

Sistema de Detección de Emociones en Audio (SER)

Asignatura: Procesamiento del habla,
visión e interacción multimodal

Daniel Marchena Jiménez, Javier Arias Fuentes,
Zohair Mouhim Gharafi.



Índice

1. Reto
2. De Audio a Datos
3. Metodología en Orange
4. Flujo de Orange
5. Comparativa
6. Conclusiones

El Reto

- **Objetivo:** Clasificar emociones (Positiva, Neutra, Negativa) usando solo la física del audio.
- **Enfoque SER:** Sin transcripción (ASR) ni procesamiento de texto (NLP).
- **Herramientas:** Extracción con Python (Librosa) y modelado en Orange Data Mining.

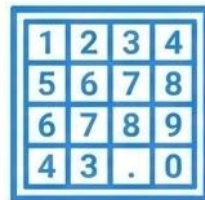
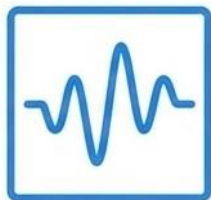


Audio



Texto (NLP)

De Audio a Datos (Librosa)



Key Feature:

MFCCs (40 coeficientes).
Replican la audición humana y capturan el timbre.

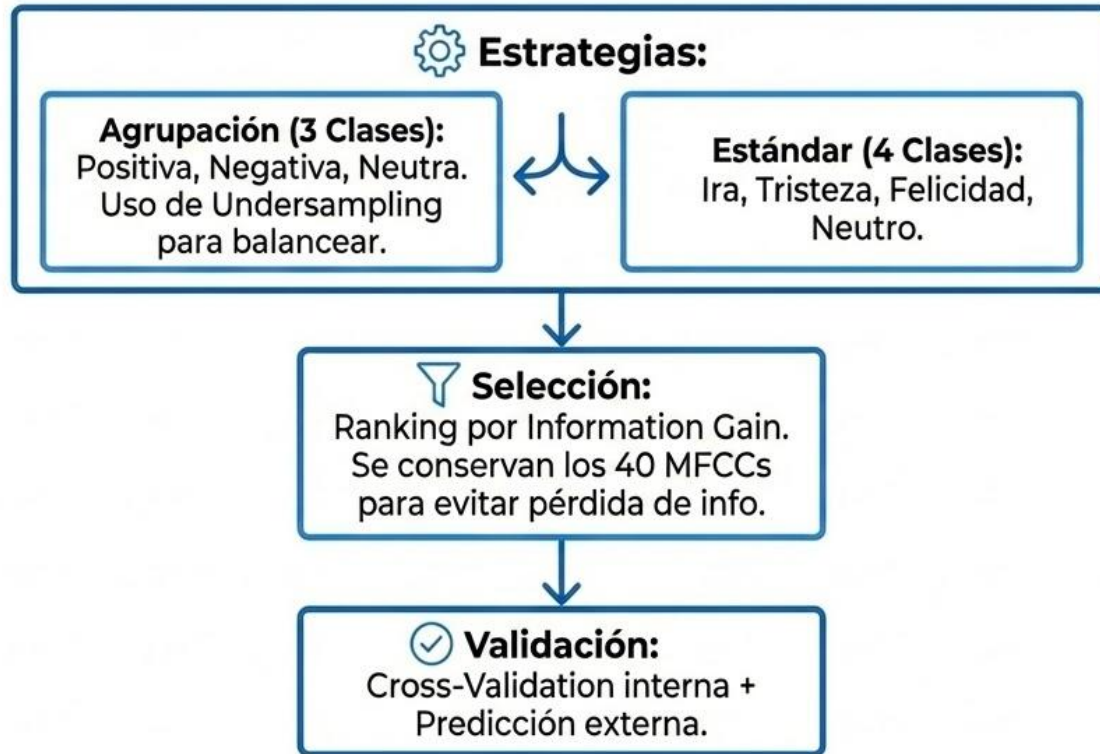
Secondary Features:

Chroma (Tono) y Mel Spectrogram (Energía).

Proceso:

Conversión de señal de audio no estructurada a vectores numéricos (medidas estadísticas).

