

Informe Proyecto 2 entrega 7

Juan Luis Solórzano (carnet: 201598)

Micaela Yataz (carnet: 18960)

2025-01-20

https://github.com/JusSolo/Mineria_Proyecto2.git

git: https://github.com/JusSolo/Mineria_Proyecto2.git

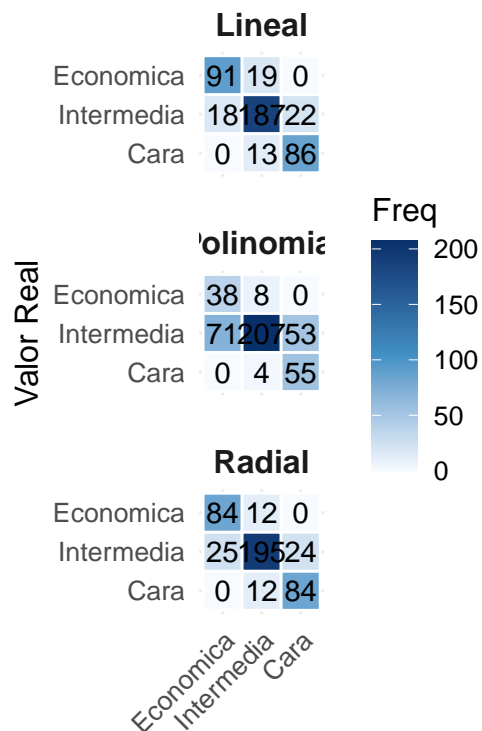
Introducción:

A lo largo del semestre hemos usado diferentes modelos de aprendizaje supervisado tanto en su versión de clasificación como de regresión. Para este último informe del proyecto 2 se pretende probar máquina de vectores de soporte con diferentes topologías para clasificar los precios de las casas en categorías y otras redes para predecir el precio de las mismas. Por otro lado se desea comparar todos los modelos anteriores.

Modelo de Clasificación con redes Neuronales

Comparacion de primeros 3 modelos con distintos kernels

Matrices de Confusión por Modelo



Observando las matrices de confusión se puede concluir que los 3 modelos se equivocan más clasificando la categoría de las casas intermedias. El modelo radial y lineal parecen ser los mejores, Siendo el radial más equilibrado y el lineal mejor en la clasificación entre las casas económicas e intermedias.

Table 1: Comparación de métricas macro-promediadas para los 3 modelos SVM usando datos de prueba

	Accuracy	Kappa	Sensitivity	Precision	F1
Linear	0.835	0.733	0.828	0.840	0.834
Radial	0.833	0.725	0.813	0.850	0.828
Polynomial	0.688	0.442	0.601	0.795	0.634

Table 2: Comparación de métricas macro-promediadas para los 3 modelos SVM usando datos de entrenamiento

	Accuracy	Kappa	Sensitivity	Precision	F1
Linear	0.890	0.822	0.882	0.895	0.888
Radial	0.929	0.885	0.922	0.935	0.928
Polynomial	0.800	0.651	0.734	0.900	0.776

Comparado ambas tablas el único modelo que parece sobre ajustado es el polynomial, pues es que tiene mayores diferencias en las métricas de desempeño con los datos de prueba y entrenamiento.

Modelo ajustado

Entre los 3 modelos anteriores se decidió ajustar el modelo radial, pues es más flexible que el lineal y fue un poco pero que el lineal. Posiblemente al ajustarlo su desempeño mejore.

Mejores hiperparámetros:

```
##      sigma      C
## 21 0.001 100
```

Table 3: Métricas macro-promediadas del modelo SVM Radial tuneado (entrenamiento vs prueba)

	Accuracy	Kappa	Sensitivity	Precision	F1
Tuned Radial (Train)	0.902	0.842	0.892	0.911	0.901
Tuned Radial (Test)	0.823	0.713	0.813	0.831	0.821

