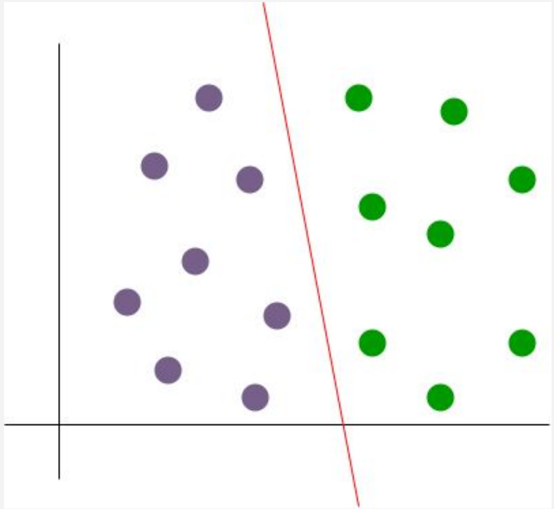


Support Vector Machine (SVM)

Vessa Rizky Oktavia
KC 01-02-03

Support Vector Machine

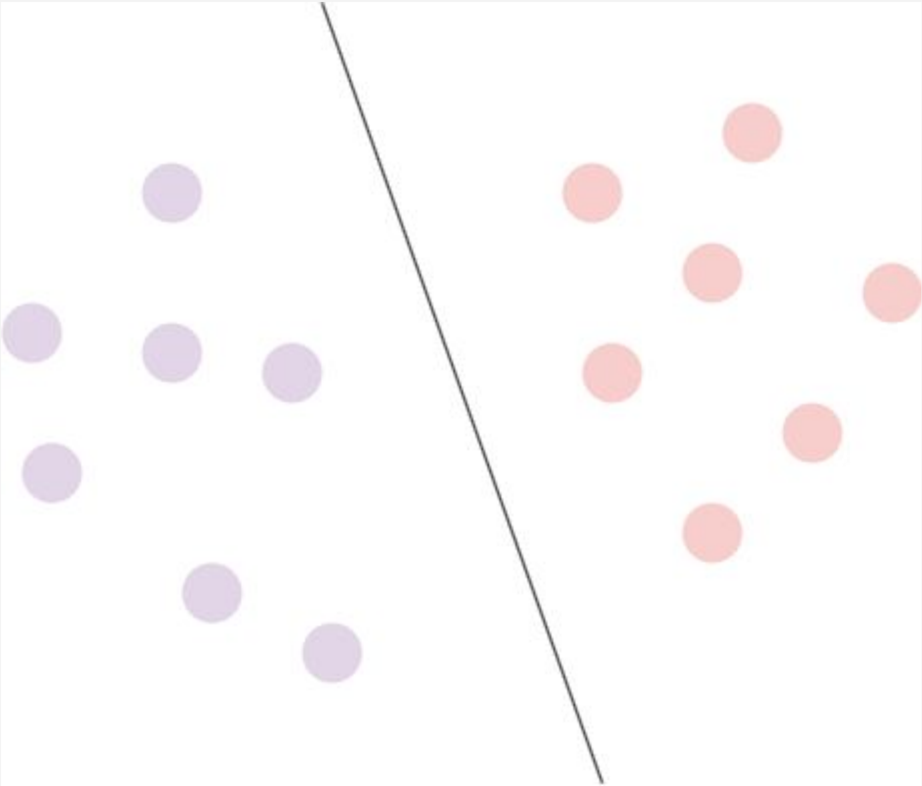
In machine learning, support vector machines (SVM) are **supervised learning** models with associated learning algorithms that analyze data for **classification** and **regression** analysis.

