Diplomado en ML Cloud - UCB CBBA

Módulo 4: Machine Learning Cloud MLOps

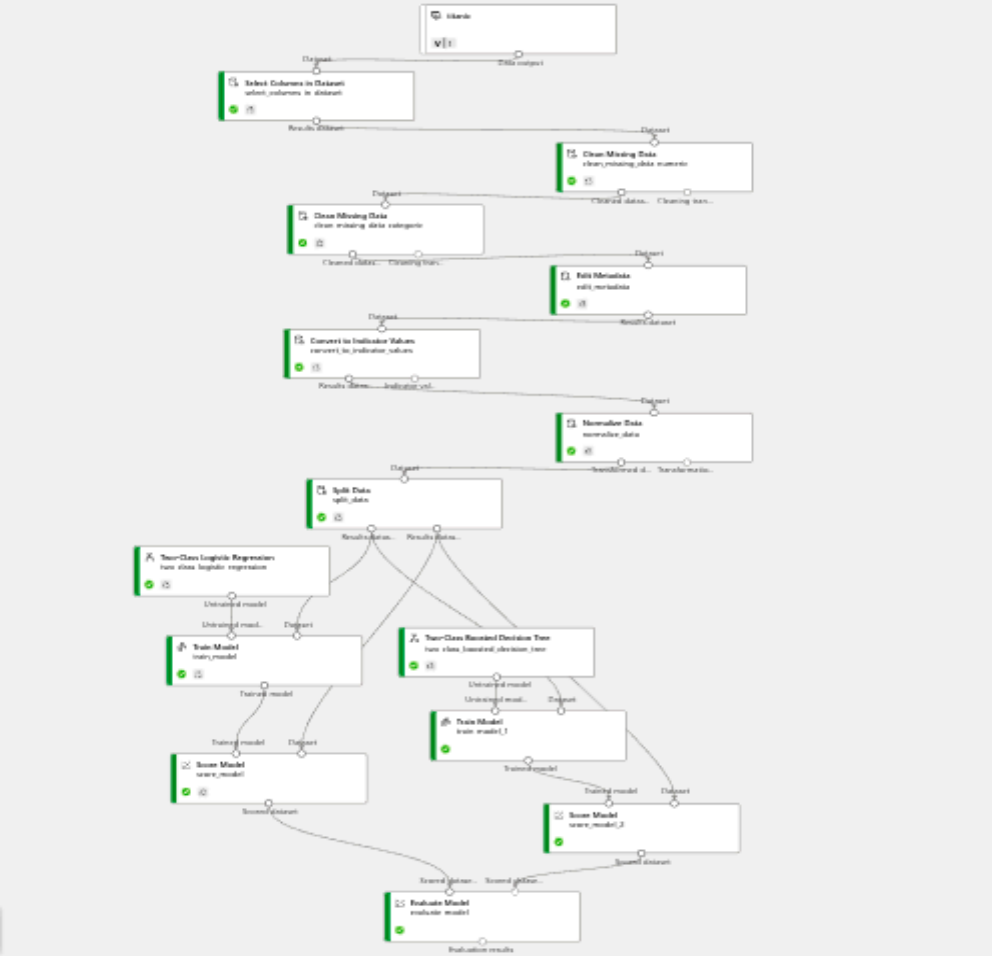
Docente: Ing. Mauricio Alejandro Quezada

