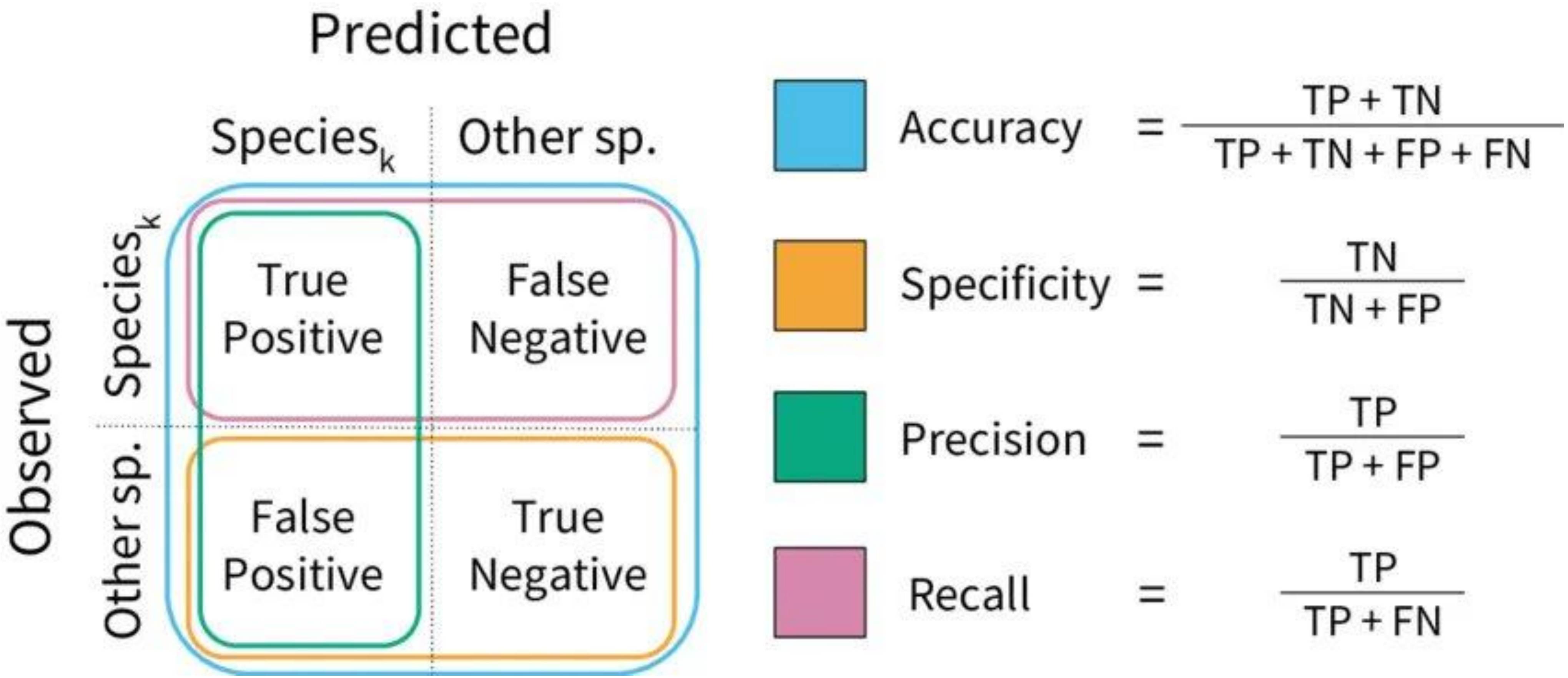
Evaluation Metrics

Evaluation Metrics

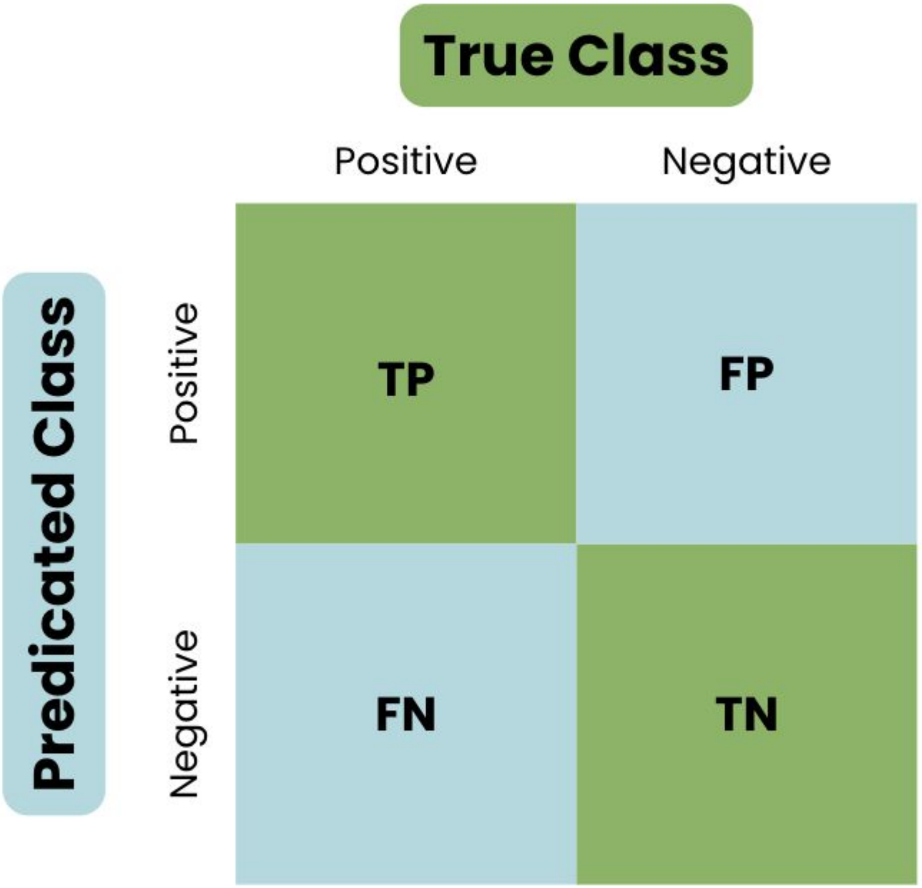
Evaluation metrics allow us to **quantify** how good or bad our model is, beyond just accuracy. They guide us in:

