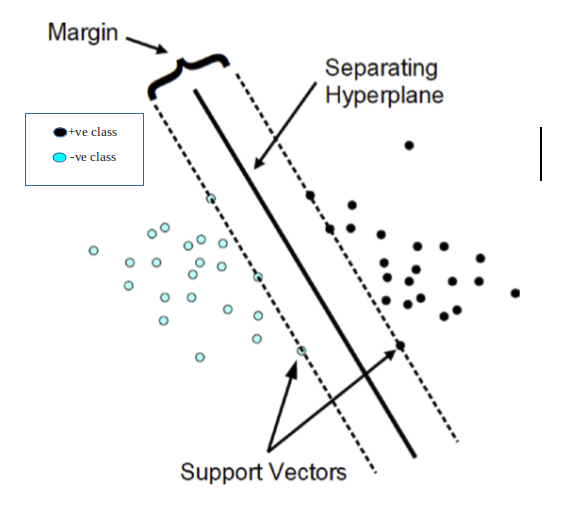
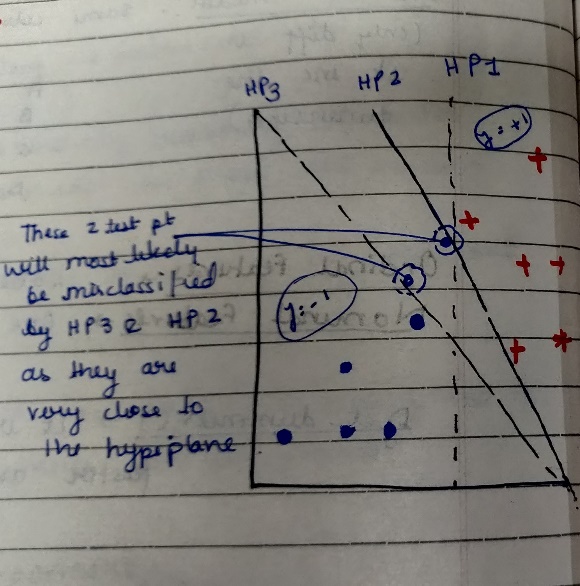
 Our aims are to provide a basic understanding of the SVM, the optimization that is happening behind the scenes, and knowledge about its parameters along with its implementation in Python.



  Content

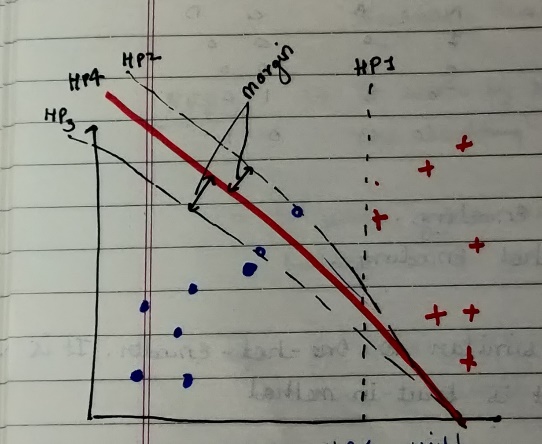
Support Vector Machine is a supervised learning algorithm that can be used for both classification and regression problems. It is mostly used for classification problems.

We should keep in mind that the main task of the classification problem is to find the best separating hyperplane/ Decision boundary. We can have the ‘n-1’ hyperplane which can be either linear or nonlinear.

Class labels are denoted as -1 for negative class and +1 for positive class in Support Vector .

We can clearly see from the above figure that both the Hyperplane (HP2 & HP3) are not able to correctly classify the test-data points because they are very close to the hyperplane.

Such data points are called Support vectors which are simply feature values in vector form.



From the above figure, we can see that Hyperplane (HP4) is the best as it is able to correctly classify all the data points including support vectors.

#### This brings us to think what exactly Margins are.

Margins are generally defined by the closest data points (called support vectors) on either of the hyperplanes.

Note: Another point to note from the above figure is that the further the data points are from the margins the more correctly they are classified.

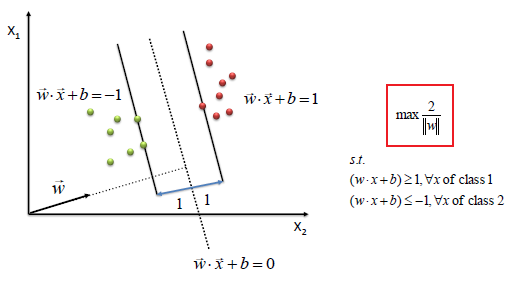
## 2. Optimization Technique used in SVM

The core of any Machine learning algorithm is the Optimization technique that is happening behind the scenes.

