



**Data Science
Academy**

www.datascienceacademy.com.br

Processamento de Linguagem Natural

Mini-Projeto 3

Reconhecimento da Fala
Detectando Emoções em Arquivos de Áudio
com Inteligência Artificial

Parte 1



Mini-Projeto 3

Reconhecimento da Fala - Detectando Emoções em Arquivos de Áudio com Inteligência Artificial

Parte 1 - Preparação dos Dados, Treinamento e Avaliação do Detector de Emoções com Machine Learning

Este é um Mini-Projeto especial.

Vamos trabalhar com uma das tarefas mais complexas em Inteligência Artificial: extrair emoções a partir da voz em arquivos de áudio.

Detectar emoções é uma das estratégias de Marketing mais importantes no mundo de hoje. Você pode personalizar aplicações para fornecer tratamentos diferentes para um indivíduo de acordo com a emoção detectada na voz de uma pessoa. Esse tipo de aplicação é um dos pilares para uma solução completa de IA.

Alguns exemplos de aplicações desse tipo de solução, incluem:

- Uma central de atendimento que toca músicas diferentes de acordo com a emoção detectada na voz do cliente.
- Um carro autônomo que desacelera quando alguém está com raiva ou com medo.
- Assistente pessoal que reage de acordo com a emoção detectada na voz.
- Aplicações de Marketing que oferecem diferentes produtos ou opções de acordo com a emoção do cliente.
- Assistente Virtual, que pode ser uma Enfermeira Virtual ou mesmo um Professor de Inglês, e que reage de acordo com a voz do interlocutor.

Entre outros exemplos.

Usaremos a biblioteca Librosa em Python para processar e extrair recursos dos arquivos de áudio. Librosa é um pacote Python para análise de música e áudio. Ele fornece os componentes necessários para criar sistemas de recuperação de informações musicais.

Usando a biblioteca librosa, conseguimos extrair recursos através do MFCC (Mel Frequency Cepstral Coefficient). Os MFCCs são coeficientes amplamente usado no reconhecimento automático de fala.



Também separamos a voz de mulheres e homens usando os identificadores fornecidos no dataset, como forma de deixar o modelo de reconhecimento de voz ainda mais preciso e personalizado.

Cada arquivo de áudio fornece muitos recursos, que são basicamente uma matriz de muitos valores. A esses recursos, iremos atribuir rótulos especificando o gênero da voz no áudio e a emoção detectada.

Como este projeto é um grande esforço de trabalho, ele foi dividido em 3 partes:

- Parte 1 - Preparação dos Dados, Treinamento e Avaliação do Detector de Emoções com Machine Learning
- Parte 2 - Preparação dos Dados, Treinamento e Avaliação do Detector de Emoções com Deep Learning
- Parte 3 - Detecção e Classificação de Emoções em Arquivos de Áudio

Este projeto pode ser facilmente adaptado aos seus próprios projetos. Tudo que você precisa é providenciar arquivos de áudio gravados com pessoas com diferentes emoções. Usaremos um dataset público para nosso trabalho.

Esta é a Parte 1. Aprenda e divirta-se.

As Partes 2 e 3 serão disponibilizadas no Capítulo 12 sobre API de voz.