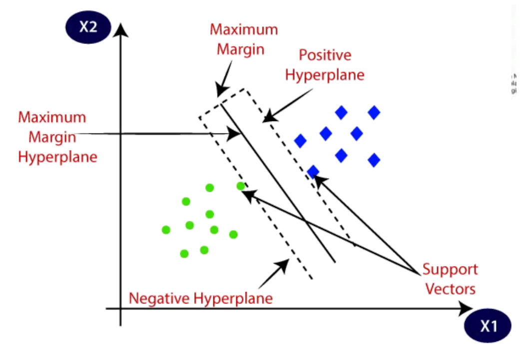
Linear SVM: Example

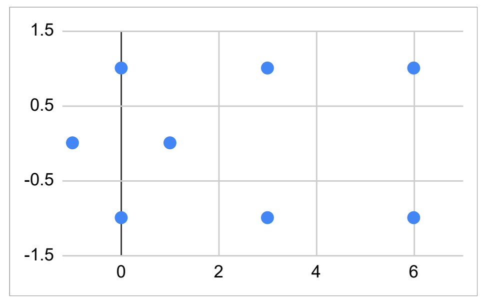
SVM is a supervised machine learning algorithm
It is used for both classification and regression.
It is mainly used in area of classification.
It helps to create best fit line or boundary such that
-the relevant data point fall in correct class for future prediction.
The best decision boundary is called hyper-plane.
SVM choses the extreme points/vectors to helps in creating a hyper-plane.
These extreme cases are called as support vectors.
Hence algorithm is termed as support vector machine.



SVM Type 1: Linear SVM
Linear SVM is used for linearly separable data
i.e a dataset points can be classified into two classes using a single straight line
Such arrangement of data is termed as linearly separable data.
The Linear SVM classifier is useful for such datasets.

Question:
Class 1: +ve labelled data points are (3,1); (3,-1); (6,1); (6,-1)
Class 2: -ve labelled data points are (1,0); (0,1); (0,-1); (-1,0)
Solution:
Output is +1 for all +ve labelled output is 1.
Output is -1 for all -ve labelled output is -1.

x	y	Class
3	1	+
3	-1	+
6	1	+
6	-1	+
1	0	-
0	1	-
0	-1	-
-1	0	-



Now adding 1 to all points

$s_1 = (3,1) \Rightarrow s_1` = (3,1,1) \mid s_2 = (3,-1) \Rightarrow s_2` = (3,1,-1)$
$s_3 = (6,1) \Rightarrow s_3` = (6,1,1) \mid s_4 = (6,-1) \Rightarrow s_4` = (6,1,-1)$
$s_5 = (1,0) \Rightarrow s_5` = (1,0,1) \mid s_6 = (0,1) \Rightarrow s_6` = (0,1,1)$
$s_7 = (0,-1) \Rightarrow s_7 = (0,-1,1) \mid s_8 = (-1,0) \Rightarrow s_8` = (-1,0,1)$

From the graph we can see there is one negative point $\alpha_1 = (1,0,1)$ and
two positive points $\alpha_2 = (3,1,1)$ & $\alpha_3 = (3,-1,1)$ form support vectors.

Generalized equation:

$\alpha_1 * s_1` * s_1` + \alpha_2 * s_1` * s_2` + \alpha_3 * s_1` * s_3` = -1 \rightarrow 1$
$\alpha_1 * s_2` * s_1` + \alpha_2 * s_2` * s_2` + \alpha_3 * s_2` * s_3` = 1 \rightarrow 2$
$\alpha_1 * s_3` * s_1` + \alpha_2 * s_3` * s_2` + \alpha_3 * s_3` * s_3` = 1 \rightarrow 3$

On solving these equations:

$\alpha_1 * (1,0,1) * (1,0,1) + \alpha_2 * (1,0,1) * (3,1,1) + \alpha_3 * (1,0,1) * (3,-1,1) = -1$; Eq 1
$\alpha_1 * (3,1,1) * (1,0,1) + \alpha_2 * (3,1,1) * (3,1,1) + \alpha_3 * (3,1,1) * (3,-1,1) = 1$; Eq 2
$\alpha_1 * (3,-1,1) * (1,0,1) + \alpha_2 * (3,-1,1) * (3,1,1) + \alpha_3 * (3,-1,1) * (3,-1,1) = 1$; Eq 3

From Eq 1, Eq 2, Eq 3:

$2\alpha_1 + 4\alpha_2 + 4\alpha_3 = -1 \rightarrow 4$
$4\alpha_1 + 11\alpha_2 + 9\alpha_3 = 1 \rightarrow 5$
$4\alpha_1 + 9\alpha_2 + 11\alpha_3 = 1 \rightarrow 6$

On solving these equations 4,5 and 6 we get,

$\alpha_1 = -3.5$, $\alpha_2 = 0.75$ and $\alpha_3 = 0.75$

To find hyper-plane:

$W` = \Sigma \alpha_i * s_i$
$W` = -3.5 * (1,0,1) + 0.75 * (3,1,1) + 0.75 * (3,-1,1)$
$W` = (1,0,-2)$
$y = W`x + b$
$W` = (1,0)$ and $b = 2$

So, the best fit line or hyper plane is at (0,2) which splits the data points into two classes.