Support Vector Machines

Overview

1 What are SVMs?

SVMs are supervised learning models used for classification and regression.

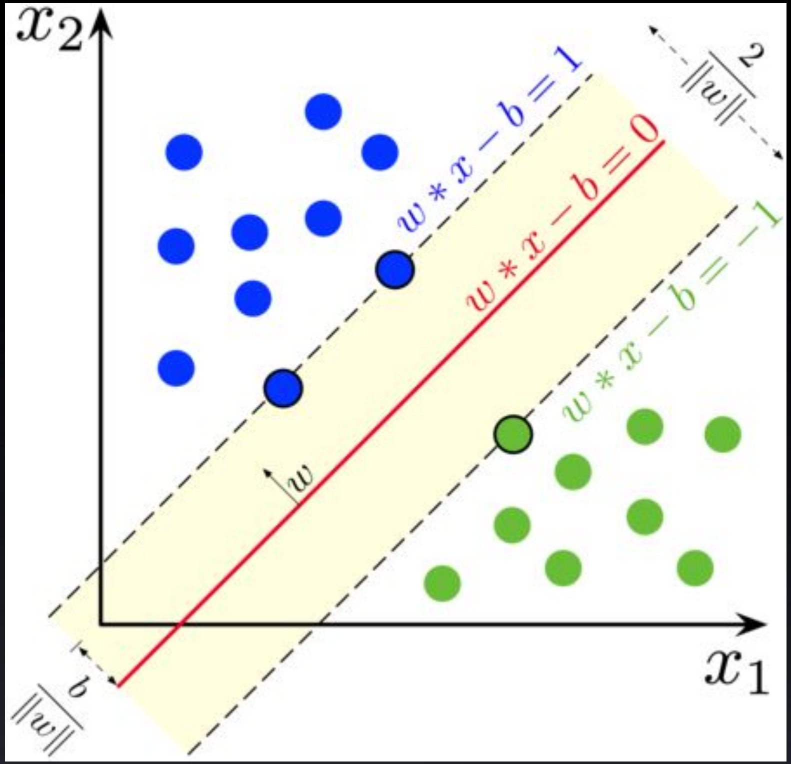
Known for effectiveness in high-dimensional spaces.

2 Applications

Text classification, image recognition, bioinformatics.

How SVM Works

- **Key Concept:** Separating hyperplane - a boundary that separates different classes.
- **Goal:** Maximize the margin between classes by finding the optimal hyperplane.
- **Support Vectors:** Data points closest to the hyperplane, influencing its position and orientation.



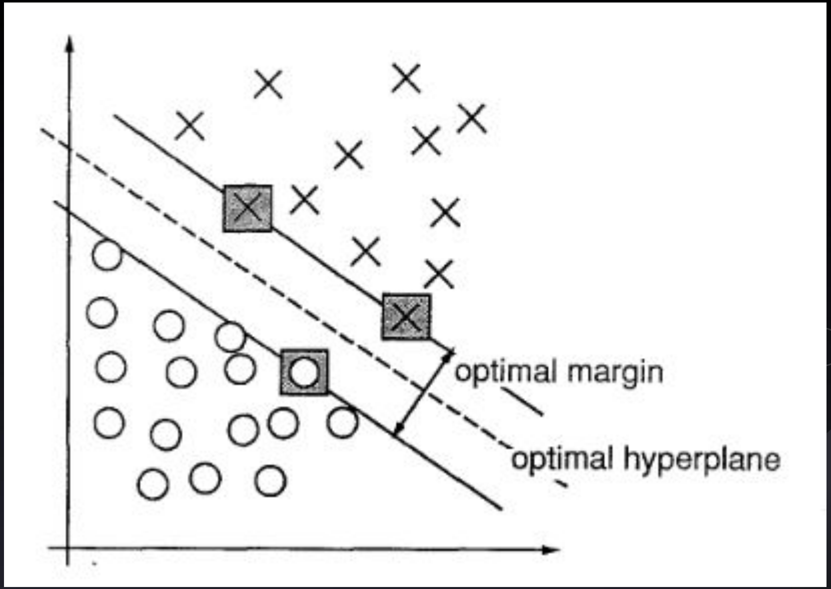
Understanding the Margin and Hyperplane

Margin

Distance between the closest points of each class to the hyperplane.

Optimal Hyperplane

The hyperplane with the maximum margin.



Mathematical Formulation of SVM

- **Objective Function:** Introduce the optimization problem for finding the hyperplane.
- **Linear SVM:** Formulated as a convex optimization problem with constraints.
- **Kernel Trick:** A technique for handling non-linear relationships with different kernels

$$\min_{\mathbf{w},b} \frac{1}{2}\|\mathbf{w}\|^2$$

Types of Kernels in SVM

Linear Kernel: Suitable for linearly separable data.

$$K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i \cdot \mathbf{x}_j$$

Polynomial Kernel: Suitable for polynomial relationships.

