**Support vector machine- (using so mush algebraic multiplication)**

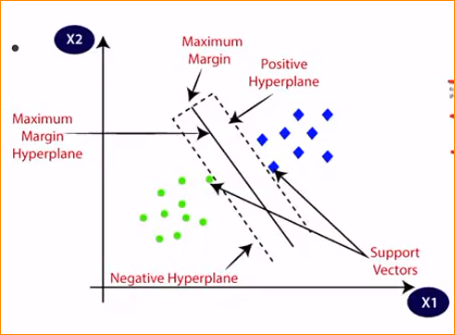
Using for the classification probelm

In here we are going to find the hyper plate or margin, that observation going to class 01, class 02 or multiclass (more than 2 class labels- difficulty of exam [hard easy or moderate]).

Example –

Data set – 2 input features (height, weight) base one these tow data predict the person has Knee Injury or not.

Draw a graph with hight(X axis ) weight (Y axis )



Blue point – Yes(Have injury)

Green point – No(Haven’t injury)

So in here you have to create a decision line (straight line) – Hyperplane

In the left side of the hyperplane No injury right side Have injury.

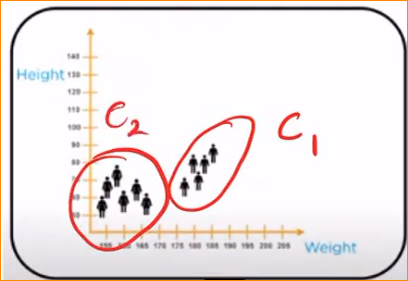
Lets we have new observation – 150 height and 60 weight (large Green dot) – this mean No injury

Hyper plane – Some time it line (2D), Some time it plane.

So this SVM call a ‘Large margin classifier’ as it tries to maximize the margin between the positive and negative support vectors.

So the Objective – **Fine the Hyper plane**

Example –



In this example we have to find out which is the suitable straight line, which can be separate these two classes.

If there is no rule mean we can create infinite lines-



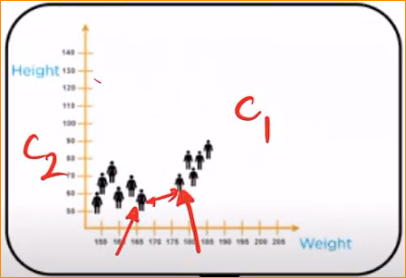
So how to fine most suitable line ?

In here the **support Vector coming to play.**

How to draw a hyper plane

1. Identify the support vectors -

Support vector – is simplify observation or data point in your plot or vector space, which are closer to each other, but they are two observation in two different classes.



Support vector in Class 01

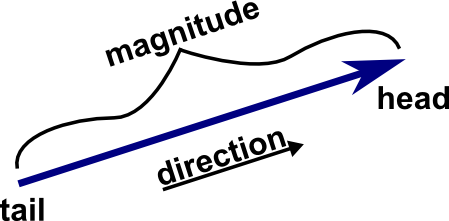
Support vector in Class 02

Shorted length.

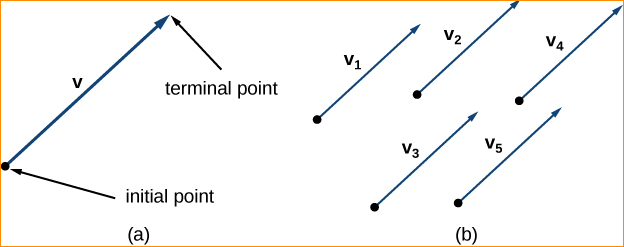
Why call vector for the observation –‘

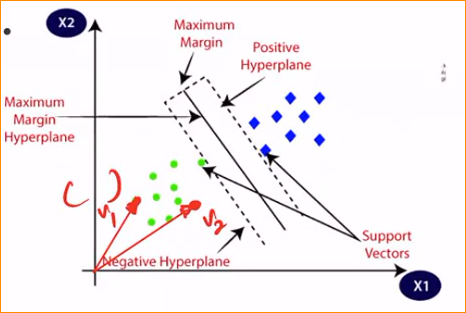
**Note -**

vectors representation



In the plot we can represent as a vector –





Like that we can represent – that why we call vectors – in here vectors mean observations