- Choosing between models
- Tuning models
- Understanding model behavior (e.g., bias, variance, imbalance)

Evaluation Metrics



COVID-19 detection



True Positive (TP): test says Positive, and reality Positive

False Positive (FP): test says Positive, and reality Negative

False Negative (FN): test says Negative, and reality Positive

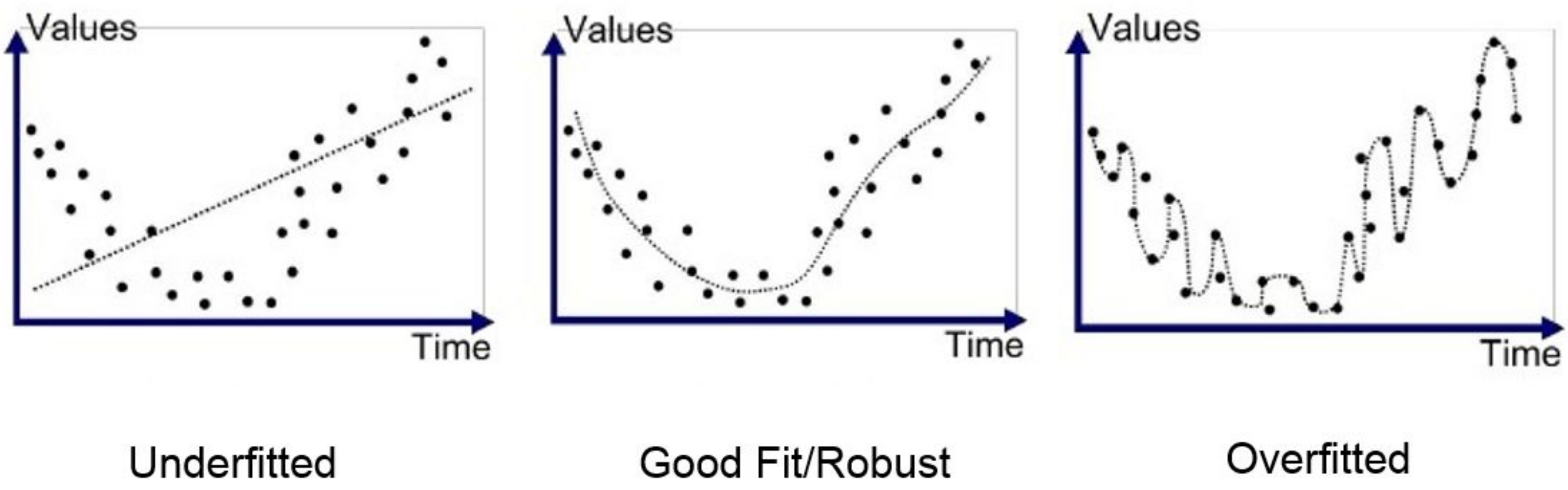
True Negative (TN): test says Negative, and reality Negative

F1 Score

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

- **Harmonic mean of precision and recall**
- **Useful when you need balance between precision and recall**

Overfitting and Underfitting



Overfitting

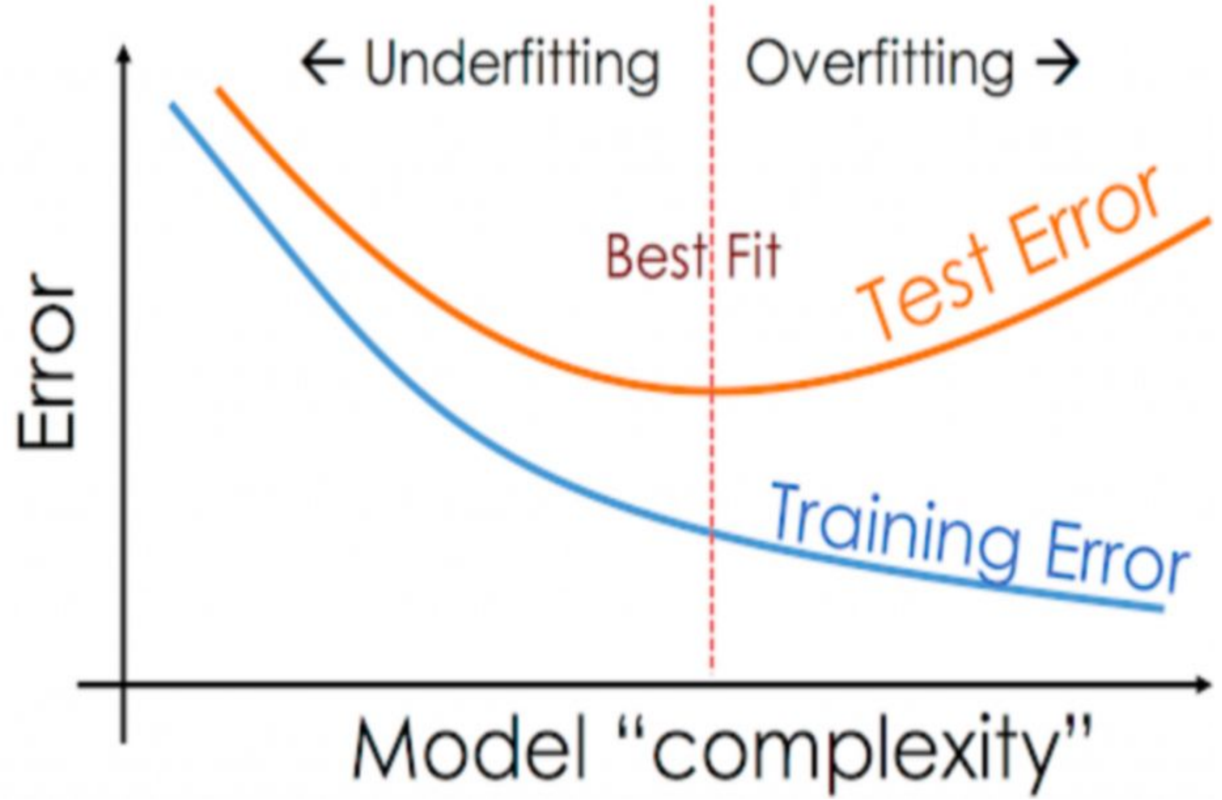
- Memorizes training data.
- Performs well on training set.
- Performs poorly on test/validation set.

