

# Fouille de Données

# Data Mining

## **Classification - Partie 5**

# Plan du cours

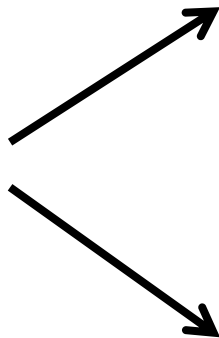
1. Les Machines à Vecteurs de Support - SVM.
2. SVM linéaire.
3. Implémentation des SVM.
4. Exemple déroulement.

# Classification : Algorithmes

**SAVOIR - PREDIRE - DECIDER**



Données

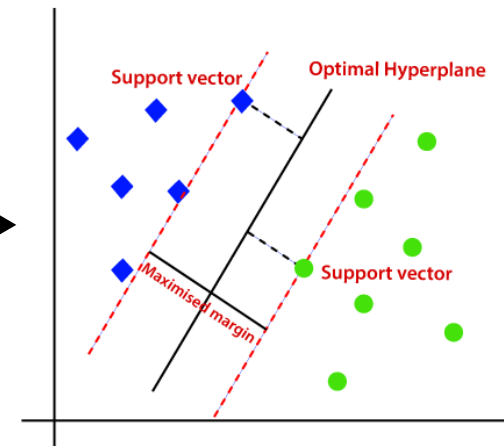


#114983146

**SVM**



$$H(X) = W^T X + b$$



Connaissances

# Machines à Vecteurs de Support

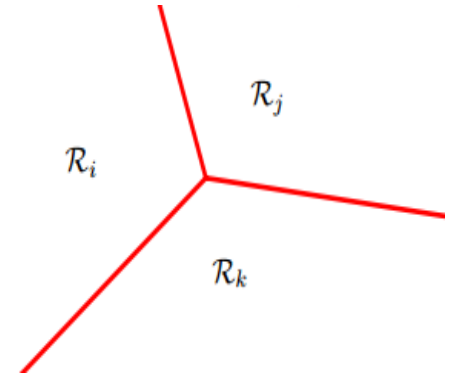
➤ Méthode de classification de données **linéaires** et **non linéaires**.

➤ Introduite par Vladimir Vapnik au début des années 90.

➤ Machine => Algorithme. Supervisé.

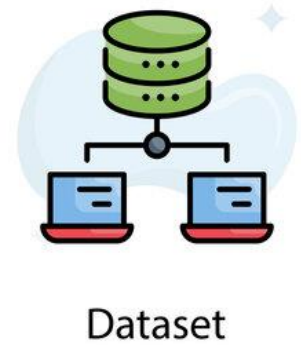
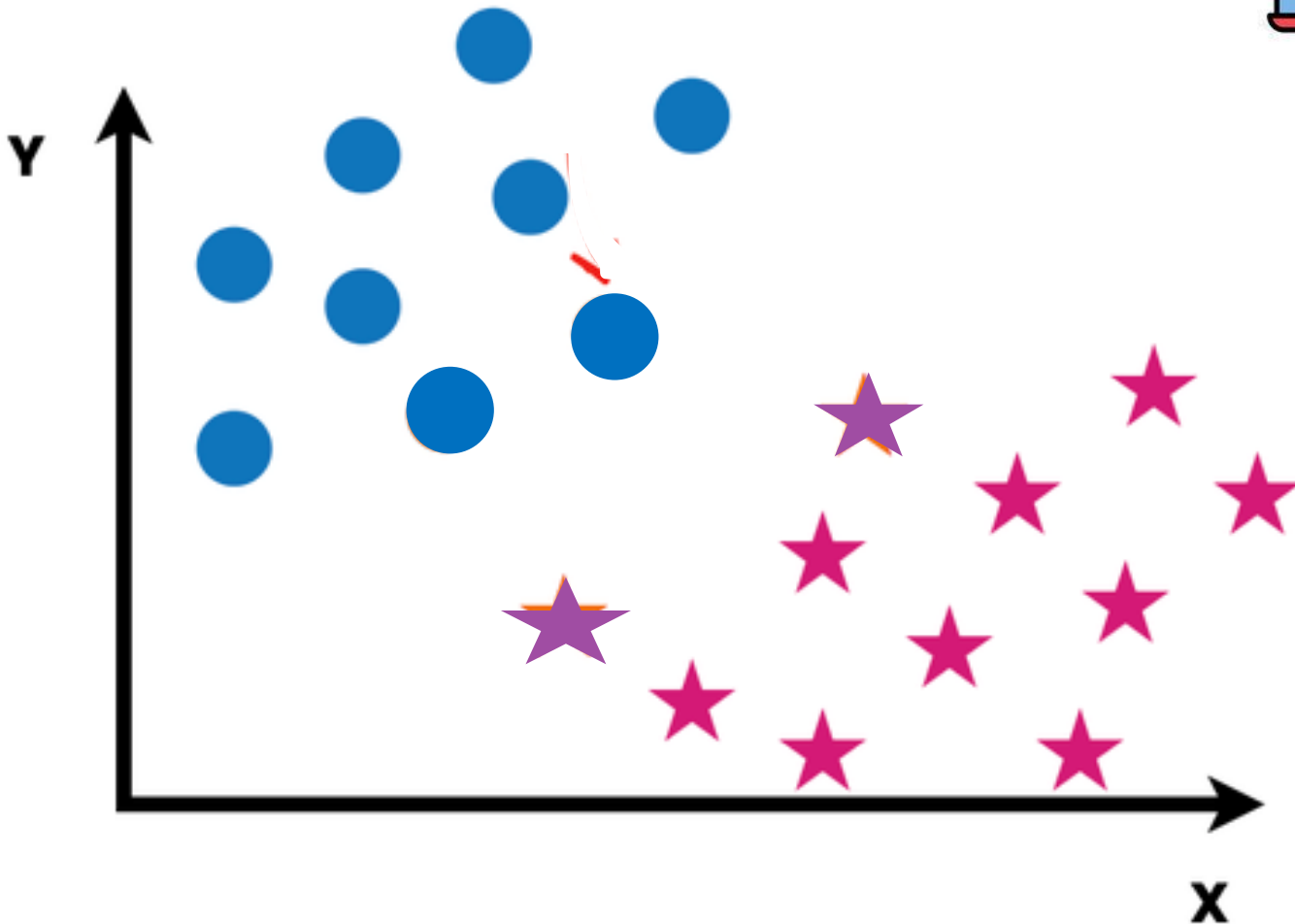
➤ **Principe** de base :

1. Rechercher un **hyperplan** qui sépare le mieux l'espace des exemples en différentes **régions de décisions**.
2. Une **frontière de décision** séparant les exemples d'une classe à une autre. Chaque région est associée à une classe.
3. Une recherche qui s'appuie uniquement sur quelques exemples de la base d'entraînement, appelés **Vecteurs Support**, ainsi que la **marge** entre eux.



# Machines à Vecteurs de Support

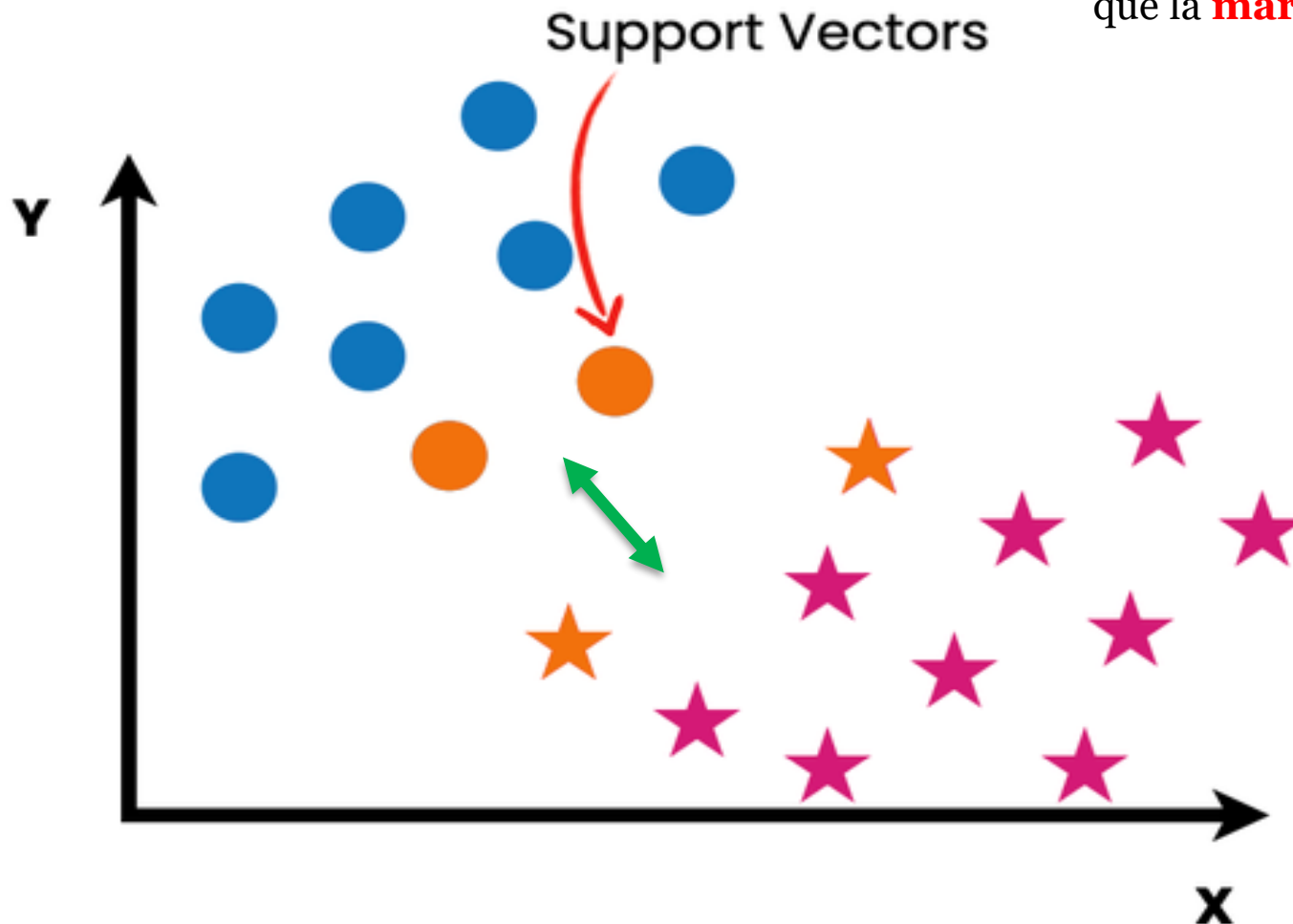
## Principe de base:



# Machines à Vecteurs de Support

- **Principe de base:** Comment ?

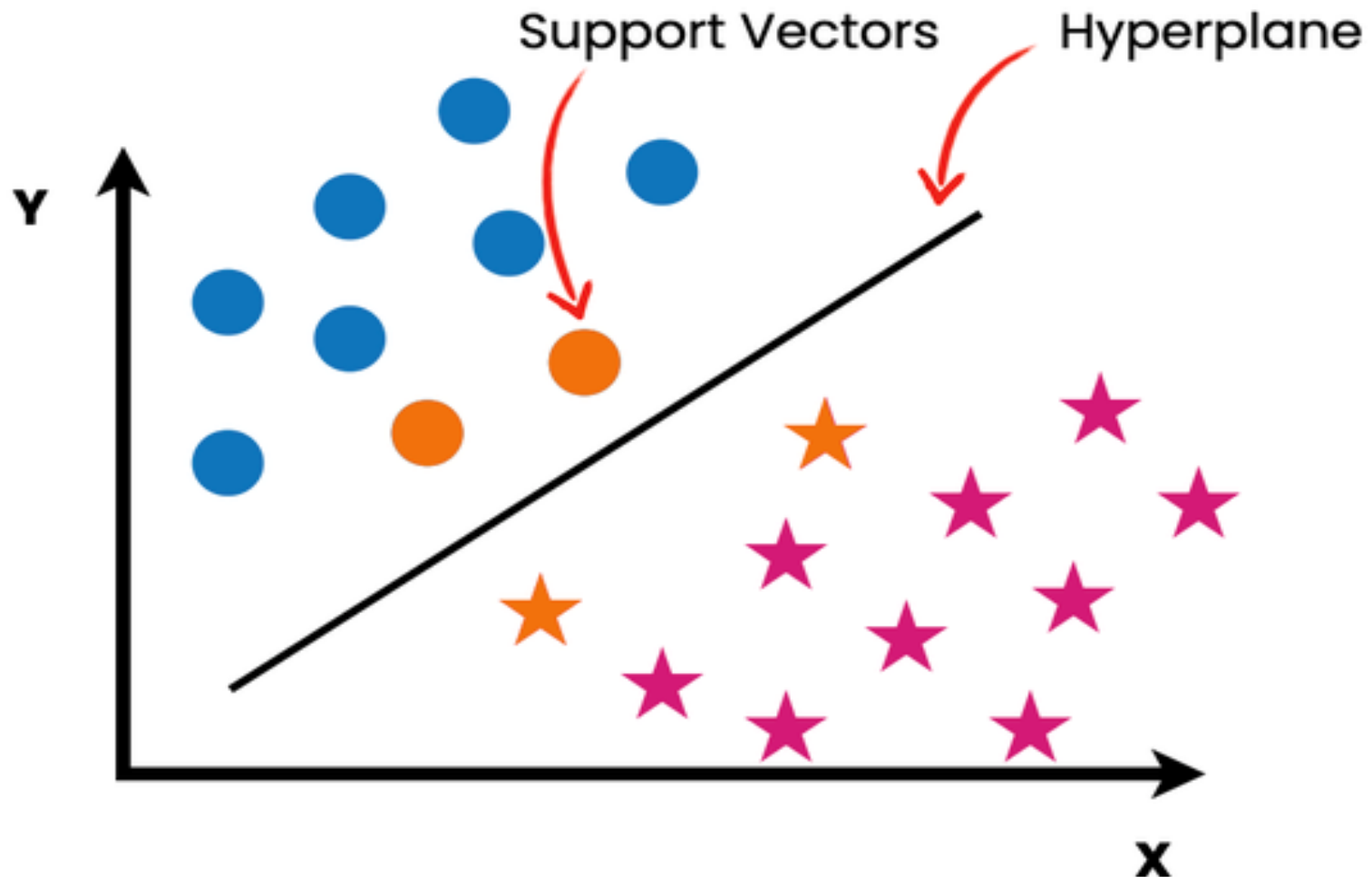
S'appuie sur quelques exemples de la base d'entraînement, appelés **Vecteurs Support**, ainsi que la **marge** entre eux.



# Machines à Vecteurs de Support

Rechercher un **hyperplan** qui sépare le mieux l'espace des exemples en différentes **régions de décisions**.

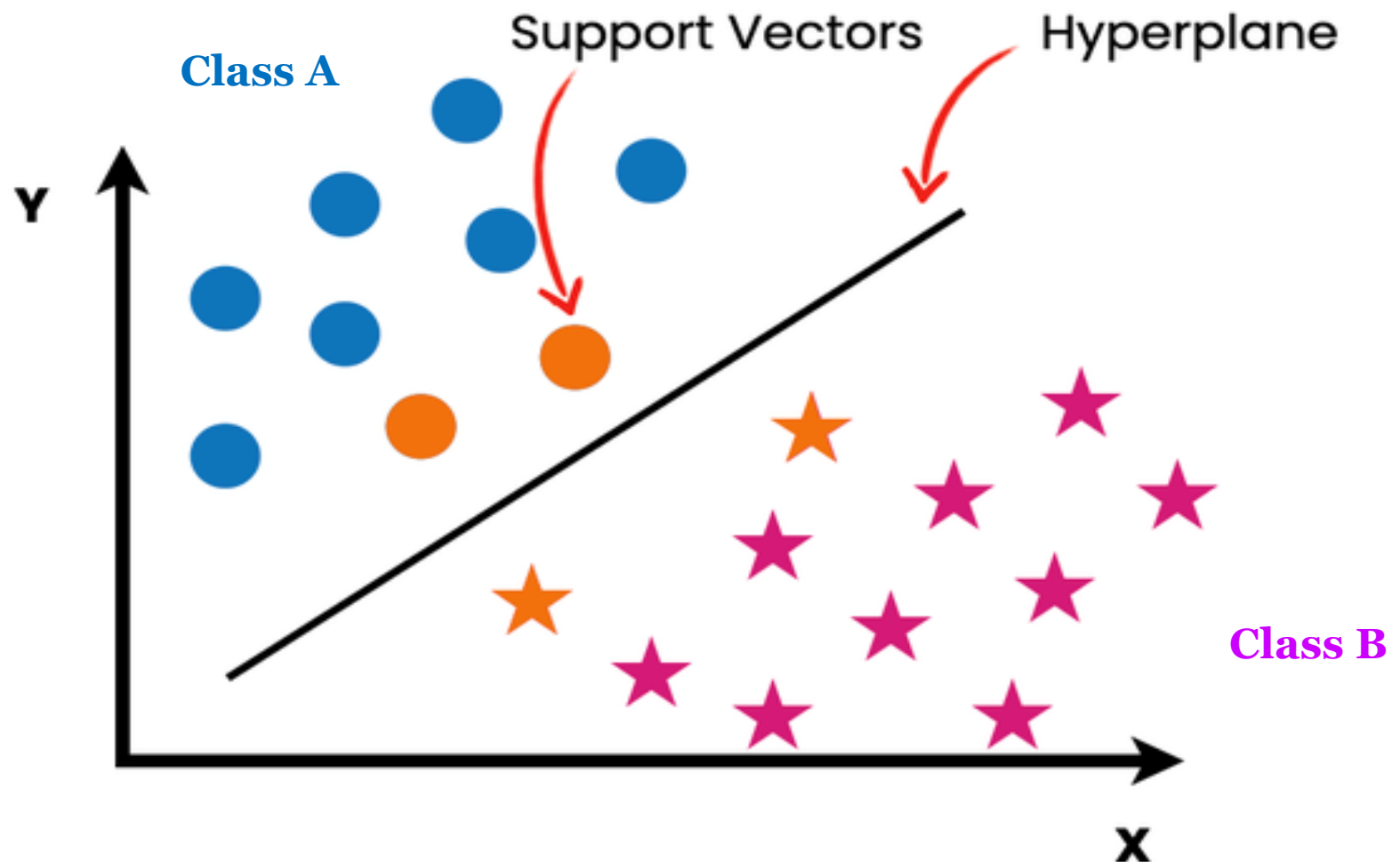
- **Principe de base:** Pourquoi ?