Estudiante: Jose Carlos Iriarte

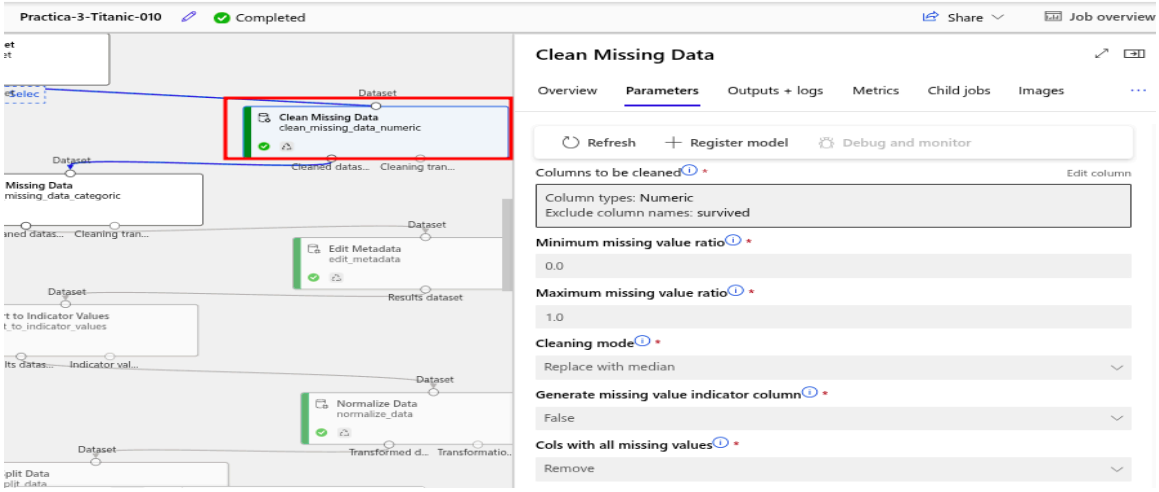
Fecha : Agosto del 2025

Laboratorio 3

Captura del DAG completo del pipeline



Captura de configuración de Clean Missing Data



