



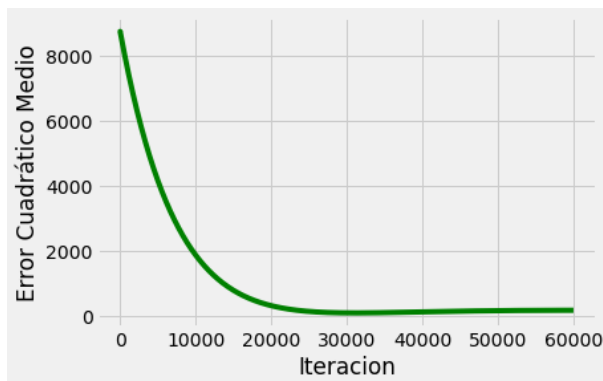
## Momento de Retroalimentación: Módulo 2 Implementación de una técnica de aprendizaje máquina sin el uso de un framework

La implementación realizada es una regresión lineal simple de 2 variables numéricas, el set de datos utilizado es una relación de edad y presión sanguínea.

Los parámetros iniciales plantean un set de entrenamiento con el 80% de los datos, con un 20% para pruebas. Tiene un learning rate de 0.0002 (alpha), realiza 60000 iteraciones, a menos que antes logre una diferencia de error de 0.0000001.

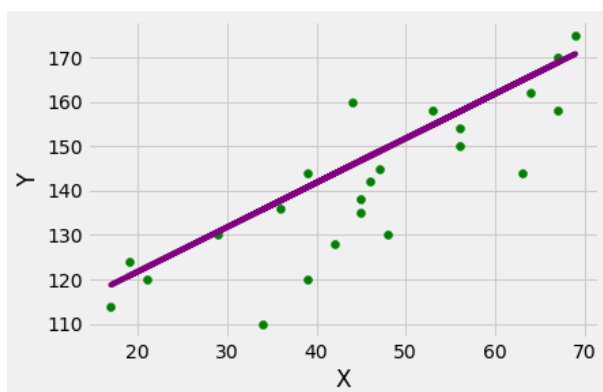
El set de datos es pequeño, por lo que los resultados del modelo para los valores de la regresión dependen de la selección de la parte de train.

En una primera iteración, obtenemos los siguientes resultados:

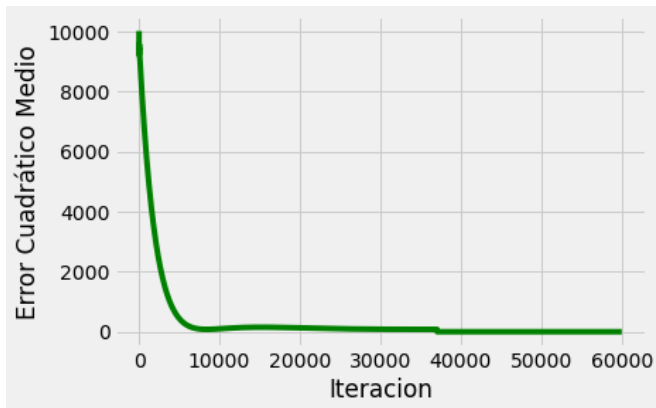


Se utilizan todas las iteraciones para reducir el error y en la gráfica de la regresión podemos ver su ajuste a los valores, el error se redujo de 7565.673 en la iteración 1000 a 172.0018 en la 60000. Que si bien no es cero, si redujo el error a 1/43 partes aproximadamente del inicial.

Iteración 1000	Iteración 60000
m: 2.8 b: 3.1	m: 1.4 b: 73.3
error: 7565.673	error: 172.018