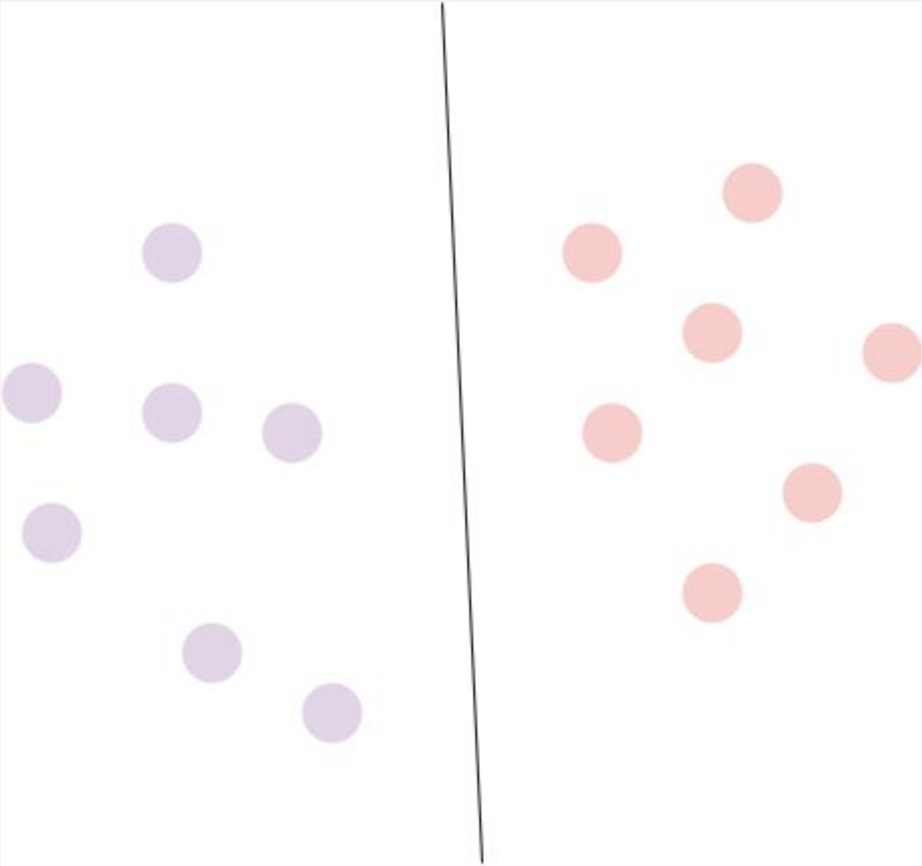
Wikipedia

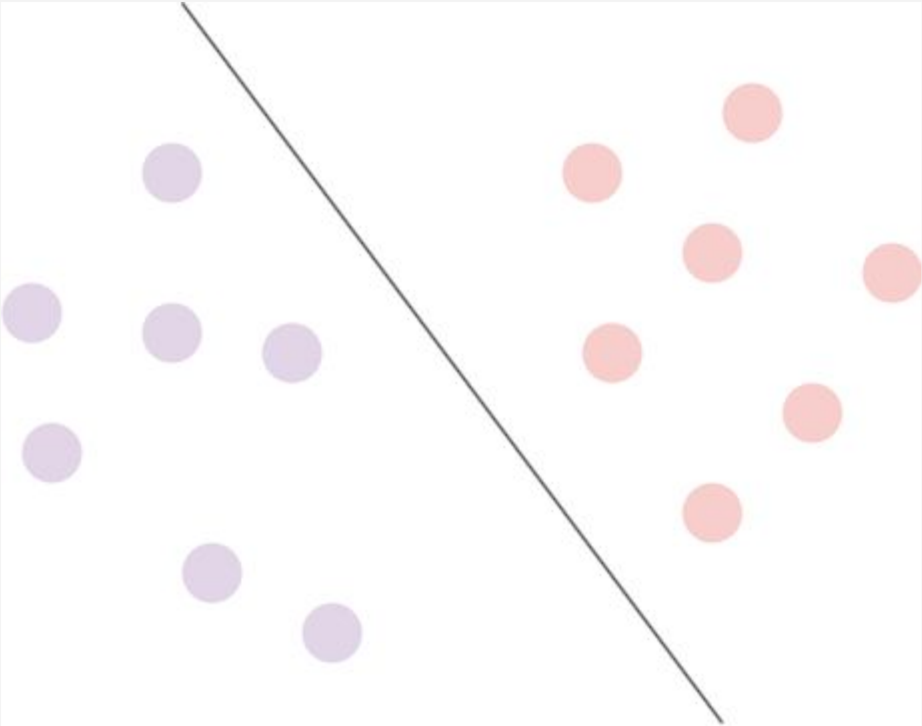
Supervised learning uses a **training set to teach models** to yield the desired output. This training dataset **includes inputs and correct outputs**, which allow the model to learn over time. The algorithm measures its **accuracy** through the **loss function**, adjusting until the **error** has been sufficiently **minimized**.

