MACHINE LEARNING

Machine Learning is the computer program which learns from experience E with respect to task T and some performance measure P, if its performance on task T, as measured by P, improves with time.

Common mathematics terminology in machine learning:

1. Mean
2. Median
3. Mode
4. Variance
5. Standard Deviation
6. Covariance
7. Correlation

Types of Machine Learning:

1. Supervised Learning.
2. Unsupervised Learning.
3. Reinforcement Learning.

1. Supervised Learning: It is the branch of machine learning which deals with the data in which the output is real and continuous or discrete values.
   Based on there output or dependent variable supervised learning is classified into two types:
   a. Regression
   b. Classification

a. Regression: In regression we have one or more independent variable and one dependent variable. The dependent variable is real or continuous.

   Further linear regression is classified into three types:
   - i. Simple linear regression
   - ii. Multiple linear regression
   - iii. Polynomial linear regression

   Simple linear regression has one dependent variable and one independent variable. The dependent variable is real or continuous. It is also called univariate linear regression. In this our goal is given a training set to learn a function h: x = y so that h(x) is the good predictor of y. For some historical reason h is called hypothesis.

Simply linear regression can be represented as y = mx + c.

m is slope and c is intercept but in more elaborative both of them are considered as parameters. C can be replaced with β0 and m can be replaced with β1.

Initially we consider some random values for weight or can be found using normal equation. If we find the weight using normal equation then we can find the best fine in one shot but if we initialize random value to weights there will be some error. We calculate the error value using Mean square error. Further to reduce this error update our weights value by iterating it over gradient descent. Gradient descent will find global minima after many iterations in which weights are updated. Gradient descent takes steps based on learning rate.

Same thing continues with multiple linear regression but multiple linear regression has multiple independent variable and one independent variable.

Polynomial regression comes into picture when the dependent variable has polynomial values with respect to previous independent variable or may be our plot of independent vs dependent variable is getting hard to find the best fit line, there we introduce extra feature to support our data and that feature can be containing polynomial values depending on previous features. That feature can contain any mathematical function value but should be elaborated form of its pervious feature.

b. Classification: In classification problem we have one or more independent variable and one dependent variable which have discrete values.
To attempt classification, one method is to use linear regression and map all predictions greater than 0.5 as a 1 and all less than 0.5 as a 0. However, this method doesn't work well because classification is not actually a linear function.
There are two types of classification problem.
   1. Binary classification: The dependent variable has only two discrete values.
   2. Multiple classification: The dependent variable has more than two discrete values.

Let’s talk about Binary classification. Here the hypothesis equation is mapped to sigmoid function. The sigmoid function classifies all the value >= 0.5 as 1 and

< 0.5 as 0. The cost function is calculated for both the cases i.e. when y == 1 or

y ==0. Then both the cost functions are added to find the total cost function but the matter of fact is when y == 0, then the cost function for y(0) is automatically neglected as per the calculation of formula and when y = 1 , then y(1) is automatically neglected as per the calculation of formula.

Later, the cost is reduced by updating parameter values as following the same process linear regression .