

La evaluación de modelos mide su rendimiento en tareas específicas. Se utilizan métricas como accuracy, precision, recall, F1-score, AUC o perplexity. Es crucial definir un conjunto de prueba que represente adecuadamente los datos reales. Además, técnicas como validación cruzada, matriz de confusión o análisis de error ayudan a refinar los modelos. Para modelos generativos, métricas como BLEU o ROUGE permiten evaluar la calidad del texto producido. La evaluación continua es vital para modelos...