

Preguntas para Guía de Aprendizaje: Preguntas sobre Modelos de Clasificación

1. ¿Cuál es la variable objetivo y cómo se codifica para asegurar que el modelo de clasificación opere con datos limpios y coherentes?
2. ¿Qué estrategia de limpieza y tratamiento de valores faltantes se implementará para garantizar integridad y calidad en el pipeline de datos?
3. ¿Cómo se seleccionarán y evaluarán las características (feature selection) para maximizar la relevancia y minimizar la redundancia?
4. ¿Qué método de partición de datos (hold-out, k-fold cross-validation, stratified sampling) se empleará para validar la robustez del modelo?
5. ¿Qué métricas de performance (precisión, recall, F1-score, AUC-ROC) se definirán como KPI primarios y por qué?
6. ¿Cómo se detectará y mitigará el sobreajuste (overfitting) y subajuste (underfitting) durante el entrenamiento?
7. ¿Qué técnica de ajuste de hiperparámetros (grid search, random search, Bayesian optimization) se utilizará y cómo se evaluará su impacto en la eficiencia?
8. ¿Cómo se interpretará la importancia de las variables y se explicarán las predicciones al negocio (SHAP, LIME, coeficientes)?
9. ¿Qué plan de monitoreo y mantenimiento en producción se diseñará para asegurar estabilidad y performance continua?
10. ¿Cuál es la estrategia de despliegue (batch vs. real-time) y cómo se integrará el modelo en la arquitectura existente?