Underfitting

- High training error
- High validation/test error.
- Model too simplistic.

Handling Overfitting and Underfitting

- **Overfitting**
  - Increasing the model complexity
  - Reducing regularization
  - Adding features to training data
- **Underfitting**
  - Adding more data
  - Data augmentation
  - Regularization
  - Removing features from data



In partnership with

RÉPUBLIQUE FRANÇAISE

Liberté
Égalité
Fraternité

AFD
AGENCE FRANÇAISE DE DÉVELOPPEMENT

apptrainers
Translating ideas to reality through innovative education

Classification

efe
Jordan

التعليم لأجل التوظيف
education for employment

Classification

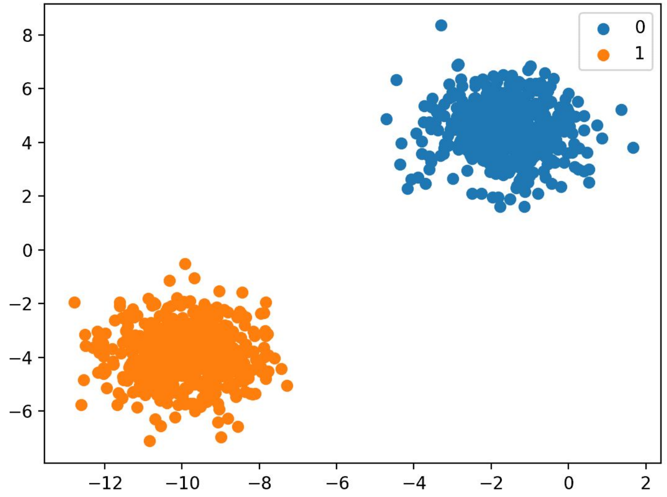
Classification is a fundamental task in supervised machine learning where the goal is to assign discrete labels to input instances based on their features by building a decision boundary that separates the data points into distinct classes based on the input features.

Key Classification Techniques:

- Logistic regression
- K nearest neighbors
- Support vector classification (SVM)
- Naïve-Bayes

Binary Classification

- Binary classification involves predicting between **two possible outcomes**.
- This is the simplest form of classification, with examples such as email **spam detection** (spam or not spam) or disease prediction (positive or negative).
- The model learns to **separate the data points into two distinct classes** based on their features.