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + c)^d$$

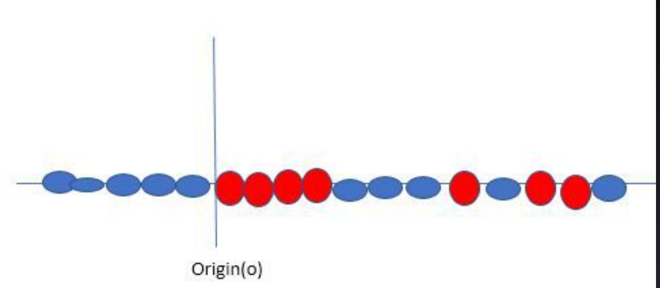
Radial Basis Function (RBF) Kernel: Used for non-linear relationships, popular in image classification.

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2)$$

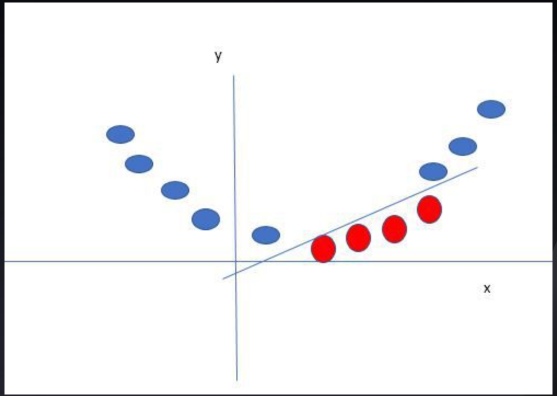
Sigmoid Kernel: Often used for neural networks.

Non-Linear SVM Using the Kernel Trick

- **Kernel Trick:** Efficient computation of transformed feature space without explicit transformation.

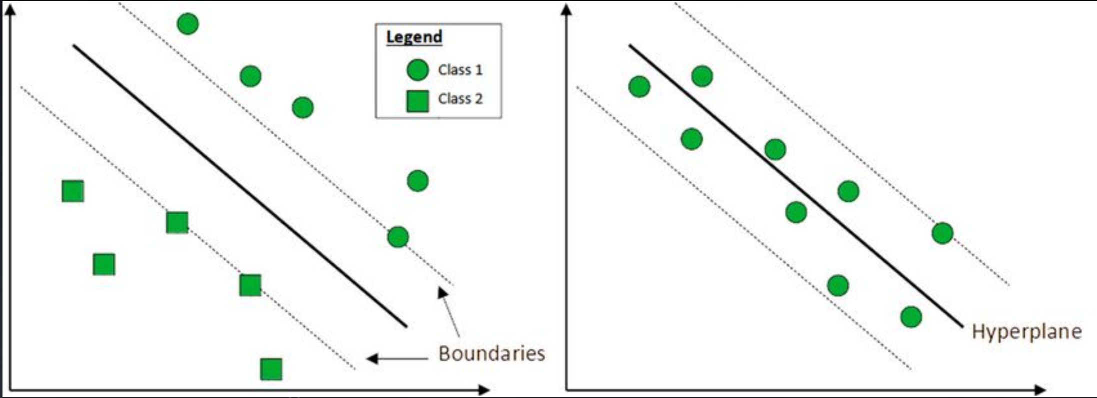


Create new variable y as a function of distance from the origin



SVM for Classification and Regression

- **Classification:** SVM's primary use, finding the decision boundary between classes.
- **Support Vector Regression (SVR):** Find the best function that predicts an output given an input



Advantages and Limitations of SVM

Advantages:

- ❑ Effective in High-Dimensional Spaces
- ❑ Robust to Overfitting
- ❑ Versatile - Can be used for classification, regression, and outlier detection

Limitations:

- ❑ Computational Complexity
- ❑ Choice of Kernel - Performance can be sensitive to the choice of kernel function and its parameters
- ❑ Can be harder to interpret compared to simpler models

Hyperparameter Tuning in SVM

- Common Hyperparameters:
  - C (Regularization): Controls the trade-off between maximizing the margin and minimizing classification error.
  - Gamma (for RBF Kernel): Controls the influence of a single training example.
- Techniques: Grid search and cross-validation for optimal tuning.

Practical Applications of SVM

Image Classification: Face recognition, object detection.

Text Classification: Spam detection, sentiment analysis.

Bioinformatics: Protein classification, gene expression data analysis.

Time Series Prediction: Financial market forecasting with SVR.

Visualizing SVM in Action