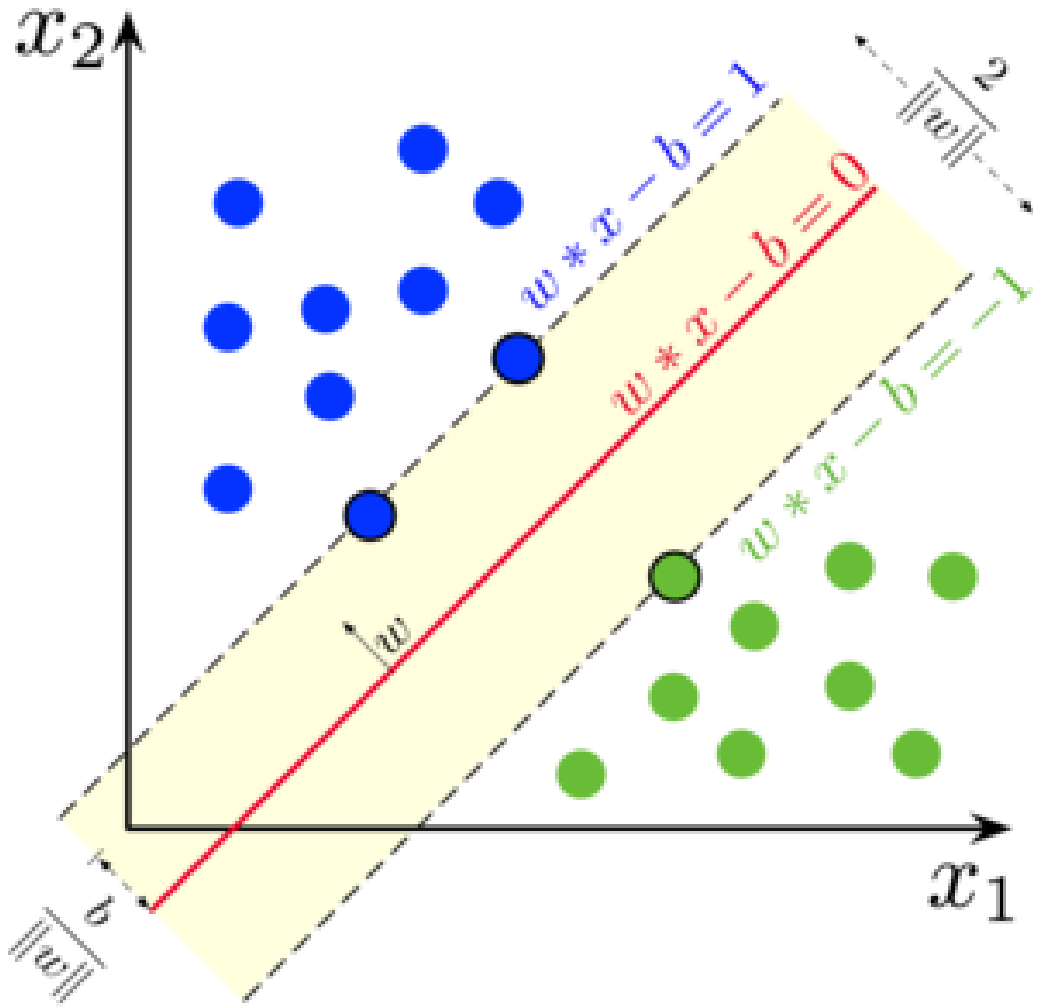
Support Vector Machines (SVM)

What is an SVM?

- SVM is a supervised machine learning algorithm used for classification and regression tasks.
- It works by finding the optimal hyperplane that separates data points into different classes.
- The algorithm aims to maximize the margin between classes for better generalization.

