

# PDDetection

## Aplicación de técnicas de minería de datos para la detección del Parkinson

**Autor:**

Adrián Arnaiz Rodríguez

**Tutores:**

Dr. Jose Francisco Díez Pastor

Dr. César Ignacio García Osorio

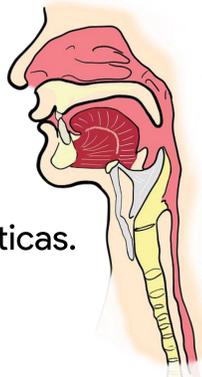
Julio 2019



Universidad de Burgos

## INTRODUCCIÓN

- La enfermedad del Parkinson produce alteraciones en el habla → DISARTRÍA.
- Investigación: predecir y monitorizar el Parkinson a través de la voz.
- En anteriores estudios, se extraen muchos tipos de diferentes de características.
- Reducir tiempos de consulta y fallos en el diagnóstico.



## OBJETIVOS

- Investigar y condensar el estado del arte: **BASE SÓLIDA** para la investigación.
- Recopilación y creación de una **BASE DE DATOS** con la que realizar el proyecto de investigación.
- Experimentación con distintos tipos de características y clasificadores → crear un **CLASIFICADOR** con el mejor rendimiento posible.
- **APLICACIÓN** que asista al facultativo médico ayudándole en la predicción del Parkinson.

## DATASET

AUDIOS EQUILIBRADOS: 1 frase, 5 vocales y 5 palabras



Características extraídas de los audios



- **DISVOICE** → **FÍSICAS**: articulación, fonación y prosodia
- + edad y sexo
- Separando por sexo

- **VGGish** → **EMBEDDINGS** y espectros
- 108 Conjuntos totales de características
- 50-300 instancias por conjunto → pocas

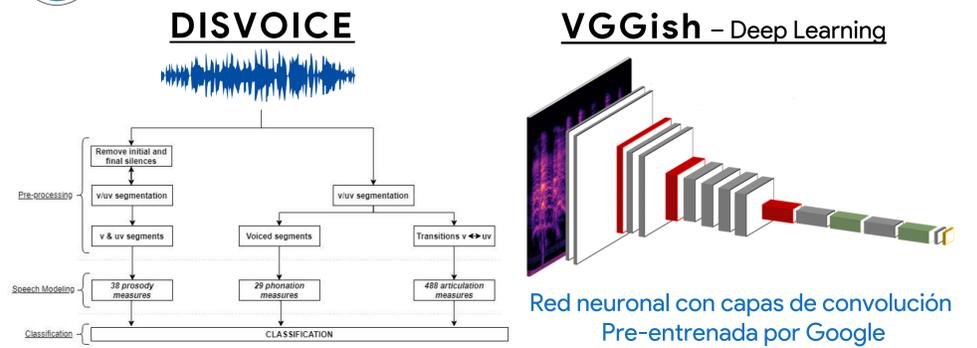
Se repite para cada audio de cada tipo

...	PD	Sano

Conjunto de datos

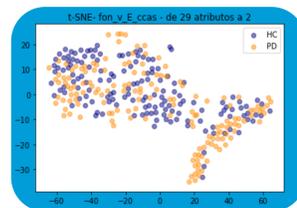
## DESARROLLO

### 1 Extracción de características

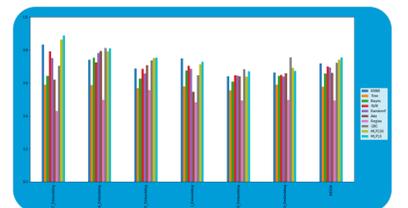


### 2 Experimentos con clasificadores

- SVM, Boosting, Bagging, MLP, Bayesianos, árboles
- Selectores características: *KBest*, Varianza límite
- Búsqueda de parámetros: Grid Search



Proyección de características



Visualizaciones de resultados

## RESULTADOS

Dividido por sexos

**0,908**  
AUC

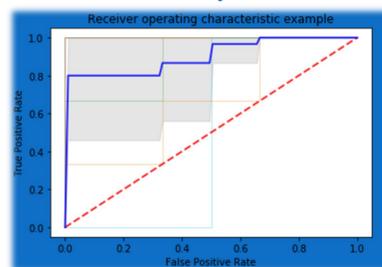
Select KBest + AdaBoost  
Dataset:  
(articulación, palabra campana, mujeres)

Sin división por sexo

**0,888**  
AUC

MLP (10 neuronas por capa)  
Dataset:  
(VGGish embeddings, frase)

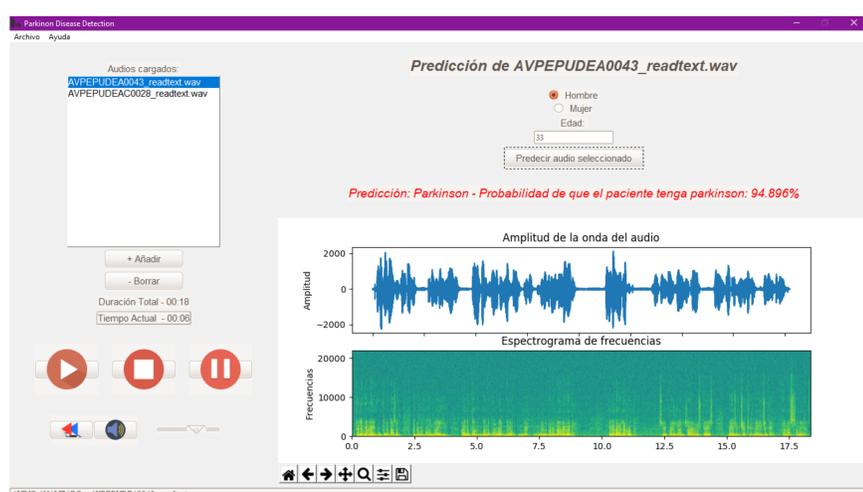
Curva ROC mejor resultado



¿Qué tipo de audio es el mejor?



## APLICACIÓN



- ✓ **PREDICCIÓN**: probabilidad de Parkinson
- ✓ Reproducción de audios
- ✓ Amplitud de onda del audio
- ✓ Espectrograma de frecuencias



aar0062@alu.ubu.es  
/AdrianArnaiz/TFG-Neurodegenerative-Disease-Detection

Grado en Ingeniería Informática  
Trabajo de fin de grado