

UNIVERSIDAD DE GUADALAJARA



CENTRO UNIVERSITARIO DE CIENCIAS EXACTAS E INGENIERÍAS

Sistemas Inteligentes IV

Actividad 3. Regresión polinómica y generalización

Pacheco Quintero Marco Antonio 213535019

26 de marzo de 2021

Semestre 2021A

Sección D01

Objetivo

- Para la parte 1, realizar un programa en Python para aplicar una regresión polinómica a los datos de los archivos adjuntos. Utilizar las herramientas sklearn para el modelo de regresión.
- Para la parte 2, realizar un programa en Python para aplicar una regresión polinómica a los datos del archivo adjunto. Utilizar las herramientas sklearn para el modelo de regresión.
- Para cada caso utilizar las métricas de coeficiente R^2 para demostrar que el entrenamiento tiene una buena generalización, es decir aplicar la métrica tanto a los datos de entrenamiento como a los datos de generalización y comparar.

Resultados

El resultado de la regresión polinómica para los datos del archivo adjunto “df_regresion_lineal_1.csv” se muestra en la Figura 1.

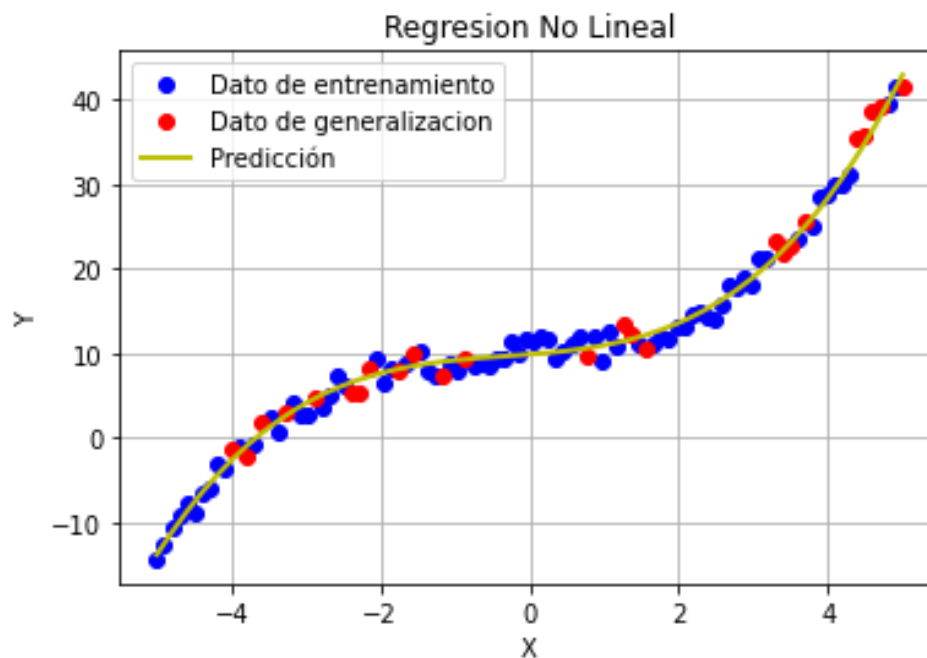


Figura 1. Regresión lineal de los datos del archivo adjunto 1.

Los resultados de las métricas de coeficiente R^2 fueron:

- Coeficiente R^2 para los datos de entrenamiento

$$R^2 = 0.9908411775045374$$

- Coeficiente R^2 para los datos de generalización

$$R^2 = 0.9906171099862217$$

El resultado de la regresión polinómica para los datos del archivo adjunto “df_regresion_lineal_2.csv” se muestra en la Figura 2.