**Close two vectors in two classes each other call – support vectors.**

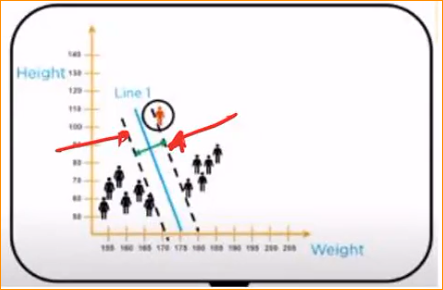
1. Identify the best hyper plane- best line

We can split data by choosing any of following two lines ?

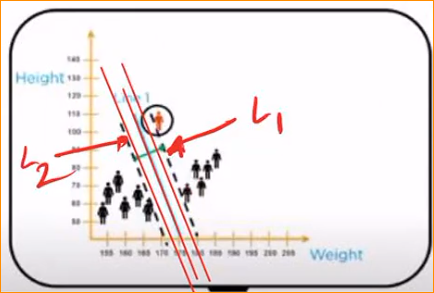


* This line has the maximum space that separates the two classes.

Now we know those line that we finally selected which go through our support vectors and have maximum distance between them.



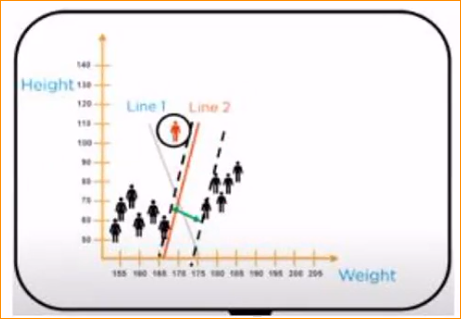
1. Identify the best line that two class separate



Find the maximum distance between the support vectors. That mean in middle.

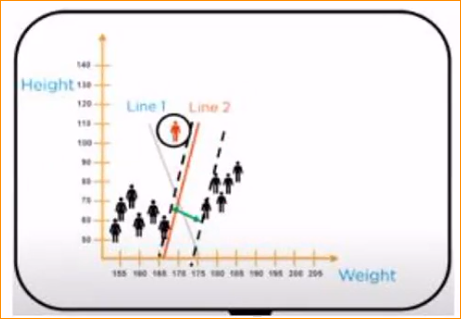
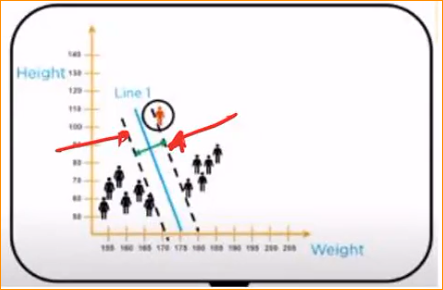
Find the mid-point in L1 and L2

Why if we draw this kind of lines – High distance right?



Example – now we have observation Find what is correct in this person has knee injury or not?

New Observation



In here this ploy saying the person has not knee injury

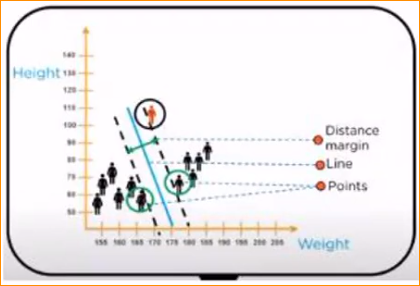
In here this ploy saying the person has knee injury

What is correct?

* That is why this line is important.

How to drawing correct line –

* Distance between the points and the line should be far as possible.