# Machines à Vecteurs de Support

Rechercher un **hyperplan** qui sépare le mieux l'espace des exemples en différentes **régions de décisions**.

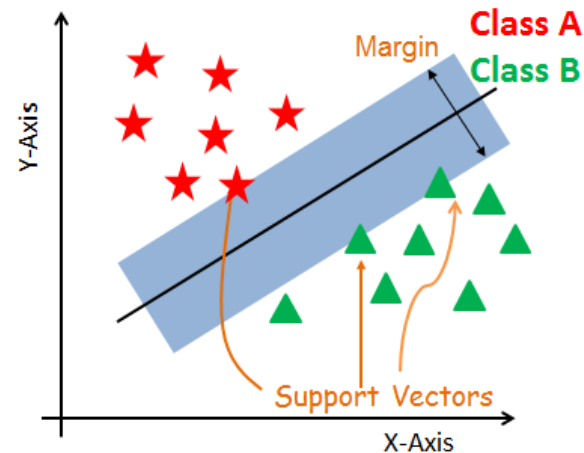
- **Principe de base:** Pourquoi ?





# Machines à Vecteurs de Support

## ➤ Concepts de base en SVM :



- **Marge** - La marge est l'écart entre l'hyperplan et les vecteurs de support.
- **Hyperplan** - Les hyperplans sont des frontières de décision qui aident à classer les points de données.
- **Vecteurs de support** - Les vecteurs de support sont les points de données qui se trouvent sur ou à proximité de l'hyperplan et influencent la position de l'hyperplan. Les points les plus proches de la frontière.
- **Fonction noyau** - Ce sont les fonctions utilisées pour déterminer la forme de l'hyperplan et de la frontière de décision.

# Machines à Vecteurs de Support

- **Principe de base:**



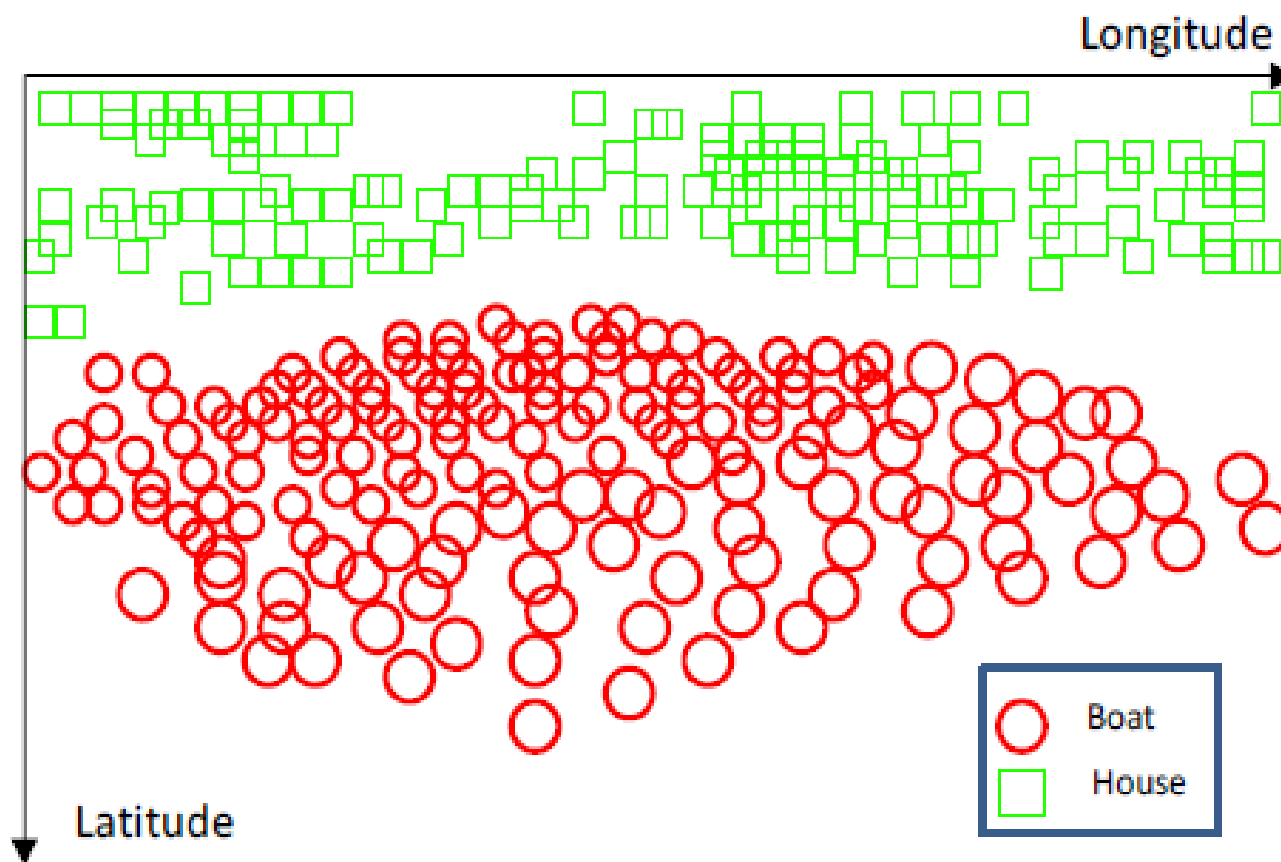
# Machines à Vecteurs de Support

- **Principe de base:**



# Machines à Vecteurs de Support

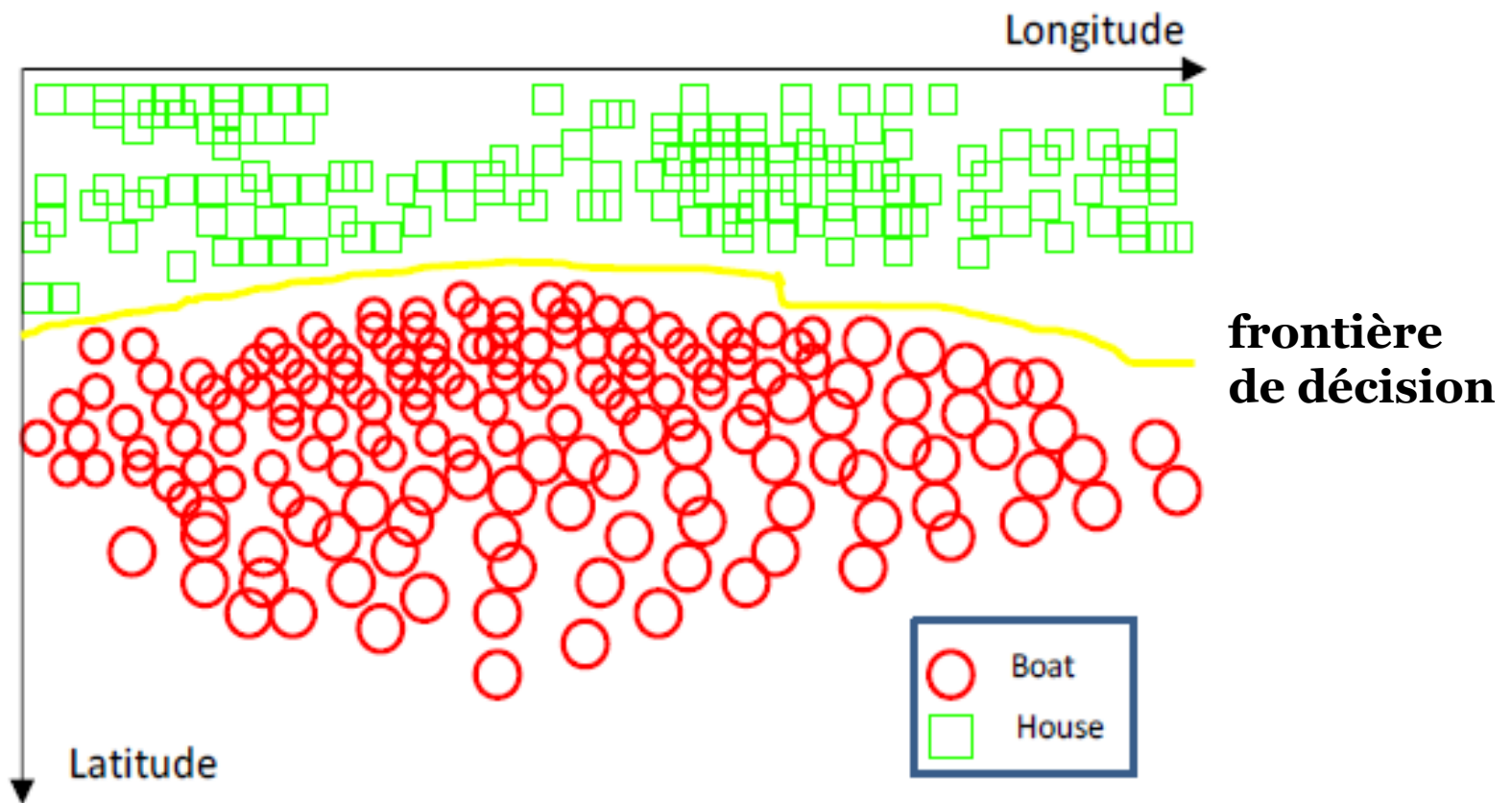
- **Principe de base:**





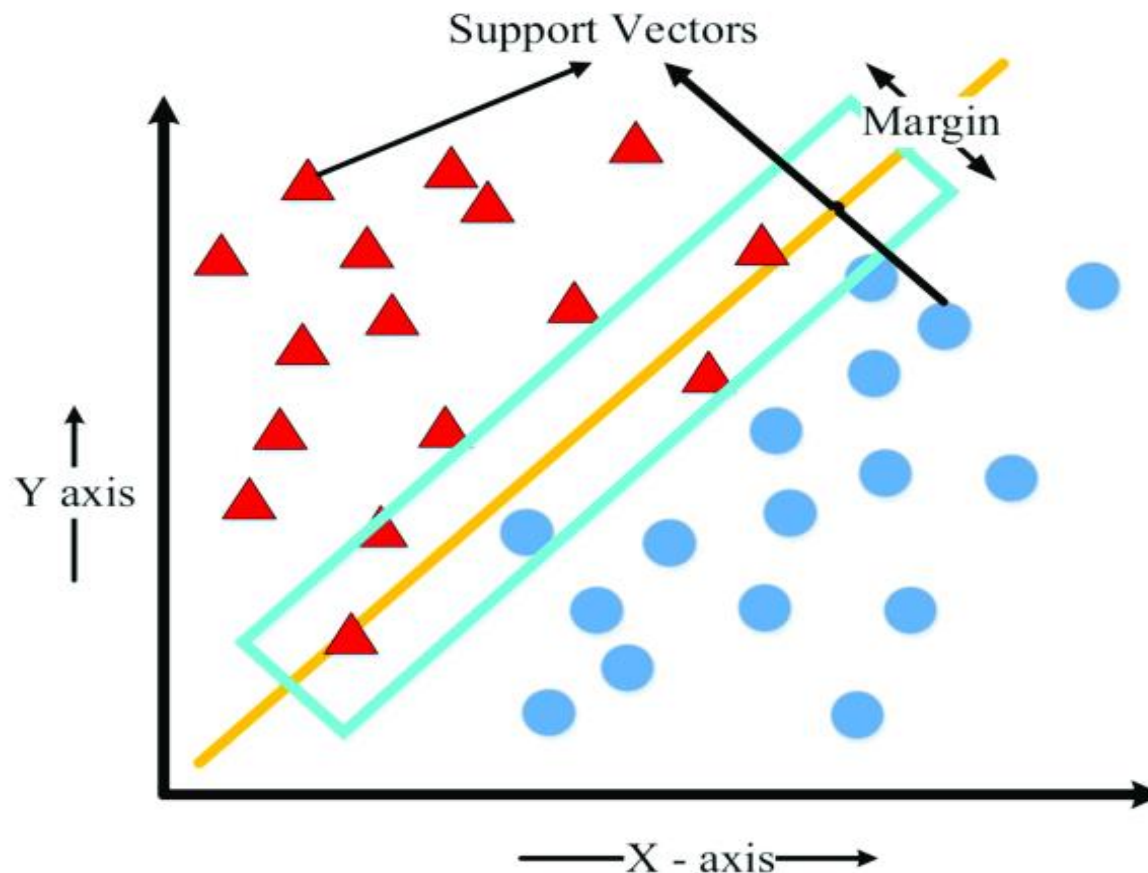
# Machines à Vecteurs de Support

- **Principe de base:**



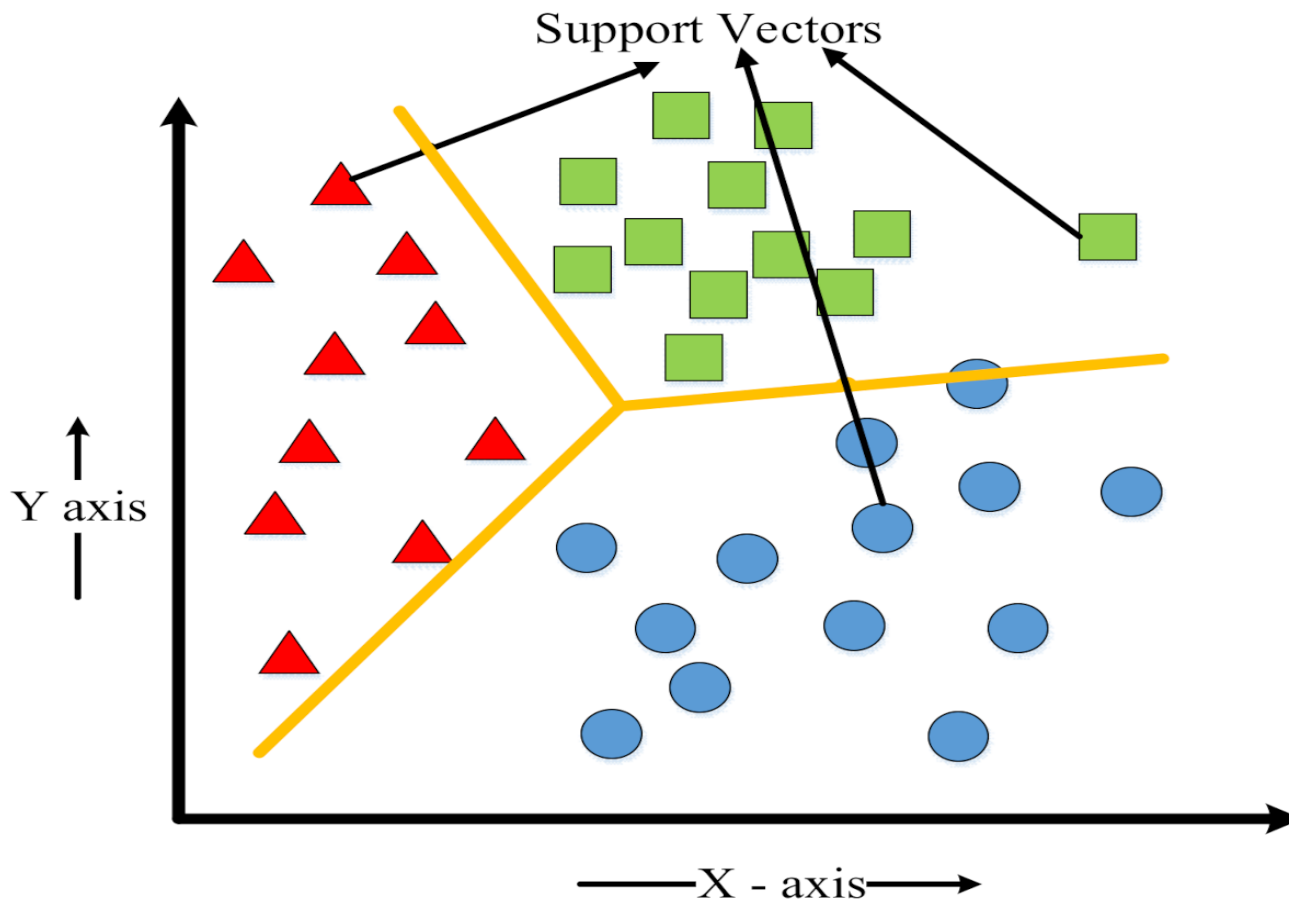
## Machines à Vecteurs de Support

- **L'hyperplan:** Cas simple d'une classification **binaire** à deux attributs =>  
**Plan 2D, 2 régions de décision, séparateur = droite**