SVM maximizes the margin by learning a suitable decision boundary/decision surface/separating hyperplane.



It can mathematically be written as:



#### Points to note from the above Figure:

a. We can clearly see that SVM tries to maximize the margins and thus is called Maximum Margin Classifier.

b. The Support Vectors will have values exactly either {+1, -1}.

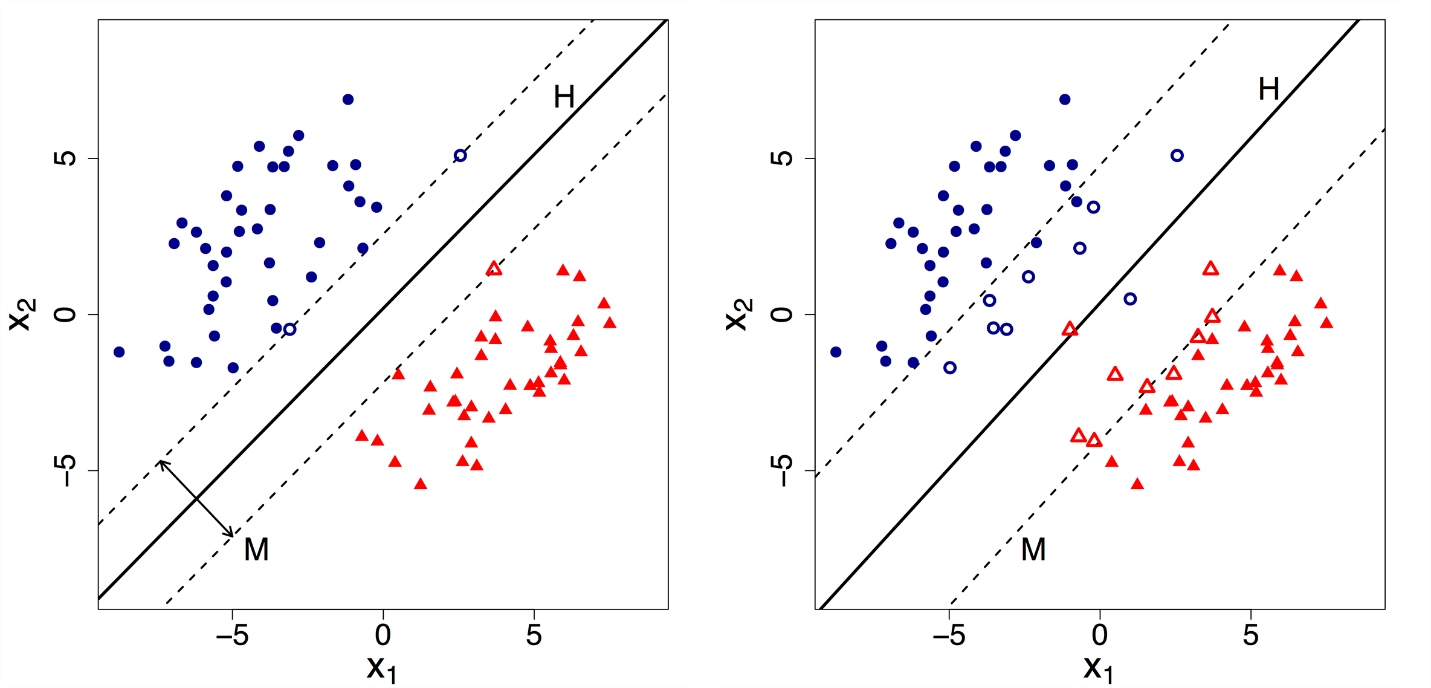
c. The more negative the values are for the green data points the better it is for classification.

d. The more positive the values are for the red data points the better it is for classification.

For more in-depth knowledge regarding the maths behind Support Vector Machine refer to this [article](https://towardsdatascience.com/demystifying-maths-of-svm-13ccfe00091e)

## 3. How to choose the Correct SVM

Choosing the correct classifier is really important. Let us understand this with an example.



Suppose we are given 2 Hyperplane one with 100% accuracy (HP1) on the left side and another with >90% accuracy (HP2) on the right side. Which one would you think is the correct classifier?

Most of us would pick the HP2 thinking that it is because of the maximum margin. But it is the wrong answer.

But Support Vector Machine would choose the HP1 though it has a narrow margin. Because though HP2 has maximum margin but it is going against the constraint that: each data point must lie on the correct side of the margin and there should be no misclassification. This constraint is the hard constraint that Support Vector Machine follows throughout.