



Tecnológico de Monterrey

Inteligencia Artificial Avanzada para la Ciencia de Datos

Análisis y reporte sobre el desempeño del modelo

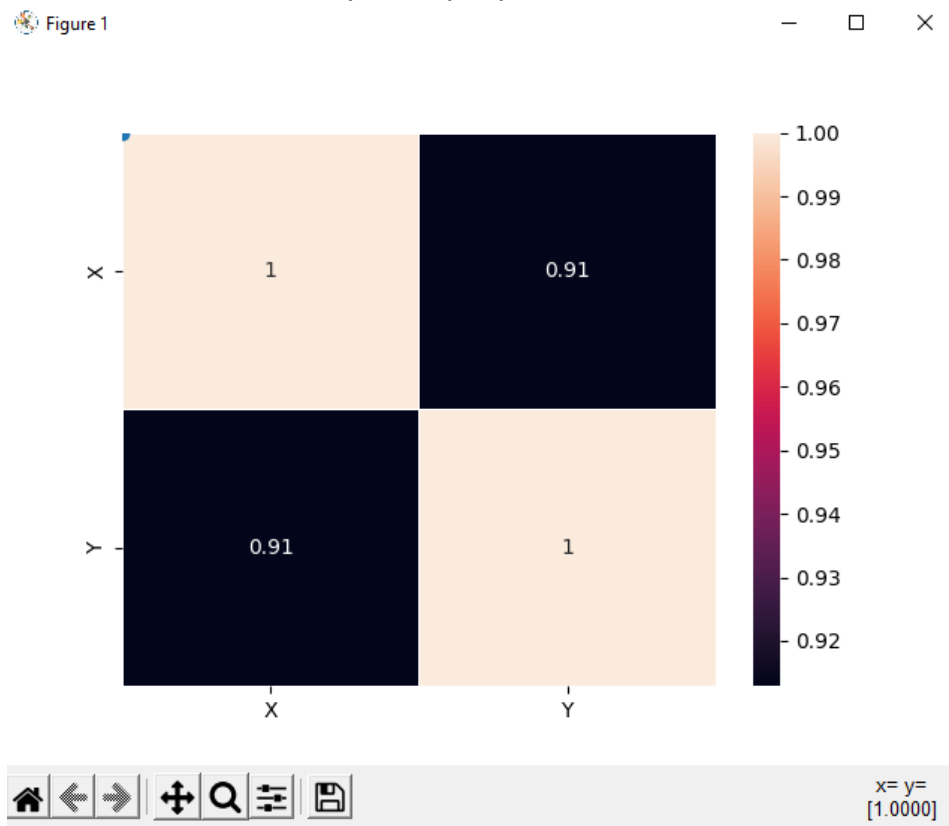
Renata de Luna Flores

A01750484

Profesor: Jorge Ramírez Uresti

Utilizando el modelo implementado a mano con un nuevo dataset, probamos y validamos el modelo.

En primer lugar, obtenemos un heatmap para poder saber si las variables sí están relacionadas entre sí. Con el heatmap generado, podemos ver que las variables tienen una alta correlación, por lo que podemos crear el modelo con estas.



Realizamos la separación de los datos en Train/Test/Validation

```
X_train, X_test, y_train, y_test = train_test_split(df_x, df_y,
    test_size=0.2, shuffle = True, random_state = 8)

X_train, X_val, y_train, y_val = train_test_split(X_train, y_train,
    test_size=0.25, random_state= 8)
```

Los datos de train y validación tienen el mismo tamaño para que se pueda trabajar con ellos correctamente.