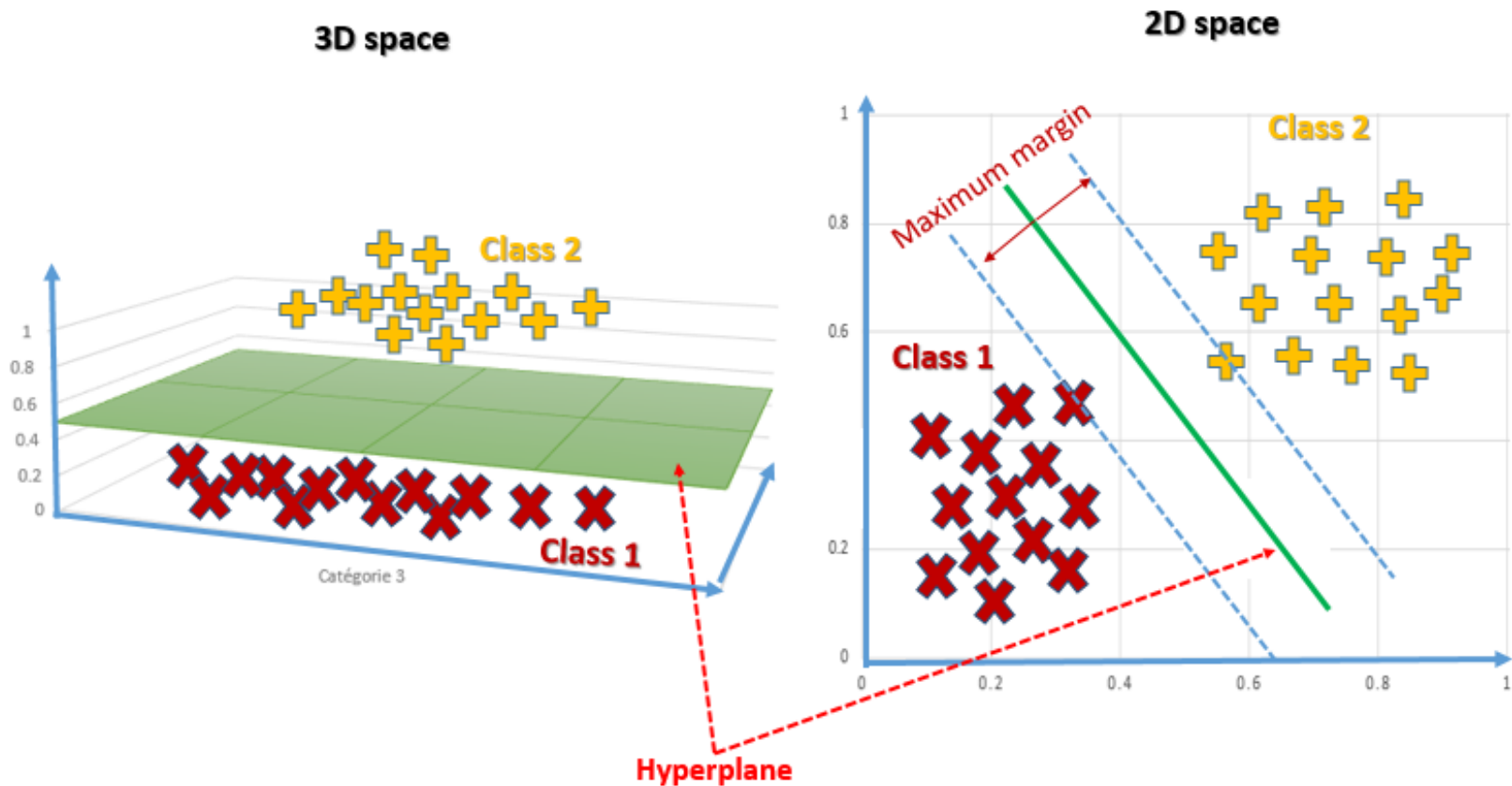
# Machines à Vecteurs de Support

- **L'hyperplan:** Cas d'une classification **multi-classes** à deux attributs =>  
**Plan 2D, 3 régions**



# Machines à Vecteurs de Support

- **L'hyperplan:** Cas d'une classification **binaire** à deux attributs => **Plan**  
**3D, 2 régions, séparateur = plan.** > 3D : hyperplan,

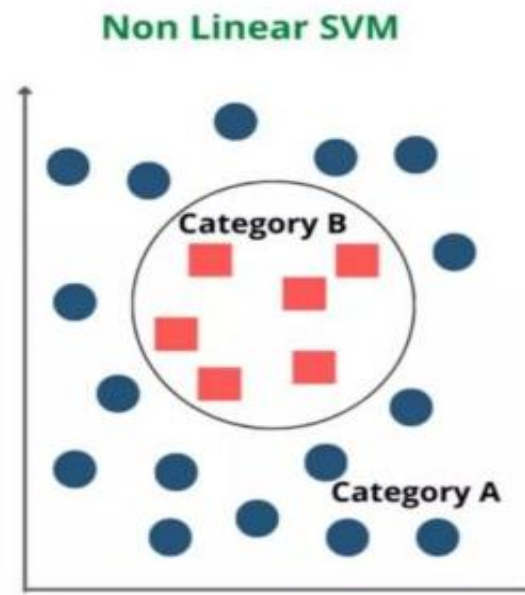
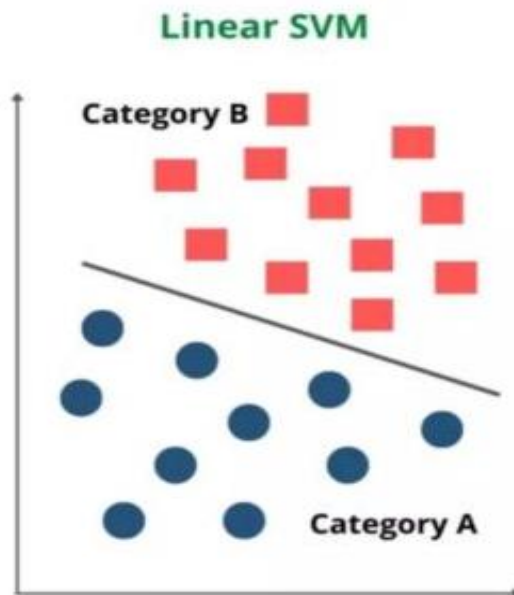




# Machines à Vecteurs de Support

## ▪ Deux Types:

- **SVM linéaire** – Les points de données peuvent être facilement séparés par une ligne linéaire.
- **SVM non linéaire** - Les points de données ne peuvent pas être facilement séparés par une ligne linéaire.

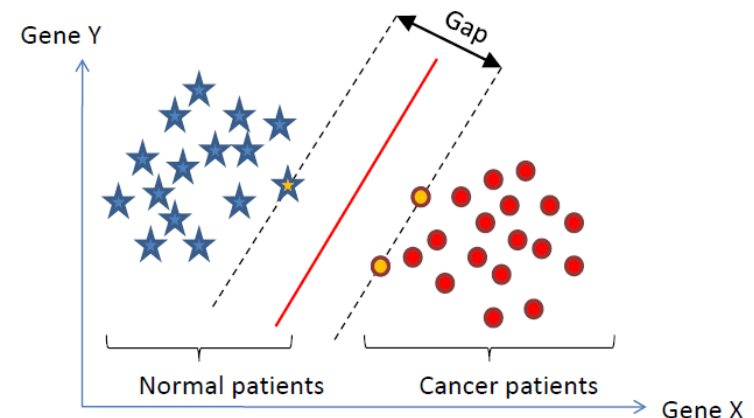
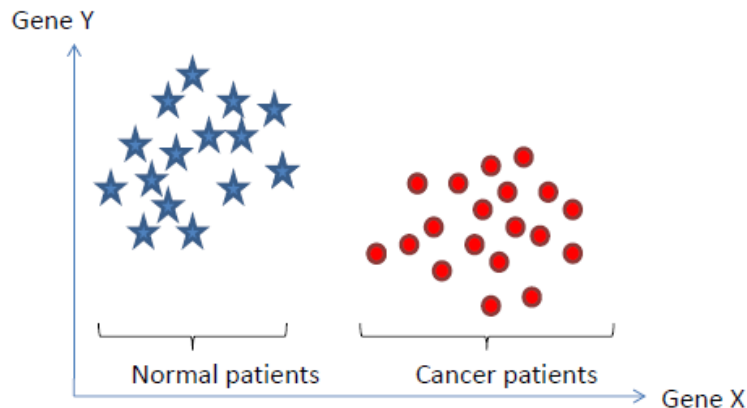
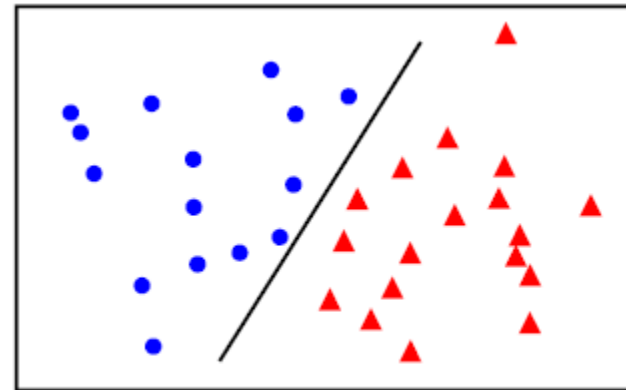
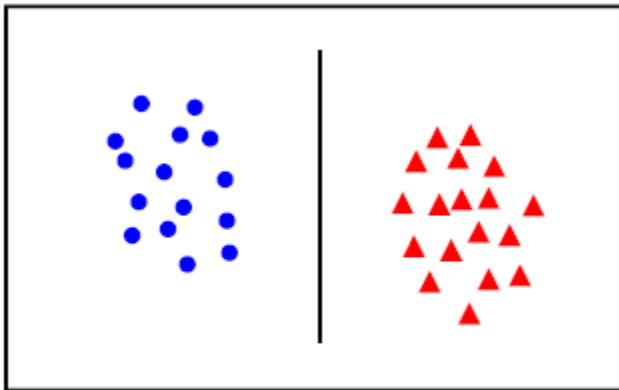


# Machines à Vecteurs de Support

Deux cas du problème SVM posé :

- **Séparable linéairement.**
- Non séparable linéairement.

On dit qu'un problème est **linéairement séparable** si une surface linéaire permet de classifier parfaitement.

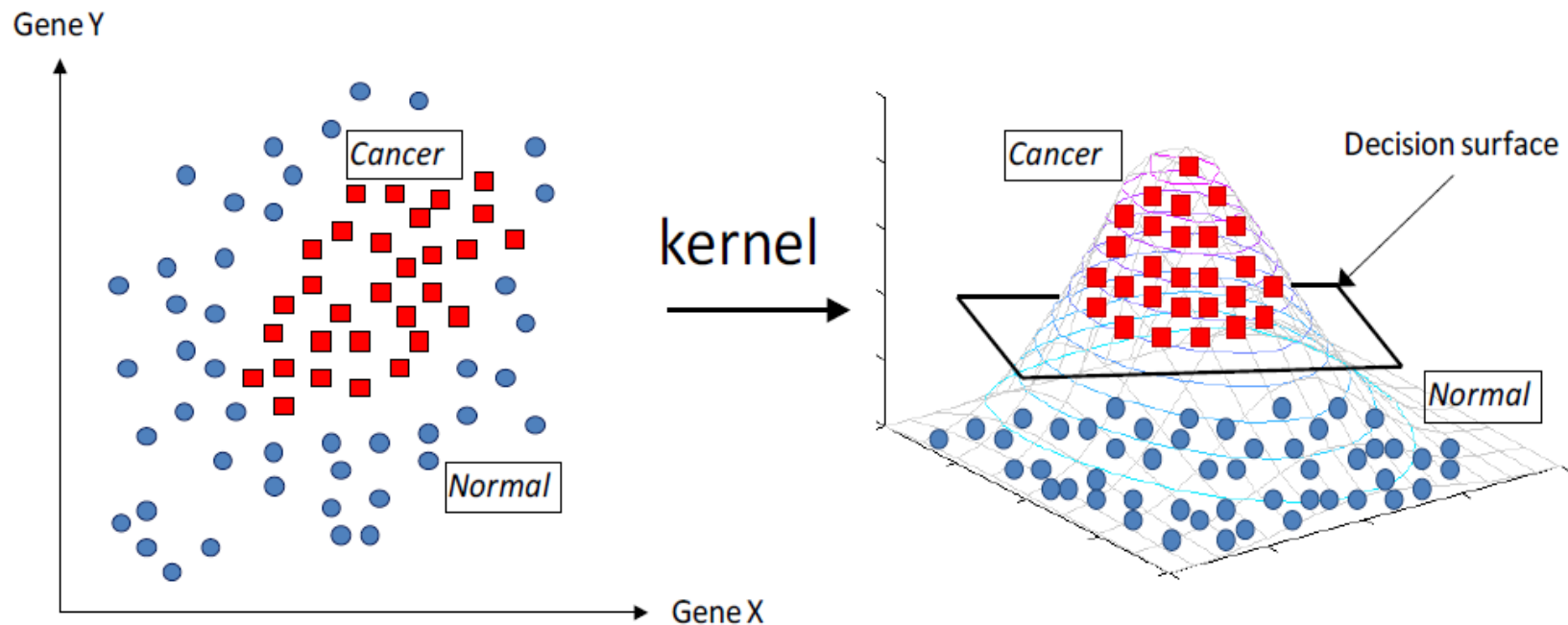


# Machines à Vecteurs de Support

Deux cas du problème posé :

- Séparable linéairement.
- **Non séparable linéairement.**
  - **Kernel Trick (noyau)**

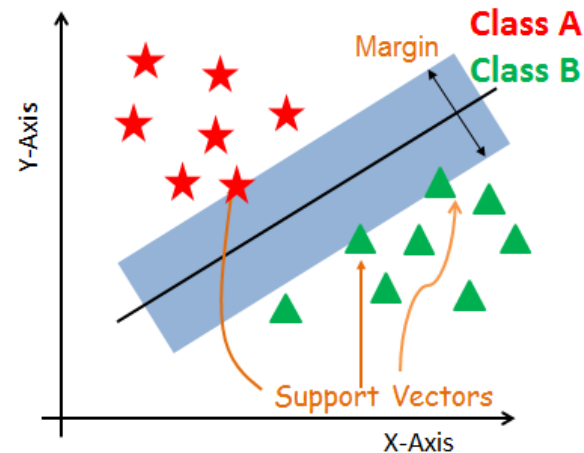
On dit qu'un problème est **linéairement séparable** si une surface linéaire permet de classifier parfaitement.



from the 2D nonlinear input space into a 3D output space using kernel functions

# Machines à Vecteurs de Support

## ➤ Concepts de base en SVM :



- **Marge** - La marge est l'écart entre l'hyperplan et les vecteurs de support.
- **Hyperplan** - Les hyperplans sont des frontières de décision qui aident à classer les points de données.
- **Vecteurs de support** - Les vecteurs de support sont les points de données qui se trouvent sur ou à proximité de l'hyperplan et influencent la position de l'hyperplan.
- **Fonction noyau** - Ce sont les fonctions utilisées pour déterminer la forme de l'hyperplan et de la frontière de décision.

## SVM Linéaire

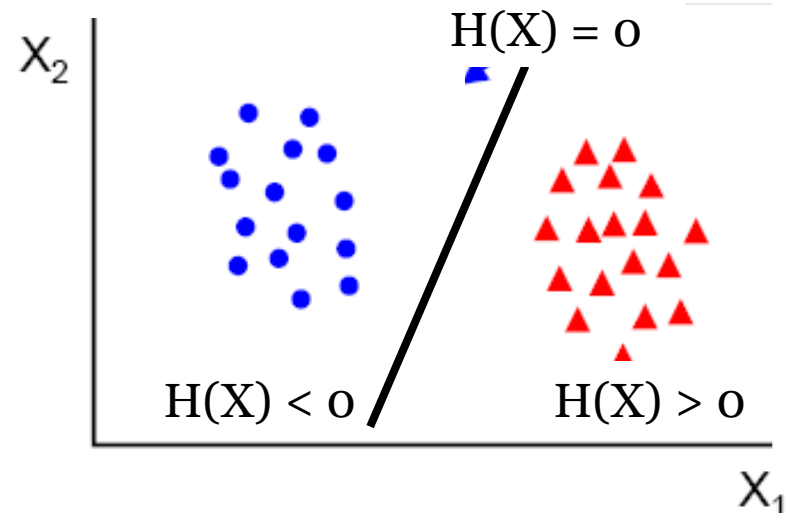
### Cas de séparabilité linéaire : **Linear SVM**

- Supposons une classification binaire, deux classes :
  - $Y_i = \{-1, +1\}$  : **Négative** et **Positive**
- Chaque exemple **X** est décrit par des attributs.
- Chaque exemple est un vecteur.
- L'**hyperplan** séparateur (la droite en 2D) a cette forme :

$$H(X) = W^T X + b$$

**W**: Vecteur de poids

**b** : biais



# SVM Linéaire

## Cas de séparabilité linéaire

➤ Supposons deux classes :

- $Y_i = \{-1, +1\}$

➤ Chaque exemple  $\mathbf{X}$  est décrit par des attributs.

