

Predicción de la enfermedad de Parkinson a partir del análisis de lenguaje utilizando técnicas de aprendizaje autónomo supervisado

PROBLEMA

Muchos de los pacientes no conocen que tienen la enfermedad de Parkinson hasta que los síntomas más graves aparecen. En ocasiones deben esperar de meses o incluso años para obtener el diagnóstico correcto, debido a que muchos de los síntomas que padecen tienden a confundirse con otras enfermedades.

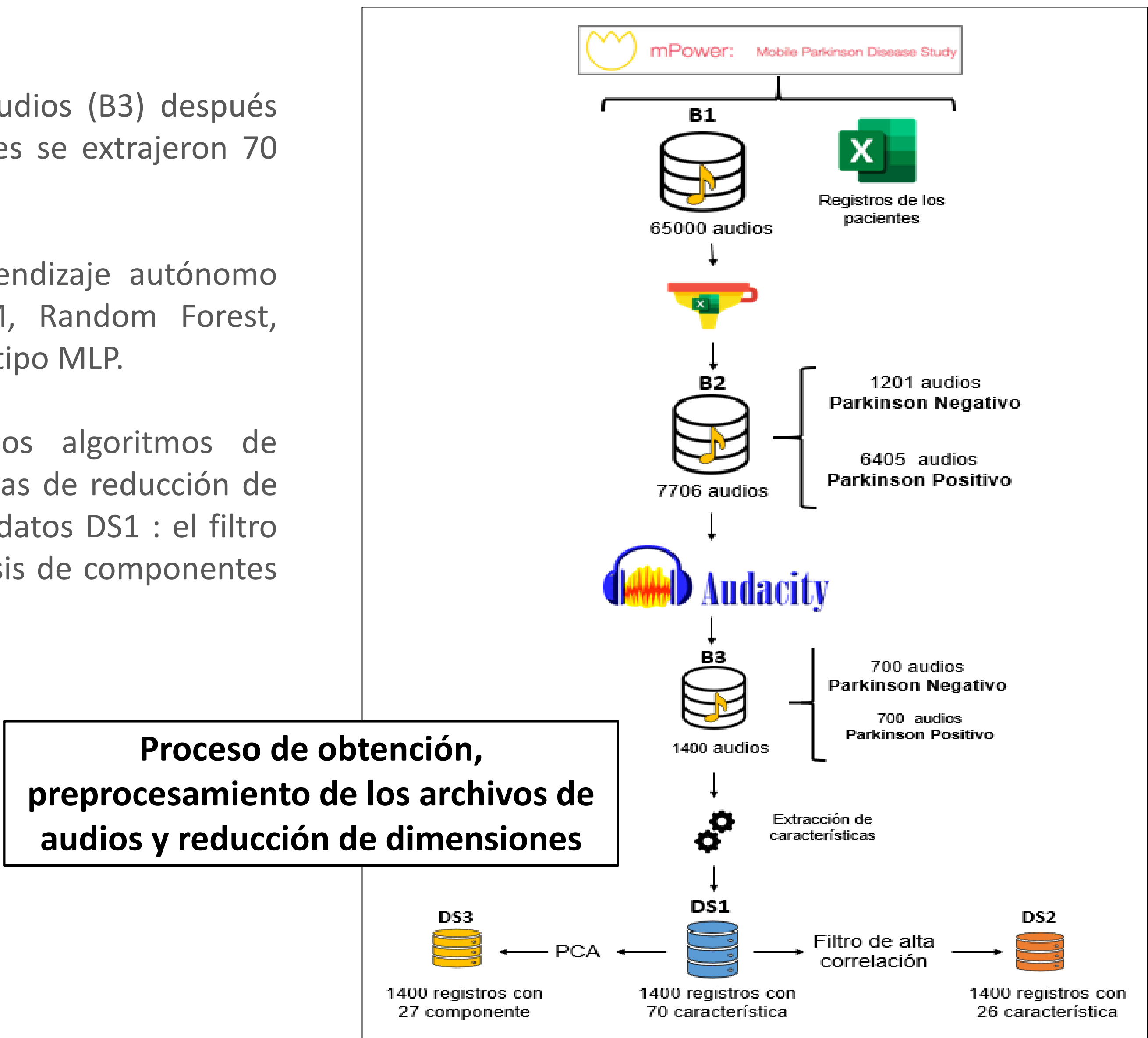
OBJETIVO GENERAL

Crear un aplicación web que permita predecir la presencia o no de los síntomas relacionados a la enfermedad de Parkinson, mediante el análisis de los patrones en la señal de voz de los pacientes.