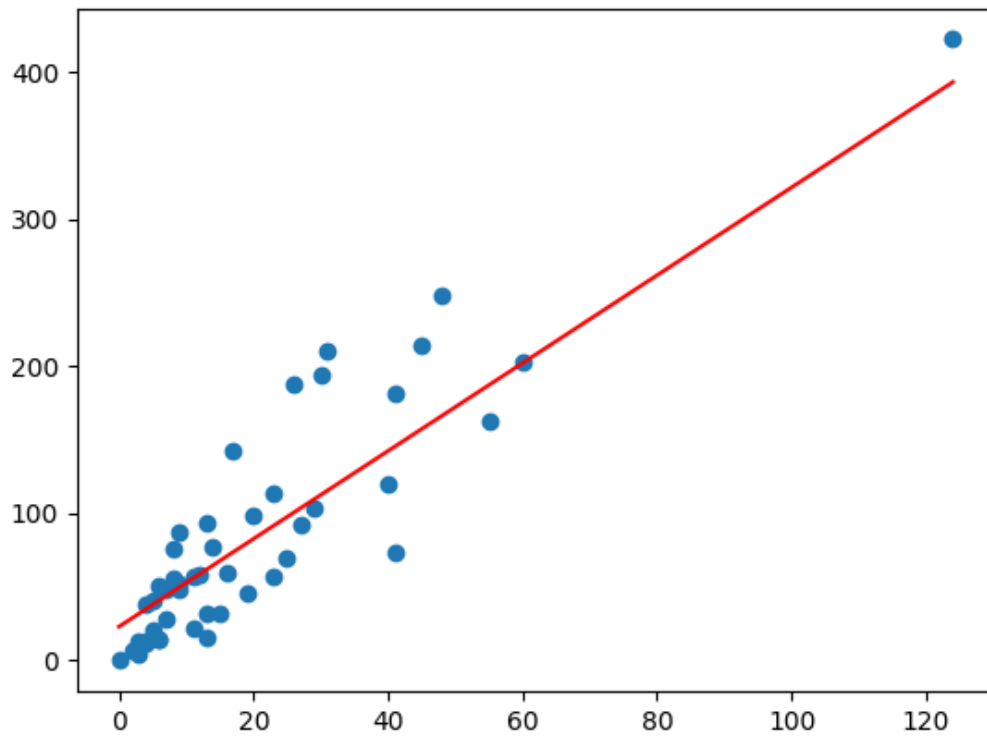
Ejemplo del modelo generado con train y predicciones con test.

Learning rate = 0.00001

Epochs = 100

Figure 1

— □ ×



Example with test data

```
1      81.580169
28     40.347878
59     122.812460
30     16.295708
12     54.091975
31     102.196314
25     40.347878
62     105.632338
55     115.940411
32     36.911853
23     26.603781
36     225.893187
16     98.760290
```

```
1      46.2
28     27.9
59     209.8
30      0.0
12     23.5
31     69.2
25     77.5
62     187.5
55     133.3
32     14.6
23     13.2
16     134.9
```

```
Name: Y, dtype: float64
MAE: 431.22748727289967
RMSE: 147.38919578334797
MAPE: inf
SSE: 21699.999162336742
Model's R^2: 0.7086602824568334
```

Como podemos ver con los resultados, teniendo un learning rate bajo y pocas epochs, el modelo no tiene un valor de R^2 muy alto, por lo que el bias y la varianza son medias. El error con estos hiper parámetros son más altos. A pesar de esto, podemos decir que el ajuste del modelo es bueno, ya que no caemos en un overfitting ni en underfitting.