➤ Chaque exemple est un vecteur.

➤ L'**hyperplan** séparateur (la droite en 2D) a cette forme :

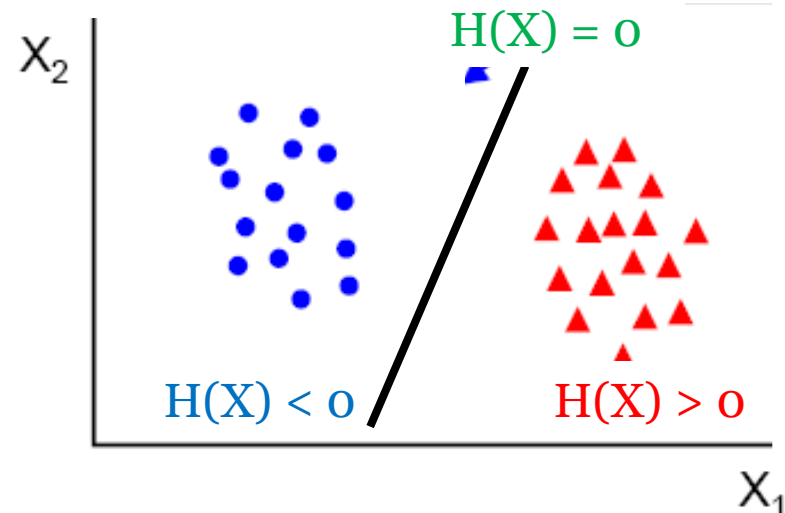
$$H(\mathbf{X}) = \mathbf{W}^T \mathbf{X} + b$$

**W**: Vecteur de poids

**b** : biais

**Problème** : rechercher une **fonction discriminante**  $H(\mathbf{X})$  qui prend  $\mathbf{X}$  en entrée et donne sa classe  $C_k$  en sortie.

=> **Entraînement** : Rechercher les meilleures valeurs de **W** et de **b**.

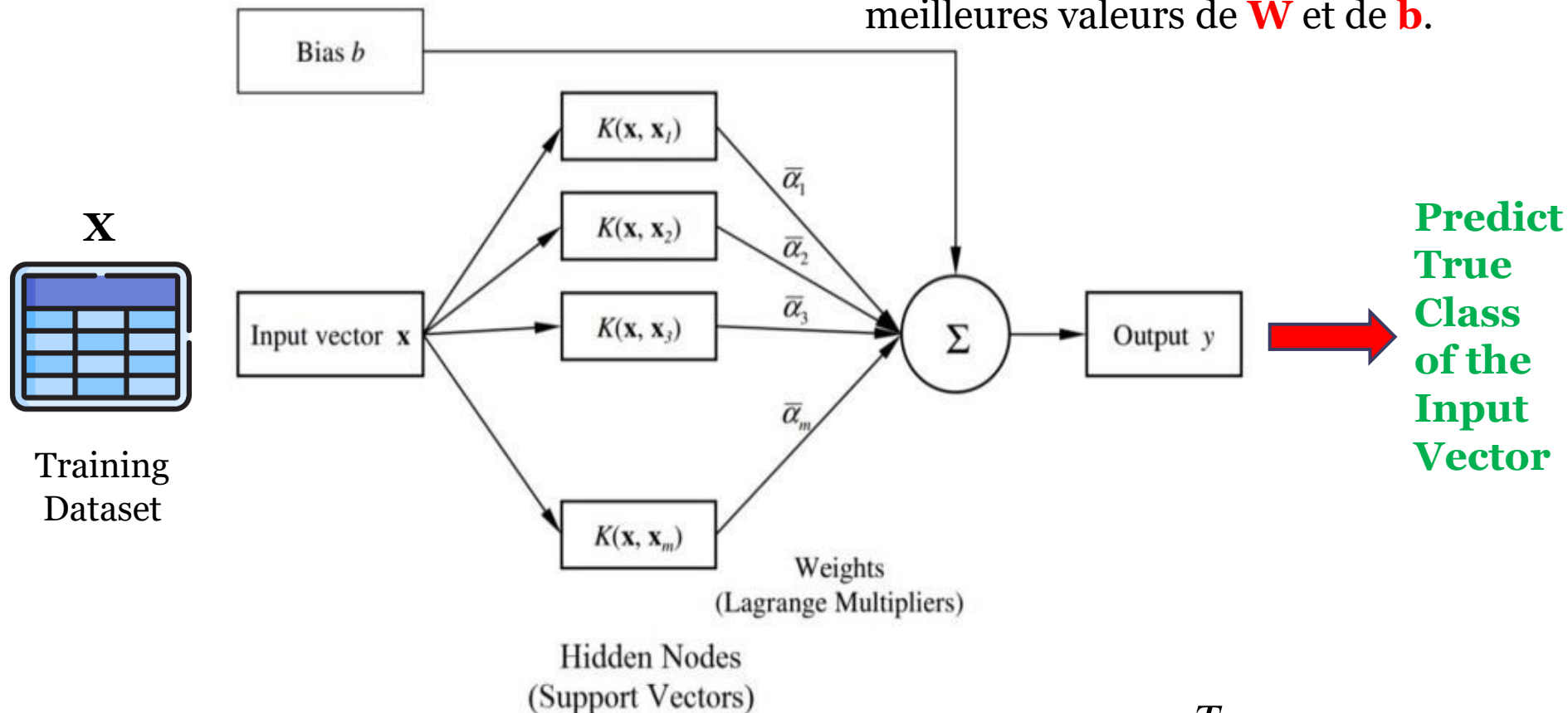


# SVM Linéaire

## Cas de séparabilité linéaire

**Problème** : rechercher une **fonction discriminante  $H(X)$**  qui prend  $X$  en entrée et donne sa classe  $C_k$  en sortie.

=> **Entraînement** : Rechercher les meilleures valeurs de  **$W$**  et de  **$b$** .



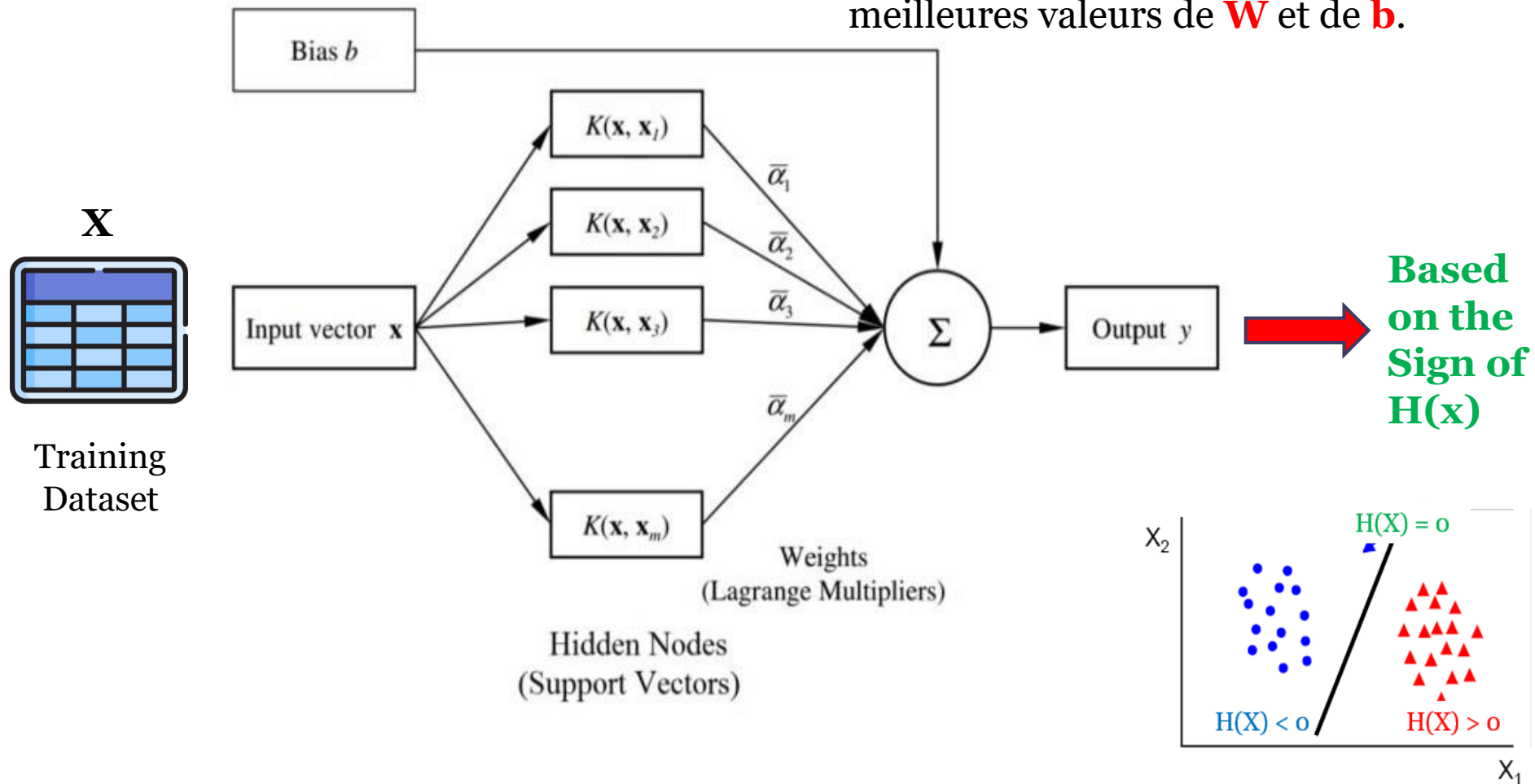
$$H(X) = W^T X + b$$

# SVM Linéaire

## Cas de séparabilité linéaire

**Problème** : rechercher une **fonction discriminante  $H(X)$**  qui prend  $X$  en entrée et donne sa classe  $C_k$  en sortie.

=> **Entraînement** : Rechercher les meilleures valeurs de  **$W$**  et de  **$b$** .





## SVM Linéaire

### Cas de séparabilité linéaire

➤ Supposons deux classes :

- $Y_i = \{-1, +1\}$

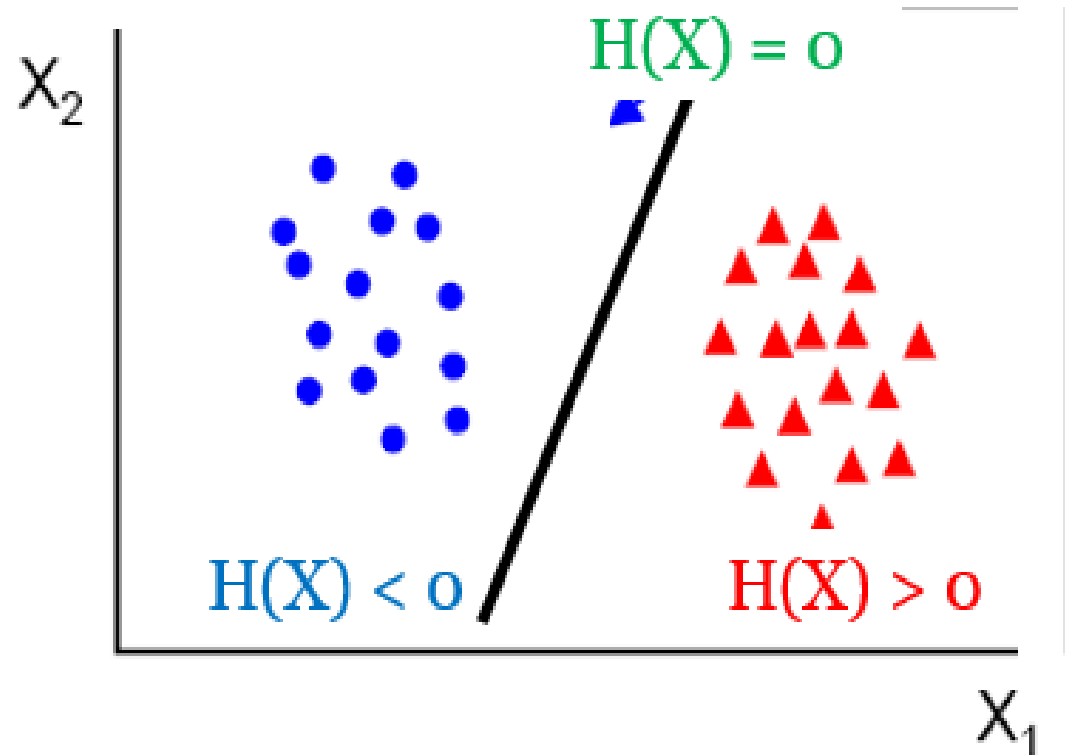
$$H(X) = W^T X + b$$

**W**: Vecteur de poids

**b** : biais

**Problème** : rechercher une **fonction discriminante**  $H(X)$  qui prend  $X$  en entrée et donne sa classe  $C_k$  en sortie.

=> **Entraînement**: Rechercher les meilleures valeurs de **W** et de **b**.



## SVM Linéaire

### Cas de séparabilité linéaire

$$H(X) = W^T X + b$$

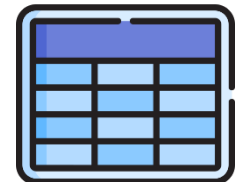
$$\left\{ \begin{array}{ll} C_k = +1 & \text{Si } H(X) > 0 : H+ \\ C_k = -1 & \text{Si } H(X) < 0 : H- \end{array} \right. \quad \text{Déterminer la classe selon le **signe (H(x))**}$$

➤ Aucun exemple ne se situe sur l'hyperplan, donc :

$$\left\{ \begin{array}{ll} C_k = +1 & \text{Si } H(X) \geq 1 \\ C_k = -1 & \text{Si } H(X) \leq -1 \end{array} \right. \quad \Rightarrow \text{Rechercher les meilleures valeurs de } \mathbf{W} \text{ et de } \mathbf{b}.$$

➔ SVM doit satisfaire : - i.e. doit correctement classifié tous les  $\mathbf{X_i}$  (les datapoints dans le dataset) pour être **correctement entraîné** :

$$y_i (W^T x_i + b) \geq 1, i = 1..n$$

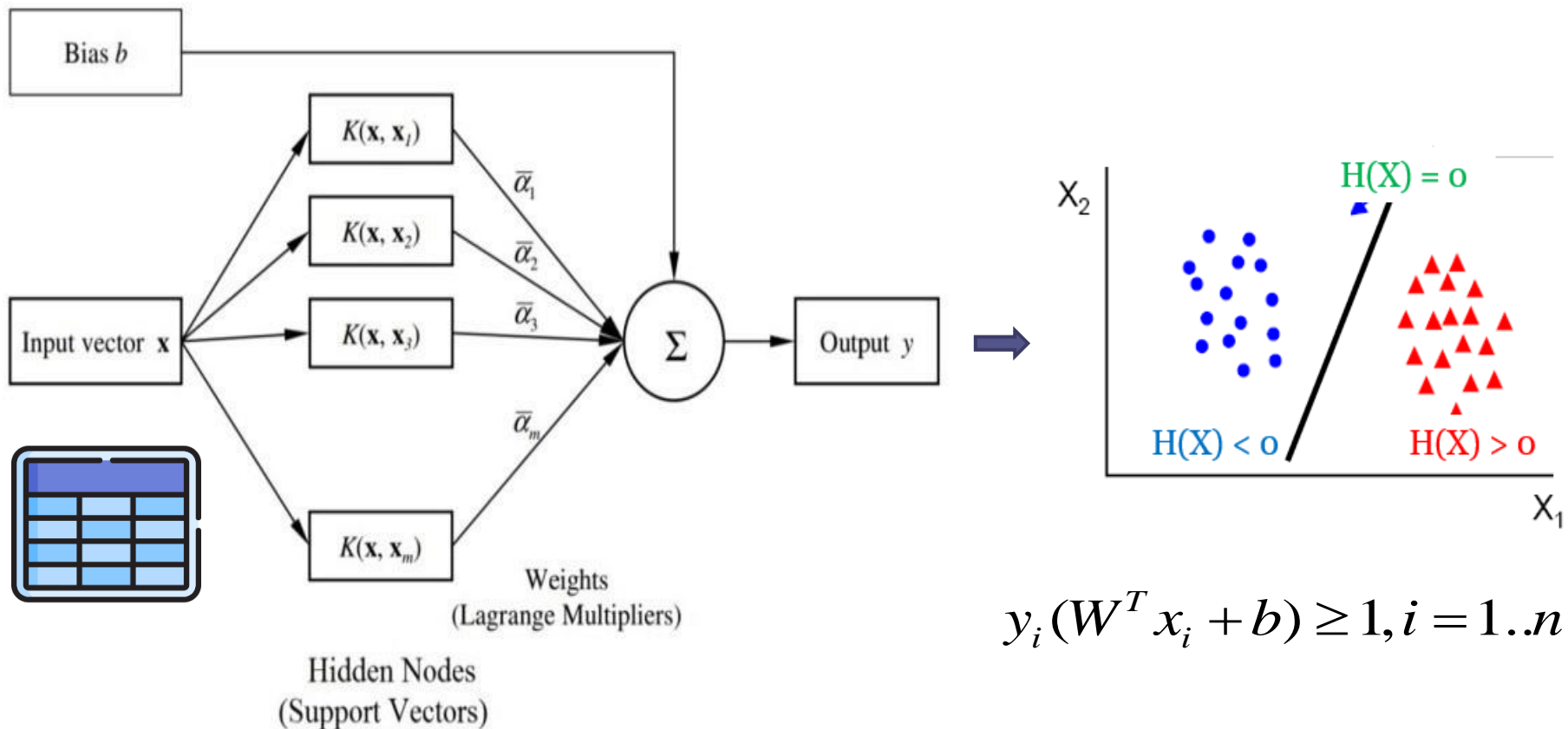


# SVM Linéaire

## Cas de séparabilité linéaire

$$H(X) = W^T X + b$$

- SVM doit satisfaire : - i.e. doit correctement classifié tous les  $\mathbf{X}_i$  (les datapoints dans le dataset) pour être correctement entraîné :



