# Support Vector Machine Algorithm (SVM)

❖ SVMs are a supervised learning algorithm used for classification and regression problems.
❖ SVMs are primarily used for classification problems.
❖ The goal of SVMs is to create a hyperplane that separates data points into different classes.
❖ The hyperplane is chosen to maximize the margin between the two classes.
❖ Support vectors are the data points that are closest to the hyperplane.
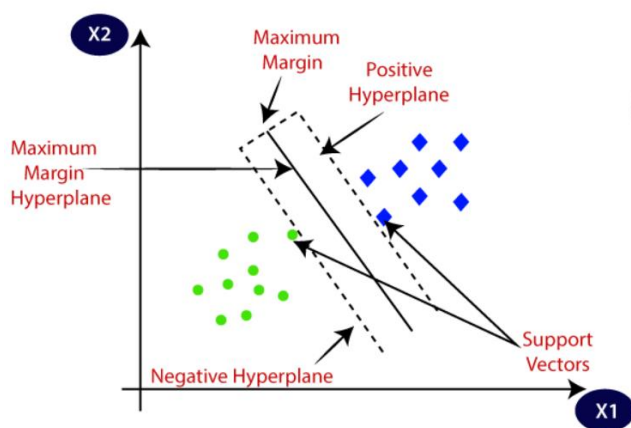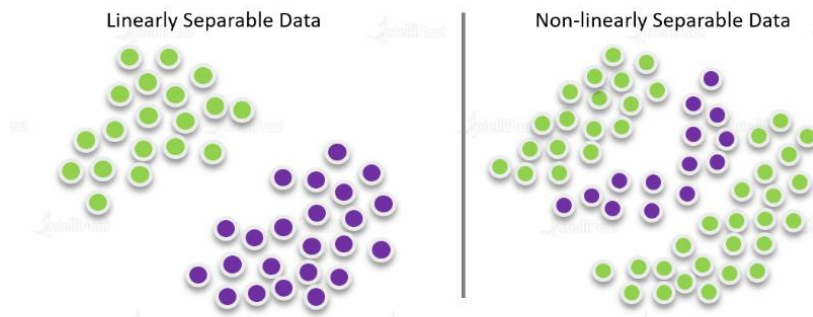

Types of SVM

SVM can be of two types:

Linear SVM:

- For linearly separable data.
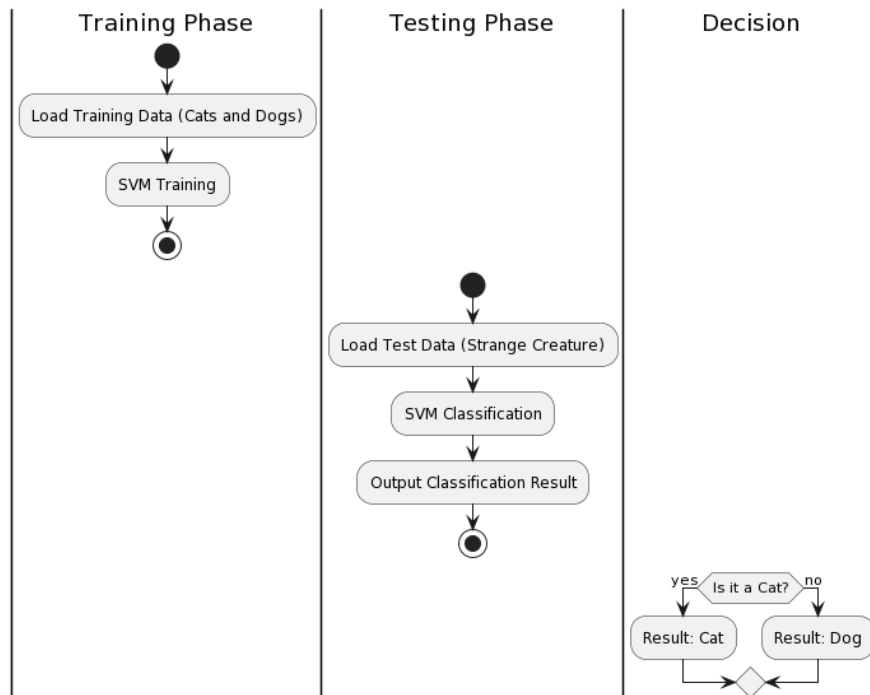- Classifies using a straight line.

Non-Linear SVM:

- For non-linearly separable data.
- Uses complex curves for classification.

SVM Figure 1: Linearly Separable and Non-linearly Separable Datasets

How Model Work:

Applications of SVM

➢ **Face Detection:** Classifies images of people's faces by creating a bounding box around them.

➢ **Bioinformatics:** Classifies genes to differentiate between proteins, identify biological problems, and detect cancer cells.

➢ **Text Categorization:** Classifies documents into different categories based on their content.

➢ **Generalized Predictive Control (GPC):** Provides control over industrial processes.

➢ **Handwriting Recognition:** Recognizes handwritten characters by matching them against pre-existing data.

➢ **Image Classification:** Classifies images into different categories.

Advantages of SVM

- It has a high level of accuracy
- It works very well with limited datasets
- Kernel SVM contains a non-linear transformation function to convert the complicated non-linearly separable data into linearly separable data
- It is effective on datasets that have multiple features
- It is effective when the number of features are greater than the number of data points
- It employs a subset of training points in the decision function or support vectors, making SVM memory efficient
- Apart from common kernels, it is also possible to specify custom kernels for the decision function

Disadvantages of SVM

- Does not work well with larger datasets
- Sometimes, training time with SVMs can be high
- If the number of features is significantly greater than the number of data points, it is crucial to avoid overfitting when choosing kernel functions and regularization terms
- Probability estimates are not directly provided by SVMs; rather, they are calculated by using an expensive fivefold cross-validation
- It works best on small sample sets due to its high training time