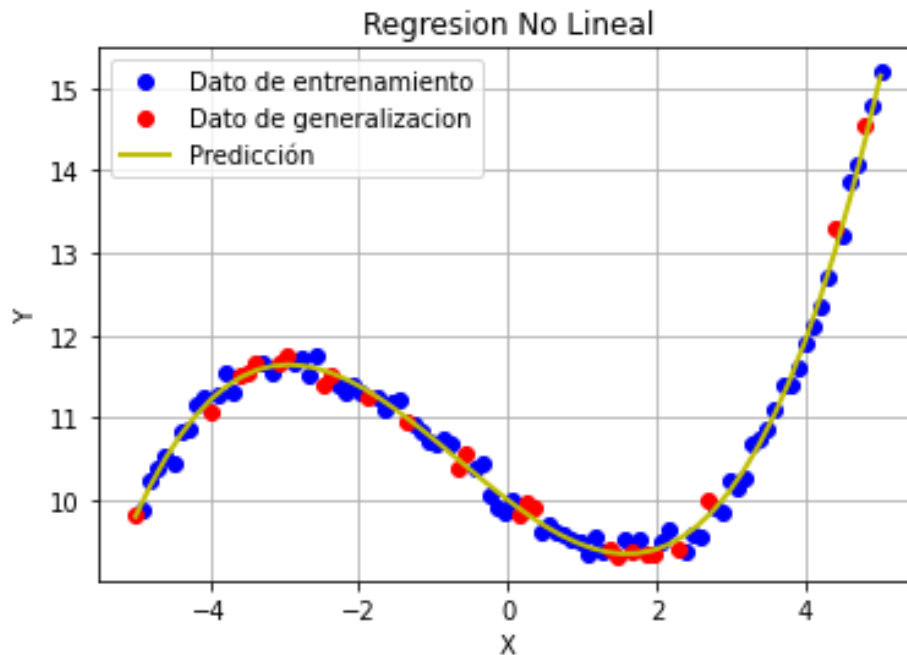


Figura 2. Regresión lineal de los datos del archivo adjunto 2.

Los resultados de las métricas de coeficiente R^2 fueron:

- Coeficiente R^2 para los datos de entrenamiento

$$R^2 = 0.9925681948977977$$

- Coeficiente R^2 para los datos de generalización

$$R^2 = 0.9924832796871269$$

El resultado de la regresión polinómica para los datos del archivo adjunto “df_regresion_lineal_3.csv” se muestra en la Figura 3.

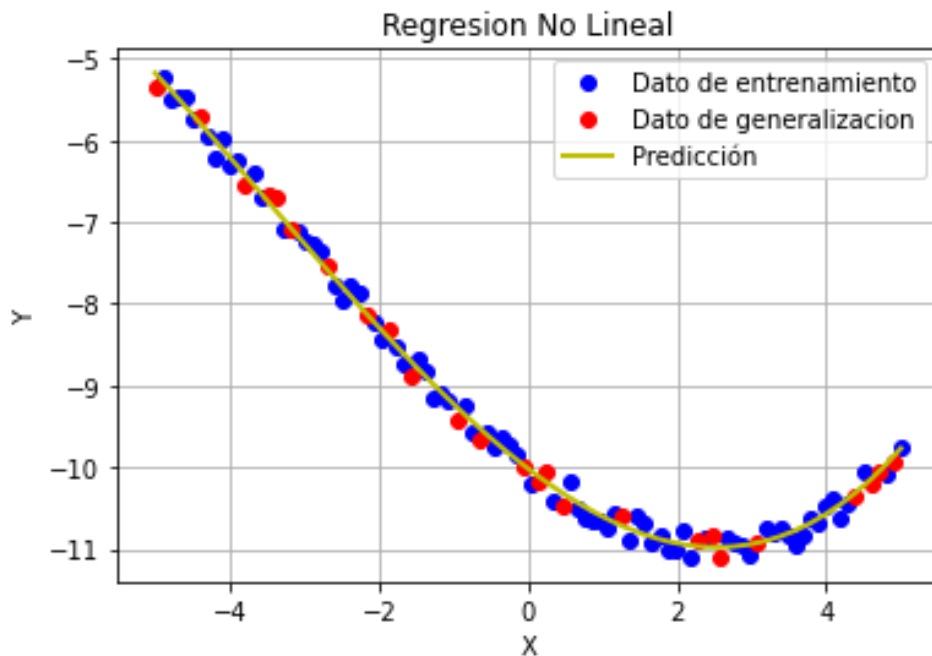


Figura 3. Regresión lineal de los datos del archivo adjunto 3.