# SVM Linéaire

## Cas de séparabilité linéaire

- Le but d'un SVM **lors de l'entraînement** est donc de trouver :
  - Un vecteur de poids **W** et un biais **b** tel que :
  - Pour tout  $x_i$  ayant comme classe  $y_i$  appartenant aux données d'entraînement :

$$y_i (W^T x_i + b) \geq 1, i = 1..n$$

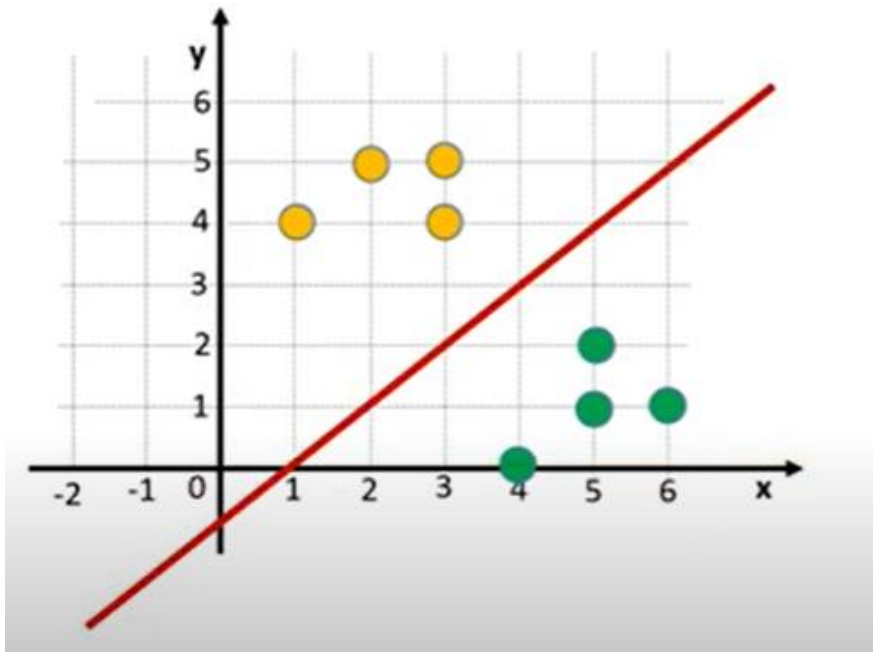
### ➤ Etapas SVM:

- 1** - Générer et tester des hyperplans qui séparent les classes, en utilisant et testant différentes valeurs de **W** et de **b**.
  - Problème : Il **existe plusieurs hyperplans** possibles. Lequel choisir ?

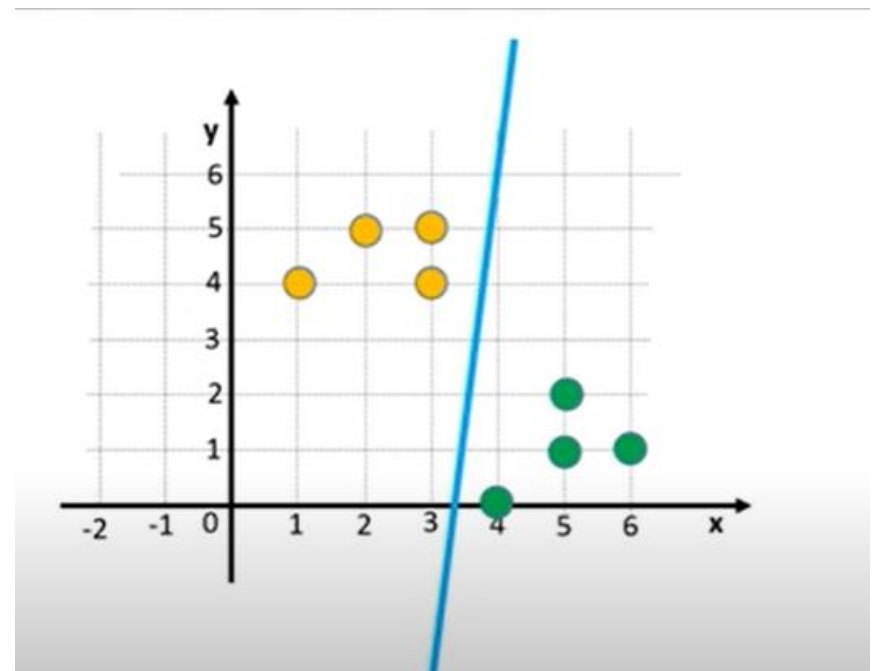
# SVM Linéaire

## Cas de séparabilité linéaire

- Problème : Il **existe plusieurs hyperplans** possibles. Lequel choisir ?
- Exemple :



Hyperplan 1

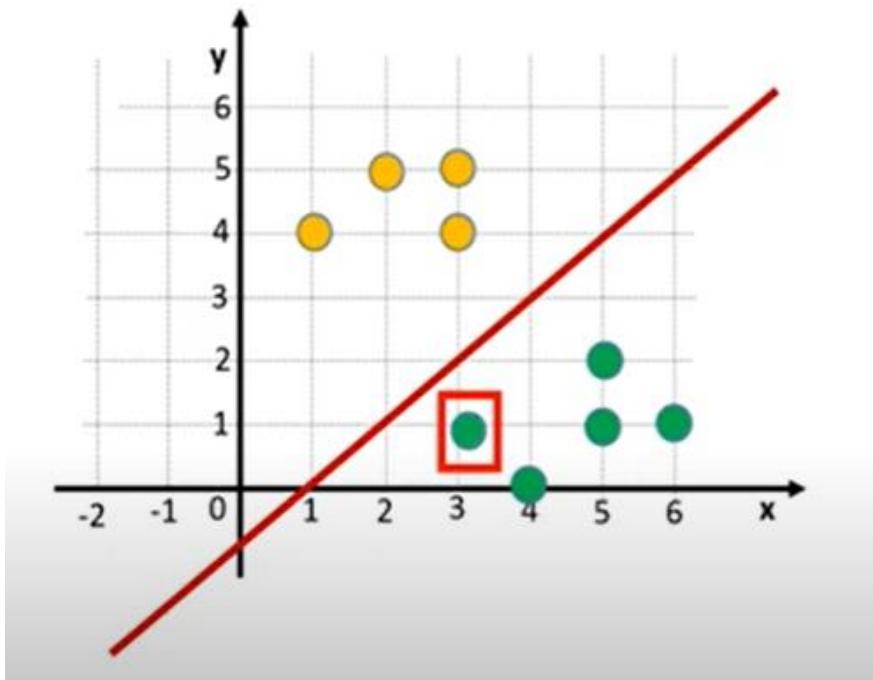


Hyperplan 2

# SVM Linéaire

## Cas de séparabilité linéaire

- Problème : Il **existe plusieurs hyperplans** possibles. Lequel choisir ?
- Exemple :



Prédire un nouvel exemple : correct

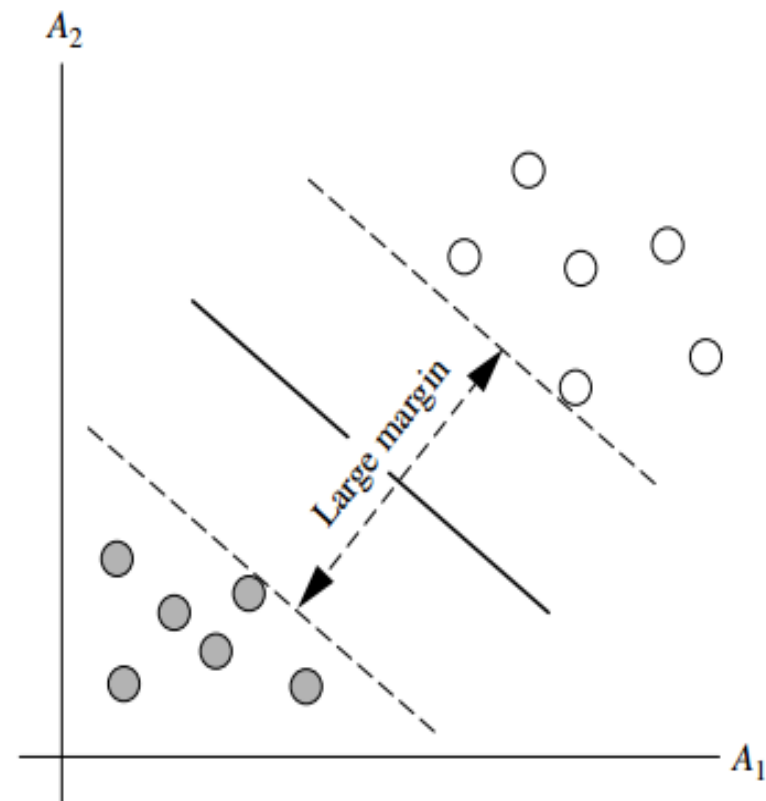
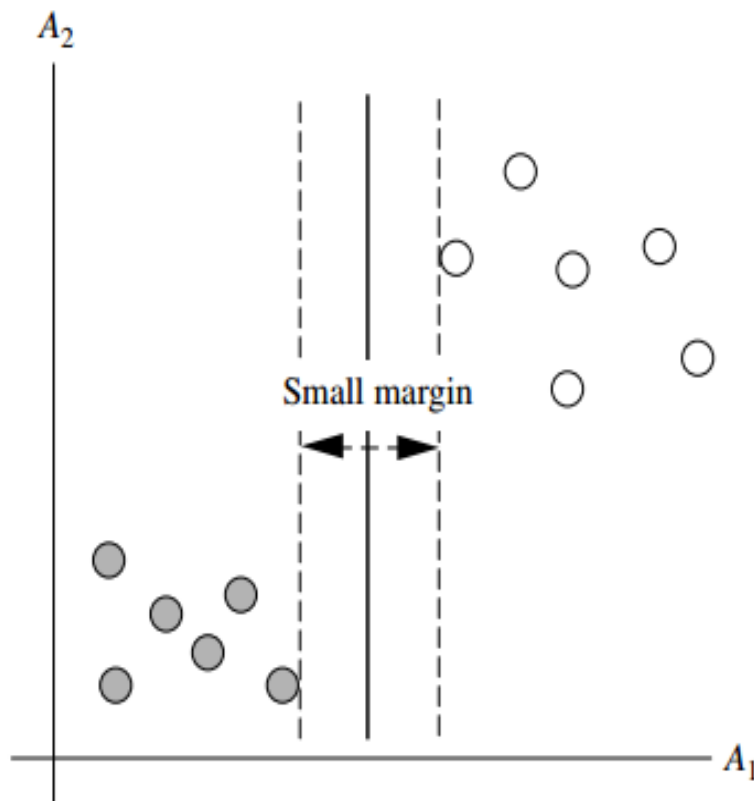


Prédire un nouvel exemple : faux

# SVM Linéaire

## Cas de séparabilité linéaire

- Quand s'arrêter avec l'étape 1 ? Quel est l'**hyperplan optimal** ? – i.e. qui sépare le mieux les données. Comment savoir que **W** et **b** trouvées sont les meilleures valeurs ?

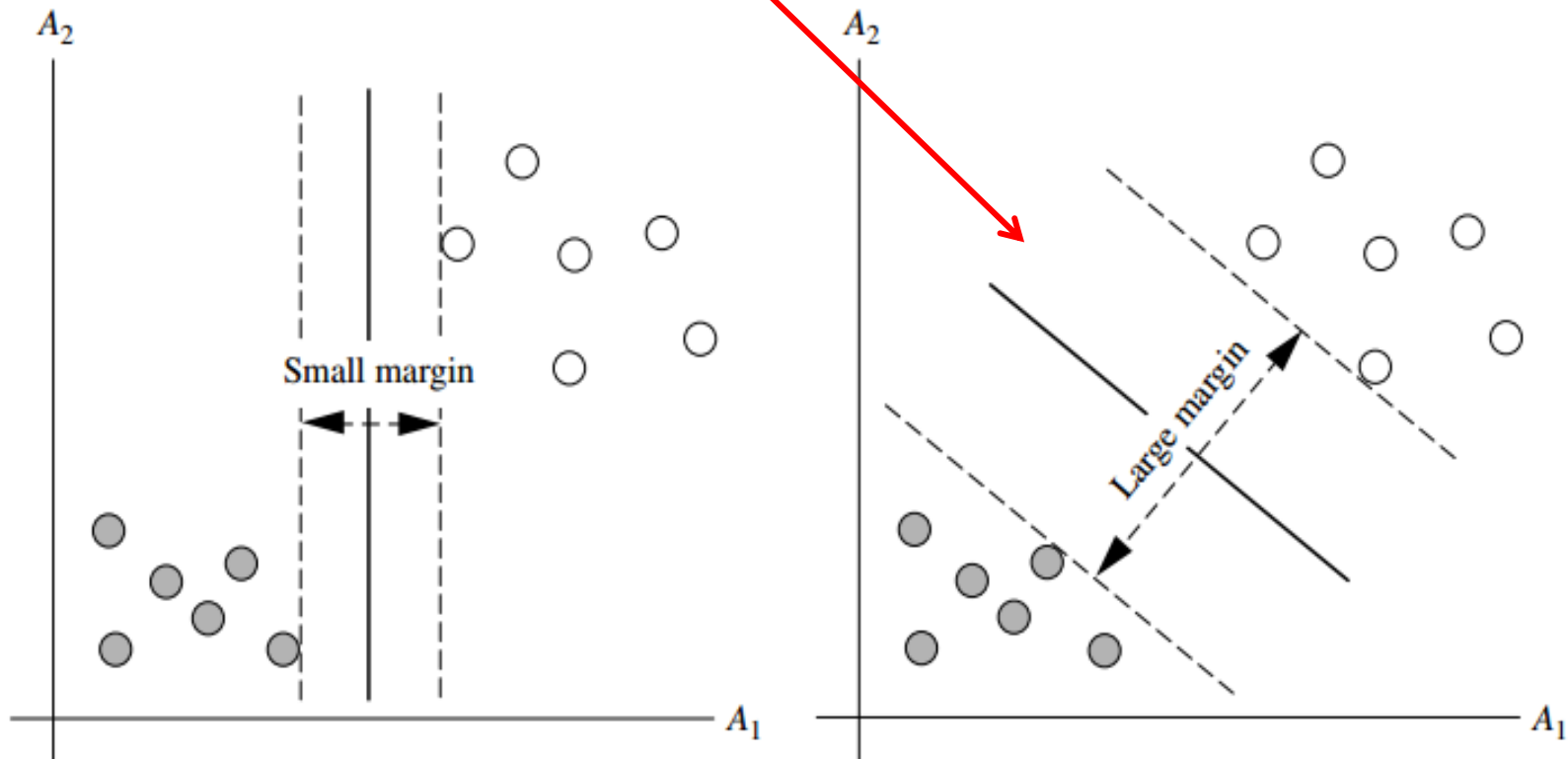


# SVM Linéaire

## Cas de séparabilité linéaire

- Quel est l'hyperplan **optimal** ? – i.e. qui sépare le mieux les données.

