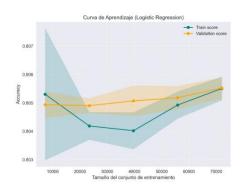
Reporte Técnico Diagnóstico de Overfitting/Underfitting

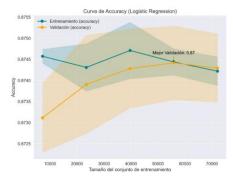
Objetivo:

Analizar fenómenos de overfitting y underfitting en un modelo de Regresión Logística aplicado a un dataset de señales cardíacas, estableciendo una línea base confiable para futuros modelos más complejos.



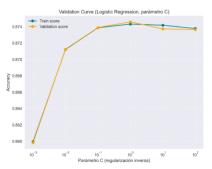
1. Metodología

- Modelo: Regresión Logística con regularización L2.
- Datos: División estratificada para preservar la representatividad de clases.
- Evaluación: curvas de aprendizaje y curvas de validación con el hiperparámetro C.
- Métricas: Accuracy y Log-loss.



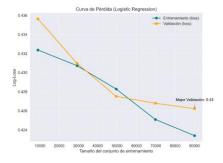
2. Resultados del diagnóstico

- Accuracy estable ≈ 87% y log-loss ≈ 0.43.
- Valores bajos de C → underfitting.
- Rango intermedio (1–10) → mejor equilibrio sesgo-varianza.
- Valores altos de C → tendencia a overfitting.



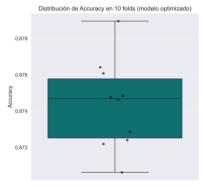
3. Análisis de curvas de aprendizaje

- La brecha entre entrenamiento y validación es reducida → buena generalización.
- La pérdida disminuye y la precisión se estabiliza → modelo confiable.



5. Estrategias implementadas

- Optimización de C y tipo de penalización (L1 vs L2).
- Validación cruzada estratificada (10 folds) para mayor robustez.
- Selección y regularización de características.
- Preparación para comparar con modelos más complejos.



6. Conclusiones y recomendaciones

- La regresión logística es un baseline sólido, equilibrando sesgo y varianza.
- C entre 1 y 10 es óptimo.
- Limitaciones: linealidad, desbalance de clases y variabilidad fisiológica.
- Futuro: explorar CNNs y aplicar técnicas de re-muestreo (SMOTE, etc.).