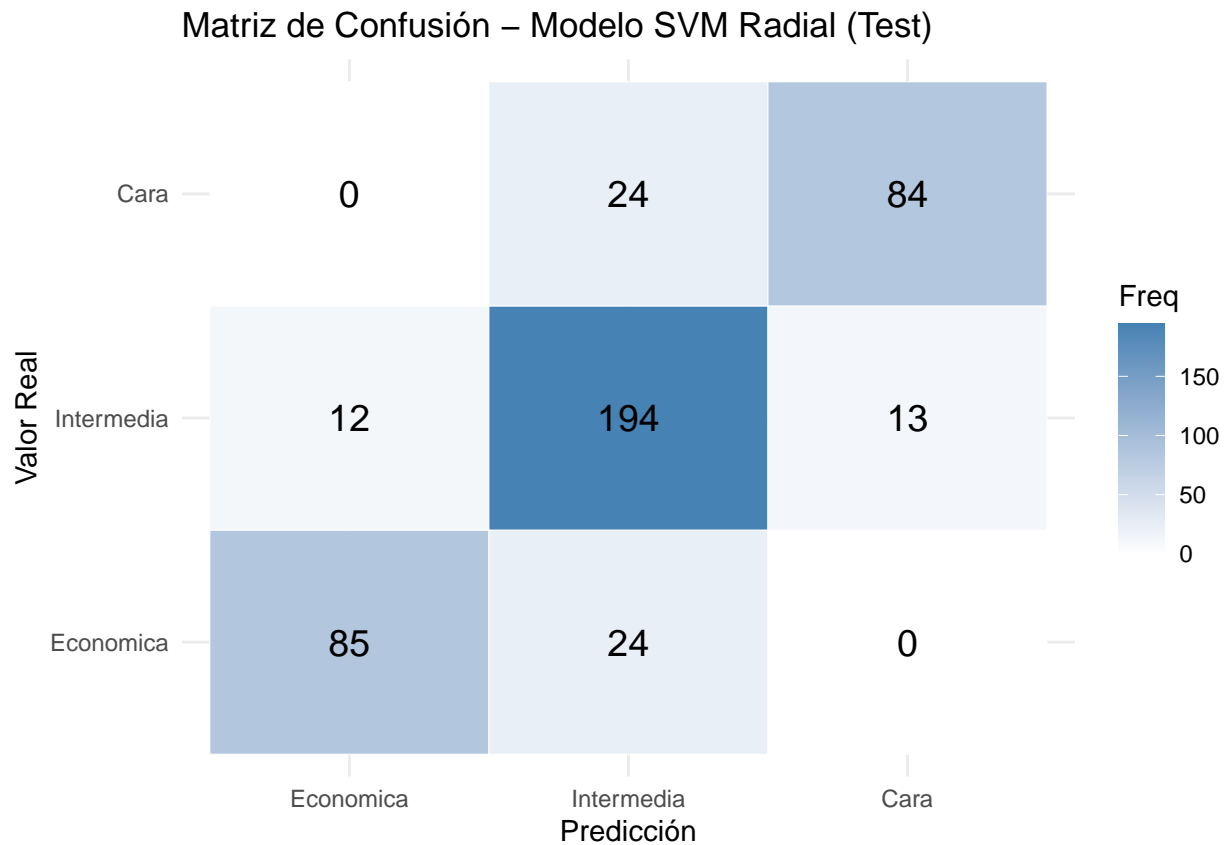
El modelo tuneado tiene claramente sobre ajuste y es peor, por lo que vamos a volver a ajustarlo pero separando los datos de entrenamiento en 2 entrenamiento y validación para evitar sobre ajuste.

Mejores hiperparámetros:

```
##      sigma      C
## 3 0.001 100
```

Table 4: Métricas del modelo SVM Radial tuneado (con train)

	Accuracy	Kappa	Sensitivity	Precision	F1
Validación (subtrain)	0.896	0.830	0.883	0.907	0.894
Prueba (test final)	0.833	0.726	0.814	0.848	0.829



Viendo las métricas y las matrices de confusión y las métricas el modelo tuneado radial parece ser un poco mejor que el lineal, pero por muy poco.

Comparacion de todos los modelos de clasificacion usados a lo largo del proyecto

Modelo de Regresion con redes Neuronales

Comparacion de todos los modelos de regresion usados a lo largo del proyecto