Solution : **Maximiser la marge**



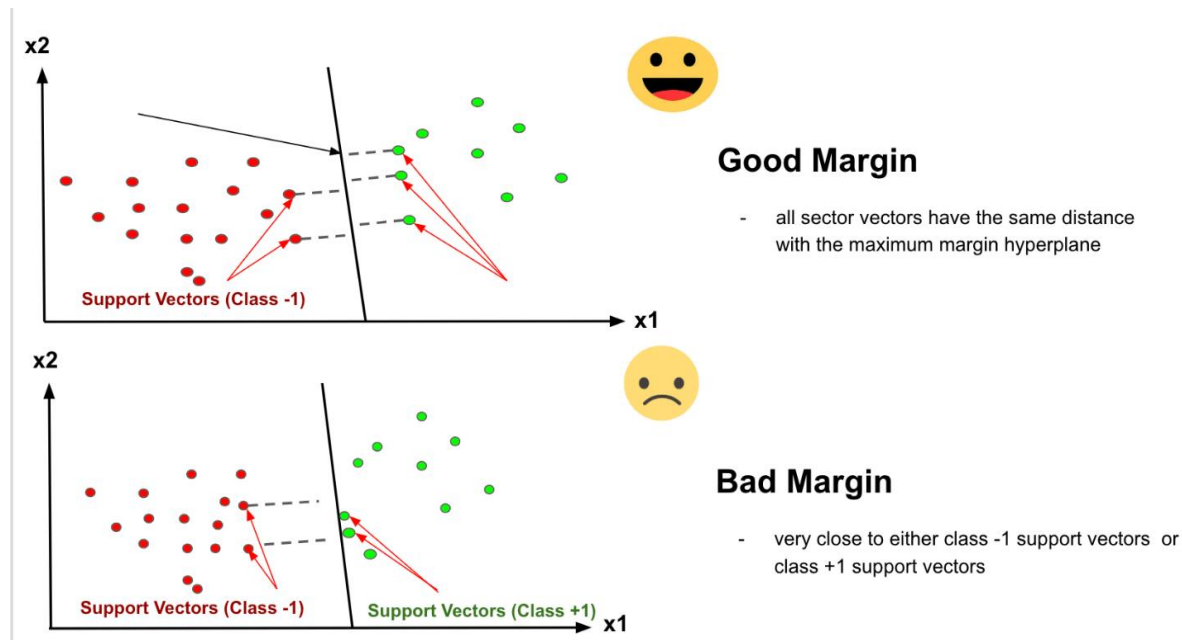


# SVM Linéaire

## Cas de séparabilité linéaire

### ➤ Etapes SVM:

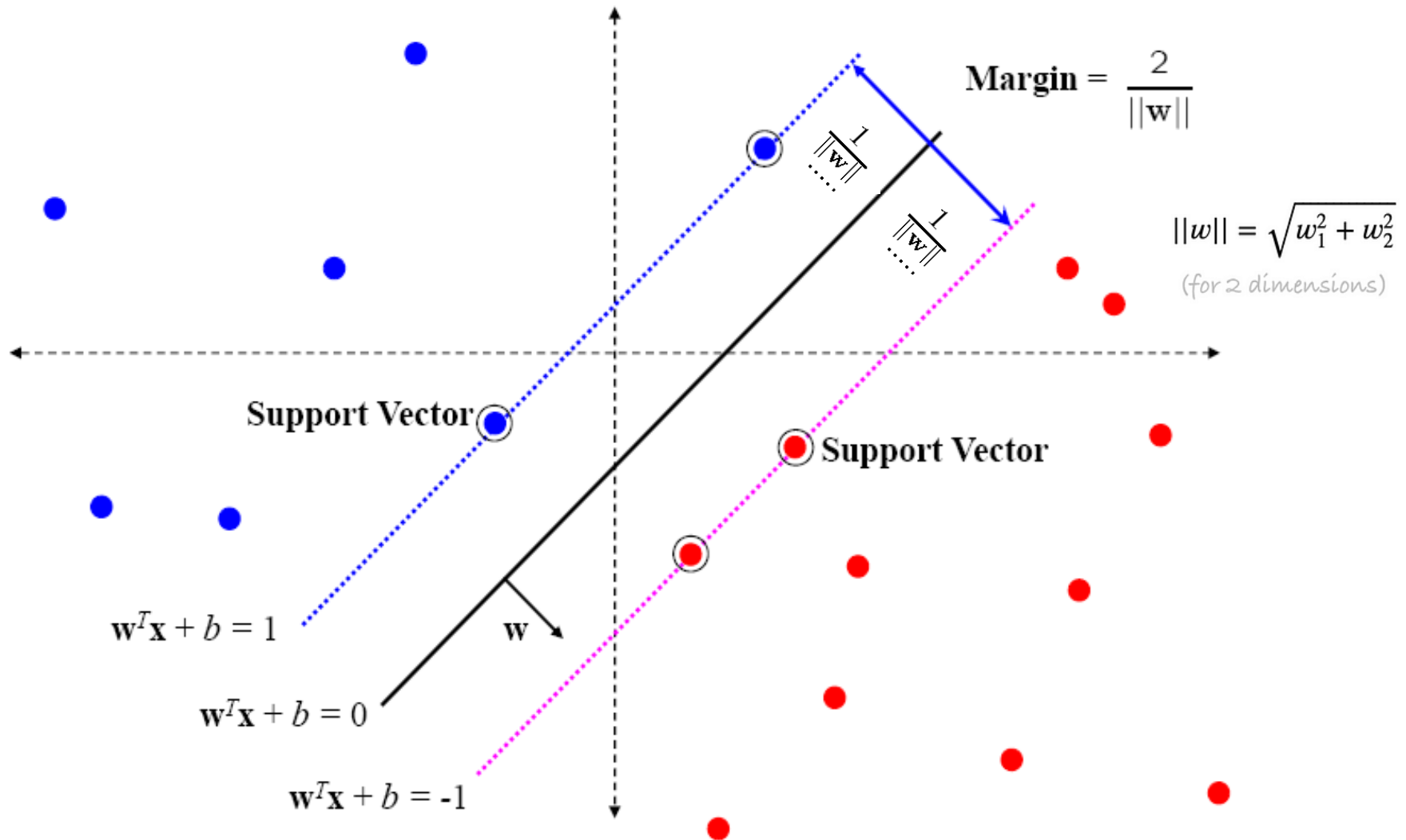
**2** – **Calculer la marge** avec chaque hyperplan généré et sélectionné l'hyperplan qui a une **marge maximale** séparant le mieux les données (voire les **vecteurs supports** seulement, trouvés lors de l'entraînement).



## SVM Linéaire

### Cas de séparabilité linéaire

$$H(X) = W^T X + b$$



## SVM Linéaire

$$H(X) = W^T X + b$$

### Cas de séparabilité linéaire

- Marge : distance entre deux hyperplans parallèles
- Marge =  $2/||\mathbf{W}||$ , avec  $||\mathbf{W}||$  : norme du vecteur  $\mathbf{W}$  ( $w_1, w_2$ )

$$||w|| = \sqrt{w_1^2 + w_2^2}$$

(for 2 dimensions)

- Maximiser la marge → **Maximiser**  $\frac{2}{||\mathbf{w}||}$

➔ Pour trouver l'hyperplan séparateur qui maximise la marge :

- Déterminer le **vecteur  $\mathbf{W}$**  qui possède la **norme euclidienne minimale** ;
- Et qui vérifie la contrainte de l'équation de **bonne classification** des exemples d'entraînement.

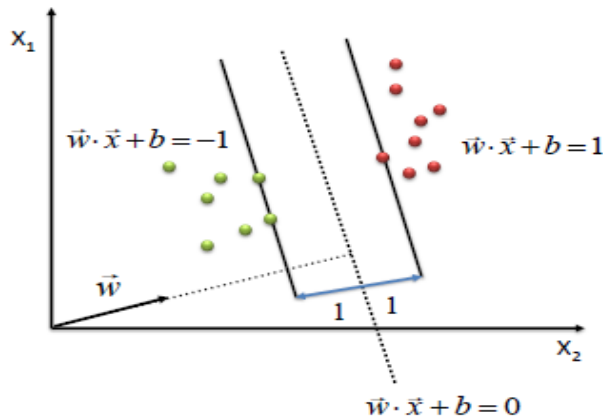
# SVM Linéaire

## Cas de séparabilité **linéaire**

➤ **Donc**, l'hyperplan séparateur **optimal** peut être obtenu en résolvant le problème de l'équation :

- Maximiser la marge  $2/\|W\|$
- Sous contraintes de :

- $W \cdot X_i + b \geq 1$ , pour classe +1
- $W \cdot X_i + b \leq -1$ , pour classe -1



$$\max \frac{2}{\|w\|}$$

s.t.  
 $(w \cdot x + b) \geq 1, \forall x \text{ of class 1}$   
 $(w \cdot x + b) \leq -1, \forall x \text{ of class 2}$

$$H(X) = W^T X + b$$

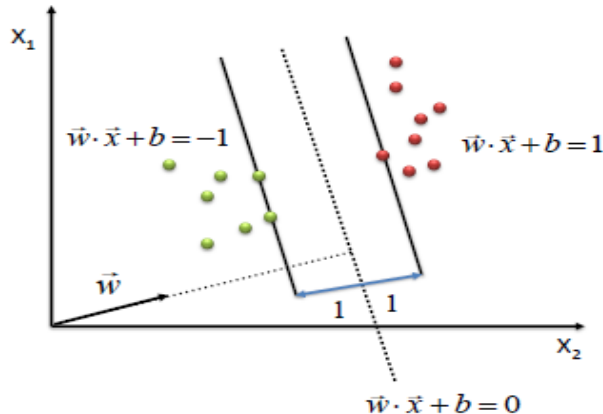
# SVM Linéaire

## Cas de séparabilité **linéaire**

➤ **Donc**, l'hyperplan séparateur **optimal** peut être obtenu en résolvant le problème de l'équation :

- Maximiser la marge  $2/||W||$
- Sous contraintes de :

- $W \cdot X_i + b = 1$ , pour vecteurs supports de classe +1
- $W \cdot X_i + b = -1$ , pour vecteurs supports de classe -1



$$\max \frac{2}{||w||}$$

s.t.  
 $(w \cdot x + b) \geq 1, \forall x \text{ of class 1}$   
 $(w \cdot x + b) \leq -1, \forall x \text{ of class 2}$

$$H(X) = W^T X + b$$

# SVM Linéaire

## Cas de séparabilité **linéaire**

➤ **Donc**, l'hyperplan séparateur **optimal** peut être obtenu en résolvant le problème de l'équation :

- Maximiser la marge  $2/||W||$  : **fonction objective**
- Sous **contraintes** de :

- $W.X_i + b = 1$ , pour vecteurs supports de classe +1
- $W.X_i + b = -1$ , pour vecteurs supports de classe -1

➤ Le problème de l'équation est un **problème de programmation quadratique sous contraintes linéaires**.

➤ **Convex optimization problem**. Problème d'optimisation convexe.

# SVM Linéaire

## Cas de séparabilité linéaire

### ➔ Résolution :

- Utiliser **l'équation de Lagrange** et introduire ses **multiplicateurs**.
- L'hyperplan séparateur optimal peut être obtenu en maximisant **le problème dual** :

$$\left\{ \begin{array}{l} \text{Maximiser } W(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j \\ \text{Sous contraintes de: } \sum_{i=1}^n \alpha_i y_i = 0 \\ \alpha_i \geq 0 \end{array} \right.$$

# SVM Linéaire

## Cas de séparabilité linéaire

### ➔ Résolution :

- Utiliser **l'équation de Lagrange** et introduire ses **multiplicateurs**.
- L'hyperplan séparateur optimal peut être obtenu en maximisant **le problème dual**.
- On déduit que :

$$\text{Avec } \left\{ \begin{array}{l} W = \sum_{i=1}^n \alpha_i y_i x_i \\ \sum_{i=1}^n \alpha_i y_i = 0 \end{array} \right. \quad \text{Pour les non vecteurs supports}$$



# SVM Linéaire

## Cas de séparabilité linéaire

- Le problème de l'équation est un problème de programmation quadratique sous contraintes linéaires.
- En utilisant la formule de Lagrange, **la fonction de décision** devient :

$$H(X) = W^T X + b$$

Nombre de vecteurs support  $\longrightarrow l$

$$H(x) = \sum_{i=1}^l \alpha_i y_i x_i x^T + b$$

Exemple à tester  $\longleftarrow x$

Les classes des vecteurs supports  $\longleftarrow y_i$

Supports vecteurs  $\longleftarrow x_i$

Multiplicateurs de Lagrange  $\longleftarrow \alpha_i$

Exemple à tester  $\longleftarrow x^T$

# SVM Linéaire

## Cas de séparabilité linéaire

- En utilisant la **formule de Lagrange**, la fonction de décision devient :

$$H(x) = \sum_{i=1}^l \alpha_i y_i x_i x^T + b$$

- **Prédiction** : La classe de x selon l'hyperplan optimal :

$$\left\{ \begin{array}{ll} \text{La classe de } X = +1 & \text{Si } H(X) > 0 \\ \text{La classe de } X = -1 & \text{Si } H(X) < 0 \end{array} \right. = \text{sign}\left(\sum_{i=1}^N \alpha_i y_i \vec{x}_i \cdot \vec{x} + b\right)$$

- La zone  $-1 < H(x) < 1$  est appelée la zone de généralisation.

# SVM Linéaire

## Cas de séparabilité linéaire



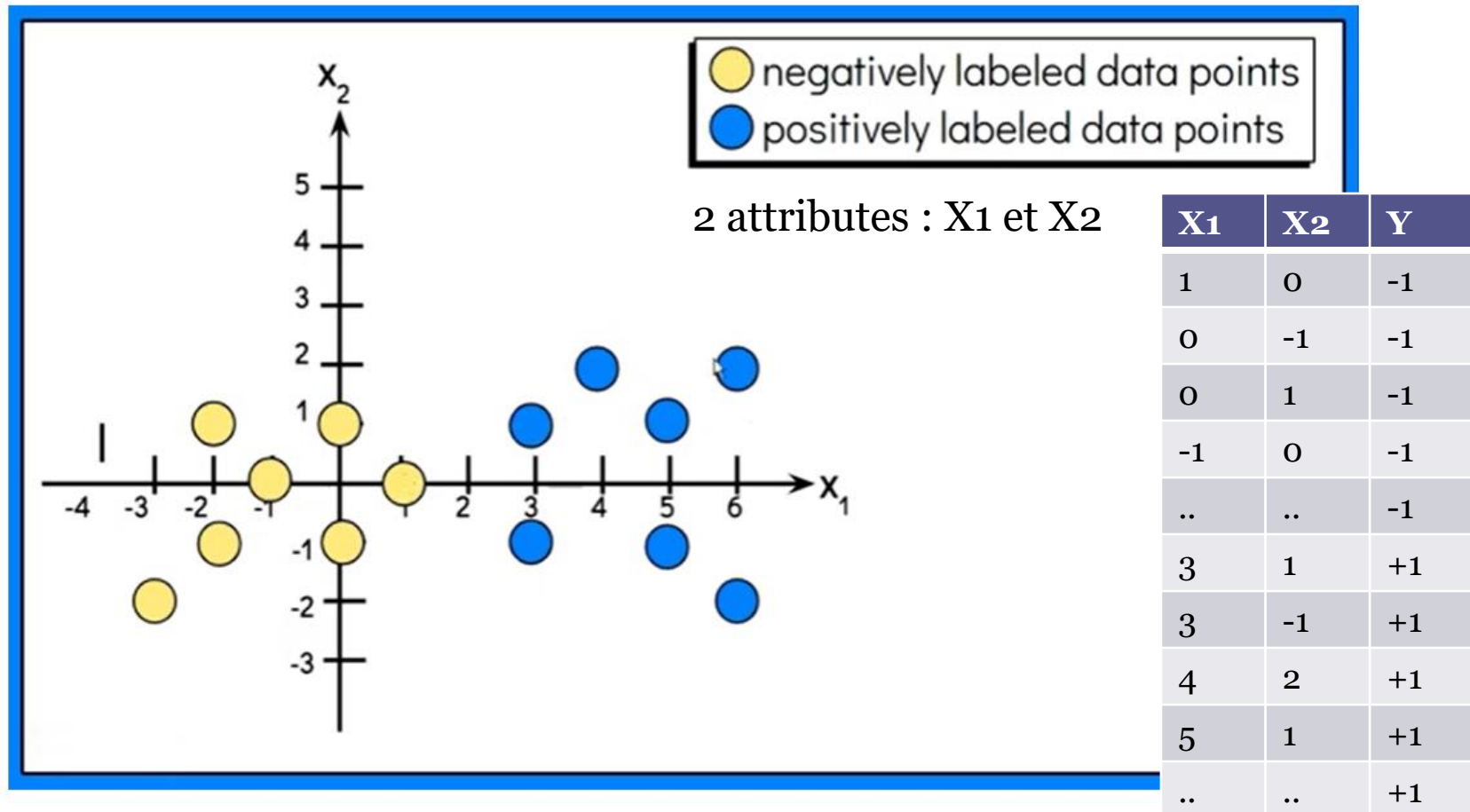
- **Implémentation des SVM:** pour la classification binaire consiste à la résolution du problème dual de programmation quadratique.
- Résolution ➔ déterminer les multiplicateurs Lagrangiens optimaux.
- Si nombre d'exemples est important => problème de temps et de mémoire.
- Solution : Méthode de Shnuking. Entraînement sur quelques exemples choisis par une heuristique, puis ajout des autres itérativement.
- L'algorithme SMO : Sequential Minimal Optimisation. (*Platt et all., 1999*)
  - Librairie LibSVM; SVC/LinearSVC class dans scikit-learn.
- Native Python Implementation : SGDClassifier class dans scikit-learn.

