

PROYECTO “TITANIC”

01

Objetivo del trabajo

Predecir la supervivencia de pasajeros del Titanic aplicando técnicas de machine learning y comparar modelos de red neuronal.

02

Dataset y Variables

- Características seleccionadas:
- Categoricals: Sexo, clase, puerto de embarque
 - Numéricas: Edad, tarifa, cantidad de familiares a bordo

03

Pipeline aplicado

- 1. **Preprocesamiento:**
 - Imputación de valores faltantes (mediana o moda)
 - Eliminación de columnas irrelevantes
- 2. **Procesamiento:**
 - Codificación one-hot para variables categoricas
 - Estandarización de variables numéricas
- 3. **Postprocesamiento:**
 - Balanceo de clases mediante sobremuestreo aleatorio

Arquitectura de RNA

Modelo	Arquitectura	Activación	Optimizador	LR Init	Épocas	Batch
Modelo A	(16,)	relu	adam	0,001	200	32
Modelo B	(32, 16)	tanh	sgd	0,010	300	64
Modelo C	(64, 32, 16)	relu	adam	0,0005	250	16

Resultados del entrenamiento

Modelo	Accuracy	Precision	Recall	F1-Score	Tiempo (s)
Modelo A	0,7727	0,8191	0,7000	0,7549	0,8200
Modelo B	0,7818	0,7981	0,7545	0,7757	0,9400
Modelo C	0,7955	0,8283	0,7455	0,7847	3,7600

Conclusión

El modelo C es el más eficiente y preciso para el objetivo planteado, logrando el mayor F1-score y precisión, y por lo tanto, se recomienda su uso para predecir la supervivencia de los pasajeros.