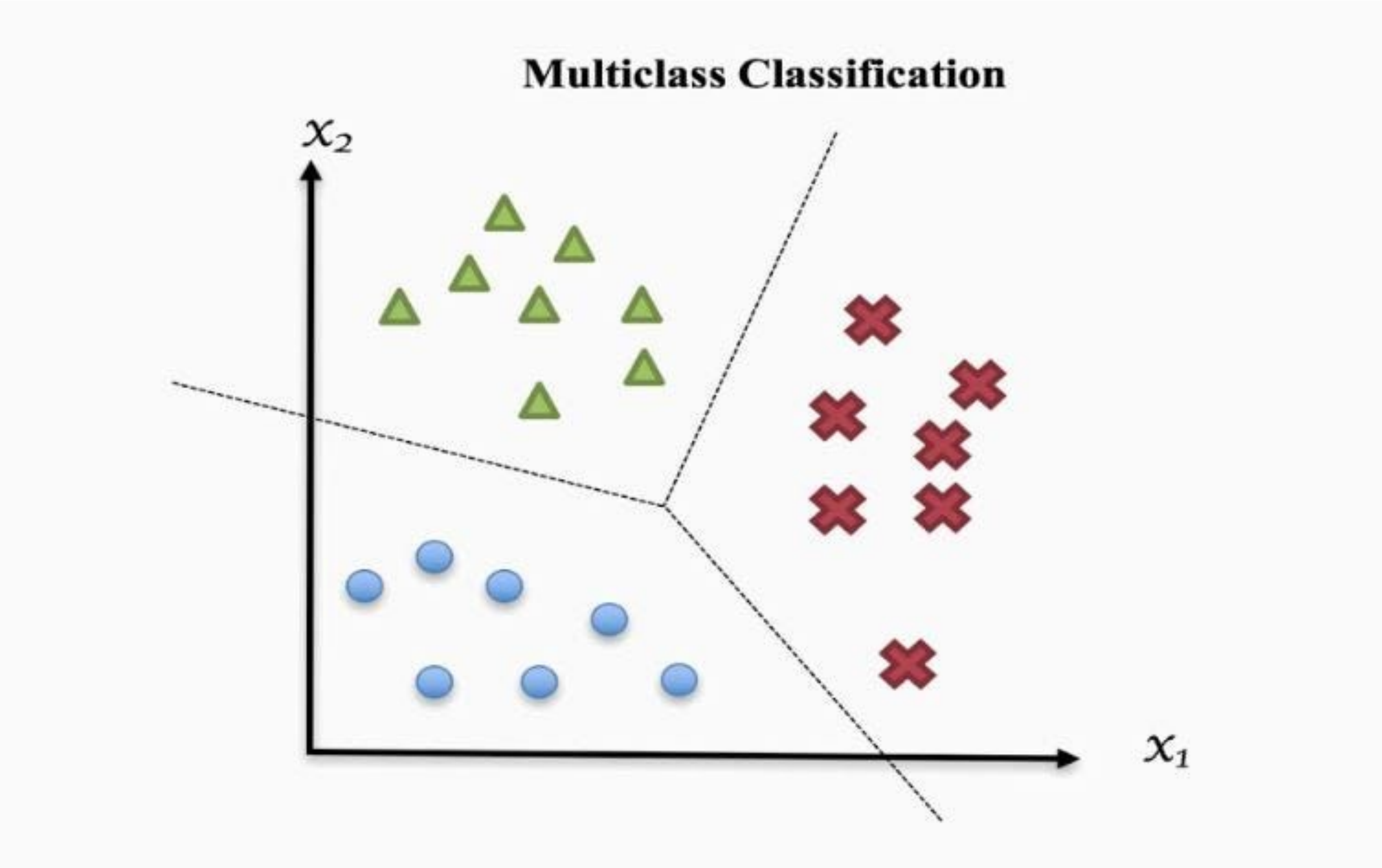
Multi-Class Classification

Each instance (data point) belongs to exactly one class out of multiple possible classes.

Example:

- Classifying types of fruit: apple, banana, orange — each fruit is only one type.
- Handwritten digit recognition (0-9), each image is exactly one digit.
- Output: One class label per instance.

Multi-Class Classification



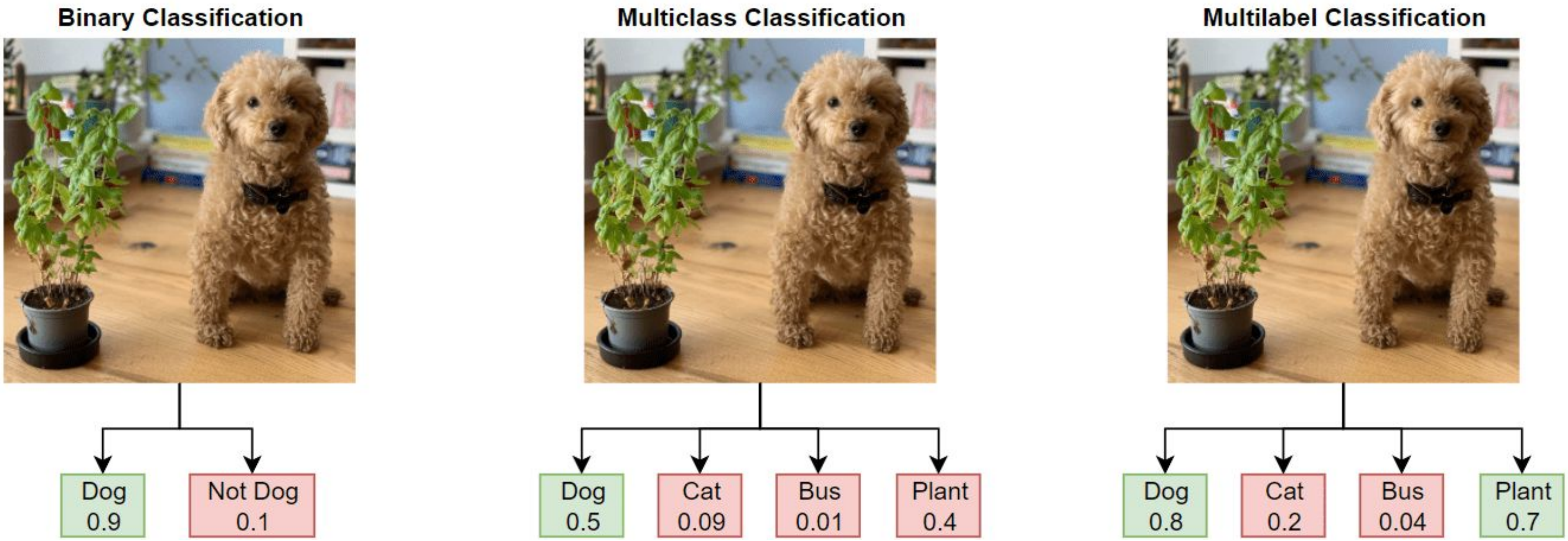
Multi-Label Classification

Each instance can belong to multiple classes simultaneously.

Example:

- Tagging a news article with multiple topics like politics, sports, and economy at the same time.
- Detecting objects in an image: an image can contain both a dog and a cat.

Output: A set (or vector) of labels per instance.

