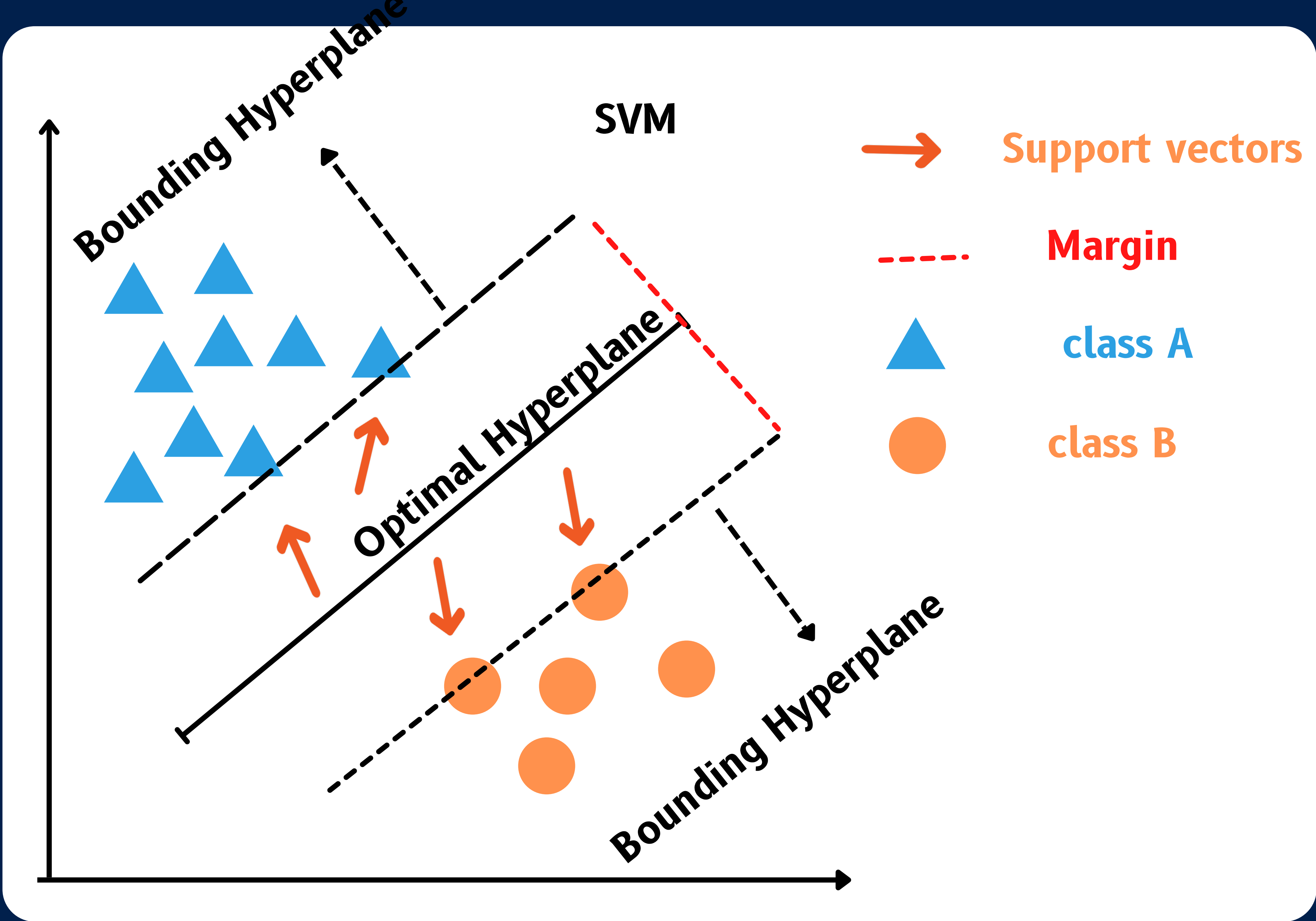
SVM (Support Vector Machine)

SVM (Support Vector Machine)

SVM is a supervised machine learning algorithm offers very high accuracy compare to other classifiers such as Logistic Regression and Decision Tree. It knows for its kernel tricks, to handle Non-Linear input spaces. It can easi--ly handle multiple contiguous and categorical variables. SVM constructs a Hyperplane in Mul--ti Dimensional space to separate different cla-sses.

- SVM generate optimal hyperplane in a iter-ative manner. Which is used to minimize er-ror. The Core idea is to find a Maximum Marginal Hyperplane(MMH) that divides the dataset best.



Support Vectors: These are the data points, which are closest to the Hyperplanes. These points will define the separating line by calculating margins. These points are more relevant to construction of the classifier.

Hyperplane: A hyperplane is a decision boundary that differentiates the two classes in SVM.

Margin : A margin is a gap between two lines on the closest class points. This calculated as perpendicular distance from line to support vectors or closest points. If the margin is larger in between the classes, Then it considered a good margin, a smaller margin considered as bad margin. If the margin is so small that the model becomes prone to overfitting or being too sensitive to outliers.

