

# PROYECTO 1

## UNEPA

Leonardo Rangel  
Santiago Molina  
François Morales



# Entendimiento del Negocio

**El Fondo de Poblaciones de las Naciones Unidas (UNFPA) busca automatizar el análisis de opiniones ciudadanas en relación con los ODS 3, 4 y 5. El proyecto tiene como objetivo desarrollar un modelo analítico que, mediante una aplicación web o móvil, relacione automáticamente estas opiniones con los ODS, permitiendo un análisis más eficiente y preciso a gran escala.**

# Entendimiento de los datos

Realizamos un entendimiento de los datos donde concluimos que: con 3,239 observaciones, los textos son 100% únicos y la variable ODS está moderadamente balanceada (ODS 5: 36.3%).

Luego, realizamos el perfilamiento de los datos evaluando completitud, unicidad y consistencia. Los datos están completos, sin duplicados, y todos los valores categóricos están dentro del rango esperado, garantizando su calidad.

3239 opiniones

ODS 3: 30.1%  
ODS 4: 33.6%  
ODS 5: 36.3%

0 duplicados

# Preparación de los datos

**Limpieza**

**Tokenización**

**Normalización**

**Separación y  
Selección**

# Implementación de los Algoritmos



Se desarrollaron 3 modelos predictivos utilizando los siguientes algoritmos:

- Random Forest
- Naive Bayes
- LDA

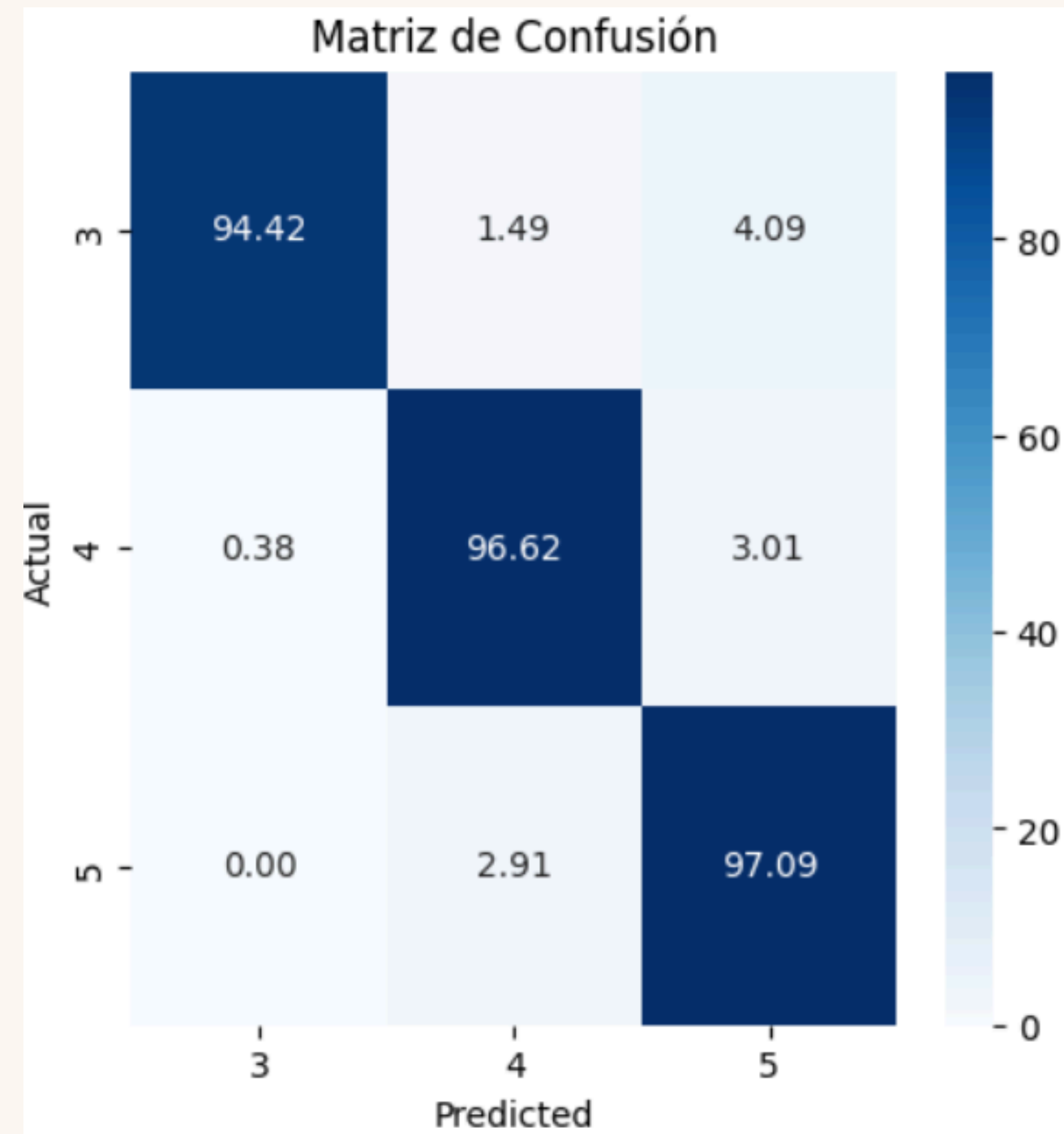
# Resultados

## Naive Bayes

Exactitud train: 0.9932

Exactitud test: 0.9605

	precision	recall	f1-score	support
3	1.00	0.94	0.97	269
4	0.96	0.97	0.96	266
5	0.93	0.97	0.95	275
accuracy			0.96	810
macro avg	0.96	0.96	0.96	810
weighted avg	0.96	0.96	0.96	810



