



Proyecto reconocimiento comandos de voz

Presentado por Info Próceres



Aplicaciones Comunes

Asistentes de Voz: Siri, Alexa, Google Assistant.

Dispositivos inteligentes:

Integración en sistemas domésticos, teléfonos, entre otros.

Herramientas Útiles

- **Librosa** para el procesamiento de audio.
- Frameworks de machine learning como **TensorFlow** y **PyTorch**.

permiten la creación de modelos personalizados para tareas específicas, como la clasificación de comandos de voz. Algunos enfoques recientes se centran en la eficiencia y la adaptabilidad de estos modelos, haciéndolos capaces de funcionar en dispositivos con recursos limitados.

Estado del arte



Avances Clave

Los avances en redes neuronales profundas y técnicas de aprendizaje automático como:

- **Red Neuronal Recurrente**
- **Red Neuronal Convolucional**
- **Transformers**

han mejorado considerablemente la precisión y la eficiencia de estos sistemas.

Objetivo

El objetivo de este proyecto es entrenar un modelo de machine learning capaz de clasificar comandos de voz simples (como, por ejemplo, "arriba", "abajo", "izquierda", "derecha").

01.

Interacción con videojuegos y dispositivos sin contacto.

02.

Permitir a las personas con discapacidades físicas o visuales interactuar con dispositivos electrónicos utilizando comandos de voz.

03.

Controlar luces, termostatos, electrodomésticos y otros dispositivos domésticos mediante comandos de voz.

04.

Permitir a los profesionales de la salud interactuar con sistemas de registro y recuperación de información utilizando comandos de voz.

05.

Permitir a los estudiantes interactuar con sistemas de aprendizaje en línea utilizando comandos de voz para realizar búsquedas, hacer preguntas y recibir respuestas.

Resultado esperado

Preciso

El sistema debe Funcionar bien con diversas voces, acentos y ruido de fondo.

Rápido

Con capacidad de reconocer comandos en tiempo real.

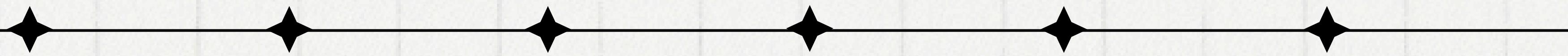
Eficiente

Capaz de funcionar en dispositivos con recursos limitados, como teléfonos móviles.

Escalable

Capaz de manejar un gran volumen de comandos de voz y adaptarse a nuevas palabras o frases.

Métodología

- 
- 01**
Recolección de datos: Se utilizará un conjunto de datos como el Google Speech Commands Dataset para entrenar y evaluar el modelo.
 - 02**
Preprocesamiento: Usando librosa, se realizará el preprocesamiento de las señales de audio para extraer características relevantes.
 - 03**
Entrenamiento del modelo: Se probarán diferentes arquitecturas de redes neuronales y se utilizará frameworks de machine learning como TensorFlow o PyTorch.
 - 04**
Evaluación: Se evaluará el rendimiento del modelo utilizando métricas como precisión, recall y matriz de confusión.
 - 05**
Optimización: Se realizará la optimización del modelo ajustando los hiperparámetros. Esto se hará mediante distintas técnicas con el objetivo de mejorar el rendimiento del modelo en términos de precisión y eficiencia.
 - 06**
Validación y despliegue: Una vez que el modelo ha sido entrenado y optimizado, se realizará una validación final utilizando datos de prueba independientes.

**¡Muchas
gracias!**

Info Próceres