Ajuste del modelo: fitt

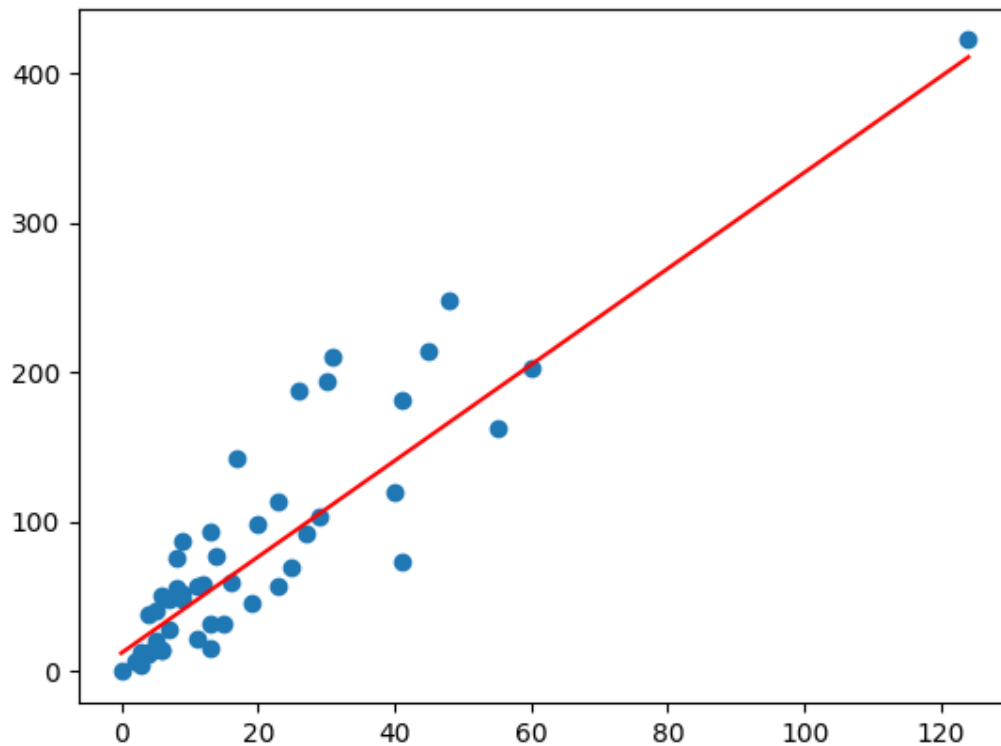
Sesgo: medio

Varianza: baja

Ejemplo 2 con con los datos de validación

Learning rate = 0.00006

Epochs = 500



Example 1

```
61 202.002920
57 91.941858
0 410.739418
43 141.279576
5 217.183757
36 232.364593
16 91.941858
12 42.604140
25 27.423304
60 53.989768
55 110.917903
9 38.808931
40 50.194558
13 88.146649
34 84.351440
49 12.242468
```

```
61 244.6
57 137.9
0 392.5
43 152.8
5 170.9
36 217.6
16 134.9
12 23.5
25 77.5
60 95.5
55 133.3
9 65.3
40 89.9
13 39.6
34 161.5
49 39.9
```

```
MAE: 574.9439719639939
RMSE: 158.26009120389966
MAPE: 6.592211207099577
SSE: 25046.256467866635
Model's R^2: 0.8070067820643232
```

En este caso, con un learning rate y épocas más grandes, el valor de R^2 aumenta, por lo que el sesgo y la varianza disminuyen. Como podemos ver en la gráfica, la mayoría de los valores predichos está muy cercanos a los valores originales, lo que también nos muestra una varianza y bias bajo.

