DATE: / /
PAGE:

Machine Learning

- What is Machine Learning ?

Study of computer algorithms that improve automatically through experience and by use of data.

- What is data science ?

Turning arbitrary data into valuable one. (can be used to get insights from a data which can be of more value).

- Linear Regression

It is a form of supervised ML technique used when data shows continous relationship.
In general, we have learned that

$$y = mx + c$$

This is an equation for a straight line in graph. In ML this equation is described as

$$h_\theta(x) = \theta_0 + \theta_1 x$$

($h_\theta(x)$ = y, θ_0 = c, θ_1 = m)

$h_\theta(x)$ = Hypothesis
θ_0 = Intercept or constant

In linear data, we can easily plot a best fitted line. However in any other synario like movie box office revenue prediction, we can have multiple line going through the points.

In order to tackle this situation, we can use a method where we change θ_1 and θ_0 to minimize MSE (Mean squared error).

Mathematically,

We can say that for $h_\theta(x) = \theta_0 + \theta_1 x$, we have to choose θ_0 and θ_1 in such a way that it minimize

$$(y^{(1)} - h_\theta(x^{(1)}))^2 + (y^{(2)} - h_\theta(x^{(2)}))^2 + \cdots$$

$$\Rightarrow \sum_{i=1}^{n} (y^{(i)} - h_\theta(x^{(i)}))^2$$

- But why Linear Regression ?

In Machine learning, we have 2 type of learning. Namely supervised and unsupervised learning.

Supervised learning => We use this when we know what all values to predict from the data. For example weather a following image is of a cat or dog, what is movie earning w.r.t its budget, etc.

Various techniques present are linear regression and classification broadly.

Linear Regression => when we have continous data eg movie prediction.

Classification => when our data is discrete, for eg weather a review is fake or not, etc.

DATE : / /
PAGE

we have more advance model in supervised learning such as knn, forest, etc which will come later.

Unsupervised learning => In this methodology, data outcome is not defined rather relational conclusions are drawn.
Eg => if a person buys pen, it is more likely to buy notebooks.

This is used in pattern detection, recomender system, etc. We have methods like KNN, K-means clustering, etc.