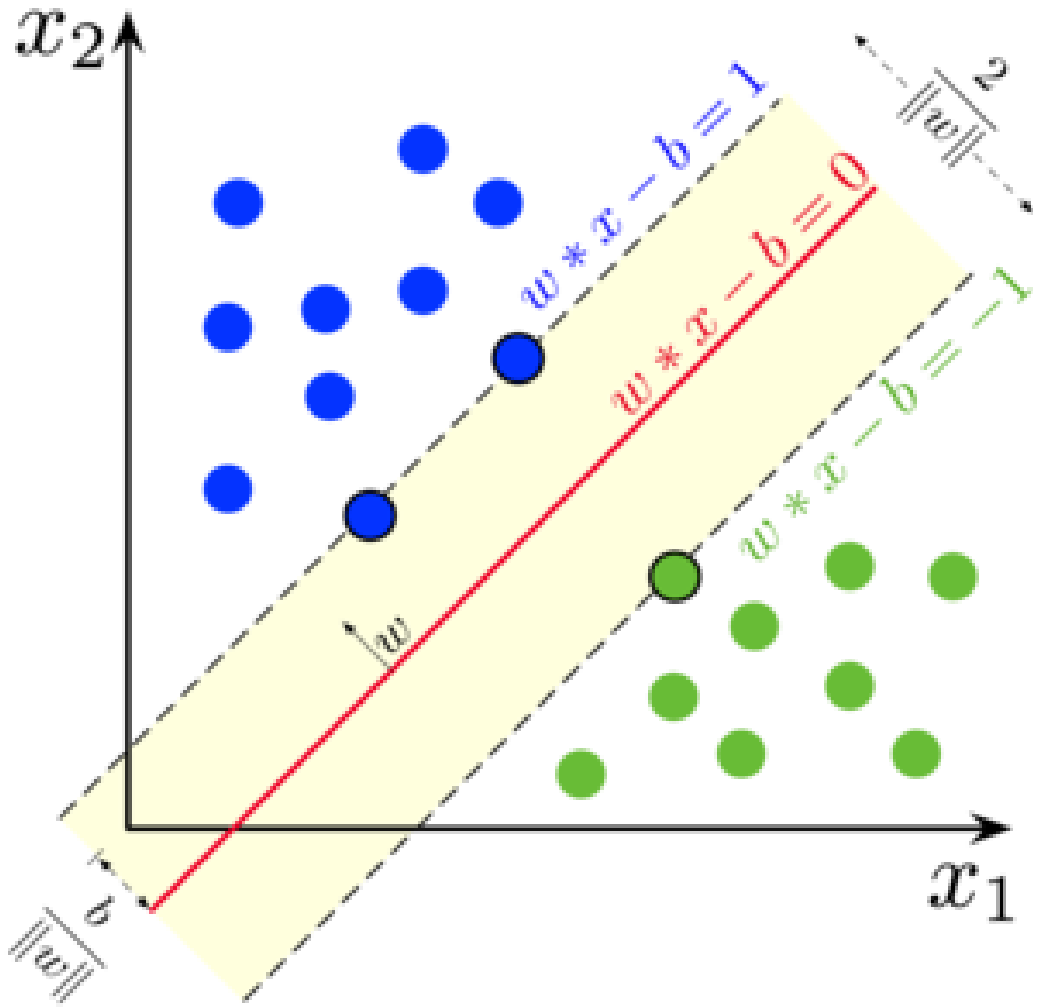


What are Margins?

•**Margins** are the distances between the hyperplane and the closest data points from each class.

•These closest points are called **Support Vectors**, and they determine the position of the hyperplane.

•SVM seeks to maximize the margin to improve the model's robustness and accuracy.



Why Wider Margins Matter:

- Wider margins reduce the model's sensitivity to small variations in the data.
- A larger margin ensures better generalization to unseen data.
- SVM optimizes the hyperplane to achieve the widest possible margin while correctly classifying the data.

What is a Kernel?

A kernel transforms data into a higher-dimensional space to make it separable by a hyperplane.

Linear Kernel:

Best for linearly separable data.

Simple and efficient.

RBF Kernel (Radial Basis Function):

Handles complex, non-linear patterns.

Effective for most scenarios.

Polynomial Kernel:

Captures curved relationships.

Useful for higher-degree decision boundaries.

```
svc = svm.SVC(kernel='rbf',C=0.1 ,gamma=0.5)
```

TECHNION
Azrieli Continuing Education and External Studies Division

What is C in SVM?

C controls the trade-off between achieving a larger margin and minimizing classification errors.

High C:

Focuses on minimizing errors, leading to a tighter fit to the training data.

Risk of overfitting.

Low C:

Allows more misclassifications to achieve a larger margin.

Promotes generalization but can lead to underfitting.

```
svc = svm.SVC(kernel='rbf',C=0.1 ,gamma=0.5)
```

What is Gamma in SVM?

Gamma defines how far the influence of a single data point reaches.

High Gamma:

Each data point has a small influence.

Leads to more complex decision boundaries (**risk of overfitting**).

Low Gamma:

Each data point has a broader influence.

Results in simpler decision boundaries (**risk of underfitting**).

```
svc = svm.SVC(kernel='rbf',C=0.1 ,gamma=0.5)
```

- **SVM** finds the optimal hyperplane to classify data by maximizing margins.
- **Margins** improve model robustness and reduce sensitivity to data variations.
- **Kernels** allow SVM to handle both linear and non-linear data structures.
- **C** and **Gamma** are key hyperparameters to balance complexity and generalization.