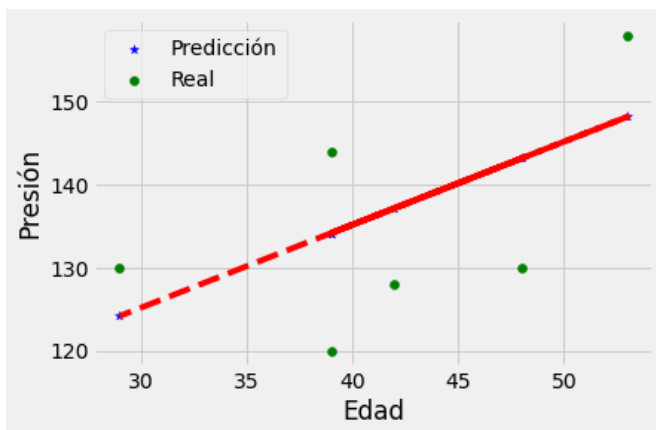
Es posible que con más iteraciones el error continúe disminuyendo y eventualmente llegue a la limitación de error establecida, pero otra forma de lograrlo es aumentando el 'learning rate', en tal caso es posible que con la cantidad actual de iteraciones logre mejores resultados antes, como podemos ver en la siguiente corrida. (Los resultados varían debido a que se vuelve a partir el dataset con datos diferentes).



En esta gráfica podemos observar cómo al aumentar el valor de  $\alpha$  de 0.0002 a 0.0004, es decir, duplicado, con alrededor de 36000 iteraciones logra converger y detiene el cálculo del gradiente para la regresión. También podemos notar que el error disminuye de 5599.7 a 78.559, es decir a 1/71 del valor inicial.

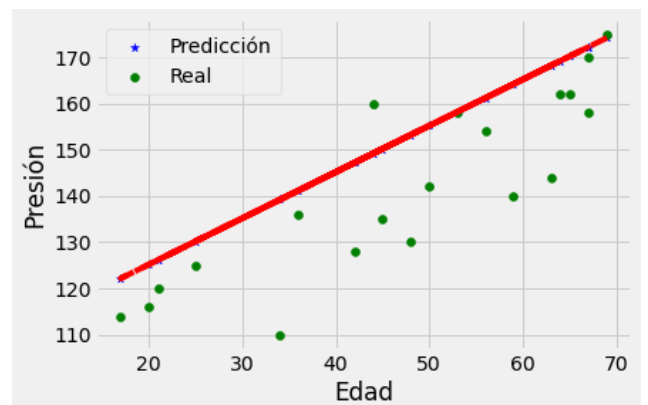
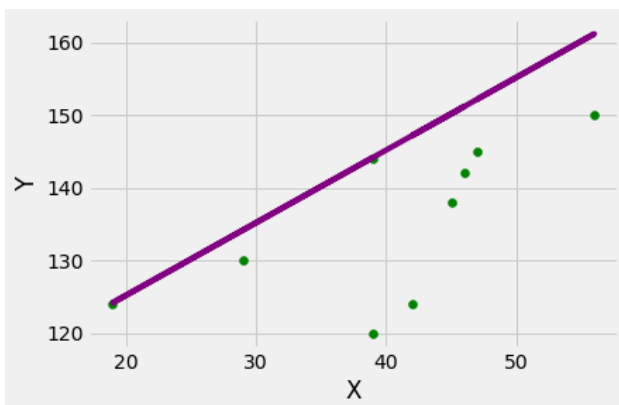
Iteración 1000  
m: 2.7 b: 7.8  
error: 5599.700

Iteración 36000  
m: 1.0 b: 94.1  
error: 78.559



Aún así podemos ver que en las predicciones hay una variación al valor real, y esto deriva de que el ajuste nunca va a ser exacto a todos los datos, ya que como vemos en la gráfica, hay valores de presión distintos para una misma edad.

Otra variación que podemos hacer para mejorar es variar la cantidad de los datos de prueba y entrenamiento, si proporcionamos menos datos de entrenamiento que de prueba la variación y bias del modelo son muy altos y el error de los datos de igual forma es un valor elevado, de ahí que no sea idóneo hacer una partición de este tipo y menos aún con un set con pocos datos. He aquí un ejemplo de 30/70 (30 entrenamiento, 70 prueba).



Finalmente si aceptamos un error mayor y pasamos de 0.0000001 a 0.0001, las iteraciones se reducen ya que logra llegar a esa diferencia antes, lo que pasa con este cambio es que estamos aceptando un margen más alto de variación en el modelo, pero manteniendo control sobre el error.

