



Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Estado de México

**Momento de Retroalimentación: Módulo 2 Uso de framework o biblioteca de aprendizaje máquina para la implementación de una solución.
(Portafolio Implementación)**

Inteligencia artificial avanzada para la ciencia de datos I

Daniel Alejandro Martínez Cienfuegos - A01745412

Grupo 101

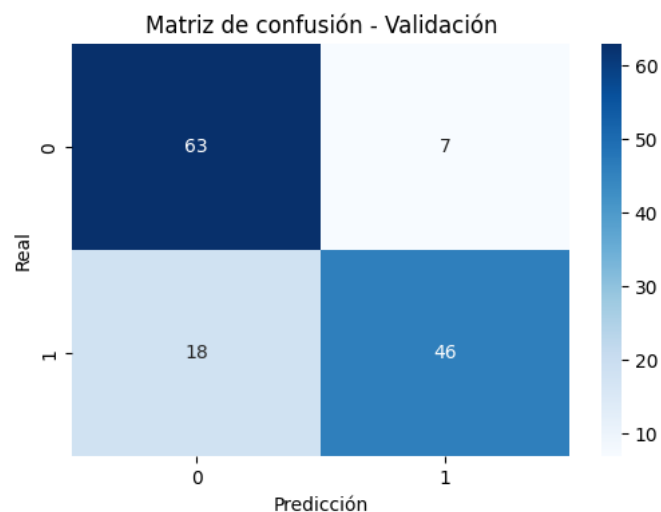
09/09/2024

Análisis de Resultados - Modelo DecisionTreeClassifier

El modelo utilizado fue un árbol de decisión con la configuración de $\text{max-depth} = 5$, y un $\text{random_state} = 42$, que ayuda a limitar la profundidad del árbol para evitar el sobreajuste y hace el experimento reproducible.

Al evaluar los resultados obtenidos en la matriz de confusión, se puede observar que el modelo clasifica correctamente a 63 pasajeros que no sobrevivieron, lo que indica que estos fueron etiquetados correctamente como no sobrevivientes. Sin embargo, comete siete errores al clasificar como sobrevivientes a pasajeros que en realidad no lo fueron. Del mismo modo, el modelo identifica correctamente a 46 pasajeros que sobrevivieron, pero también deja de identificar correctamente a 18 sobrevivientes, clasificándolos erróneamente como no sobrevivientes. En cuanto a las métricas de evaluación, la exactitud del modelo es del 81%, lo que sugiere un desempeño sólido en términos de predicción general. La precisión, que mide la proporción de predicciones positivas correctas entre todas las predicciones positivas, refleja que el modelo tiene pocos falsos positivos. Esto implica que la mayoría de las veces que el modelo predice que alguien sobrevivió, está en lo correcto. La sensibilidad o recall, que evalúa cuántos de los sobrevivientes reales fueron correctamente identificados, muestra que el modelo logra capturar al 71% de ellos. Por último, el F1 Score, una métrica que equilibra precisión y recall, presenta un valor de 0.78, lo que refleja un buen balance entre la capacidad de encontrar sobrevivientes y de evitar clasificaciones incorrectas.

```
Accuracy: 0.8134
Precision: 0.8679
Recall: 0.7188
F1 Score: 0.7863
```



La visualización del árbol de decisión muestra que la primera característica clave utilizada por el modelo para tomar decisiones es el sexo de los pasajeros. Esta característica refleja la alta probabilidad de que los hombres no sobrevivieran, mientras que la mayoría de las mujeres sí lo hicieron. A partir de esta división inicial, el modelo utiliza otras características como la edad y la tarifa pagada para tomar decisiones adicionales. Para las mujeres, las pasajeras más jóvenes y aquellas que pagaron tarifas más altas parecen tener una mayor probabilidad de supervivencia. En el caso de los hombres, factores como la edad, el número de acompañantes y la clase del pasajero influyen notablemente en la predicción de supervivencia.