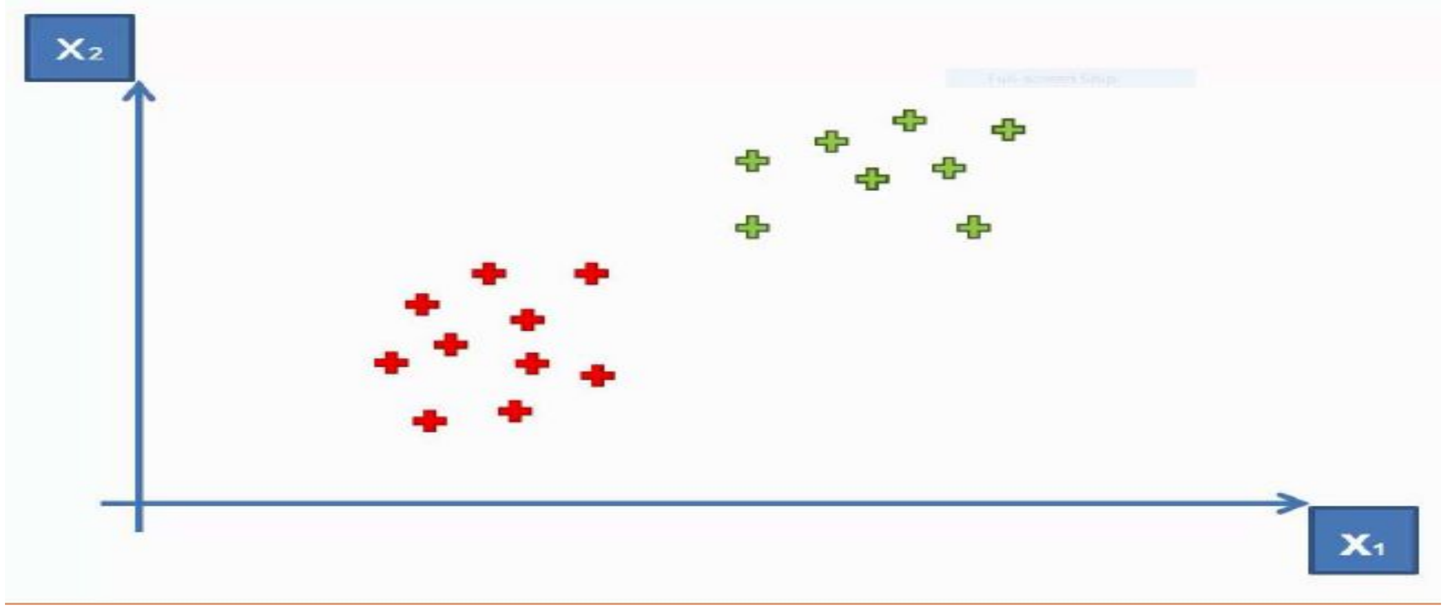

Binary Classification
Multiclass Classification
Multilabel Classification
Dog 0.9
Not Dog 0.1
Dog 0.5
Cat 0.09
Bus 0.01
Plant 0.4
Dog 0.8
Cat 0.2
Bus 0.04
Plant 0.7

Support Vector Machine

SVM

Support Vector Machine (SVM) is a **supervised** machine learning algorithm used for **classification** and sometimes regression problems.

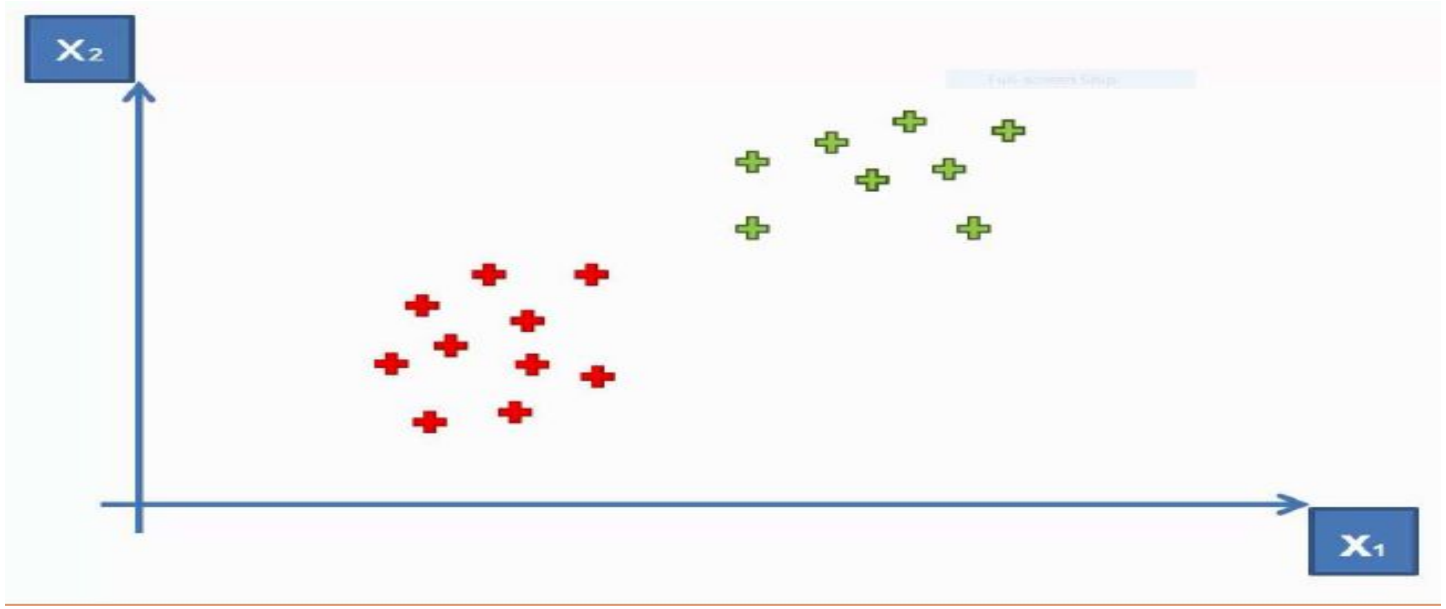
Its **main idea** is to find the best boundary (hyperplane) that separates different classes..