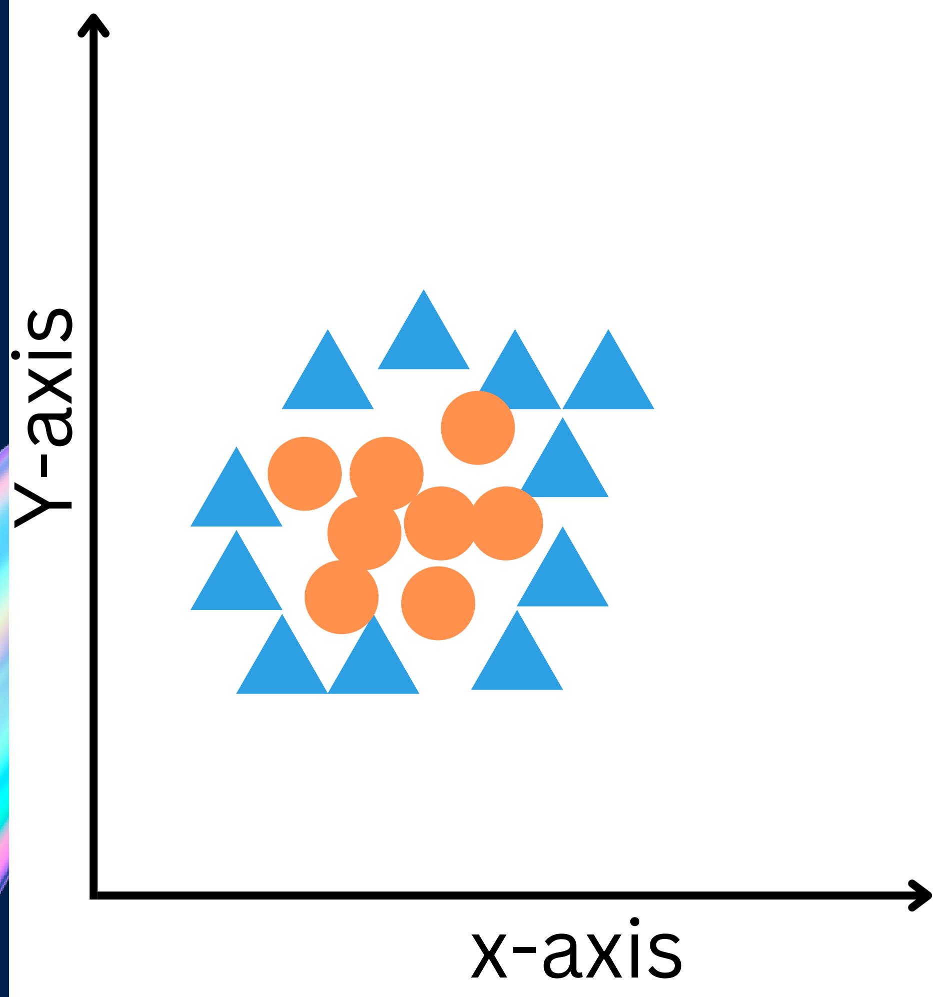
Dealing with non-linear and inseparable planes:

Some Problems can't solve using hyperplane, In such situations, SVM uses a kernel tricks to transform the input space to a higher dimensional space, The data points are plotted on the x-axis and z-axis (z is the squares of both x and y).

Z = x ** 2

Z = y ** 2

NEXT



SVM Kernels: A kernel transforms an input data space into the require form, SVM uses a technique called the kernel trick. The kernel takes a low-dimensional space. You can say that it converts non-separable problem to separable problem by adding more dimensions to it. It is most useful in non-linear separation problem. Kernel tricks can help to built more accurate classifier.

Linear Kernel: A linear kernel can be used as the normal dot product of any two given observation. The product between. Two vectors is the sum of multiplication of each pair of input values.

K(x, xi) = SUM(x * xi)

Polynomial Kernel: A polynomial kernel is used more generalized form of the linear kernel, The polynomial kernel distinguish curved or non-linear input space.

K(x, xi) = 1 + SUM(x * xi) ^ d **#where is d is polynomial**

Radial Basic Function Kernel (RBF): **RBF** kernel is a popular kernel function commonly used in SVM, RBF can map an input space in infinite dimensional space.

K(x, xi) = exp(-gamma*SUM(x-xi2))**

Here gamma is a parameter, which ranges from 0 to 1. A higher value of gamma will perfectly fit the training data lets to overfitting, **Gamma=0.1** consider to be a good default value.

Hyperparameter tuning

Regularization : **"C"** parameter used to maintain regularization, **"C"** is the penalty parameter represents misclassification of enor term.

Gamma: A lower value of gamma will loosely fit the training dataset. Where the higher value will exactly fit the training dataset. It leads to overfitting, Lower value of gamma to consider only nearby points in calculating the separation line.