máximo possível de naturalidade e realismo tanto na entonação e ritmo da fala.

Exploração do Conjunto de Dados:

A fonte de dados escolhida primeiramente será o projeto Common Voice da Mozilla, que disponibiliza muitos gigas de áudios de fala humana, todos verificados por outros usuários, o tipo de dado será a fala humana e o formato de dado que provavelmente será usado é o WAV (o Common Voice fornece áudios em MP3 que serão posteriormente convertidos se necessário).

Próximos Passos:

Os próximos passos serão ainda sobre os dados, pois além de converter eles, é necessário realizar uma triagem para garantir que eles possuam uma mesma abordagem, pois não queremos dados que fujam muito do padrão desejado, e após isso começaremos o pré-processamento de dados, convertendo os arquivos .WAV para um espectrograma.