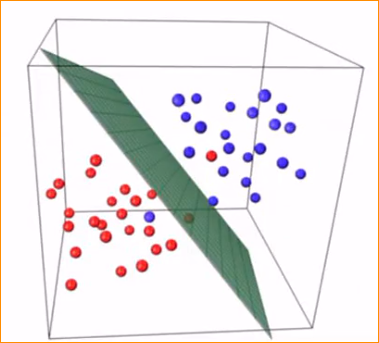
So drawing the line maximizing the distance of lines and support vectors.

Now we can use this blue line to get a prediction –

Now left side has not knee injury and right side has injury.

**If you have more than input features-**

Example of having 3 input features – 3D

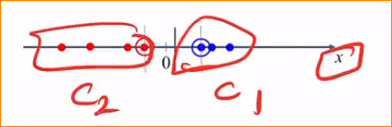


Hyper plane

Non-Linear Support vector machine

That mean only have one input feature

This data set equally separate-



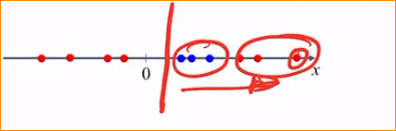
Class 01 – blue color

Class 02 – Red color

In here get support vectors and drawing a line –



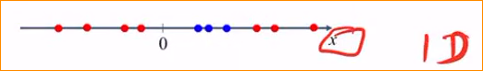
But not equally separate –



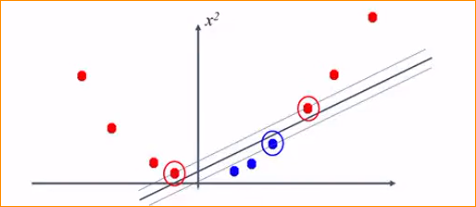
In here you cannot separate a linear hyperplane, to separate the observation in the dataset.

For this kind of problem, we are using nonlinear SVM –

* We have to map this observation into a higher dimension vector space.



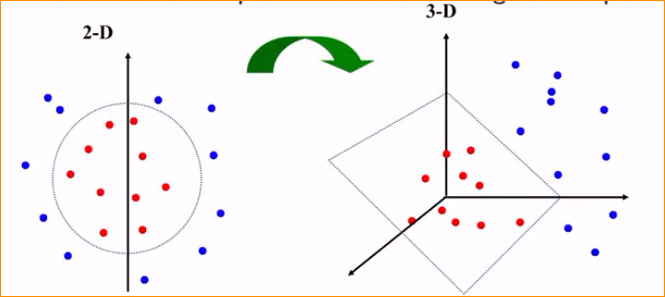
Make this one into 2D –



Now I can separate this in to two data set.

In this mapping there are so may thing s to do –

Example – 2D data set into 3D data set



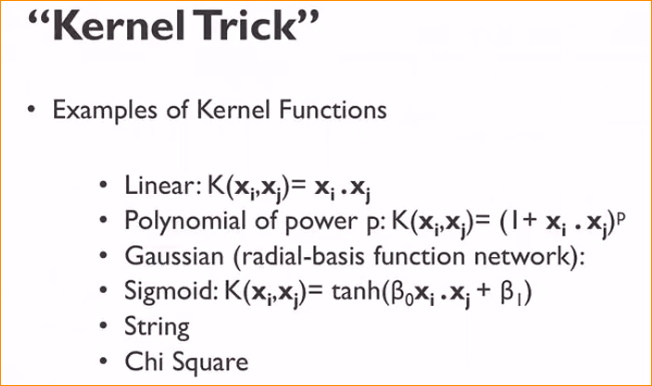
Kernel trick –

A kernel function is some function that corresponds to an inner product in some expanded feature space.

Kernel is mapping that we are going to map these original data set into high dimensional vector space.

Why use kernels –

* Make non-separable problems separable.
* Map data into better representational space.



If you have more column than the observation SVM is good. Ex- 100 variable 80 observations

**You can not say this model or any model is best fit with out train.**