Ajuste del modelo: `fitt`

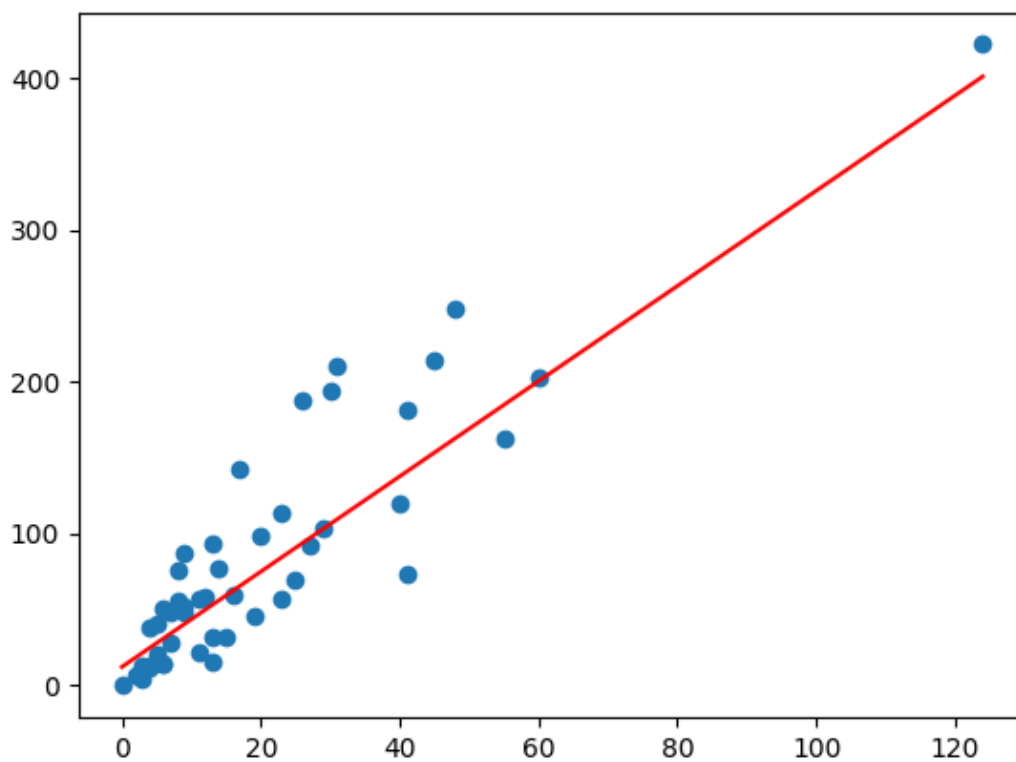
Sesgo: bajo

Varianza: baja

Ejemplo 3

Learning rate = 0.000001

Epochs = 1000



61	244.6	61	199.203490
57	137.9	57	96.973987
0	392.5	0	393.087031
43	152.8	43	142.801006
5	170.9	5	213.304111
36	217.6	36	227.404733
16	134.9	16	96.973987
12	23.5	12	51.146968
25	77.5	25	37.046347
60	95.5	60	61.722434
55	133.3	55	114.599763
9	65.3	9	47.621813
40	89.9	40	58.197279
13	39.6	13	93.448832
34	161.5	34	89.923676
49	39.9	49	22.945726

MAE: 499.38216794416223
RMSE: 143.83358730397222
MAPE: 6.167361962278909
SSE: 20688.1008367294
Model's R^2: 0.8405884265147326

