[[1]](#footnote-1)

*Speech diarization* sobre entrevistas con pacientes de Parkinson y posibilidad de predicción de algunos diagnósticos.

Edwin Andrés León Castro

edaleon.1988@gmail.com

**Resumen - El Parkinson es una enfermedad neurodegenerativa progresiva caracterizada por la pérdida de células dopaminérgicas en la sustancia negra del cerebro. Esto conduce a síntomas motores, como temblores, rigidez y bradicinesia, así como a alteraciones no motoras, como problemas cognitivos y del habla. Aproximadamente el 90 % de los pacientes con Parkinson desarrollan alteraciones en el habla conocidas como disartria hipocinética, que se manifiesta en tono monótono, baja intensidad y dificultad para articular palabras.**

**En el presente documento se plantea una metodología para abordar un problema de ciencia de datos aplicado al análisis de señales de audio de pacientes con Enfermedad de Parkinson, EP. Inicialmente, se presenta un análisis del problema, el objetivo general y un árbol de problemas. Posteriormente, se exploran trabajos relacionados con la extracción de características y diagnóstico de la EP mediante señales de audio.**

**Finalmente, se propone una metodología para procesar un conjunto de datos que incluye entrevistas de audio de 40 pacientes, con el fin de extraer información que facilite un diagnóstico.**

**Índice de Términos - Parkinson's disease, Machine learning, Voice recognition, Feature extraction.**

1. introducción

La Enfermedad de Parkinson EP es un trastorno neurodegenerativo progresivo que afecta aproximadamente al 1% de la población mundial, pero se estima que el 20% de los pacientes con Parkinson nunca son diagnosticados [1]. El proceso típico de diagnóstico de la EP lleva más de dos años.

Como resultado, se necesita un método que ayude con el proceso de diagnóstico de la EP en función de los síntomas del paciente. Aproximadamente el 89% de las personas que padecen EP presentan alteraciones del habla. Estas alteraciones del habla se manifiestan principalmente en fluctuación de frecuencia fundamental, aperiodicidad en la vibración de las cuerdas vocales, reducción de la intensidad del habla, ronquera y articulación irregular. Por lo tanto, el habla es una excelente opción para ofrecer pistas diagnósticas para la detección automática de la EP [2].

Por lo tanto, se hace necesario identificar y analizar las interacciones verbales entre personal médico y pacientes con Parkinson para extraer indicadores clínicos relevantes, identificando patrones del habla específicos de la enfermedad.

1. OBJETIVO GENERAL

Apoyar el manejo clínico de la Enfermedad de Parkinson (EP) mediante el uso de inteligencia artificial para analizar entrevistas habladas entre pacientes y personal médico, con el fin de extraer indicadores clínicos relevantes que puedan apoyar la toma de decisiones médicas.

1. OBJETIVOS ESPECÍFICOS

* Revisar las soluciones actuales y los modelos disponibles de diarización de hablantes que permitan separar voces en archivos de audio con alto nivel de ruido o condiciones acústicas no controladas.
* Implementar un modelo de diarización de hablantes que segmente entrevistas, diferenciando las intervenciones del paciente y del entrevistador.
* Revisar las soluciones actuales y los modelos disponibles para la detección de la Enfermedad de Parkinson a partir de archivos o señales de audio.
* Implementar un modelo que permita clasificar el grado de Parkinson a partir de las características del audio.

1. MATERIALES Y MÉTODOS

El conjunto de datos disponible es una serie de entrevistas de audio realizadas a pacientes con Parkinson, en total hay registro de 40 pacientes, a los cuales se les pedía pronunciar palabras con la letra A, Letra F, Letra P, Letra M, Animales y Frutas, los archivos quedaron nombrados de la siguiente manera:

|  |  |
| --- | --- |
| Ejercicio | Nombre del archivo |
| Palabras que empiezan con la letra A | [id\_paciente]\_Prueba\_Fluidez\_Verbal\_A\_Letter.wav |
| Palabras que empiezan con la letra F | [id\_paciente]\_Prueba\_Fluidez\_Verbal\_F\_Letter.wav |
| Palabras que empiezan con la letra M | [id\_paciente]\_Prueba\_Fluidez\_Verbal\_M\_Letter.wav |
| Palabras de la categoría Frutas | [id\_paciente]\_Prueba\_Fluidez\_Verbal\_Frutas\_Letter.wav |
| Palabras de la categoría Animales | [id\_paciente]\_Prueba\_Fluidez\_Verbal\_Animales\_Letter.wav |

Al revisar el dataset se encontró que 7 pacientes no cuentan con las 5 pruebas, sin embargo; al menos cuentan con 2 archivos.

Gráfico, Histograma

El contenido generado por IA puede ser incorrecto.

Al revisar la calidad de los archivos se encuentra que estos fueron grabados en ambientes no controlados, con alto nivel de ruido, sonidos de fondo, momentos en que habla paciente y entrevistador al tiempo y audios donde no habla el paciente o errores en la prueba:

Imagen de la pantalla de un computador

El contenido generado por IA puede ser incorrecto.

REFERENCIAS

[1] A. M. Anter, A. W. Mohamed, M. Zhang, and Z. Zhang, “A robust intelligence regression model for monitoring Parkinson’s disease based on speech signals.,” *Future Generation Computer Systems*, vol. 147, pp. 316-316–327, 2023, doi: 10.1016/j.future.2023.05.012.

[2] P. Warule, S. P. Mishra, and S. Deb, “Time-frequency analysis of speech signal using Chirplet transform for automatic diagnosis of Parkinson’s disease,” *Biomed Eng Lett*, vol. 13, no. 4, pp. 613-613–623, 2023, doi: 10.1007/s13534-023-00283-x.

1. Documento recibido el xx de xx de xx.

   . [↑](#footnote-ref-1)