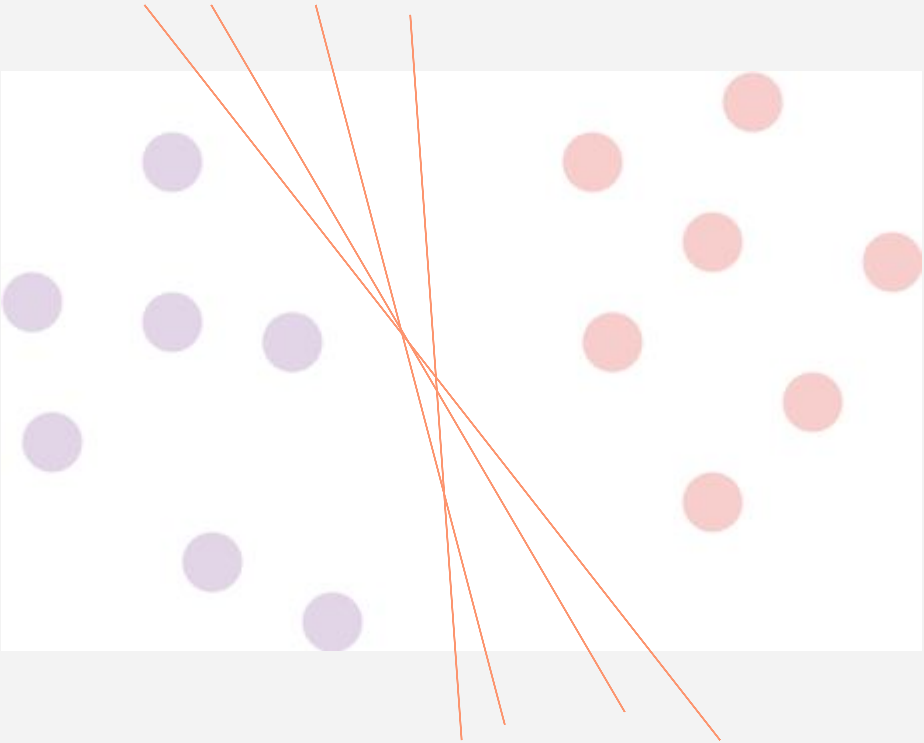
IBM

Supervised Learning







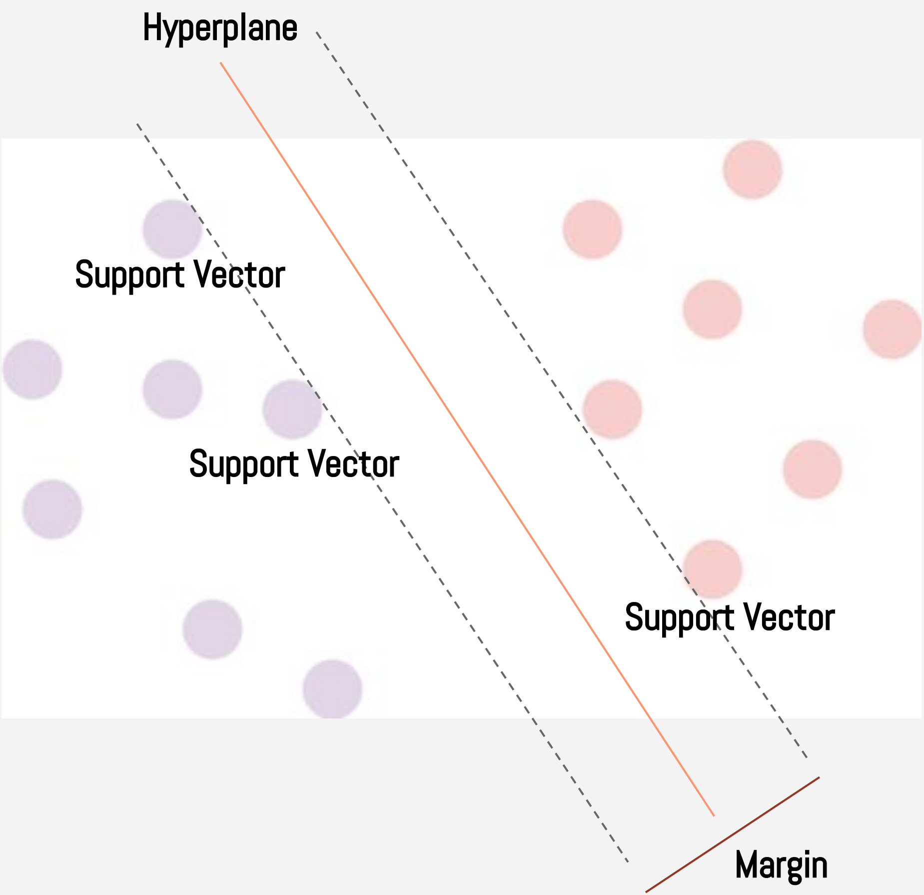


SVM aims to find a hyperplane that best separates the data into two classes.

This hyperplane should maximize the margin between the classes.

The margin is defined as the distance between the hyperplane and the nearest data points from each class.

