Regressione lineare   
Utilizzando il Boston Housing Dataset per predirre il valore di abitazioni nella zona di Boston.

|  |  |
| --- | --- |
| Metri quadri | Valore € |
| 80 | 120k |
| 150 | 300k |
| 30 | 100k |
| 20 | 70k |
| 70 | ? |
| 160 | ? |

Metri quadri = proprietà (X)

Valore = target (Y)

Dobbiamo trovare la funzione che mappa la relazione fra X e Y.

La regressione consiste nel tracciare una retta il più vicino possibile a tutti i punti di un grafico cartesiano (scatter.plot)

La retta si rappresenta con l’equazione y=q+mx

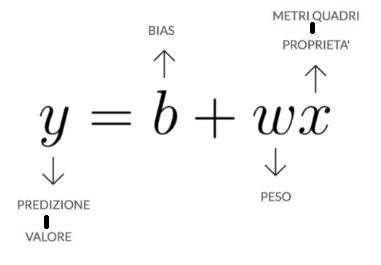


**Q** ed **M** sono i valori che bisogna ottimare per trovare la retta che meglio si adatta ad X e Y (proprietà e Target)

Q= (intercetta / Bias) Un aumento (Q+) della intercetta sposta la retta verso l’alto, una diminuzione (Q-)la sposta verso il basso

Mx =(Coefficiente angolare) Inclinazione della retta Q , la velocità con cui Y cresce al variare di X.

Un aumento di MQ porterà un aumento maggiore del Prezzo   
UnDecremento di MQ porterà ad una diminuzione del Prezzo (riduce inclinazione della retta)

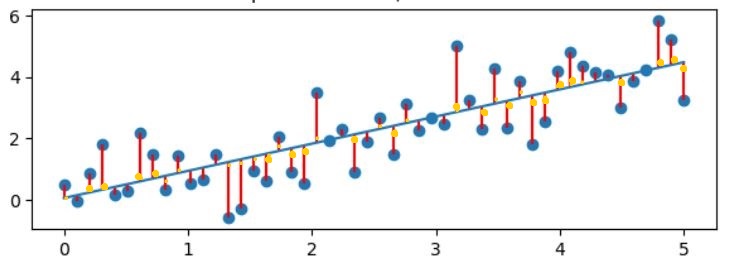


Y= valore(prezzo)  
x= MQ

------------------------------------   
Qualità di un modello

Funzione di costo : misura la qualità di un modello di Machine Learning  
 la più comune per la regressione è il metodo della Somma dei Quadrati Residui (RSS)

Sommare ogni errore del nostro modello per ogni dato disponibile.



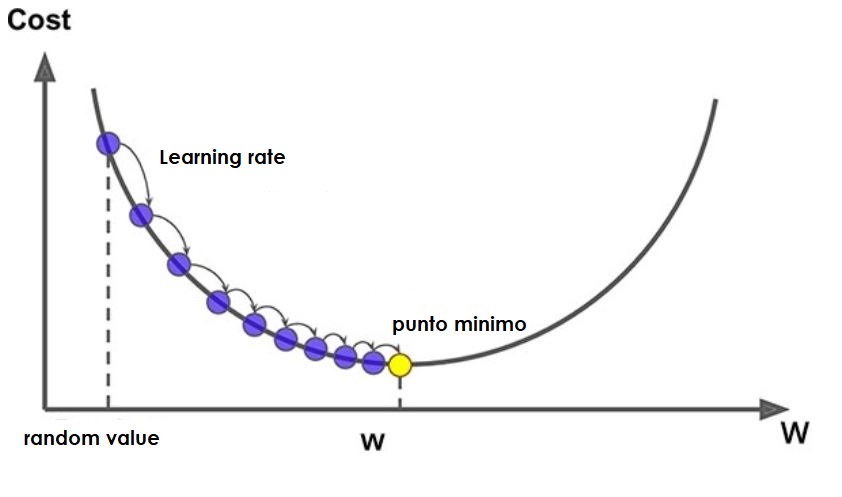
F(x)= bias+w(peso)x

Somma dei quadrati residui = somma degli errori al quadrato per ogni punto del nostro dataset  
 Un errore è la differenza fra valore corretto (pallini blu) ed   
 il valore predetto (pallini gialli)

MINORE è il risultato della funzione di costo, MAGGIORE è la qualità del nostro modello.  
 (il nostro modello ne sbaglia poche significa)

Algoritmo Gradiant Descent ( famiglia degli algoritmi di ottimizzazione )   
Serve a trovare MAX o MIN in una funzione.

Come trovare il punto di minimo? In cui la funzione assume il valore minore (punto giallo)  
Abbiamo un valore di peso casuale (w=random value) e dobbiamo spostarlo giù per la curva



Aggiornare il valore del peso (detto learning rate tra 0.0001 e 1)  
Il numero di cicli massimi viene impostato ad un valore predefinito dette EPOCHE,  
siamo noi ad impostare quante Epoche/salti dovrà fare.