# SVM Linéaire

<https://www.youtube.com/watch?v=TPVzIKJOcNo>  
<https://axon.cs.byu.edu/Dan/678/miscellaneous/SVM.example.pdf>

## Cas de séparabilité linéaire

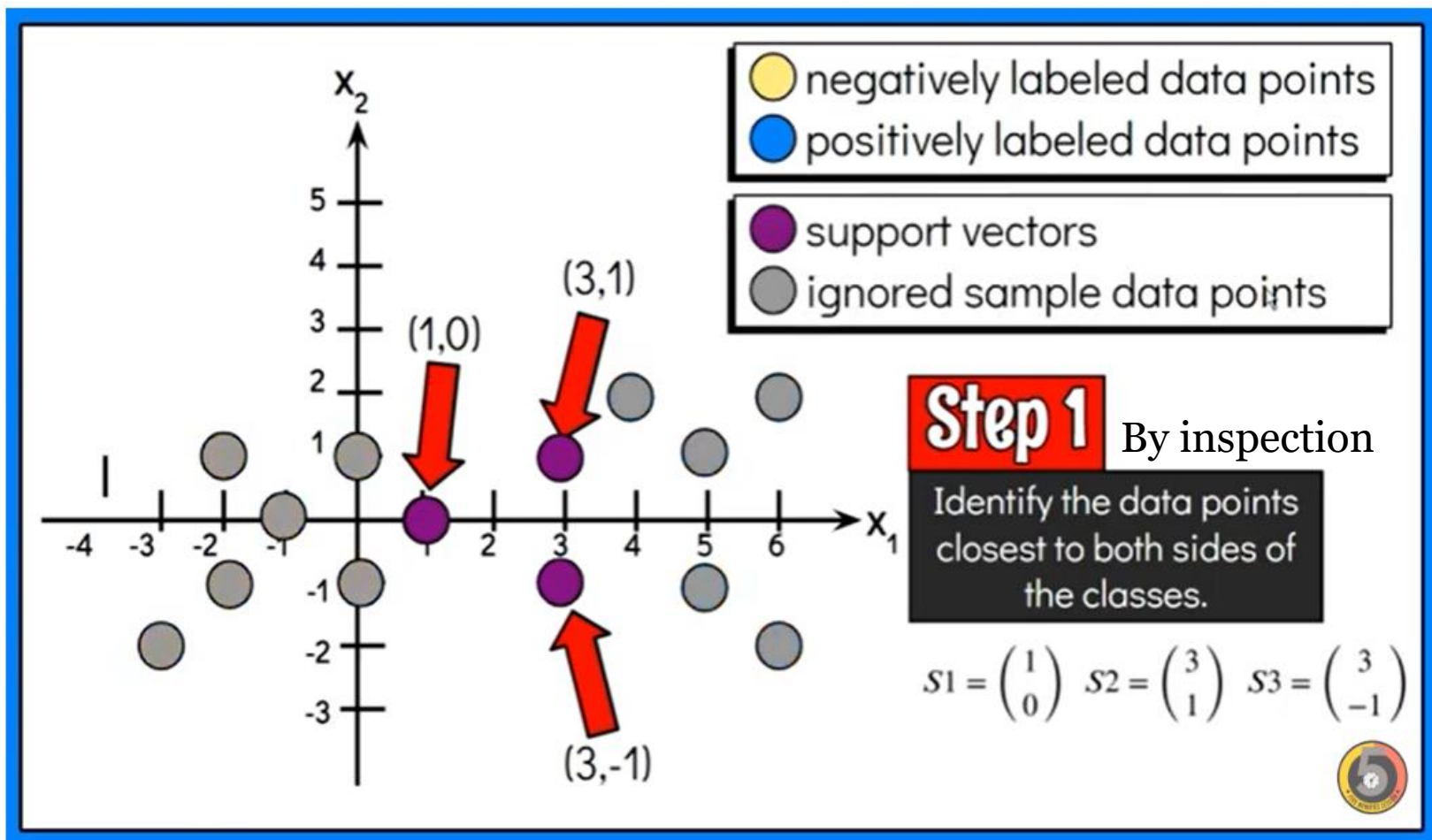
### ➤ Implémentation des SVM : Exemple - Déroulement



# SVM Linéaire

## Cas de séparabilité linéaire

### ➤ Implémentation des SVM : Exemple - Déroulement



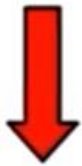
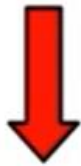
# SVM Linéaire

## Cas de séparabilité linéaire - Exemple - Déroulement

**Step 2**

Augment 1 as bias  
input to each vector.

$$s1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad s2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad s3 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$



$$\bar{s}1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \bar{s}2 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \quad \bar{s}3 = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}$$

## SVM Linéaire

$$H(x) = \sum_{i=1}^l \alpha_i y_i x_i x^T + b$$

### Cas de séparabilité **linéaire**

➤ **Donc**, l'hyperplan séparateur **optimal** peut être obtenu en résolvant le problème de l'équation :

- Maximiser la marge  $2/||W||$  : **fonction objective**
- Sous **contraintes** de :

- $W \cdot X_i + b = 1$ , pour vecteurs supports de classe +1
- $W \cdot X_i + b = -1$ , pour vecteurs supports de classe -1

➤ Le problème de l'équation est un **problème de programmation quadratique sous contraintes linéaires**.

➤ **Convex optimization problem**. Problème d'optimisation convexe.

# SVM Linéaire

## Cas de séparabilité linéaire - Exemple - Déroulement

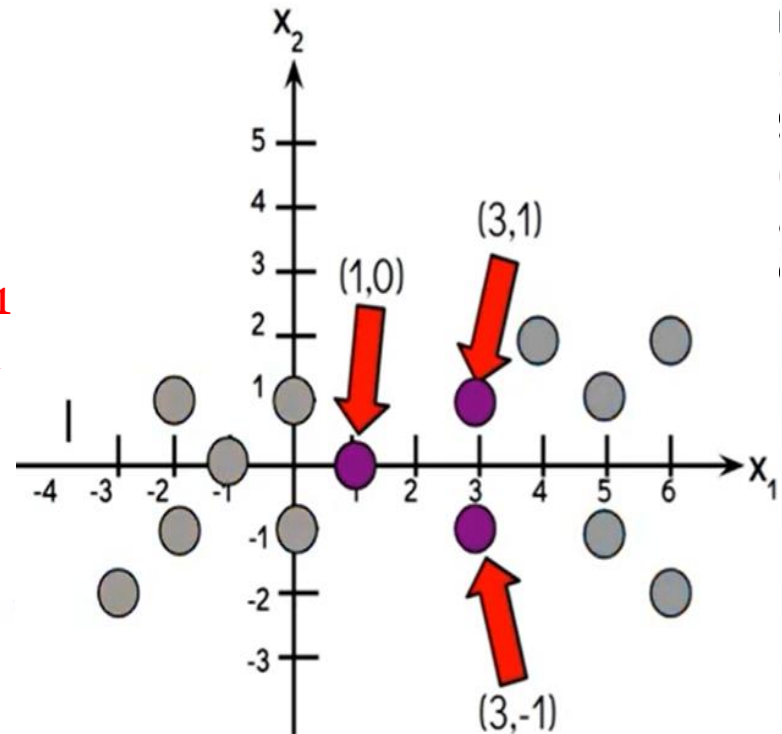
**Step 3** Write the equations needed to calculate the weighted vector.

$$W = \sum_{i=1}^n \alpha_i y_i x_i$$

$W \cdot X_i + b = 1$ , pour VS de classe +1

$W \cdot X_i + b = -1$ , pour VS de classe -1

$$\bar{S}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \bar{S}_2 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \quad \bar{S}_3 = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}$$





# SVM Linéaire

## Cas de séparabilité linéaire - Exemple - Déroulement

**Step 3**

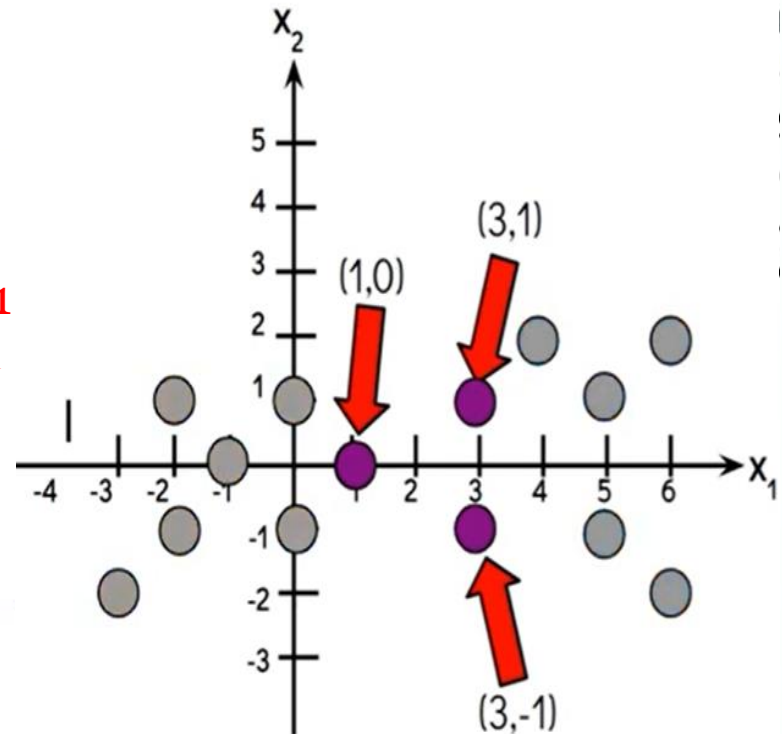
Write the equations needed to calculate the weighted vector.

$$\bar{w} = \sum_i \alpha_i \bar{S}_i$$

$W.X_i + b = 1$ , pour VS de classe +1

$W.X_i + b = -1$ , pour VS de classe -1

$$\bar{S}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \bar{S}_2 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \quad \bar{S}_3 = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}$$



## SVM Linéaire

### Cas de séparabilité linéaire - Exemple - Déroulement

$$\bar{w} = \sum_i \alpha_i \bar{S}_i$$

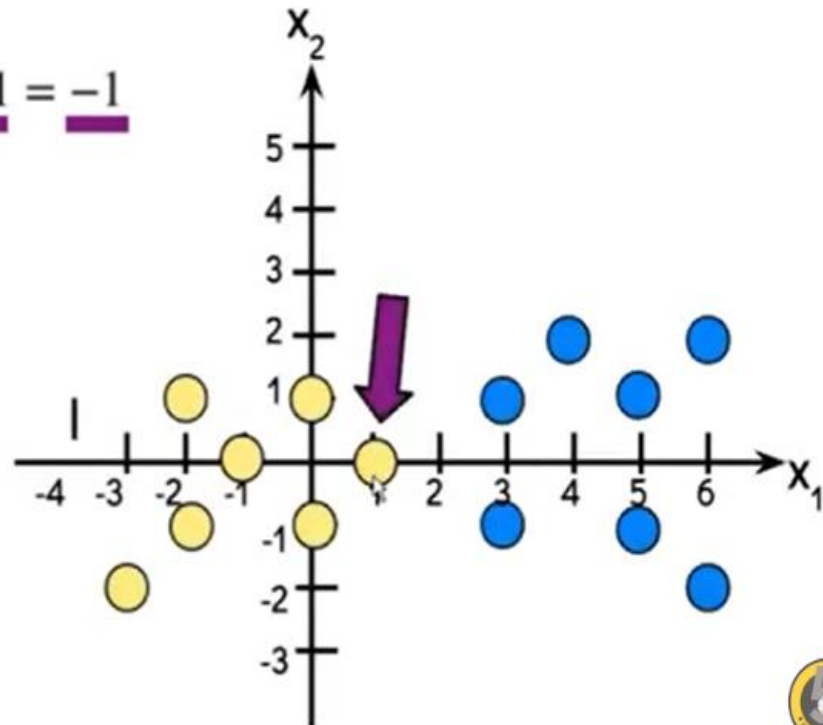
**Step 3**

Write the equations needed to calculate the weighted vector.

$$\alpha_1 \bar{S}_1 \cdot \bar{S}_1 + \alpha_1 \bar{S}_2 \cdot \bar{S}_1 + \alpha_1 \bar{S}_3 \cdot \bar{S}_1 = -1$$

$$\bar{S}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

- $W \cdot X_i + b = 1$ , pour VS de classe +1
- $W \cdot X_i + b = -1$ , pour VS de classe -1



# SVM Linéaire

## Cas de séparabilité linéaire - Exemple - Déroulement

### Step 3

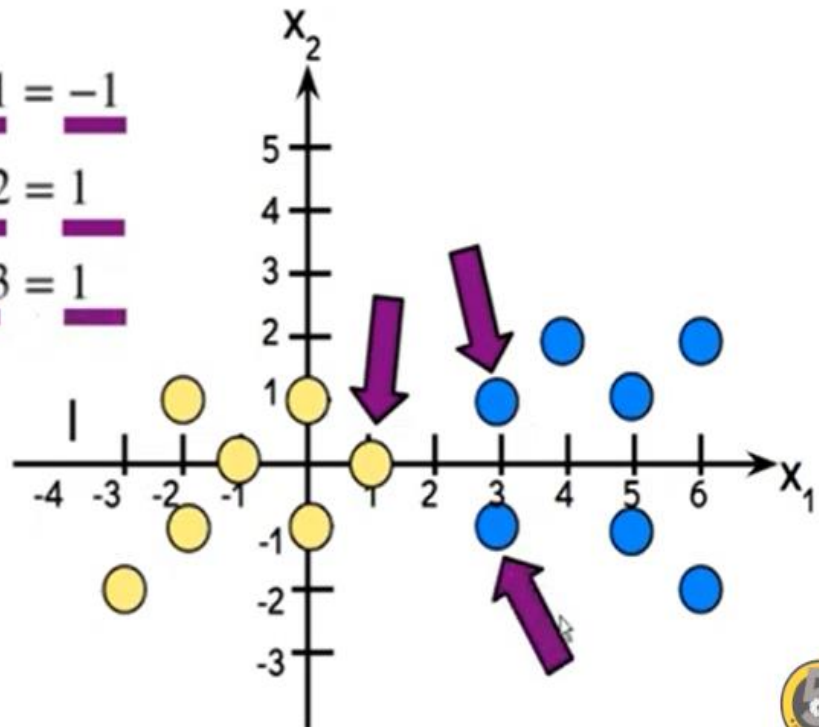
Write the equations needed to calculate the weighted vector.

$$\alpha_1 \bar{S}_1 \cdot \bar{S}_1 + \alpha_1 \bar{S}_2 \cdot \bar{S}_1 + \alpha_1 \bar{S}_3 \cdot \bar{S}_1 = -1$$

$$\alpha_1 \bar{S}_1 \cdot \bar{S}_2 + \alpha_1 \bar{S}_2 \cdot \bar{S}_2 + \alpha_1 \bar{S}_3 \cdot \bar{S}_2 = 1$$

$$\alpha_1 \bar{S}_1 \cdot \bar{S}_3 + \alpha_1 \bar{S}_2 \cdot \bar{S}_3 + \alpha_1 \bar{S}_3 \cdot \bar{S}_3 = 1$$

- $W \cdot X_i + b = 1$ , pour VS de classe +1
- $W \cdot X_i + b = -1$ , pour VS de classe -1



## SVM Linéaire

### Cas de séparabilité linéaire - Exemple - Déroulement

**Step 4**

Substitute the  
values into the  
equation.

$$\bar{S}1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \bar{S}2 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \quad \bar{S}3 = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}$$

$$\alpha_1 \bar{S}1 \cdot \bar{S}1 + \alpha_2 \bar{S}2 \cdot \bar{S}1 + \alpha_3 \bar{S}3 \cdot \bar{S}1 = -1$$



$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = -1$$

# SVM Linéaire

## Cas de séparabilité linéaire - Exemple - Déroulement

**Step 4**

Substitute the  
values into the  
equation.

$$\bar{S}1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \bar{S}2 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \quad \bar{S}3 = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}$$

$$\alpha_1 \bar{S}1 \cdot \bar{S}1 + \alpha_1 \bar{S}2 \cdot \bar{S}1 + \alpha_1 \bar{S}3 \cdot \bar{S}1 = -1$$



$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = -1$$

$$\alpha_1 \bar{S}1 \cdot \bar{S}2 + \alpha_1 \bar{S}2 \cdot \bar{S}2 + \alpha_1 \bar{S}3 \cdot \bar{S}2 = 1$$



$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} = 1$$



## SVM Linéaire

### Cas de séparabilité linéaire - Exemple - Déroulement

**Step 4**

Substitute the  
values into the  
equation.

$$\bar{S}1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \bar{S}2 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \quad \bar{S}3 = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}$$

$$\alpha_1 \bar{S}1 \cdot \bar{S}1 + \alpha_1 \bar{S}2 \cdot \bar{S}1 + \alpha_1 \bar{S}3 \cdot \bar{S}1 = -1$$



$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = -1$$

$$\alpha_1 \bar{S}1 \cdot \bar{S}2 + \alpha_1 \bar{S}2 \cdot \bar{S}2 + \alpha_1 \bar{S}3 \cdot \bar{S}2 = 1$$



$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} = 1$$

$$\alpha_1 \bar{S}1 \cdot \bar{S}3 + \alpha_1 \bar{S}2 \cdot \bar{S}3 + \alpha_1 \bar{S}3 \cdot \bar{S}3 = 1$$



$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} = 1$$



# SVM Linéaire

## Cas de séparabilité linéaire - Exemple - Déroulement

### Step 5

Perform dot product operation between two vectors.