These nearest data points are called support vectors.



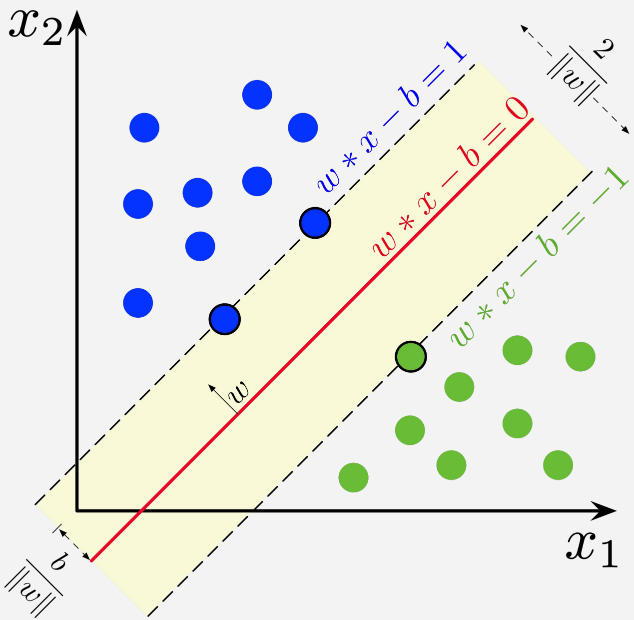
SVM aims to find a hyperplane that best separates the data into two classes.

This hyperplane should maximize the margin between the classes.

The margin is defined as the distance between the hyperplane and the nearest data points from each class.