SVM

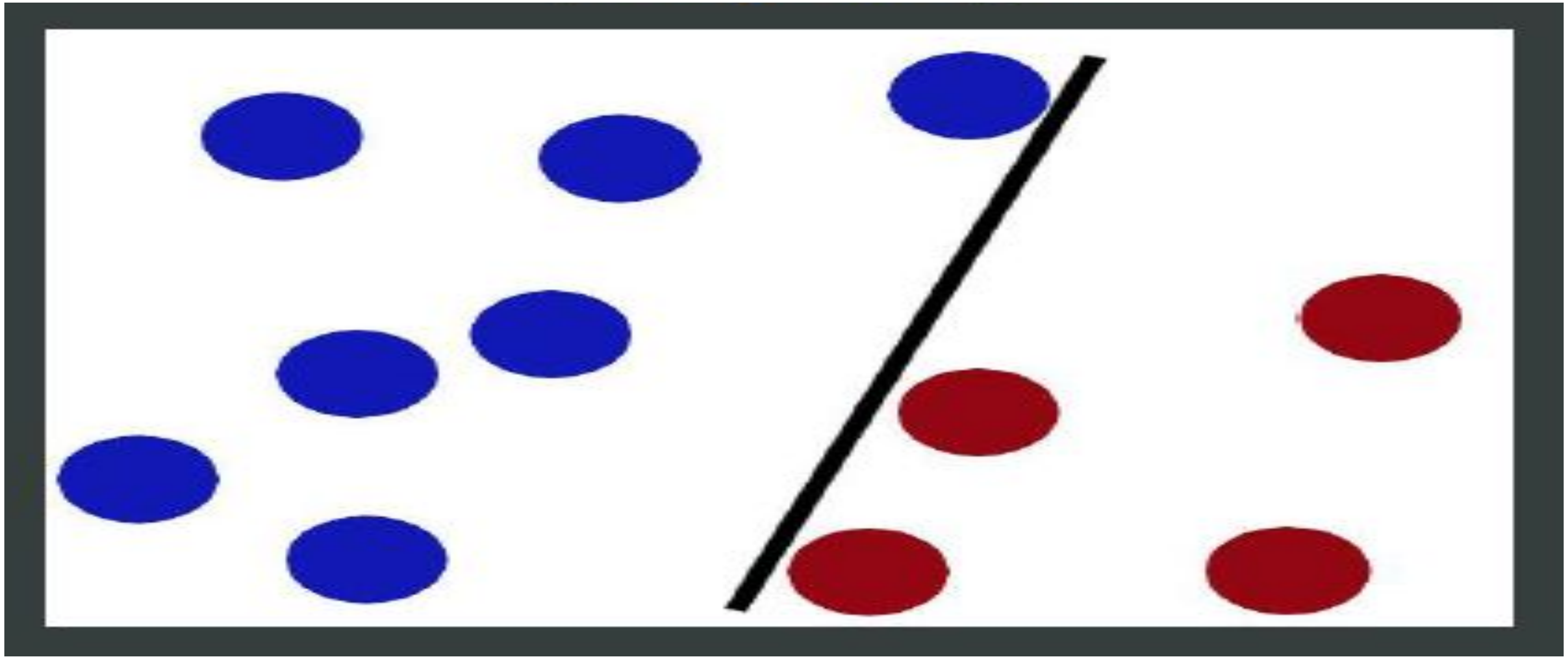
Support Vector Machine (SVM) is a **supervised** machine learning algorithm used for **classification** and sometimes regression problems.

Its **main idea** is to find the best boundary (hyperplane) that separates different classes..



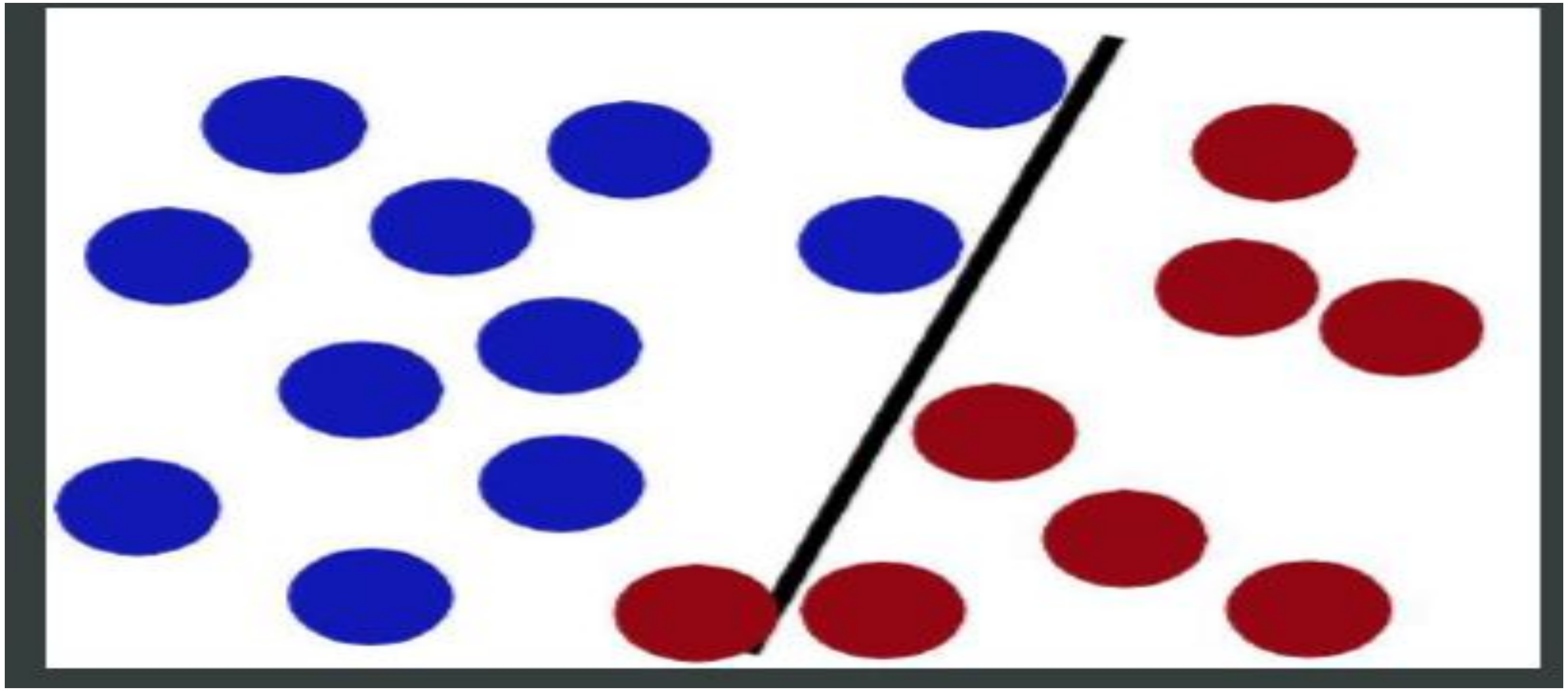
Support Vector Machine (SVM) – Example 1

- We have **2 colors of balls** on the table that we want to separate, we get a stick and put it on table, pretty well right?



