

Al final se realizaron las siguientes pruebas para conocer cómo cambiaba el desempeño del modelo al transformar los parámetros que tiene el modelo. En el primero se escogió que el tamaño del conjunto de datos de entrenamiento tuviese el 70 por ciento del total de datos. Se puede ver que los resultados no son muy buenos con este valor.

```
1 # Prueba 1
2 train_X, train_Y, test_X, test_Y = split_dataset(70, datos) # División de dataset de prueba y entrenamiento
3 test_y, resultado = regresion(test_X, test_Y, fit(train_X, train_Y)) # Predicción de y del dataset de prueba
4 resultado, MSE, MAE, RMSE, MAPE, SSE, R2 = evaluacion(resultado)
5 resultado
```

Mean Squared Error: 0.4784252664607663
Mean Absolute Error: 0.5471487170441045
Root Mean Squared Error 0.6916829233548898
Mean Percent Absolute Error 0.27988871657177916
Sum of Squared Errors 71.76378996911494
Coeficiente de determinación 0.7151342963811769

	income	happiness	Predicción happiness	Error	Error Absoluto	Squared Error
0	4.639914	3.737942	3.531611	0.206330	0.206330	0.042572
1	7.466653	5.960547	5.521604	0.438943	0.438943	0.192671
2	3.528319	2.546525	2.749060	-0.202536	0.202536	0.041021
3	2.428752	1.200786	1.974977	-0.774191	0.774191	0.599372
4	3.542748	3.078293	2.759218	0.319075	0.319075	0.101809
...
145	7.448117	5.963422	5.508555	0.454867	0.454867	0.206904
146	7.225192	4.985255	5.351618	-0.366363	0.366363	0.134222
147	4.243873	4.754168	3.252802	1.501366	1.501366	2.254101
148	5.249209	4.568705	3.960548	0.608157	0.608157	0.369855
149	3.440847	2.070664	2.687481	-0.616817	0.616817	0.380463

150 rows × 6 columns

Como segunda prueba el tamaño del conjunto de entrenamiento se escogió más grande y como se puede observar, las métricas de evaluación mejoraron en comparación de la prueba anterior.

```
1 # Prueba 2
2 train_X, train_Y, test_X, test_Y = split_dataset(90, datos) # División de dataset de prueba y entrenamiento
3 test_y, resultado = regresion(test_X, test_Y, fit(train_X, train_Y)) # Predicción de y del dataset de prueba
4 resultado, MSE, MAE, RMSE, MAPE, SSE, R2 = evaluacion(resultado)
5 resultado.head()
```

Mean Squared Error: 0.5413035373328922
Mean Absolute Error: 0.6061574065890772
Root Mean Squared Error 0.7357333330309918
Mean Percent Absolute Error 0.22451272185720014
Sum of Squared Errors 27.06517686664461
Coeficiente de determinación 0.7544831690314232

	income	happiness	Predicción happiness	Error	Error Absoluto	Squared Error
0	4.755680	2.666116	3.591761	-0.925645	0.925645	0.856819
1	1.994275	2.584729	1.620063	0.964666	0.964666	0.930581
2	3.542748	3.078293	2.725703	0.352590	0.352590	0.124320
3	3.065059	3.407980	2.384624	1.023357	1.023357	1.047259
4	5.358716	3.752659	4.022341	-0.269681	0.269681	0.072728

El cambiar el tamaño de los conjuntos de entrenamiento y prueba puede ayudar a mejorar el desempeño de los modelos pues estos tienen más datos de los cuales obtener información. Otra manera de seguir mejorando la evaluación podría ser con un cross-validation para que no entrenara con solamente unos datos, si no que vaya cambiando qué datos son de entrenamiento y cuáles son de prueba.