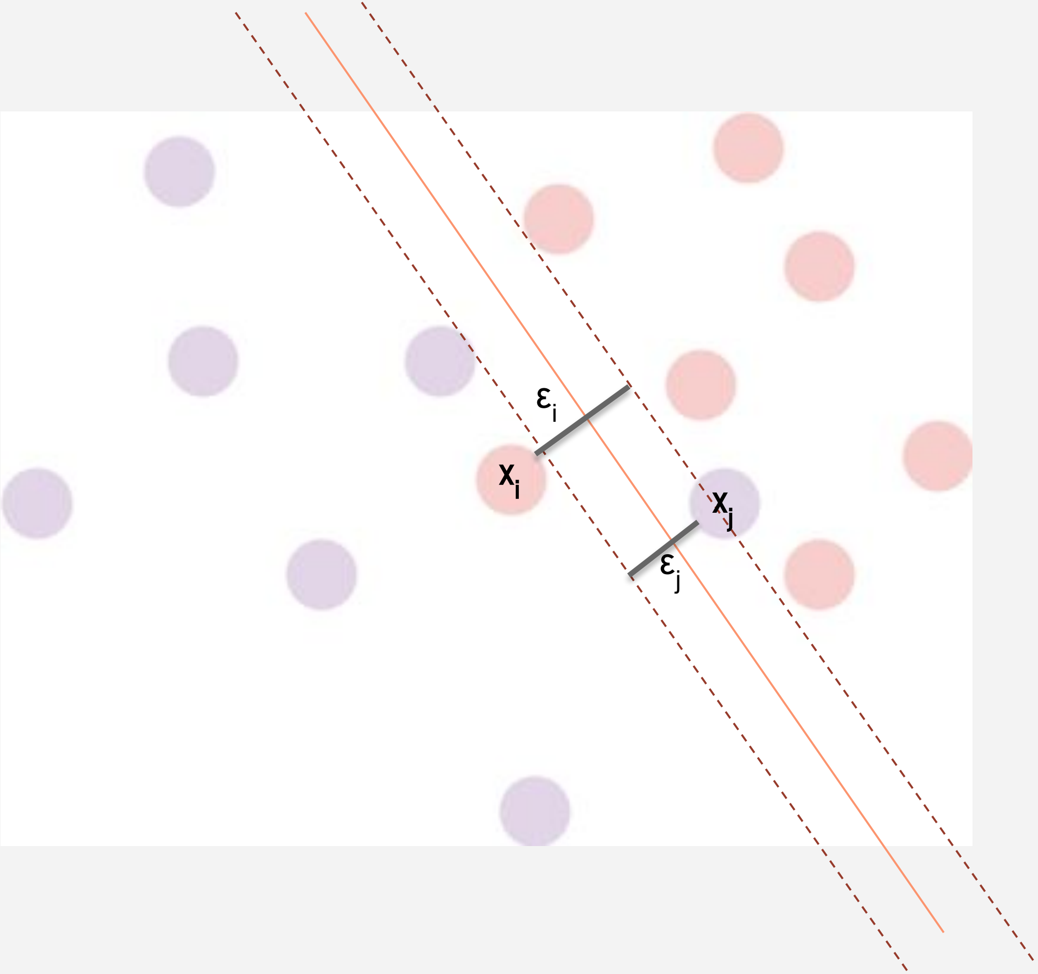
These nearest data points are called support vectors.

Linear SVM Hyperplane



Src: https://en.m.wikipedia.org/wiki/File:SVM_margin.png

How to handle data with outliers?



$$\begin{cases} \mathbf{w}^T\mathbf{x}_i + b \geq 1 - \xi_i & y_i = 1 \\ \mathbf{w}^T\mathbf{x}_i + b \leq -1 + \xi_i & y_i = -1 \\ \xi_i \geq 0 & \forall i \end{cases}$$

Parameter C

- The parameter C is a hyperparameter that determines the trade-off between maximizing the margin between classes and minimizing the classification error on the training data.
- C is a regularization parameter that controls the penalty for misclassified data points. Its value influences the behavior of the SVM model during training and can have a significant impact on the resulting decision boundary.