Ajuste del modelo: fitt

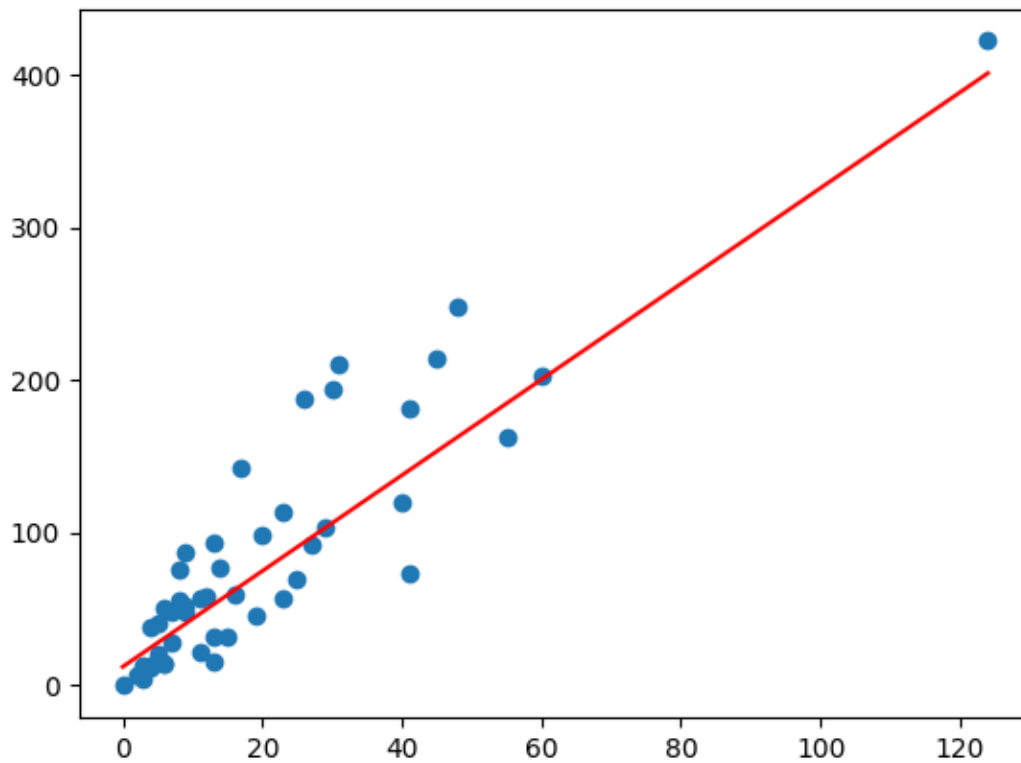
Sesgo: bajo

Varianza: baja

Ejemplo 4

Learning rate = 0.000008

Epochs = 1500



61	202.002921	61	244.6
57	91.941856	57	137.9
0	410.739423	0	392.5
43	141.279575	43	152.8
5	217.183758	5	170.9
36	232.364594	36	217.6
16	91.941856	16	134.9
12	42.604138	12	23.5
25	27.423301	25	77.5
60	53.989765	60	95.5
55	110.917902	55	133.3
9	38.808929	9	65.3
40	50.194556	40	89.9
13	88.146647	13	39.6
34	84.351438	34	161.5
49	12.242465	49	39.9

MAE: 574.9439952819235
 RMSE: 158.26009639514274
 MAPE: 6.592211338205563
 SSE: 25046.25811099987
 Model's R^2: 0.8070067694032068

