

Projeto de IA Geradora de Áudio

O objetivo deste projeto é desenvolver uma Inteligência Artificial (IA) capaz de gerar áudios de voz humana de alta qualidade com base em textos ou comandos específicos (como outro áudio). Esta IA será projetada para síntese da fala. O projeto explorará técnicas avançadas de processamento de linguagem natural (NLP) e aprendizado profundo para criar um sistema robusto e eficiente.

Resultados

Foi desenvolvido um método para automatizar a transcrição de áudios utilizando uma IA como ferramenta, apesar de apresentar algumas falhas, o método é mais eficaz que a transcrição por nós, o que agiliza o processo.

Além disso, o treinamento de um novo modelo de IA teve resultados positivos, mesmo que sendo apenas um teste simples, sendo capaz de gerar um áudio de voz humana apenas com uma entrada textual.

Dificuldades

A maior dificuldade até então tem sido o tempo, tanto o treinamento do modelo, quanto a preparação dos dados é um processo muito lento e pode apresentar falhas.

A transcrição dos áudios tem apresentado falhas, então ela está tendo que ser supervisionado e corrigida, o que vem a ser um problema ao se tratar de um grande volume de dados, então estamos procurando formas de lidar com isso.

Além da transcrição o treinamento em si, do modelo, é um processo lento. O último feito, com apenas 20 áudios curtos, levou cerca de duas horas.

Próximos passos

Agora vamos procurar por meios de automatizar tanto a manipulação de dados quanto o treinamento da IA, agilizando o treinamento do modelo.