Actividad: Elegir la métrica adecuada para tu problema

Caso 1: Mantenimiento predictivo en fábricas Una empresa de manufactura quiere anticipar fallas en sus máquinas. Cada equipo tiene sensores que registran temperatura, vibraciones y horas de uso. Se busca predecir la vida útil restante (en horas).

Caso 2: Clasificación de reseñas de clientes En un e-commerce se quiere clasificar automáticamente las reseñas de clientes en positivo, neutral o negativo para agilizar el análisis de satisfacción.

Caso 3: Detección de intrusiones en ciberseguridad Un sistema de seguridad informática recibe millones de conexiones diarias. Se busca identificar si cada conexión es legítima o un ataque.

Caso 4: Predicción de consumo energético en hogares inteligentes Una empresa eléctrica quiere predecir el consumo energético de los hogares en las próximas horas para optimizar la red y reducir costos.

|  |  |  |  |
| --- | --- | --- | --- |
| Caso de Aplicación | ¿Es clasificación o regresión ? | Posibles métricas | ¿Cúal eligirias y porque? (Justificación del error mas costoso) |
| Mantenimiento predictivo en fábricas | Regresión | * **RMSE (Root Mean Squared Error):** Penaliza más los errores grandes. * **MAE (Mean Absolute Error):** Más interpretable, útil si los errores extremos no son tan críticos. * **R² (Coeficiente de determinación):** Evalúa qué tan bien el modelo explica la variabilidad. | Los errores grandes (por ejemplo, predecir que una máquina durará 100 horas cuando en realidad fallará en 10) pueden causar **paradas inesperadas**, pérdidas económicas y riesgos operativos. RMSE penaliza más los errores grandes, lo que ayuda a entrenar modelos más conservadores y precisos en predicciones críticas. |
| Clasificación de reseñas de clientes | Clasificación multiclase | * **Accuracy:** Si las clases están balanceadas. * **F1-score (macro):** Si hay clases desbalanceadas, mide precisión y recall por clase. * **Matriz de confusión:** Para ver errores específicos entre clases. | Si el modelo clasifica mal reseñas negativas como positivas, puede ocultar problemas reales de satisfacción. Como puede haber **desbalance entre clases** (más reseñas positivas que negativas), el F1-score macro evalúa el rendimiento en cada clase por igual, evitando que el modelo ignore las clases minoritarias. |
| Detección de intrusiones en ciberseguridad | Clasificación binaria con clases desbalanceadas | * **Precision y Recall:** Muy importante evitar falsos positivos o negativos. * **F1-score:** Balance entre precisión y recall. * **AUC-ROC:** Evalúa la capacidad del modelo para distinguir entre clases. * **Matriz de confusión:** Para análisis detallado. | El error más costoso es **no detectar un ataque** (falso negativo). Por eso, el **recall** es clave: mide cuántos ataques reales se detectan. AUC-ROC complementa esto mostrando la capacidad del modelo para distinguir entre conexiones legítimas y maliciosas, útil en sistemas con millones de eventos diarios. |
| Predicción de consumo energético en hogares inteligentes | Regresión | * **MAE / RMSE:** Para evaluar precisión de la predicción. * **MAPE (Mean Absolute Percentage Error):** Útil si se quiere expresar el error en porcentaje. * **R²:** Para entender la capacidad explicativa del modelo. | La empresa necesita ajustar la red eléctrica en tiempo real. MAPE permite entender el error en **términos porcentuales**, lo que es útil para comparar entre hogares con distintos niveles de consumo. Además, es más interpretable para decisiones operativas. |