Unsupervised learning

The data is not labelled as right or wrong

The Idea is to find the structure in the data, may be into 2 clusters **(this is one type of unsupervised algorithm)**

|  |
| --- |
| How exactly is that accomplished? |

Example

This algorithm is used in google news

So basically it clusters similar news into a bunch of groups

Cluster Gene

Given Human Genome, can tell how much of different types of genes one has

Certain individuals do or do not have a specific gene

And then the individuals can be grouped

This is unsupervised because we are not telling in advance if these are type 1 people or etc.

Few other areas where you can use

Social network, market segment, astronomical and computing

We can find which 2 items have similarity and then group them so they can interact or can be served efficiently

|  |
| --- |
| Here is important point of behaviour modification |

Cocktail party algorithm

Here it is possible to cluster one person’s and another person’s voice and separate it out using a single audio recording where both are speaking

|  |
| --- |
| Understand the shit well, and think where else or how else can you use this?  I want to use this to separate the various instruments in a song  But I guess there is already a software for that which is paid  Or maybe use this remove noise from the communication  Or maybe use this into the noise cancelling phones etc. |

The idea is you keep 2 microphones are different distances to the speaker, so that one type of sounds comes more strong on one microphone and other type on other

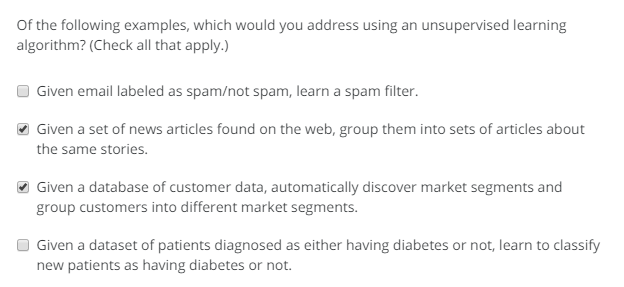
So you can differentiate

Use **Octave** programming environment and implement the algorithms

It seems you can learn much faster if you use this

This guy, who is incredibly smart and yet is able to explain things in such simple terms, is really fascinating.

I guess you need to be really very smart to explain things in very simple terms.



**Linear regression**

Why linear regression is called linear?

In LR the output can be written as

Y = C + Ax

Where C is a constant and A is a parameter and x is predictor

So if you can write anything in this form then it is linear

Non-linear are something like sine, cosine or squared etc.

Y = bo + b1X1 + b2X12

This equation is also linear, since it is of the form constant and product of parameter and predictor

Non-linear are something that does not fit linear model

<http://blog.minitab.com/blog/adventures-in-statistics-2/what-is-the-difference-between-linear-and-nonlinear-equations-in-regression-analysis>

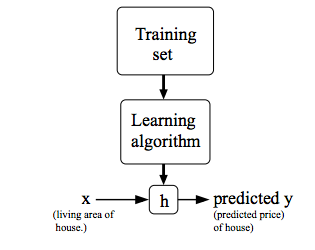
Regression = predicting real-valued output

Classification = predict discrete value output

Training set🡪Algorithm🡪hypothesis (function)

And this function will take the input and produce the output

|  |  |
| --- | --- |
| Hypothesis  Suggested explanation for a phenomenon or prediction of a phenomena  Not well tested | Theory  Well substantiated explanation for a proven hypothesis  Well tested and substantiated |



Formal definition:

Given a training set, learn the function h: x🡪y, such that h(x) is a good predictor of corresponding value of y

How to do a minimization problem

How to know if you have understood the concept by solving the problems

That is the only way

Or apply that concept somewhere else

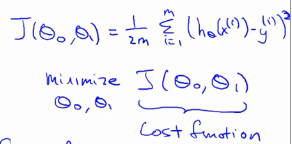
**Cost Function**

For the training set, how do you come up with the values for parameters such it corresponds to a straight line that fits the values?

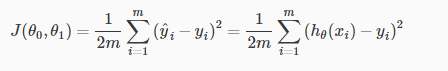
Another way to say is, get the parameter values for the best fit line for the data

Our prediction h(x) should be close to actual y for the training examples

Minimize the J is the cost function and it is also called squared error function



In other words, the accuracy of our hypothesis is computed using the cost function



Half of mean is taken for the convenience and it eventually gets cancelled out when taken the derivative