Ajuste del modelo: `fitt`

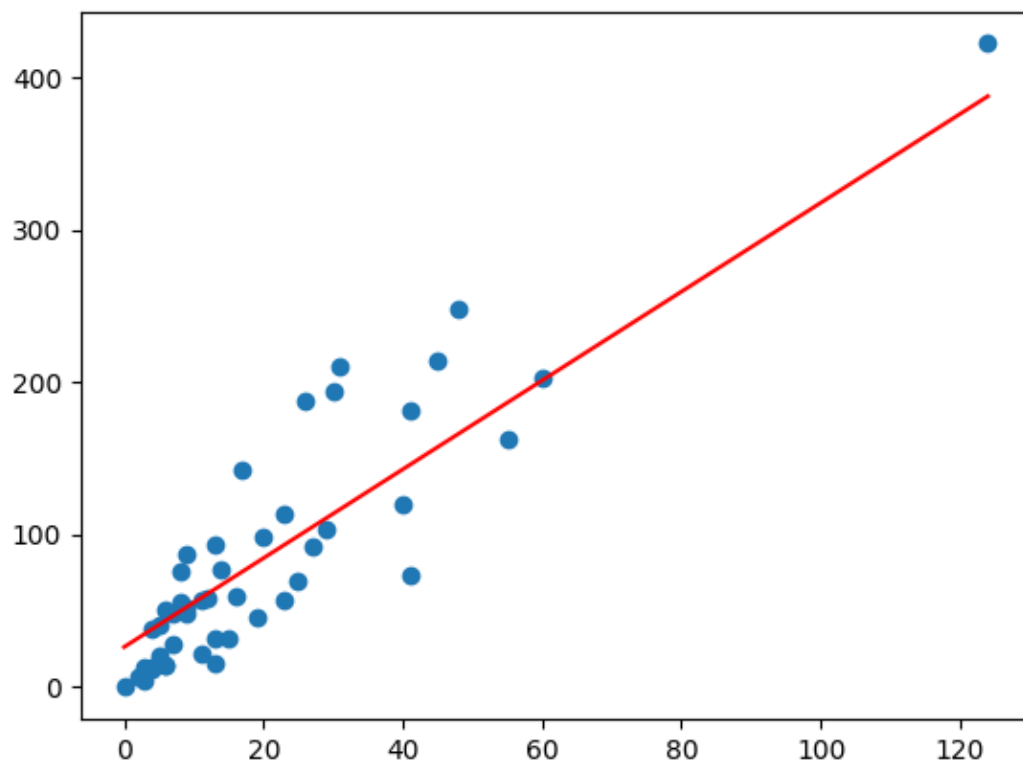
Sesgo: bajo

Varianza: baja

Ejemplo 5

Learning rate = 0.001

Epochs = 2000



61	244.6	61	244.6
57	137.9	57	137.9
0	392.5	0	392.5
43	152.8	43	152.8
5	170.9	5	170.9
36	217.6	36	217.6
16	134.9	16	134.9
12	23.5	12	23.5
25	77.5	25	77.5
60	95.5	60	95.5
55	133.3	55	133.3
9	65.3	9	65.3
40	89.9	40	89.9
13	39.6	13	39.6
34	161.5	34	161.5
49	39.9	49	39.9

```
MAE: 485.6713117699222
RMSE: 140.4477702929609
MAPE: 6.060076341310239
SSE: 19725.576180264314
Model's R^2: 0.8480051329208158
```

Ajuste del modelo: fitt

Sesgo: bajo

Varianza: baja

Este último modelo, con un learning rate superior a los anteriores y con un mayor número de épocas, es el modelo que tiene una mejor precisión. El valor de R^2 es el más alto a lo largo de las variaciones de los hiper parámetros.

Bias o sesgo

Tomando en cuenta que el sesgo mide qué tan lejos está el valor estimado de la población, podemos decir que el modelo tiene un sesgo bajo, ya que los valores que el modelo predice se encuentran muy cerca de los valores de la población original. A pesar de que el bias cambia cuando se modifican los hiper parámetros, el bias en cada una de las iteraciones es bajo, ya que los valores predichos por el modelo son muy cercanos a los valores dependientes del dataset. El bias del modelo es bajo.

Variance

Al dividir el set de datos en train/test/validation, los datos son diferentes, sin embargo, podemos decir que la varianza es pequeña, ya que los resultados y errores no varían mucho al cambiar de set de datos con el que se realizan las pruebas. Las pruebas con el dataset de test se realizan correctamente. El modelo generado puede predecir los valores a pesar de que los valores de las variables

independientes o de entrada sean diferentes a las variables con las que se realizó el modelo. Esto puede ser explicado con R^2 : al ser este valor alto, la mayoría de los valores que se quieren predecir pueden ser predichos con la ayuda de los valores de las variables independientes del modelo. Los valores predichos no varían mucho entre sí, por lo que la varianza es baja.

Ajuste del modelo

El modelo al no presentar errores muy altos, teniendo un bias y varianza baja, podemos decir que el modelo no se encuentra ni en overfitting ni en underfitting. Si tuviéramos uno de estos errores, el modelo generaría muchos errores al cambiar de set de datos (en el caso de overfitting), ya que hubiera solo aprendido los valores en lugar de generar patrones, o por otro lado, generaría errores más altos ya que no aprendió lo suficiente para poder generar patrones correctos. Podemos ver que la mayoría de los valores predichos son muy similares a los valores de salida del dataset, a pesar de que hay algunos que tienen un bias más alto. Sin embargo, al ser pocos estos valores, podríamos incluso decir que se trata de valores atípicos.

Regularización para el desempeño del modelo

Para mejorar el desempeño del modelo, podemos modificar los hiper parámetros para lograr que este sea más preciso. Sin duda, los dos hiper parámetros de suma importancia en nuestro modelo son el learning rate y el número de epochs. Al analizar los resultados variando estos hiper parámetros, podemos ver que al ser mayor el número de epochs la precisión del modelo mejora, ya que realiza más iteraciones para aprender y generar patrones con los datos. Por otro lado, es importante encontrar un valor del learning rate que permita que las predicciones sean más precisas. Como lo vimos con las distintas iteraciones realizadas, un valor muy pequeño de learning rate puede afectar nuestro modelo y provocar que tenga predicciones menos precisas. En este caso, el learning rate con el que tuvimos una mejor precisión fue con 0.001, y con valores más pequeños se generaron más errores en las predicciones.