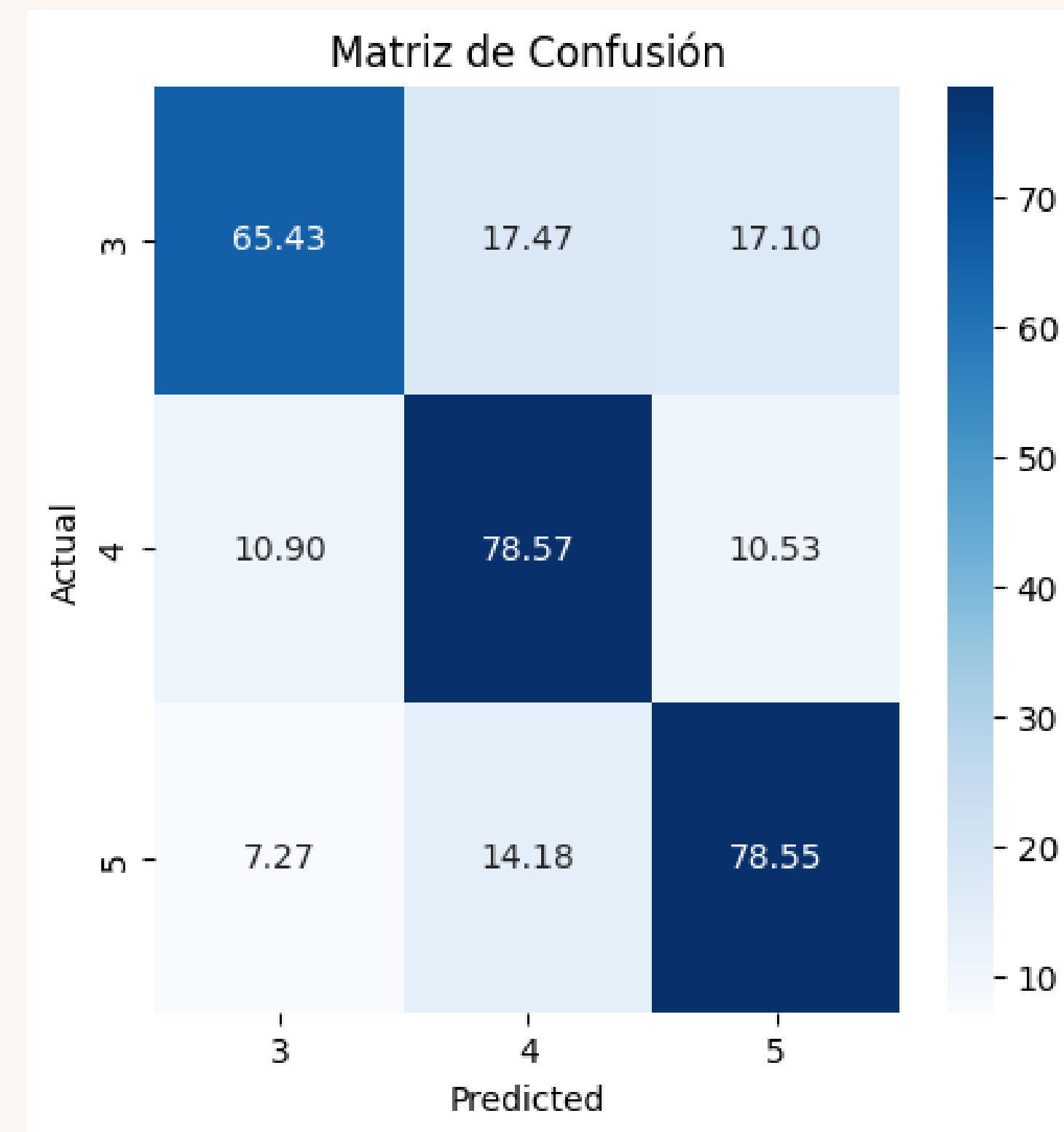
# Resultados

## LDA

Exactitud train: 0.89

Exactitud test: 0.74

	precision	recall	f1-score	support
3	0.78	0.65	0.71	269
4	0.71	0.79	0.75	266
5	0.74	0.79	0.76	275
accuracy			0.74	810
macro avg	0.75	0.74	0.74	810
weighted avg	0.75	0.74	0.74	810
Exactitud en prueba: 0.74				



# Resultados

## Random Forest

Exactitud train: 1.00

Exactitud test: 0.97

	precision	recall	f1-score	support
3	0.98	0.96	0.97	269
4	0.97	0.97	0.97	266
5	0.96	0.97	0.97	275
accuracy			0.97	810
macro avg	0.97	0.97	0.97	810
weighted avg	0.97	0.97	0.97	810



¡Muchas  
gracias!