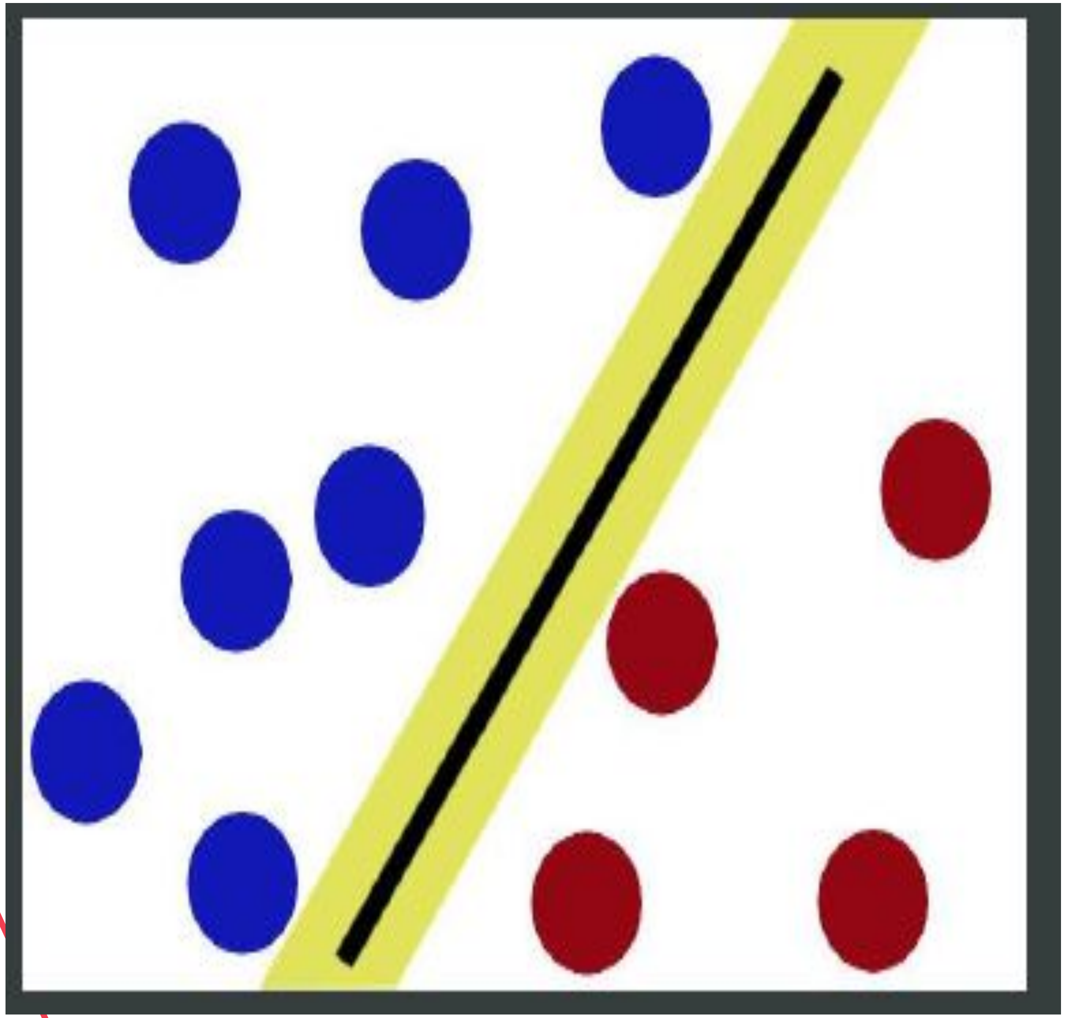
Support Vector Machine (SVM) – Example 1

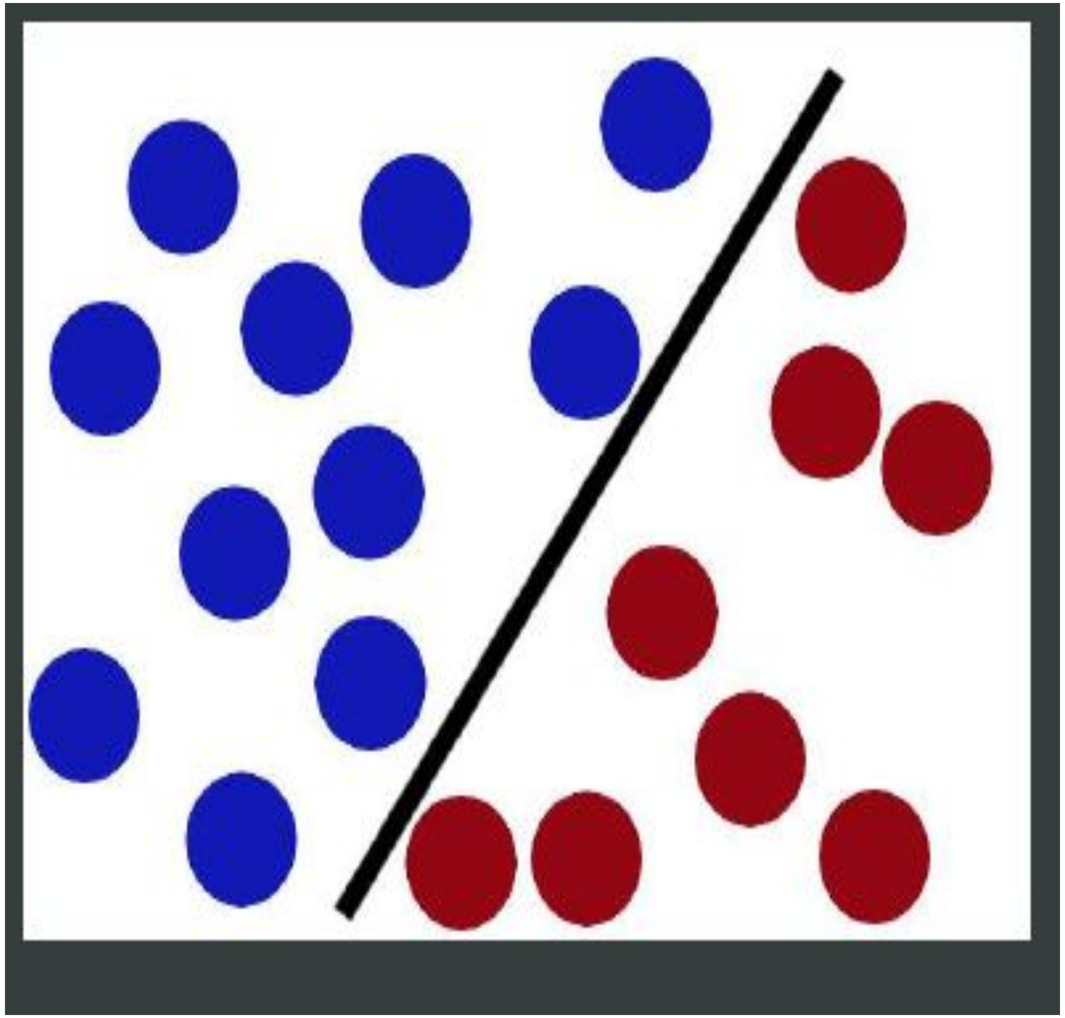
- Some **villain comes** and places more balls on the table, it kind of works but one of the balls is on the **wrong side** and there is probably **a better place to put the stick now.**