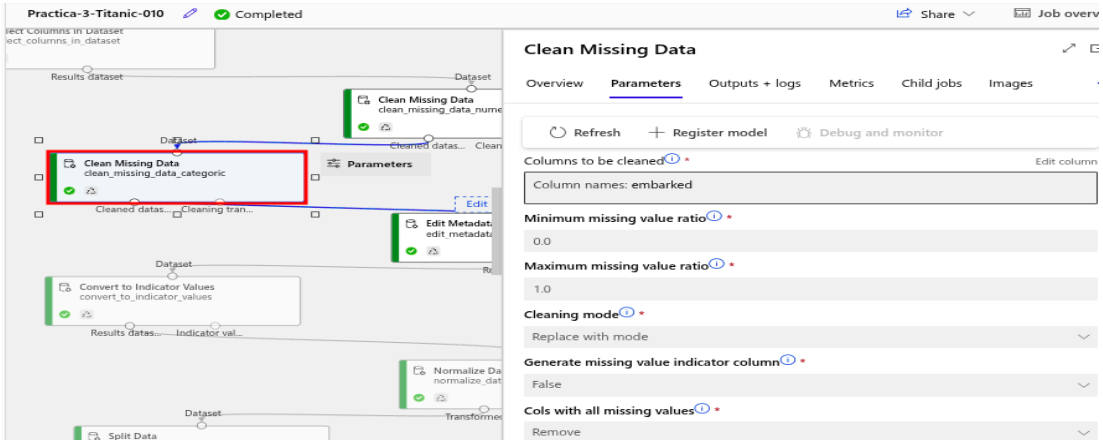
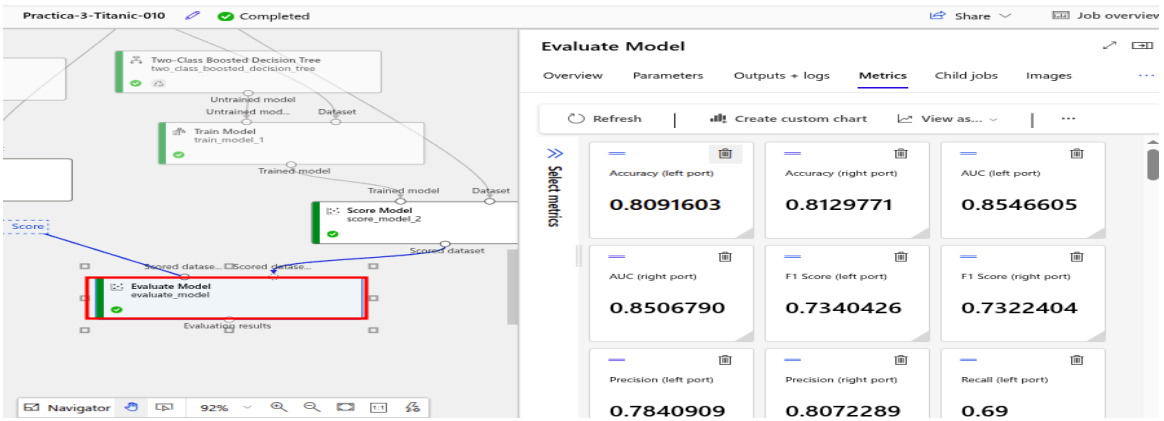
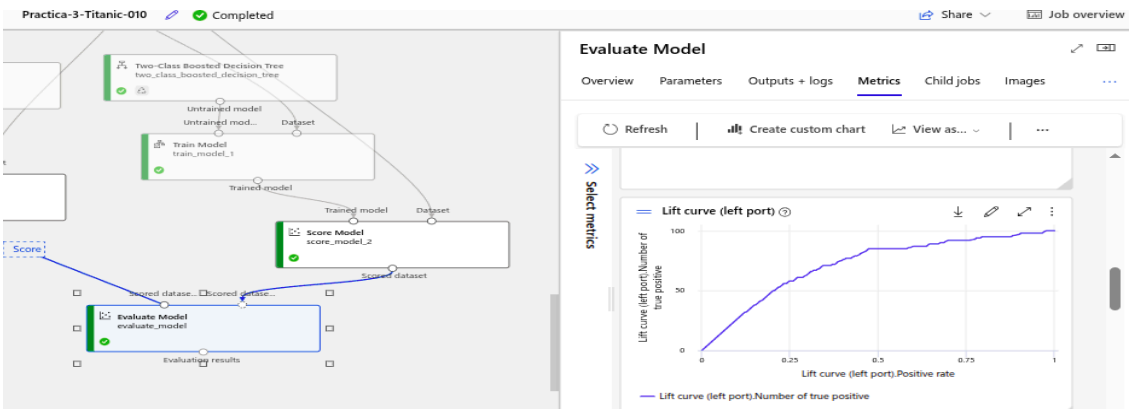
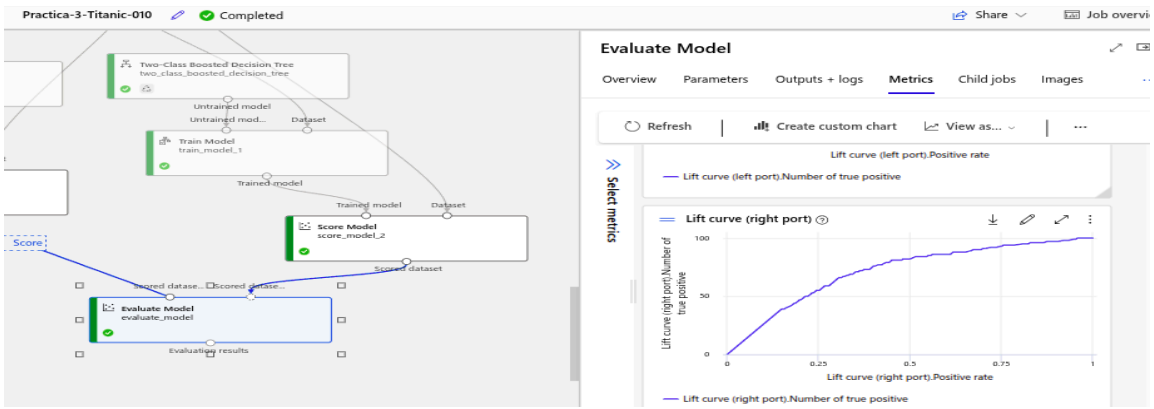
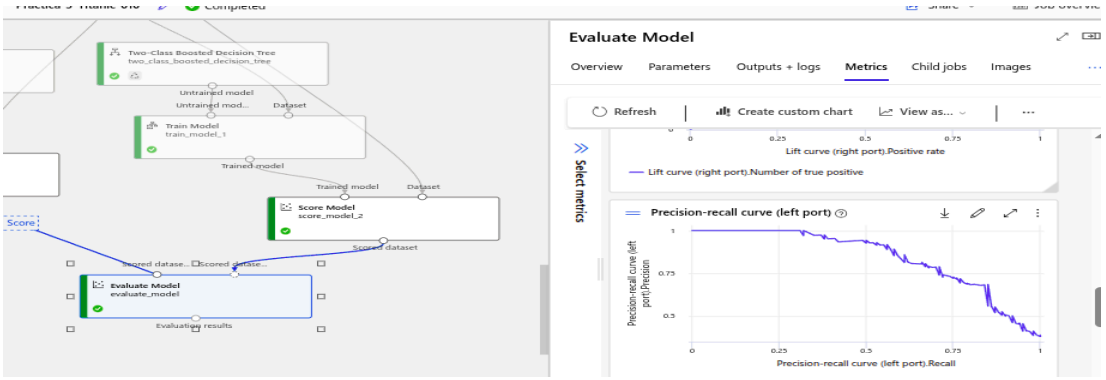


