Machine Learning

By Team Emfinity

What?

Machine Learning is an algorithm which makes the device capable of learning stuff based on its past experiences and it simultaneously improves the performance of the task.

It analyzes, predicts and sorts huge amounts of data.

There is a given set of data which then is formulated by any required ML algo to do the needful and then formulated

ML

SUPERVISED LEARNING

1) Linear regression
2) Logistic regression
3) Neural Networks

UNSUPERVISED LEARNING

1) Support vector machine (SVM)
2) K nearest neighbors

LINEAR REGRESSION

As mentioned earlier it comes under supervised learning i.e it is a ML algo that has labelled dataset as an input and gives a particular determined value as an output via curve fitting method .

It is commonly used as predictive analysis

When

Linear reg is used when we have independent variables or parameters as our dataset

It has to be specific as to what our dataset should

House pricing example

The dataset given will be for 10000s tuples for a fairly accurate result

Maths involved

Choose your constants so that h(x) is close to y for our training examples(x,y).

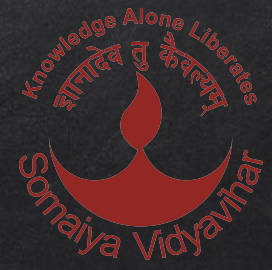
Therefore we have to minimize the constats implies minimize h(x) –y

This value is squared to avoid negative values and then summed up for all tuples present and then divided by the no of training examples

Therefore our final cost func becomes 1/m Σ((h(x)–y)^2)

Grad des is used to minimize the cost funcn J(Θ)

Start with some Θ and then keep on changing it to reduce J(Θ) until u hopefully reach a minima



Title : Logistic Reg.

By Team Emfinity

Why and when?

- Algorithms like linear regression are used to predict a specific value of a given problem
- But when we need to classify something (for eg. true or false,cat or not cat,etc) or distinguish between various different things (for eg given image of all round objects identify which ones are of a Pizza or ball or just random circle) linear regression is not useful as we don't have a specific value to calculate.
- In such cases where we have to classify or distinguish between given data we use logistic regression

What?

- Logistic regression is the appropriate regression analysis to conduct when the dependent variable is dichotomous (binary).
- Like all regression analyses, the logistic regression is a predictive analysis.
- Logistic regression is used to describe data and to explain the relationship between one dependent binary variable and one or more nominal, ordinal, interval or ratio-level independent variables.

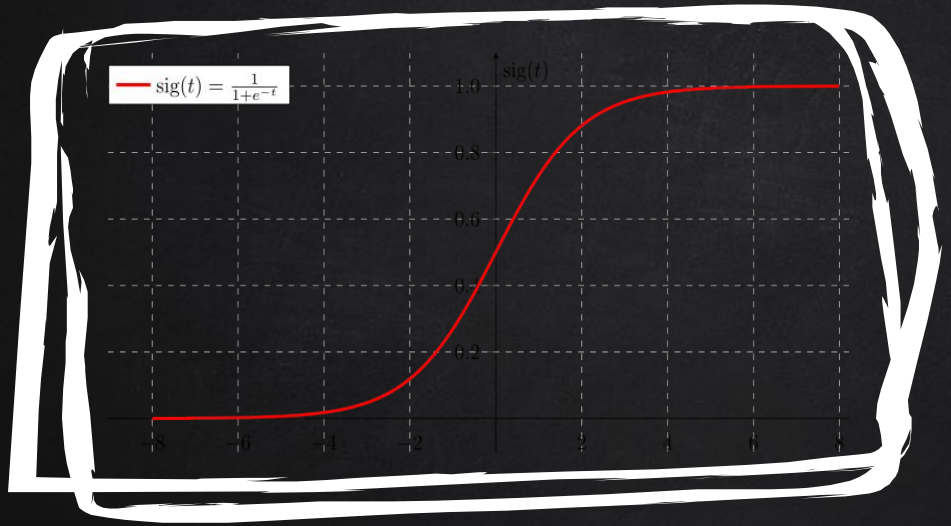
Dos and don'ts

- The dependent variable should be dichotomous in nature (e.g., presence vs. absence).
- There should be no high correlations among the predictors.
- When selecting the model for the logistic regression analysis, another important consideration is the model fit. Adding independent variables to a logistic regression model will always increase the amount of variance.However, adding more and more variables to the model can result in overfitting, which reduces the generalizability of the model beyond the data on which the model is fit.

How?

- Given input features x give $h(x) = P(y=1 | x)$ where h is hypothesis function.
- But here comes the change in definition of h(x).
- In linear regression we had to calculate a specific value os the output h(x) could be $-\infty \leq h(x) \leq \infty$.
- But here we want to predict a value of a dichotomous variable ,i.e., binary 1s and 0s. So $0 \leq h(x) \leq 1$.
- Hence we use a new function known as sigmoid function

Sigmoid Function



A sigmoid function

$\sigma(x) = 1 / (1 + e^{-x})$.

$-\infty < x < \infty$ and $0 \leq y \leq 1$

$\sigma(0) = 1/(1+e^{-0}) = 1/(1+1) = 0.5$

$\sigma(\infty) = 1/(1+e^{-\infty}) = 1/(1+0) = 1$

$\sigma(-\infty) = 1/(1+e^{-(-\infty)}) = 1/(1+\infty) = 0$

How?

- $h(x) = \sigma(w^T x + b)$.
- Here $w^T x + b$ makes the model where w, x and b are all vectors, w and b are parameters of our model and x is a vector of the input features.
- $L(h,y) = -[y\log(h) + (1-y)\log(1-h)]$ => Loss function for one element
- $J(w,b) = -(1/m)\Sigma[y\log(h) + (1-y)\log(1-h)]$ => cost function of model
- Minimization of cost function J from given dataset(also known as knowledge in terms of AI) using Gradient descent algorithm => training step.
- After training the output given by h will be between 0 and 1. We set a threshold value on the basis of which we decide if it is true or false for true if > 0.5 and the numeric value of h is the confidence of our ans.

For eg.

- Consider that we want to make a model that can predict if a patient having known his size of cancel cells and no. of days since infection or some similar independent data variable predict if his cancer is fatal or bening.
- Knowledge : Record of let's say 10000 patients (for a practical application this much no. of data is must).
- Output: Fatal or bening with confidence.
- Decision making criteria : Threshold.

Threshold and it's value

True Prediction	Too Small	Too Large
Fatal	✓	x
Bening	x	✓