Metodología en Orange



Flujo de Orange

Filter

Pythagorean Forest CN2 Rule Viewer Nomogram Scoring Sheet Viewer

Model

Constant CN2 Rule Induction Calibrated Learner kNN

Tree Random Forest Gradient Boosting SVM

Linear Regression Logistic Regression Naive Bayes Scoring Sheet

AdaBoost PLS Curve Fit Neural Network

Stochastic Gradient Descent Stacking Save Model Load Model

Evaluate

Test and Score Predictions Feature as Predictor Confusion Matrix

ROC Analysis Performance Curve Calibration Plot Permutation Plot

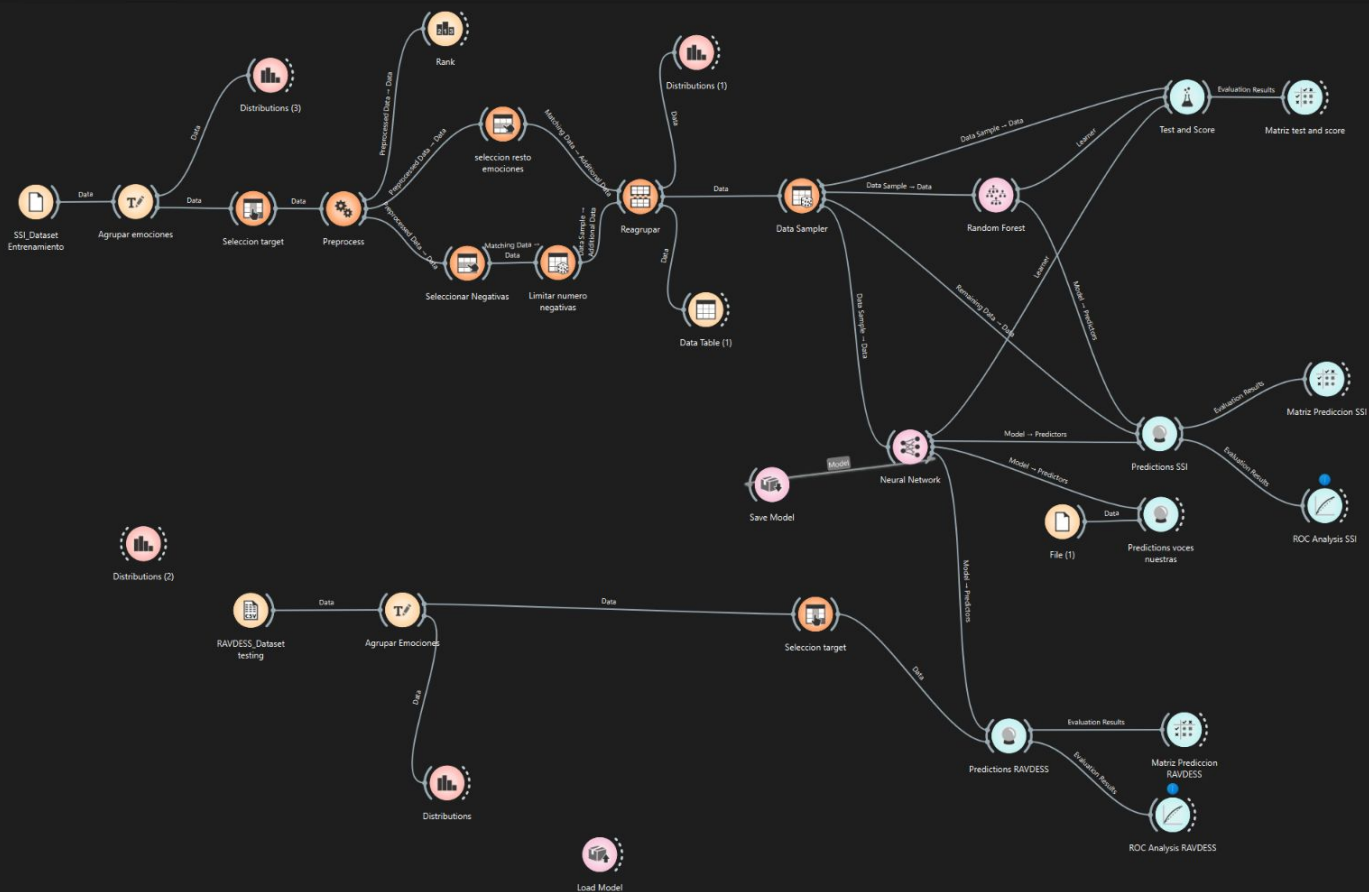
Parameter Fitter

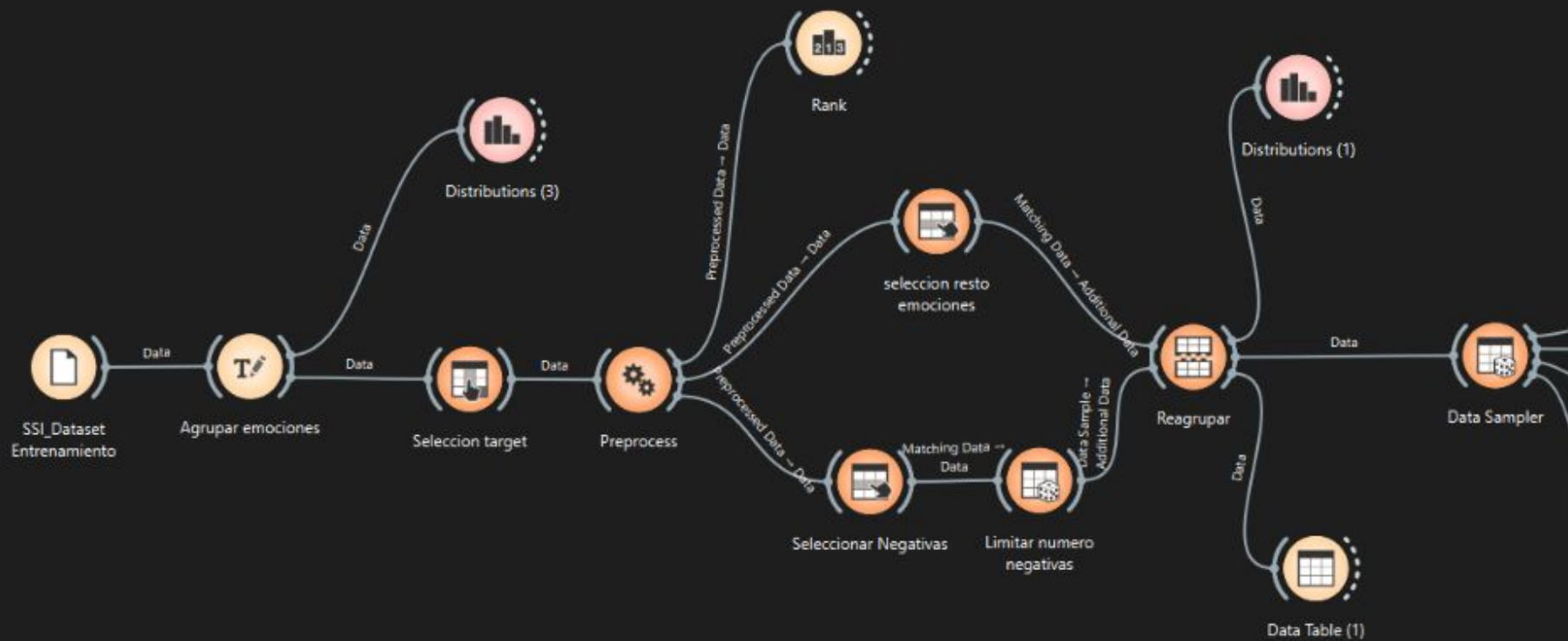
Data Sampler

Randomly draw a subset of data points from the input dataset.