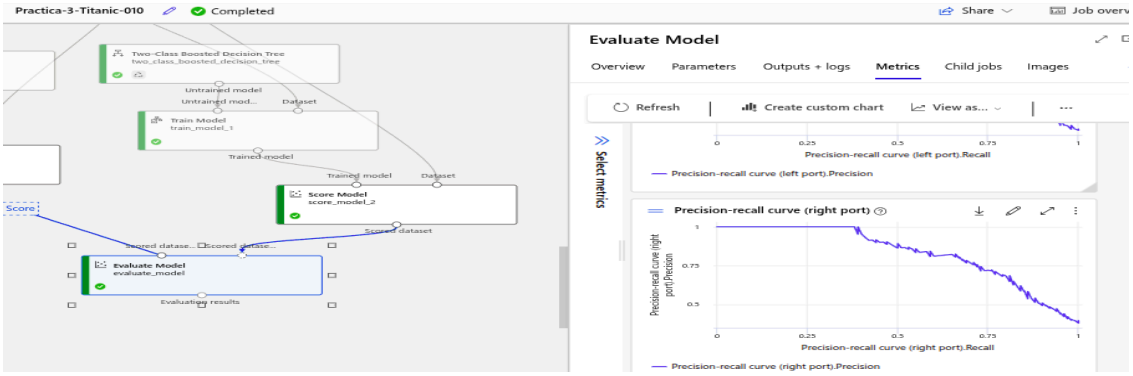
Tabla comparativa de métricas (Logistic vs Boosted)

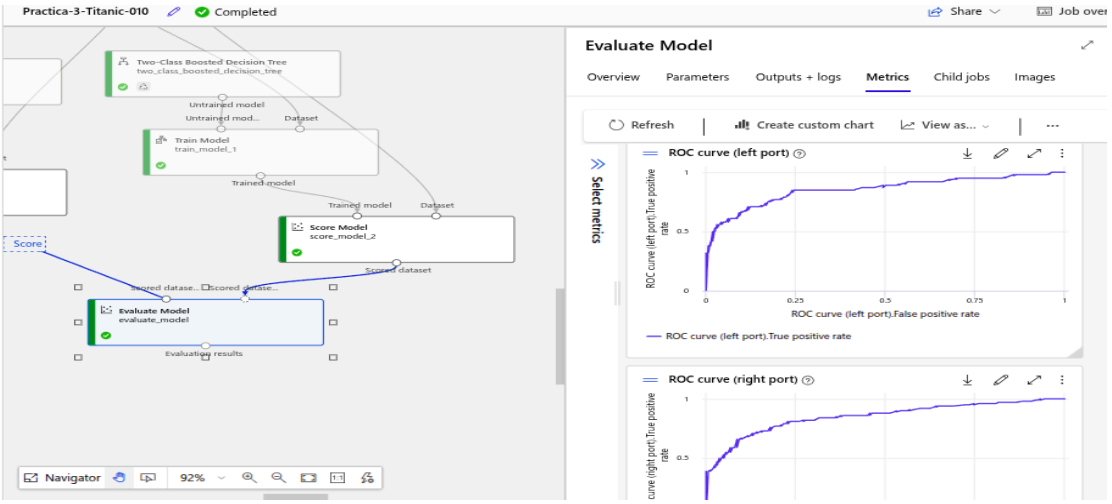












Conclusiones breves

- Ambos modelos (Logistic Regression y Boosted Decision Tree) alcanzaron un rendimiento aceptable, con accuracies alrededor de 0.81.
- El Boosted Decision Tree mostró mejor desempeño general, destacando en AUC (0.85) y precisión (≈0.80), lo que indica mayor capacidad discriminativa.
- La Logistic Regression tuvo métricas más equilibradas, pero quedó ligeramente por debajo en F1 y Recall.
- En problemas como Titanic, donde el costo de falsos negativos es relevante (no predecir sobrevivientes), un modelo con mayor Recall podría ser más útil, aunque implique sacrificar algo de precisión.