Support Vector Machine (SVM) – Example 1

- SVMs try to put the stick in the best possible place by having as **big a gap** on either side of the stick as possible.
- **Solution - Next Slide**





Support Vector Machine (SVM)

- **Hyperplane**: A decision boundary that separates classes.
- **Margin**: The distance between the hyperplane and the closest points from each class.
- **Support Vectors**: The data points that lie closest to the hyperplane — they “support” the decision boundary.

Support Vector Machine (SVM) - Example

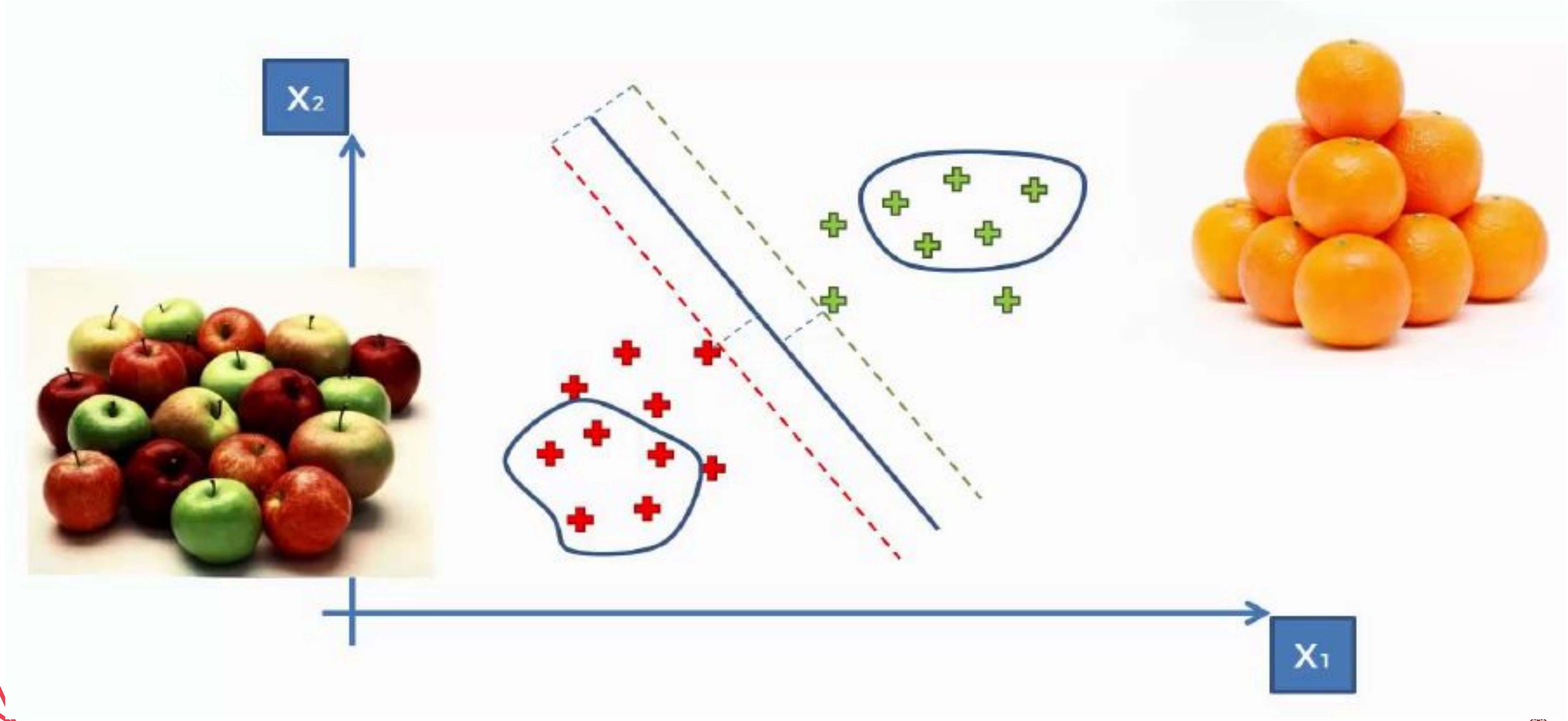
In partnership with

RÉPUBLIQUE FRANÇAISE

AFD AGENCE FRANÇAISE DE DÉVELOPPEMENT

Support Vector Machine (SVM) - Example

apptrainers



efe Jordan
التعليم لأجل التوظيف
education for employment

Support Vector Machine (SVM) - Example