Margin using C

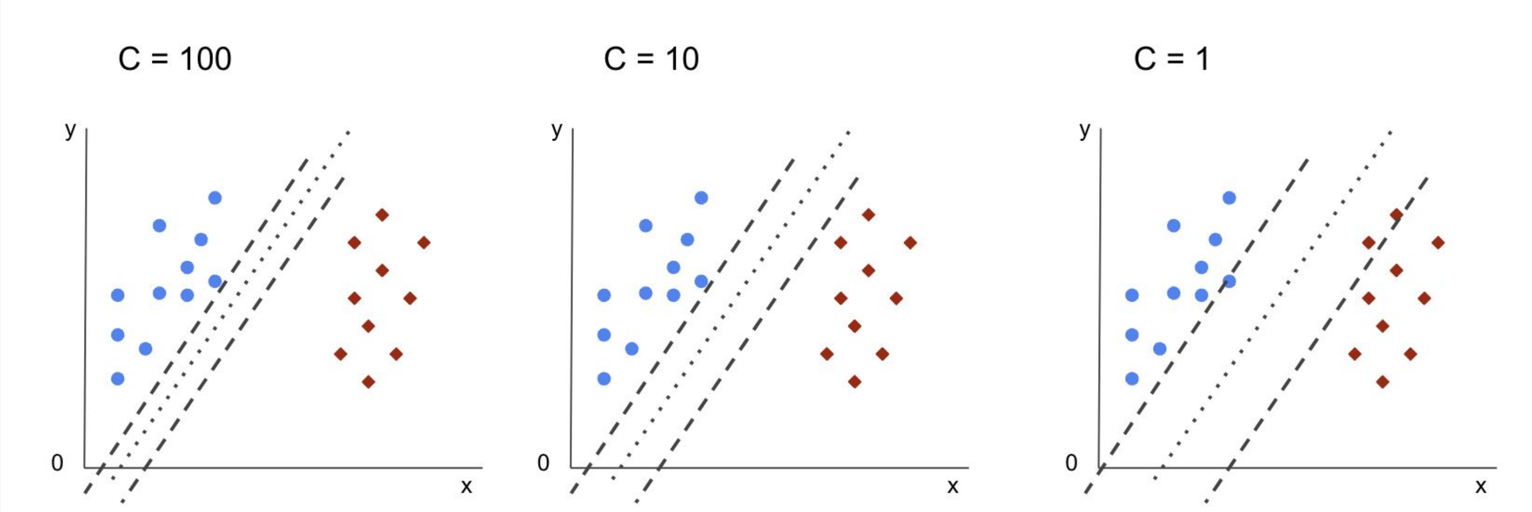
Soft Margin

- When C is small, the SVM is more tolerant of misclassifications.
- It allows for a wider margin between classes, which is desirable when the data is noisy or when there are outliers.
- The model may misclassify some training data points but is likely to generalize better to unseen data.

Hard Margin

- When C is large, the SVM imposes a higher penalty for misclassifications.
- It seeks to minimize training errors, potentially resulting in a smaller margin.
- The model may have a smaller margin and may fit the training data closely, which can lead to overfitting, especially when the data is noisy.

Margin using C



Src: https://stackabuse.com/understanding-svm-hyperparameters/

How can we separate this data?

SVM

- Linear
- Non-Linear

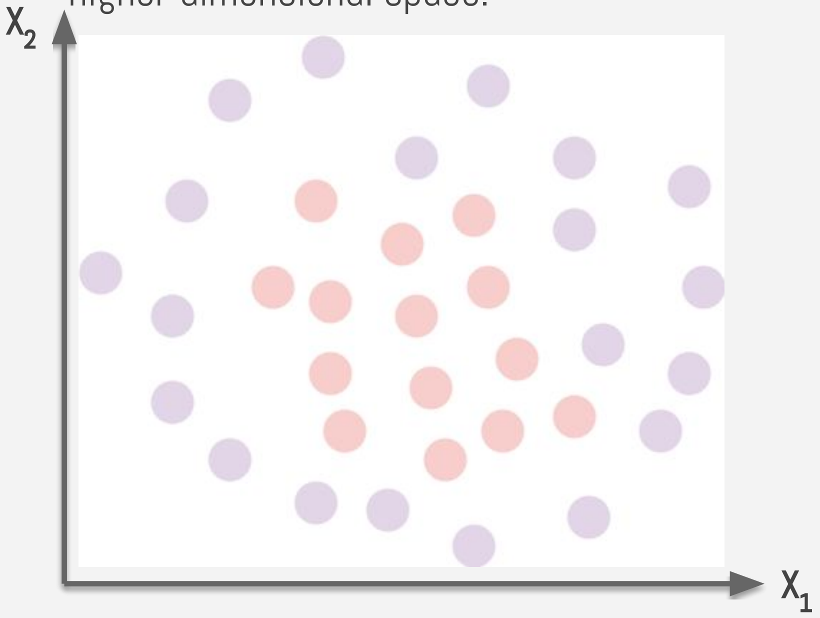
Kernel methods owe their name to the use of kernel functions, which enable them to **operate in a high-dimensional, implicit feature space** without ever computing the coordinates of the data in that space, but rather by **simply computing the inner products** between the images of all pairs of data in the feature space.

