

Práctica 04. Support Vector Machines.

Sistemas de Aprendizaje Automático.

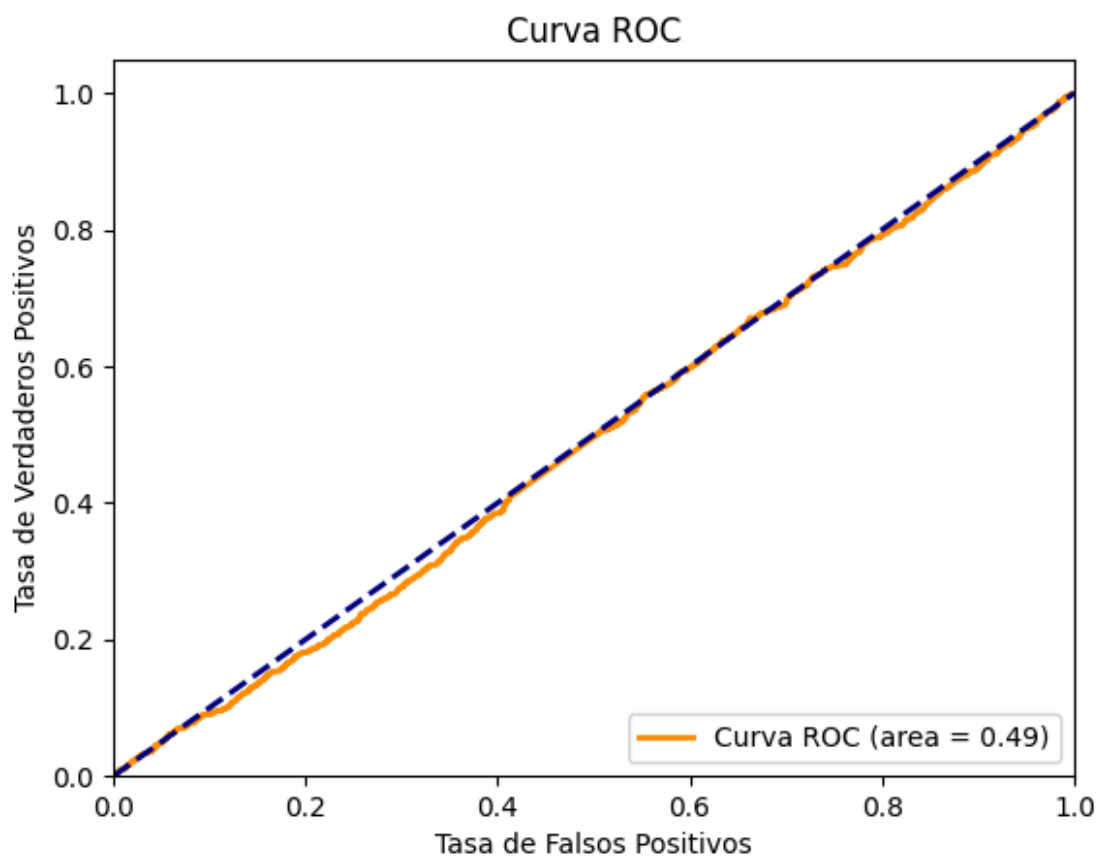
Álvaro Martínez Lineros.

Introducción.

Se desea construir un modelo de clasificación basado en support vector machines (SVM) que ayude a predecir si se va a realizar fraude en transacciones bancarias.

Resultados.

Curva ROC:



Matriz de confusión:

Verdaderos negativos	3202	Falsos positivos	2469
Falsos negativos	757	Verdaderos positivos	572

Reporte de clasificación:

	precision	recall	f1-score	support
No fraude	0.81	0.56	0.67	5671
Fraude	0.19	0.43	0.26	1329
accuracy			0.54	7000
macro avg	0.50	0.50	0.46	7000
weighted avg	0.69	0.54	0.59	7000

Factores clave.

Tras entrenar el modelo y ejecutar un test se puede sacar la conclusión de que ninguno de los factores son clave para la toma de decisiones, debido a su baja precisión de predicción.

Limitaciones del modelo.

La baja correlación entre los datos hace que la precisión del modelo sea extremadamente baja. Una posible mejora podría ser añadir otros factores correlacionados con la detección del fraude.

Aplicabilidad.

En una empresa este modelo no sería aplicable por su baja precisión y su poca generalización. Haría falta incrementar el f1-score de ambas situaciones a al menos 0.8 para poder considerar su aplicación.

Conclusión.

Este modelo no es aplicable en una situación real. Posibles soluciones serían aumentar el número de factores correlacionados con el fraude o usar modelos como RandomForest, KNN, etc.