Estimación de la Cantidad de Mentas en una Caja mediante Análisis de Audio y Aprendizaje Automático

Mateo Sanchez, Oscar Rodriguez, Maria Jose Sanchez {sebastian2221784, maria2221385, oscar2221532}@correo.uis.edu.co



Resumen

Se propone un sistema que estima la cantidad de mentas en una caja usando el sonido que se genera al agitarla. Se recolectaron 2436 grabaciones con distintos dispositivos y condiciones. Se extrajeron características como MFCCs, RMS, ZCR y Spectral Centroid. Se entrenó un modelo Random Forest que alcanzó una precisión final del 81%. El sistema predice con mayor certeza al promediar varias grabaciones.

Objetivos

General: Diseñar un sistema que estime la cantidad de mentas usando audio y aprendizaje automático.

Específicos:

- Recolectar audios para distintas cantidades.
- Extraer características relevantes.
- Entrenar y evaluar un modelo Random Forest.
- Implementar un sistema funcional en tiempo real.

Metodología

Datos: 2436 audios con agitación controlada y libre.

Características: MFCCs, RMS, ZCR, Spectral Centroid.

Modelo: Random Forest optimizado con Grid Search.

Evaluación: Precisión por clase, matriz de confusión, importancia de características.

Sistema en Tiempo Real

Sistema en Python usando sounddevice y librosa. Permite grabar 6 audios y estimar la cantidad de mentas con el modelo entrenado. La predicción final se obtiene promediando los resultados individuales.

Conclusiones

- El sistema logra estimaciones confiables cada 5 mentas.
- La forma de agitación influye en la precisión.
- Modelos más complejos no mejoraron significativamente.
- Machine learning es clave para resolver tareas de estimación no invasiva.

Repositorio y Referencias

Repositorio: https://github.com/ Matt22vL/RETOS-

Referencias clave:

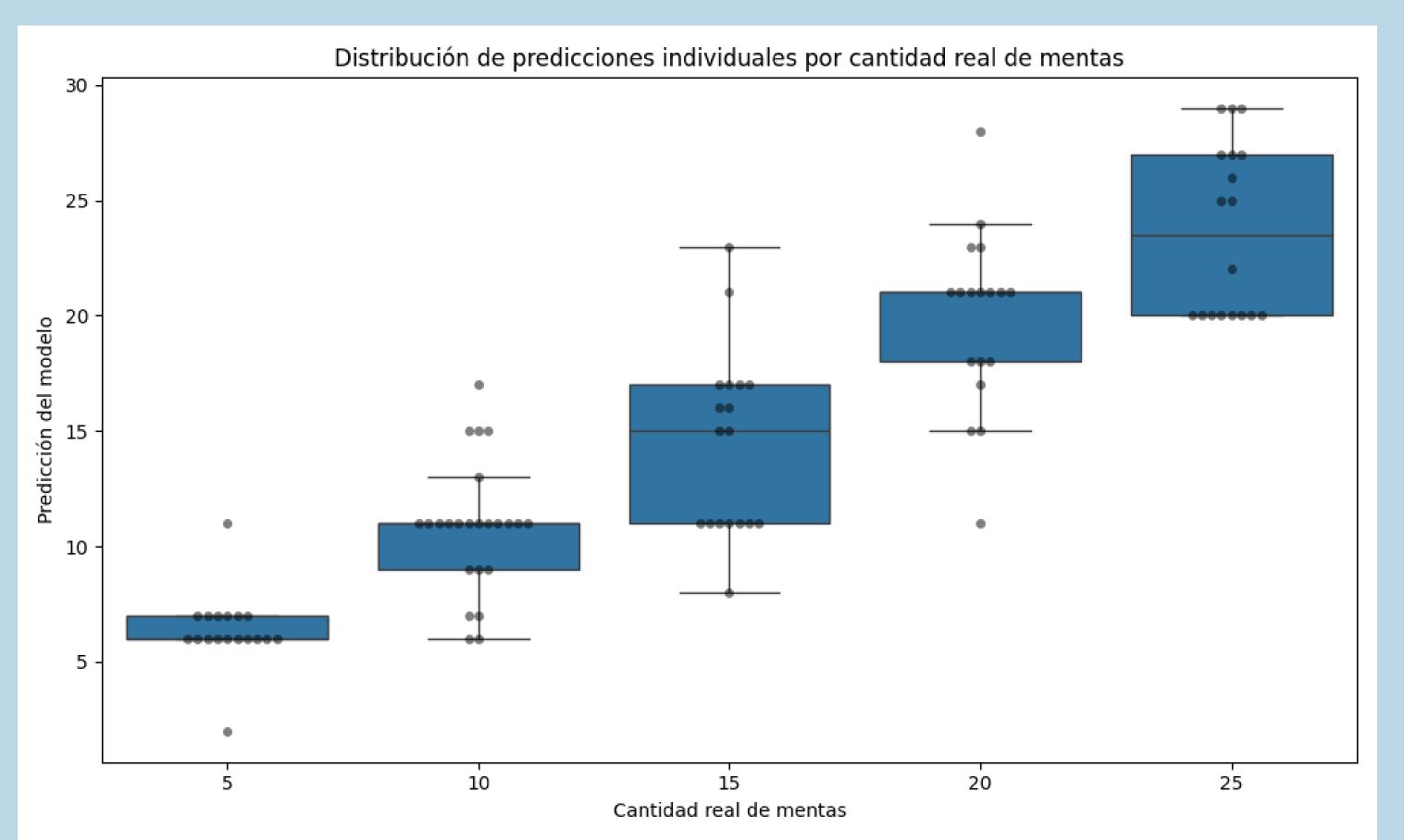
Logan, MFCCs for music modeling. Giannakopoulos, Audio Features.

