***MACHINE LEARNING***

***WEEK1***

WHAT IS MACHINE LEARNING ?

Machine learning is a method of data analysis that automates analytical model building. Using algorithms that iteratively learn from data, machine learning allows computers to find hidden insights without being explicitly programmed where to look.

## WHY IS MACHINE LEARNING IMPORTANT?

Resurging interest in machine learning is due to the same factors that have made [data mining](http://www.sas.com/en_us/insights/analytics/data-mining.html) and Bayesian analysis more popular than ever. Things like growing volumes and varieties of available data, computational processing that is cheaper and more powerful, and affordable data storage.

All of these things mean it's possible to quickly and automatically produce models that can analyze bigger, more complex data and deliver faster, more accurate results – even on a very large scale. And by building precise models, an organization has a better chance of identifying profitable opportunities – or avoiding unknown risks.

MACHINE LEARNING algorithms can be broadly divided into two categories-

1) SUPERVISED LEARNING

In supervised learning, we are given a data set and already know what our correct output should look like, having the idea that there is a relationship between the input and the output.

Supervised learning problems are categorized into "REGRESSION" and "CLASSIFICATION" problems.

a) REGRESSION PROBLEMS - In a regression problem, we are trying to predict results within a continuous output, meaning that we are trying to map input variables to some continuous function.

Example: Given a picture of a person, we have to predict their age on the basis of the given picture.

b)CLASSIFICATION PROBLEMS - In a classification problem, we are instead trying to predict results in a discrete output. In other words, we are trying to map input variables into discrete categories.

Example: Given a patient with a tumor, we have to predict whether the tumor is malignant or benign.

2) UNSUPERVISED LEARNING

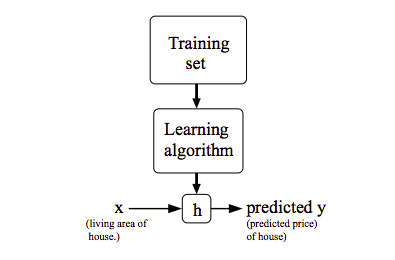
“ Here is a data set. Can u find some structure in the data.....”

-Unsupervised Learning

Unsupervised learning allows us to approach problems with little or no idea what our results should look like. We can derive structure from data where we don't necessarily know the effect of the variables.

Example: Take a collection of 1,000,000 different genes, and find a way to automatically group these genes into groups that are somehow similar or related by different variables, such as lifespan, location, roles, and so on.

MODEL REPRESENTATION



Here x is the input variable and y is the output or target variable.

A pair (x superscript(i), y superscript(i)) is called a training example.

Training Set : Set of training examples.

H in the above figure means hypothesis.

It is a good predictor of corresponding value of y.

**COST FUNCTION**

We can measure the accuracy of our hypothesis function by using a cost function. This takes an average difference of all the results of the hypothesis with inputs from x's and the actual output y's.

This function is otherwise called the "Squared error function", or "Mean squared error".

**GRADIENT DESCENT**

## Gradient Descent

Gradient descent is an optimization algorithm used to find the values of parameters (coefficients) of a function (f) that minimizes a cost function (cost).

Gradient descent is best used when the parameters cannot be calculated analytically (e.g. using linear algebra) and must be searched for by an optimization algorithm.

### ***Intuition for Gradient Descent***

Think of a cereal bowl. This bowl is a plot of the cost function (f).A random position on the surface of the bowl is the cost of the current values of the coefficients (cost).The bottom of the bowl is the cost of the best set of coefficients, the minimum of the function.The goal is to continue to try different values for the coefficients, evaluate their cost and select new coefficients that have a slightly better (lower) cost.Repeating this process enough times will lead to the bottom of the bowl and you will know the values of the coefficients that result in the minimum cost.

Its sometimes called Batch Gradient Descent because the training set involves the whole

batch of training examples.

***LINEAR REGRESSION***

Linear regression is an approach for modeling the relationship between a scalar dependent variable y and one or more explanatory variables (or independent variables) denoted X. The case of one explanatory variable (independent variable) is called simple linear regression.

**Gradient descent for Linear Regression**

When specifically applied to the case of linear regression, a new form of the gradient descent equation can be derived. We can substitute our actual cost function and our actual hypothesis function and modify the equation.