

**Cuadro comparativo entre regresión, clasificación y clustering**

Característica	Regresión	Clasificación	Clustering
<b>Definición</b>	Predice un valor numérico continuo	Asigna una etiqueta o clase a cada observación	Agrupar observaciones en clusters sin etiquetas previas
<b>Objetivo</b>	Minimizar el error entre predicciones y valores reales	Maximizar la precisión de clasificación	Agrupar observaciones similares en clusters
<b>Algoritmos Comunes</b>	- Regresión Lineal	- Regresión Logística	- K-Means
	- Regresión Polinómica	- Árboles de Decisión	- DBSCAN
	- Regresión Ridge	- Random Forest	- Agglomerative Clustering
	- Lasso	- SVM (Máquinas de Vectores de Soporte)	- Mean Shift
	- Elastic Net	- K-Nearest Neighbors (K-NN)	- Birch
	- Support Vector Regression (SVR)	- Naive Bayes	
<b>Entrada de Datos</b>	Variables independientes y variable dependiente continua	Variables independientes y variable dependiente categórica	Solo variables independientes
<b>Salidas</b>	Valor numérico continuo	Etiqueta de clase	Etiqueta de cluster
<b>Métricas de Evaluación</b>	- MSE (Mean Squared Error)	- Accuracy	- Silhouette Score
	- RMSE (Root Mean Squared Error)	- Precision, Recall, F1-Score	- Inertia (para K-Means)
		- MAE (Mean Absolute Error)	- ROC-AUC
		- R <sup>2</sup> (Coeficiente de Determinación)	- Confusion Matrix
<b>Ejemplos de Uso</b>	- Predicción de precios	- Detección de spam	- Segmentación de clientes
	- Predicción de ventas	- Clasificación de imágenes	- Agrupación de noticias
	- Predicción de demanda	- Diagnóstico médico	- Clustering de genes
<b>Ventajas</b>	- Interpretación sencilla (en regresión lineal)	- Variedad de algoritmos para diferentes necesidades	- No requiere etiquetas
	- Aplicable a muchos problemas prácticos	- Funciona bien con datos categóricos	- Descubre estructuras ocultas en los datos
	- Permite manejar relaciones lineales y no lineales	- Modelos avanzados pueden manejar datos complejos	- Puede utilizarse para reducir dimensionalidad
<b>Desventajas</b>	- Sensible a outliers (en algunos casos)	- Puede requerir gran cantidad de datos etiquetados	- Número de clusters debe ser especificado (en algunos)
	- Puede no capturar bien relaciones no lineales	- Puede ser computacionalmente intensivo	- Sensible a la escala de datos
	- Supone que la relación entre variables es fija	- Interpretabilidad puede ser un reto en modelos complejos	- No garantiza clusters significativos