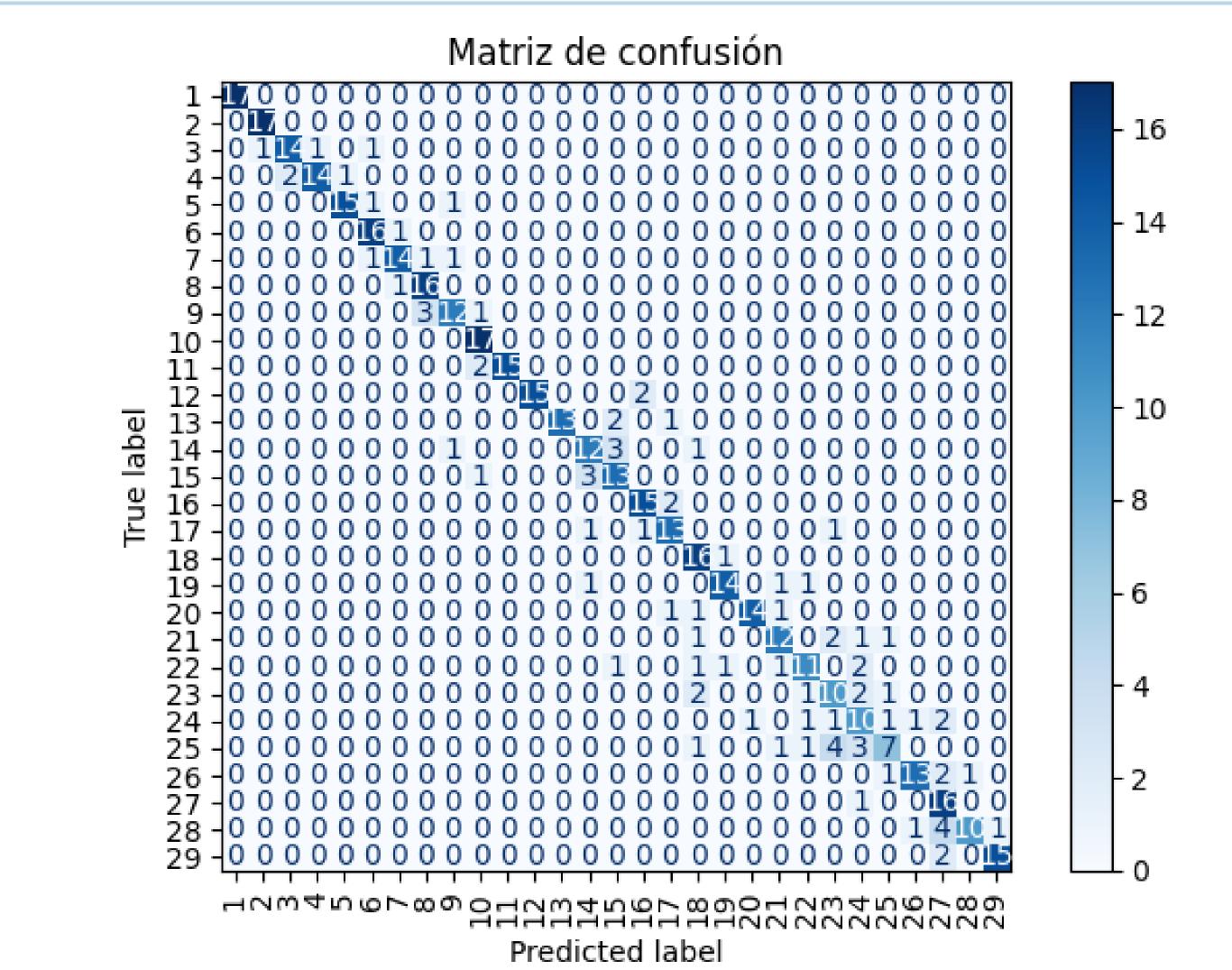
Breiman, Random Forests.