[More...](#)

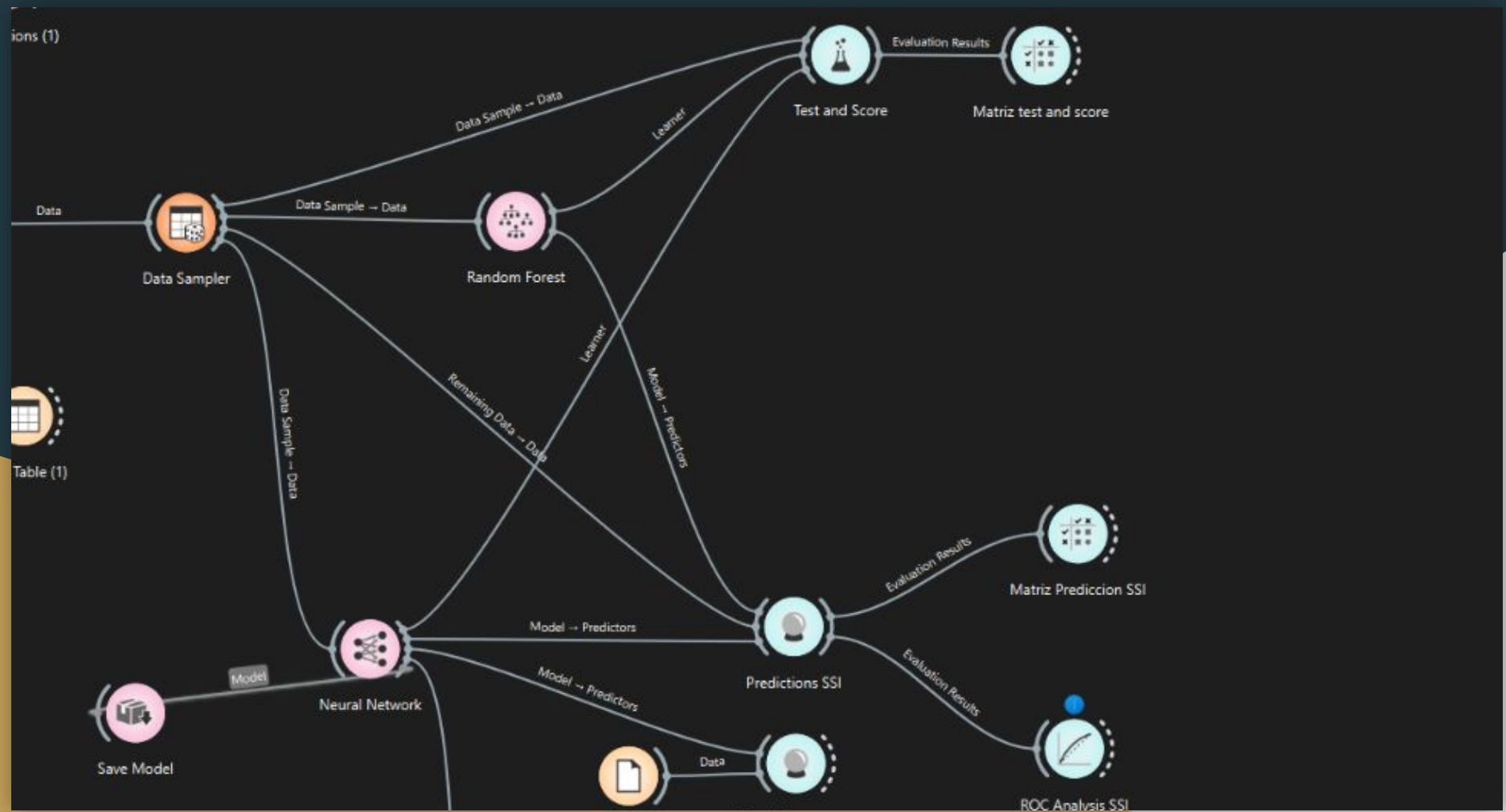
T . / ||

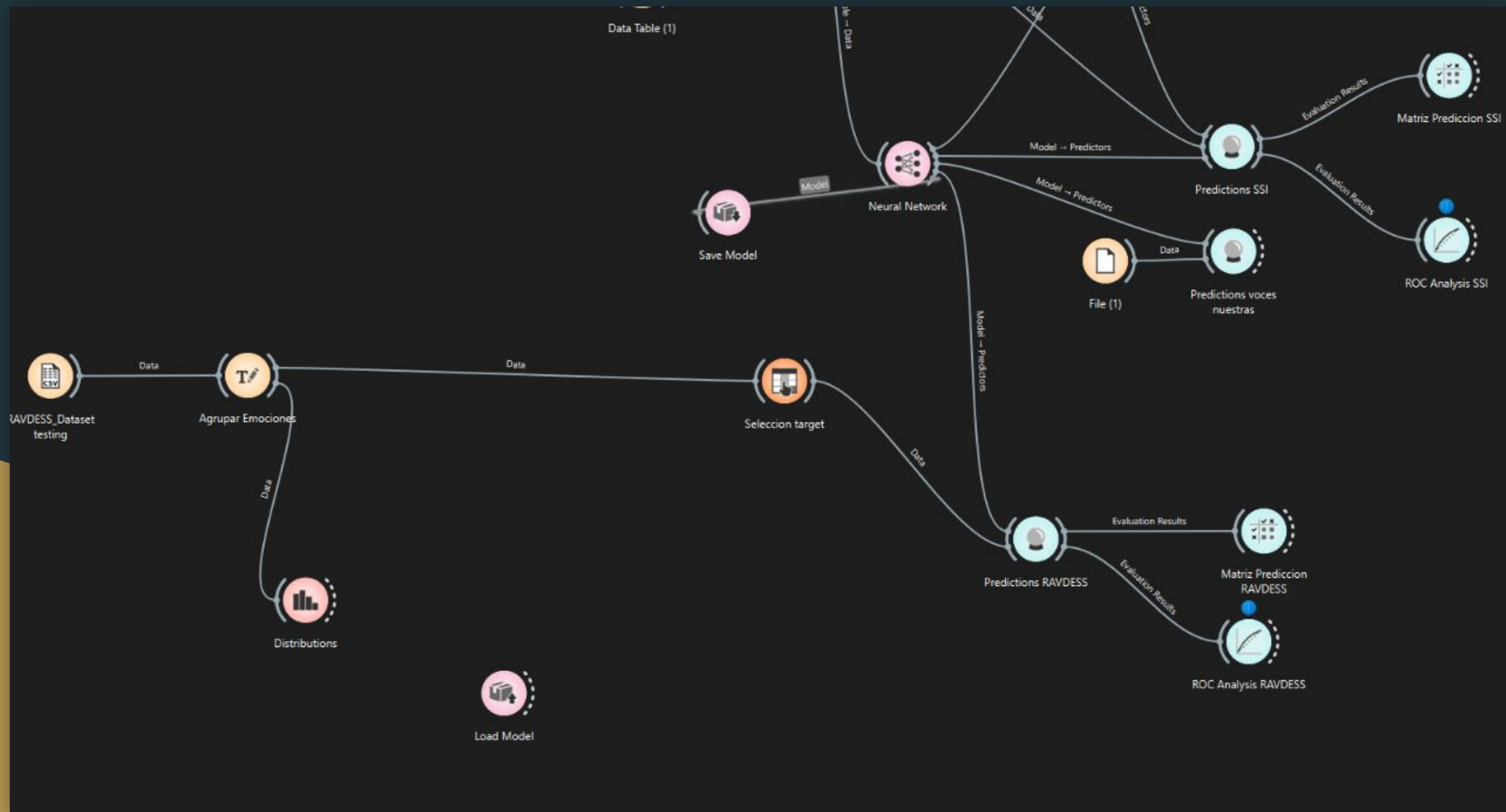




ions (1)

Table (1)





Comparativa

Random Forest:

- Más estable en datasets medianos.

Neural Network:

- Mayor potencial para patrones no lineales complejos.

Conclusiones

Viabilidad del Enfoque SER:

- Clasificación de emociones factible usando solo física del audio, validando MFCCs como huella emocional.

Eficacia Metodológica:

- La combinación de extracción con Librosa y modelado en Orange permitió una iteración ágil y efectiva.

Impacto de los Datos:

- Rendimiento excelente en datos controlados (RAVDESS), con desafíos de generalización ante voces naturales y ruido.

**Gracias por
su atención**