$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = -1$$



$$\alpha_1 (1 + 0 + 1) + \alpha_2 (3 + 0 + 1) + \alpha_3 (3 + 0 + 1) = -1$$



$$\underline{2\alpha_1 + 4\alpha_2 + 4\alpha_3 = -1}$$

$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} = 1$$



$$\alpha_1 (3 + 0 + 1) + \alpha_2 (9 + 1 + 1) + \alpha_3 (9 - 1 + 1) = 1$$



$$\underline{4\alpha_1 + 11\alpha_2 + 9\alpha_3 = 1}$$

$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} = 1$$



$$\alpha_1 (3 + 0 + 1) + \alpha_2 (9 - 1 + 1) + \alpha_3 (9 + 1 + 1) = 1$$



$$\underline{4\alpha_1 + 9\alpha_2 + 11\alpha_3 = 1}$$



# SVM Linéaire

## Cas de séparabilité linéaire - Exemple - Déroulement

**Step 6**

Solve the  
simultaneous  
equations.

$$\begin{aligned}2\alpha_1 + 4\alpha_2 + 4\alpha_3 &= -1 \\4\alpha_1 + 11\alpha_2 + 9\alpha_3 &= 1 \\4\alpha_1 + 9\alpha_2 + 11\alpha_3 &= 1\end{aligned}$$

$$\begin{aligned}\alpha_1 &= -3.50 \\ \alpha_2 &= 0.75 \\ \alpha_3 &= 0.75\end{aligned}$$

$$\begin{aligned}4\alpha_1 + 11\alpha_2 + 9\alpha_3 &= 1 \\ -(4\alpha_1 + 9\alpha_2 + 11\alpha_3) &= 1\end{aligned}$$

$$2\alpha_2 - 2\alpha_3 = 0$$

$$\begin{aligned}2(2\alpha_1 + 4\alpha_2 + 4\alpha_3) &= -1 \\ 1(4\alpha_1 + 11\alpha_2 + 9\alpha_3) &= 1\end{aligned}$$

$$\begin{aligned}4\alpha_1 + 8\alpha_2 + 8\alpha_3 &= -2 \\ 4\alpha_1 + 11\alpha_2 + 9\alpha_3 &= 1\end{aligned}$$

$$\begin{aligned}4\alpha_1 + 8\alpha_2 + 8\alpha_3 &= -2 \\ -(4\alpha_1 + 11\alpha_2 + 9\alpha_3) &= 1\end{aligned}$$

$$-3\alpha_2 - 1\alpha_3 = -3$$

$$\begin{aligned}1(2\alpha_2 - 2\alpha_3) &= 0 \\ 2(-3\alpha_2 - 1\alpha_3) &= -3\end{aligned}$$

$$\begin{aligned}2\alpha_2 - 2\alpha_3 &= 0 \\ -6\alpha_2 - 2\alpha_3 &= -6\end{aligned}$$

$$\begin{aligned}2\alpha_2 - 2\alpha_3 &= 0 \\ -(-6\alpha_2 - 2\alpha_3) &= -6\end{aligned}$$

$$\begin{aligned}8\alpha_3 &= 6 \\ \div 8 & \div 8\end{aligned}$$

$$\alpha_3 = 0.75$$

$$\begin{aligned}2\alpha_2 - 2\alpha_3 &= 0 \\ 2\alpha_2 - 2(0.75) &= 0 \\ 2\alpha_2 - 1.50 &= 0 \\ 2\alpha_2 &= 0 + 1.50 \\ 2\alpha_2 &= 1.50 \\ \div 2 & \div 2\end{aligned}$$

$$\alpha_2 = 0.75$$

$$\begin{aligned}2\alpha_1 + 4\alpha_2 + 4\alpha_3 &= -1 \\ 2\alpha_1 + 4(0.75) + 4(0.75) &= -1 \\ 2\alpha_1 + 3 + 3 &= -1 \\ 2\alpha_1 &= -1 - 3 - 3 \\ 2\alpha_1 &= -7 \\ \div 2 & \div 2\end{aligned}$$

$$\alpha_1 = -3.50$$





## SVM Linéaire

### Cas de séparabilité linéaire - Exemple - Déroulement

#### Step 7

Calculate the weighted vector.

$$\alpha_1 = -3.50$$

$$\alpha_2 = 0.75$$

$$\alpha_3 = 0.75$$

$$\bar{S}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \bar{S}_2 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \quad \bar{S}_3 = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}$$

$$\bar{w} = \sum_i \alpha_i \bar{S}_i$$

$$\begin{aligned} \bar{w} &= \sum_i \alpha_i \bar{S}_i \\ &= -3.5 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + 0.75 \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} + 0.75 \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} \end{aligned}$$



## SVM Linéaire

### Cas de séparabilité linéaire - Exemple - Déroulement

#### Step 8

Interpret the computed weighted vector to draw the hyperplane.

$$\bar{w} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$$



The last entry in the weighted vector as the hyperplane offset **b**.

$$= (w_1, w_2, b)$$



## SVM Linéaire

### Cas de séparabilité linéaire - Exemple - Déroulement

#### Step 8

Interpret the computed weighted vector to draw the hyperplane.

$$\bar{w} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$$

The last entry in the weighted vector as the hyperplane offset **b**.

$$y = \underline{w}x + \underline{b}$$

We can write the separating hyperplane equation  $y = wx + b$  with  $w = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  and  $b = -2$

$$H(\mathbf{x}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \mathbf{X} - 2$$



# SVM Linéaire

## Cas de séparabilité linéaire - Exemple - Déroulement

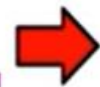
### Step 8

Interpret the computed weighted vector to draw the hyperplane.

$$\bar{w} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$$

The last entry in the weighted vector as the hyperplane offset **b**.

$$y = \underline{w}x + \underline{b}$$



We can write the separating hyperplane equation  $y = wx + b$  with  $w = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  and  $b = -2$

$$\mathbf{H}(\mathbf{x}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \mathbf{X} - 2$$

$$\text{margin} = \frac{2}{\|\mathbf{w}\|}$$

$$\|\mathbf{w}\| = \sqrt{w_1^2 + w_2^2}$$

$$\|\mathbf{W}\| = \text{sqrt}(1^2 + 0^2)$$



## SVM Linéaire

### Cas de séparabilité linéaire - Exemple - Déroulement

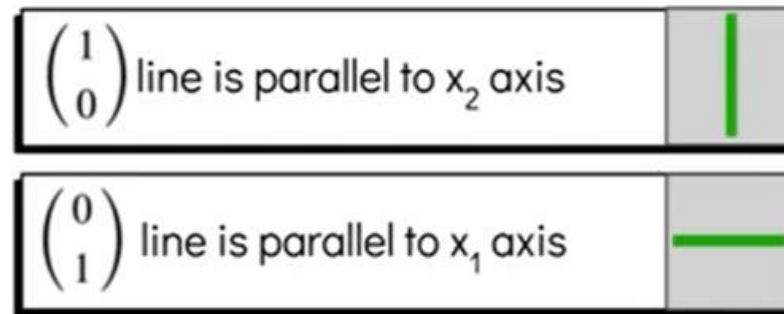
#### Step 8

Interpret the computed weighted vector to draw the hyperplane.

$$\mathbf{H}(\mathbf{x}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \mathbf{X} - 2$$

$$\bar{w} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$$

$$= (w_1, w_2, b)$$



2D-line equation

$$x_2 = -\frac{w_1}{w_2}x_1 - \frac{b}{w_2}$$

slope

intercept



# SVM Linéaire

## Cas de séparabilité linéaire - Exemple - Déroulement

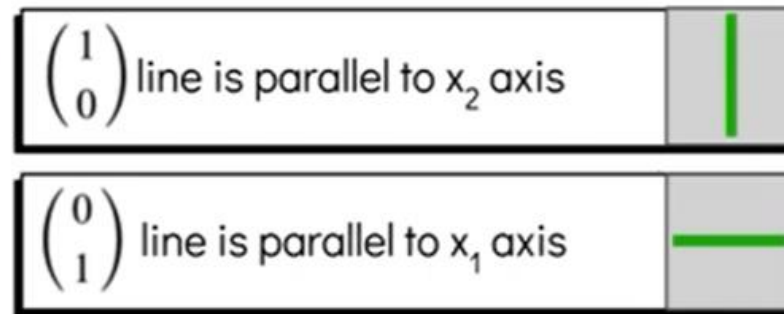
### Step 8

Interpret the computed weighted vector to draw the hyperplane.

$$\mathbf{H}(\mathbf{x}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \mathbf{X} - 2$$

$$\bar{w} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$$

$$= (w_1, w_2, b)$$



For  $x_1=0 \Rightarrow$   
 $x_2=+2$

2D-line equation

$$x_2 = -\frac{w_1}{w_2}x_1 - \frac{b}{w_2}$$

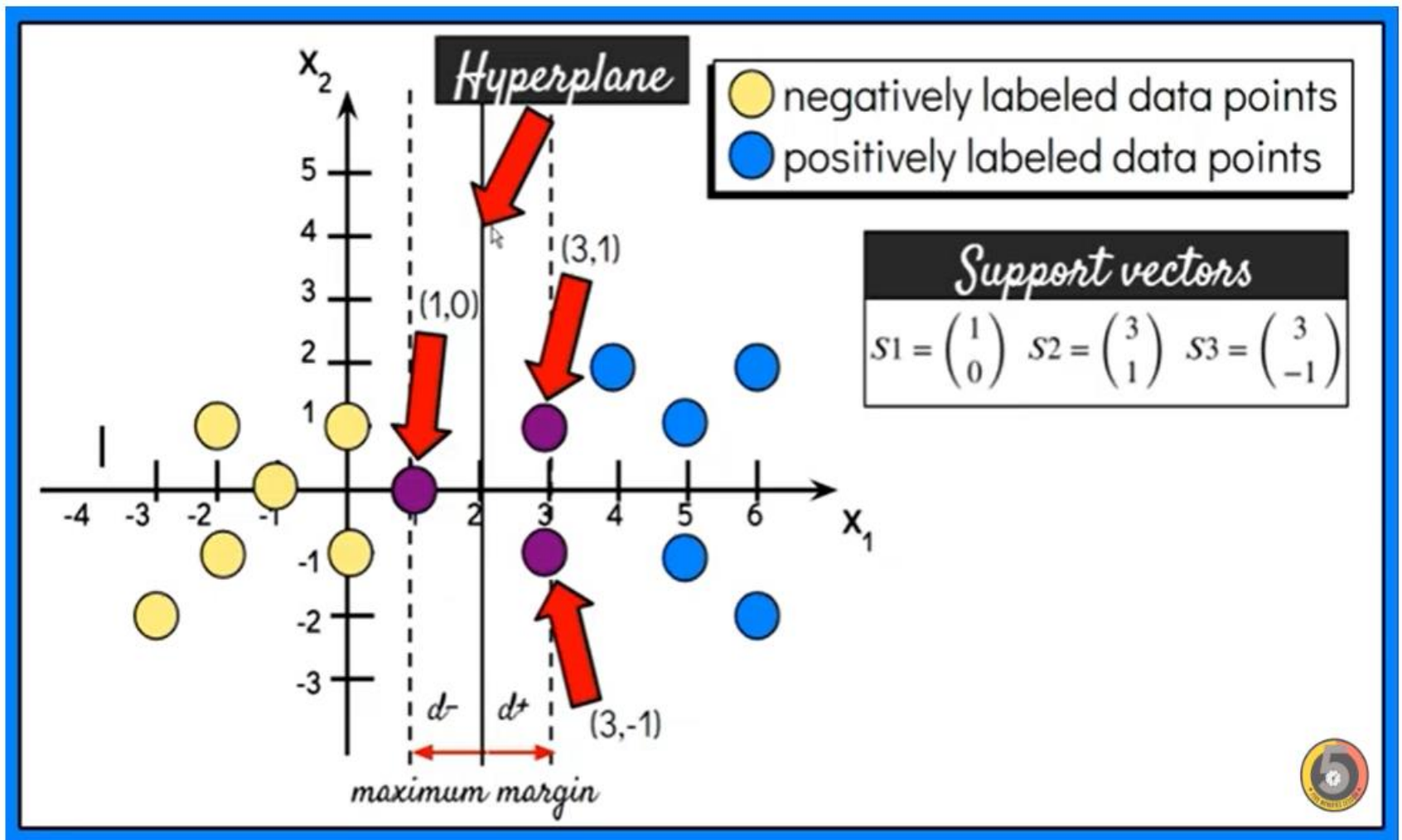
slope

intercept



# SVM Linéaire

## Cas de séparabilité linéaire - Exemple - Déroulement



## SVM Linéaire

### Cas de séparabilité linéaire - Exemple - Déroulement

**Step 8** Make Predictions for  
 $X = (x_1, x_2)$ .

$$H(x) = \sum_{i=1}^l \alpha_i y_i x_i x^T + b$$

$$H(\mathbf{x}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \mathbf{X} - 2$$

$$H(\mathbf{x}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{dot} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - 2$$

*Support vectors*

$$s_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad s_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad s_3 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

$$\alpha_1 = -3.50$$

$$\alpha_2 = 0.75$$

$$\alpha_3 = 0.75$$





## SVM Linéaire

### Cas de séparabilité linéaire - Exemple - Déroulement

**Step 8** Make Predictions for  $X = (x_1, x_2)$ .

$$H(x) = \sum_{i=1}^l \alpha_i y_i x_i x^T + b$$

$$H(\mathbf{x}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \mathbf{X} - 2$$

$$H(\mathbf{x}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{dot} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - 2$$

*Support vectors*

$$s_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad s_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad s_3 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

$$\alpha_1 = -3.50$$

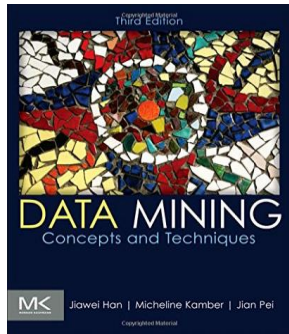
$$\alpha_2 = 0.75$$

$$\alpha_3 = 0.75$$

Ex:  $X = (4, 2) \Rightarrow H(X) = (1 \cdot 4 + 0 \cdot 2) - 2 = 2 > 0 \Rightarrow X$  class is +1

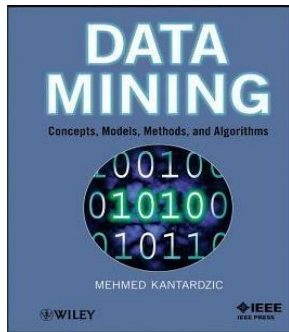


# Références



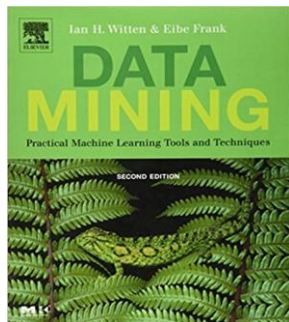
## **Data Mining : concepts and techniques, 3rd Edition**

- ✓ Auteur : Jiawei Han, Micheline Kamber, Jian Pei
- ✓ Éditeur : Morgan Kaufmann Publishers
- ✓ Edition : Juin 2011 - 744 pages - ISBN 9780123814807



## **Data Mining : concepts, models, methods, and algorithms**

- ✓ Auteur : Mehmed Kantardzi
- ✓ Éditeur : John Wiley & Sons
- ✓ Edition : Aout 2011 – 552 pages - ISBN : 9781118029121



## **Data Mining: Practical Machine Learning Tools and Techniques**

- ✓ Auteur : Ian H. Witten & Eibe Frank
- ✓ Éditeur : Morgan Kaufmann Publishers
- ✓ Edition : Juin 2005 - 664 pages - ISBN : 0-12-088407-0

# Références

Cours – Abdelhamid DJEFFAL – Fouille de données avancée

✓ [www.abdelhamid-djeffal.net](http://www.abdelhamid-djeffal.net)

WekaMOOC – Ian Witten – Data Mining with Weka

✓ <https://www.youtube.com/user/WekaMOOC/featured>

Cours - Phugo Larochelle – Apprentissage Automatique

✓ [https://www.youtube.com/playlist?list=PL6Xpj9I5qXYFD\\_rc1tttugXLfE2](https://www.youtube.com/playlist?list=PL6Xpj9I5qXYFD_rc1tttugXLfE2TcKyIO)

TcKyIO

Chris McCormick - SVM Example

✓ <http://mccormickml.com/2013/04/16/trivial-svm-example/>