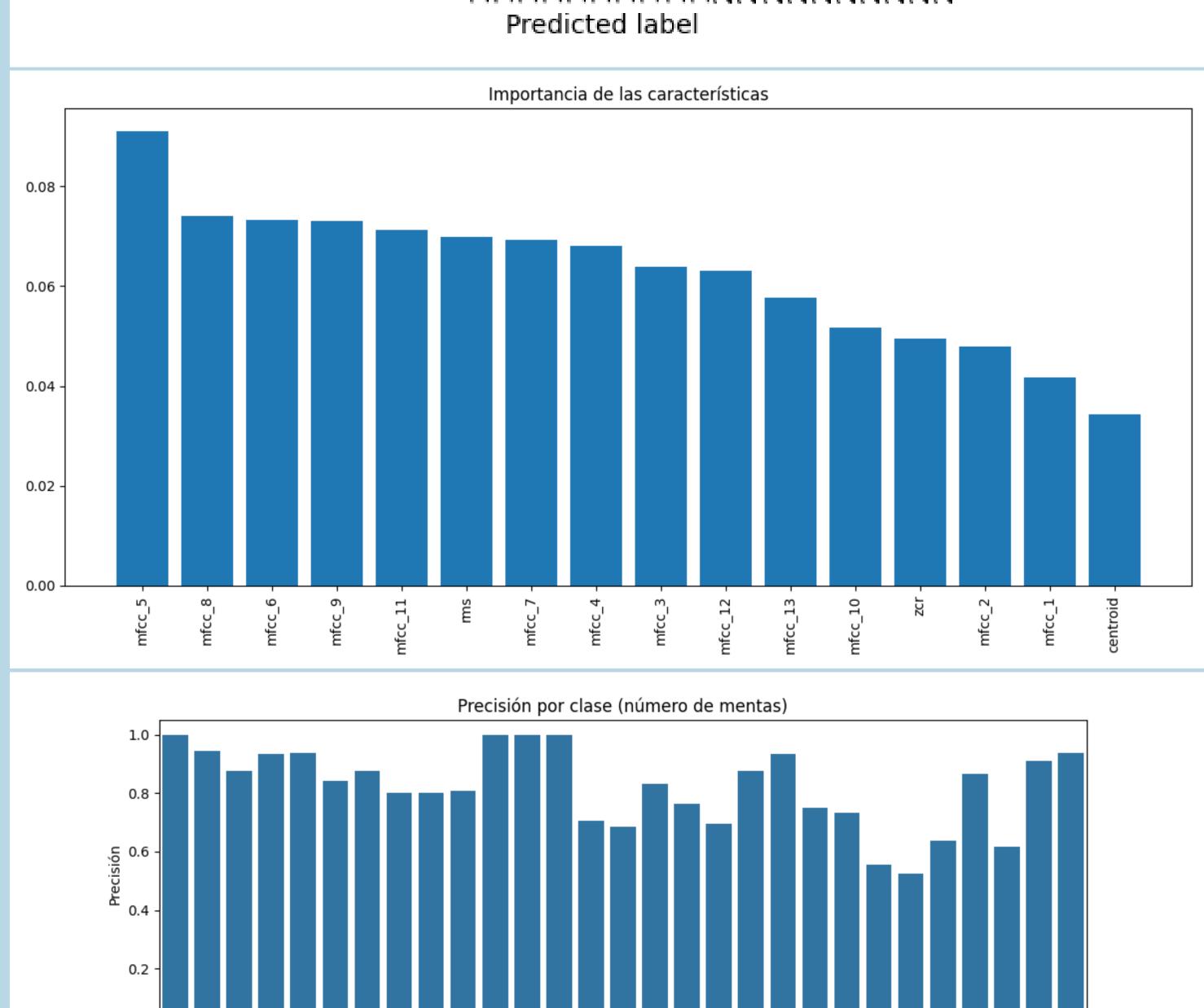
Librosa Documentation.

Resultados

- Precisión de validación cruzada: 76%.
- Precisión sobre el conjunto de prueba: 81%.
- Mejor desempeño cada 5 mentas.
- MFCCs fueron las características más informativas.
- El promedio de 6 grabaciones mejora la precisión.







1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29