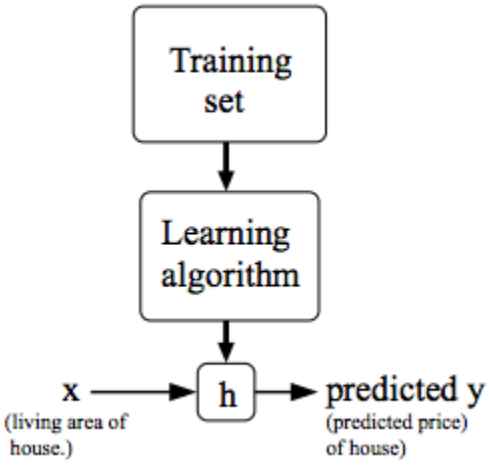
Linear Regression using Single Variable / Univariant Linear Regression :

In this we will have data set (training set) to build a model that find the relation between our input variable, X and output Variable y, so that our model predicts the value of y for unseen value of X.

Motivation Example : To predict housing price based on the size in sq. Ft.

Size in sq. Ft (X)	Price($) in 1000's (y)
2104	**460**
1416	**232**
1534	**315**
852	**178**
……..	**…….**
……..	**…….**

Model Representation:



Notations :

X → training data set

Y → output

m → size of our training data set.

$(x^{(i)}, y^{(i)})$ → i^{th} training example.

h → hypothesis (h : X → y).

Here hypothesis is the function that show how X and y values are related to each other.

Representation of h: $\mathbf{h_{theta}(x) = theta0 + theta1*X.}$

Cost Function (J) :

In the above equation of h, the theta0, theta1 are parameters(learning parameters), now the problem is how to choose these parameters.

Idea : Our goal is to have values of the theta0, theta1 such that our $\mathbf{h_{theta}(x)}$ value should be close to the y value (we can check that while we are training).

So, the difference → $\mathbf{h_{theta}(x) - y}$ should be minimum, to achieve this we can use Error functions like Mean Square Error, Root Mean Square error, ……,etc.

Mean Square Error = $J(\theta_0,\theta_1) = \frac{1}{2m} * \sum_{i=1}^{m}(h_{theta}(x^i) - y^i)^2$