Los resultados de las métricas de coeficiente R^2 fueron:

- Coeficiente R^2 para los datos de entrenamiento

$$R^2 = 0.9961683940311234$$

- Coeficiente R^2 para los datos de generalización

$$R^2 = 0.9964611890046697$$

El resultado de la regresión polinómica para los datos del archivo adjunto “df_regresion_lineal_temp.csv” se muestra en la Figura 4.

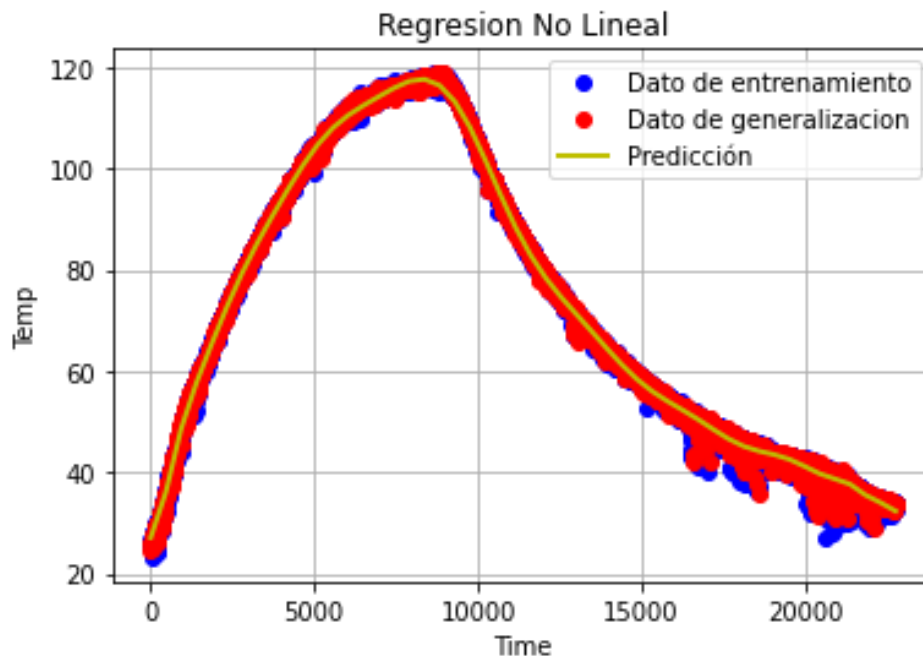


Figura 4. Regresión lineal de los datos del archivo adjunto.

Los resultados de las métricas de coeficiente R^2 fueron:

- Coeficiente R^2 para los datos de entrenamiento

$$R^2 = 0.9988365002977483$$

- Coeficiente R^2 para los datos de generalización

$$R^2 = 0.9988875858144616$$

Conclusión

Para los conjuntos de datos correspondientes a la parte 1 de la actividad no hubo mayor problema y bastó con un modelo de regresión lineal de grado 3 sencillo. Para el conjunto de datos de la segunda parte de la actividad al principio se buscó usar un modelo de regresión regularizado, sin embargo, al buscar valores para el parámetro alfa se llegó a la conclusión de que el mejor ajuste ocurría cuando alfa era cero y el grado del polinomio era relativamente alto, así que al final se optó por un modelo no regularizado. En cuanto a la comparación entre las métricas de coeficiente R^2 de datos de entrenamiento y generalización, se buscó que fueran los mejores posibles.