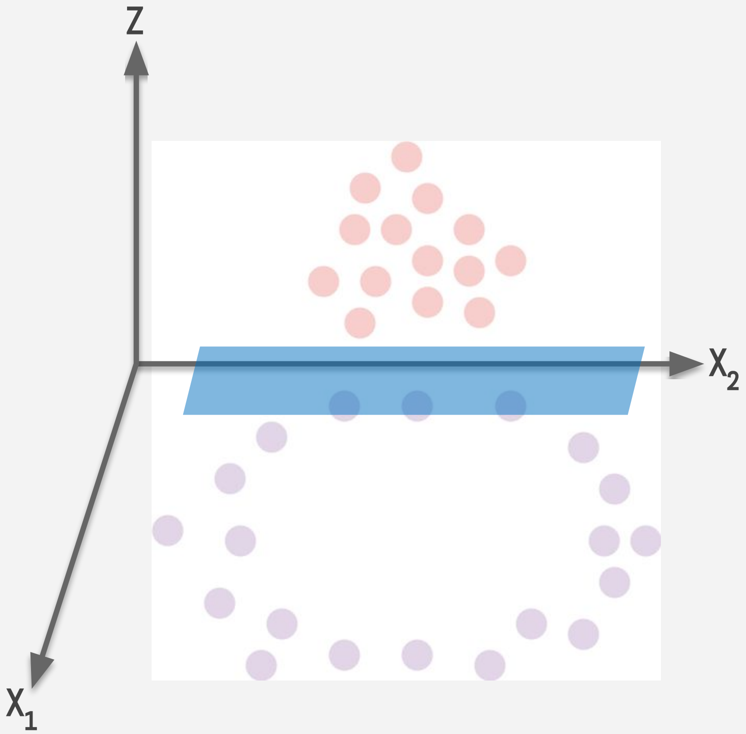
Wikipedia

Kernel Trick

In many real-world scenarios, data may not be linearly separable in its original feature space.

However, by applying a nonlinear transformation to the data, it might become linearly separable in a higher-dimensional space.





Linear

Polynomial

RBF

Sigmoid

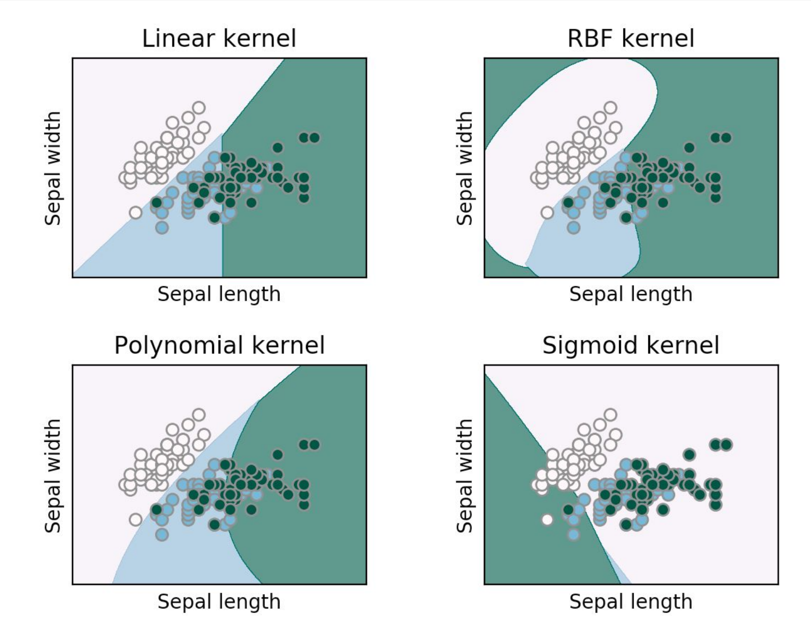
Kernel Trick

$K(x, y) = \varphi(x) \cdot \varphi(y)$

Kernel Trick

No.	Kernel Name	Kernel Function
1.	Linear Kernel	x·y
2.	Polynomial Kernel	(γ * x·y + r)^d
3.	Radial Basis Function (RBF) Kernel	exp(-γ * ||x - y||^2)
4.	Sigmoid Kernel	tanh(γ * x·y + r)

Differences



Src: https://www.analyticsvidhya.com/blog/2021/06/support-vector-machine-better-understanding/

Benefits and Drawbacks

Benefits

- Effective in High-Dimensional Spaces
- Robust to Overfitting
- Works Well with Small to Medium-Sized Datasets
- Global Optimum

Drawbacks

- Sensitivity to Kernel Choice

Thanks!

Do you have any questions?

vessa@ittelkom-sby.ac.id
+6283849878767