

INFORME FINAL: PREDICCIÓN DE SUPERVIVENCIA EN PACIENTES UCI

1. INTRODUCCIÓN

Este proyecto tuvo como objetivo desarrollar un modelo predictivo de clasificación que estime la supervivencia de pacientes ingresados a una Unidad de Cuidados Intensivos (UCI), utilizando datos simulados que contienen variables clínicas relevantes.

2. EXPLORACIÓN DE DATOS

Se realizó un análisis exploratorio del dataset, el cual contiene 5000 instancias y 15 atributos. Se identificaron valores faltantes en múltiples variables, y se analizaron la distribución y correlación entre variables mediante histogramas y mapas de calor.

3. PREPROCESAMIENTO

Se imputaron los datos faltantes usando la media para variables numéricas. Las variables categóricas fueron codificadas con Label Encoding y se estandarizaron las variables numéricas para asegurar un entrenamiento más eficiente de los modelos.

4. MODELOS IMPLEMENTADOS

Se entrenaron dos algoritmos supervisados:

- Modelo 1: Random Forest Classifier
- Modelo 2: Regresión Logística

Ambos modelos fueron entrenados y evaluados con métricas como matriz de confusión y reporte de clasificación.

5. COMPARACIÓN DE DESEMPEÑO

El desempeño fue comparado usando la métrica ROC AUC. El modelo Random Forest mostró un mejor rendimiento, indicando mayor capacidad para clasificar correctamente los casos positivos y negativos.

6. CONCLUSIONES

El modelo Random Forest fue más efectivo para predecir la supervivencia de pacientes en UCI. Se recomienda probar otros algoritmos, ajustar hiperparámetros más finamente y utilizar datos reales en futuros trabajos.