PROPUESTA

- Se escogieron 1400 archivos de audios (B3) después del preprocesamiento, de las cuales se extrajeron 70 características.
- Se evaluaron 4 algoritmo de aprendizaje autónomo supervisado las cuales son: SVM, Random Forest, Logistic Regression y Red Neuronal tipo MLP.
- Durante el entrenamiento de los algoritmos de aprendizaje se aplicaron dos técnicas de reducción de dimensiones sobre el conjunto de datos DS1 : el filtro de alta correlación (DS2) y el análisis de componentes principales - PCA(DS3).

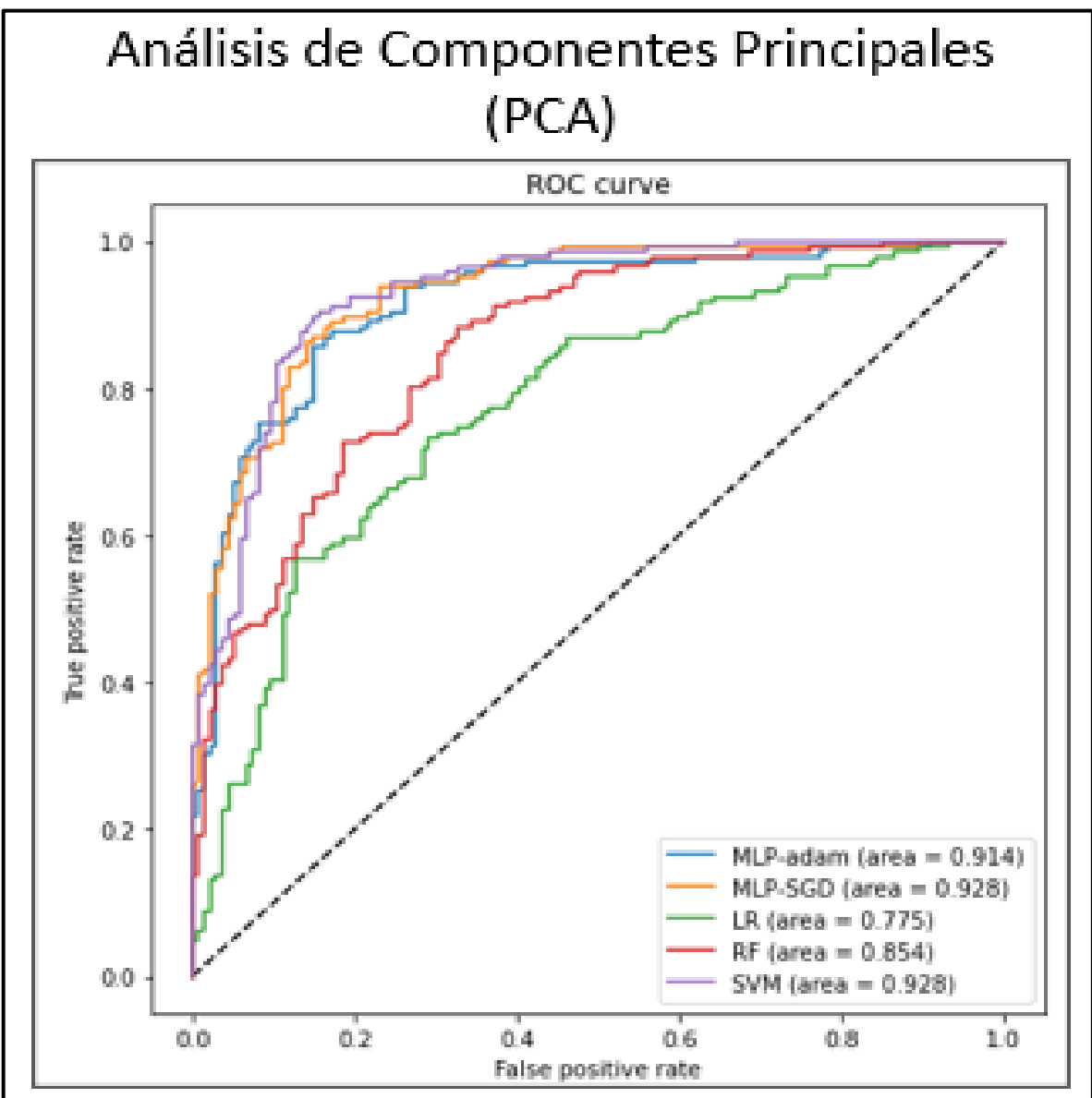
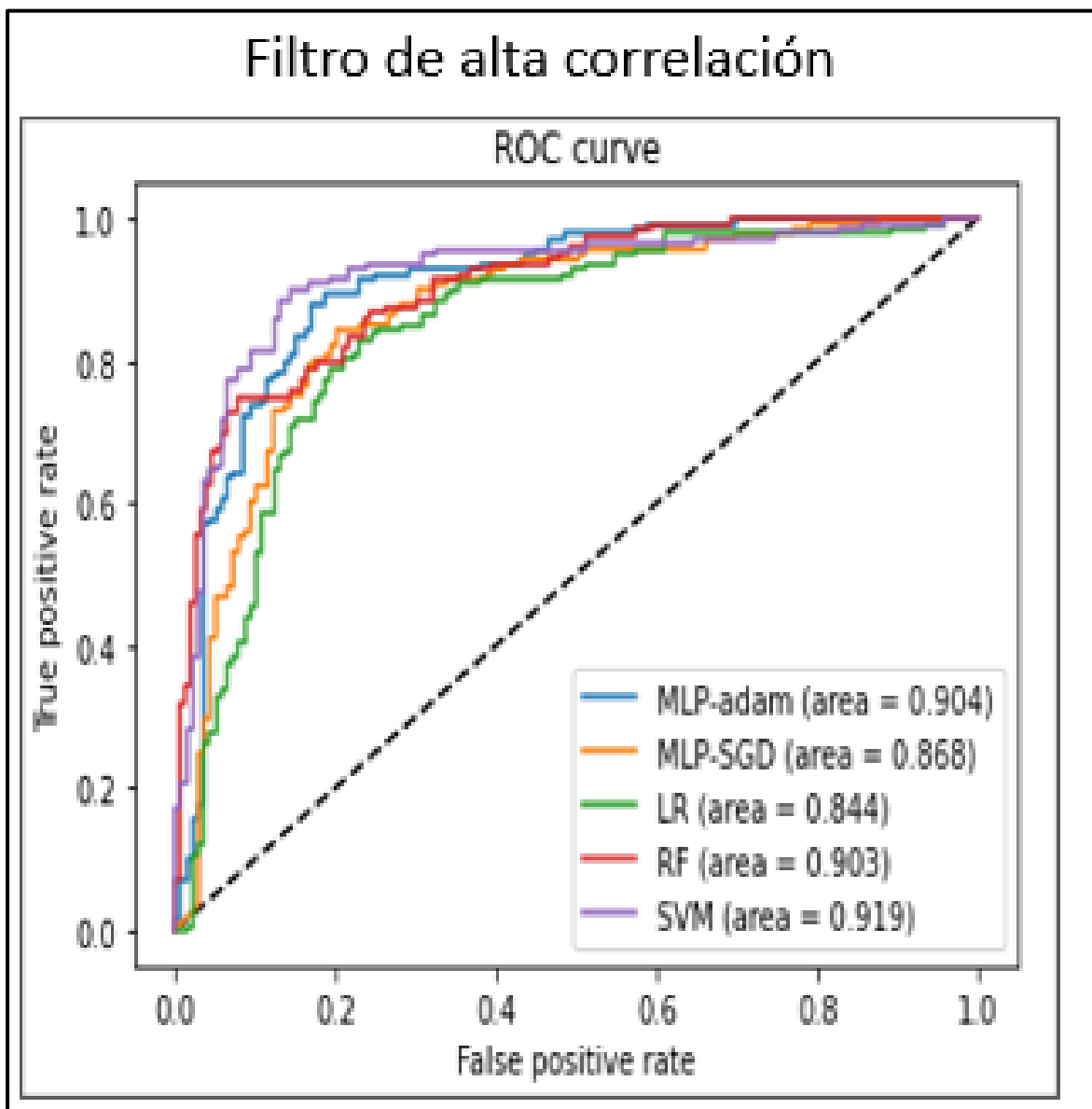


RESULTADOS

- De acuerdo a las 2 gráfica de la curva ROC, se demostró que tanto el MLP-SGD y SVM que utilizaron PCA tuvieron un área bajo la curva mayor que el resto de los algoritmos de aprendizaje supervisado.
- El SVM que usó los conjuntos de datos de filtro de alta correlación tuvo la mayor precisión entre todos los algoritmo de aprendizaje supervisado, con un 88%.

Tabla comparativa entre las precisiones obtenidas

	Filtro de alta correlación	PCA
MLP	83%	86%
RF	75%	82%
SVM	88%	87%
LR	80%	71%



CONCLUSIONES

- En la etapa del preprocesamiento de los audios, se pudo observar que los ruidos que contenían los archivos de audios presentan valores altos en la extracción de las características, en consecuencia, confunde los algoritmos de aprendizaje autónomo supervisado.
- El algoritmo Support Vector Machine (SVM), obtuvo la mejor precisión en cada una de las técnicas de reducción de dimensiones utilizada en el presente trabajo: con un 88% de precisión con la técnica del Filtro de Alta Correlación y un 87% de precisión con PCA.