Fouille de Données

Data Mining

Classification - Partie 5

Plan du cours

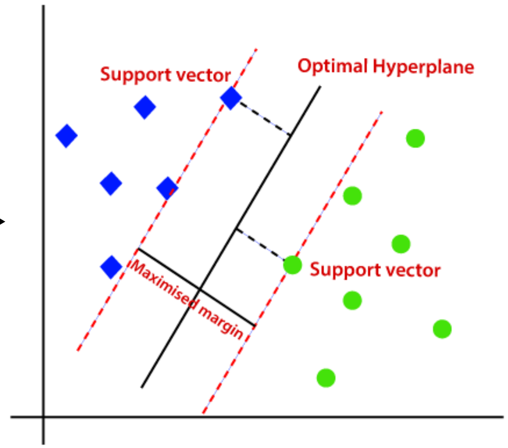
1. Les Machines à Vecteurs de Support - SVM.
2. SVM linéaire.
3. Implémentation des SVM.
4. Exemple déroulement.

Classification : Algorithmes

SAVOIR - PREDIRE - DECIDER

Données

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



Connaissances

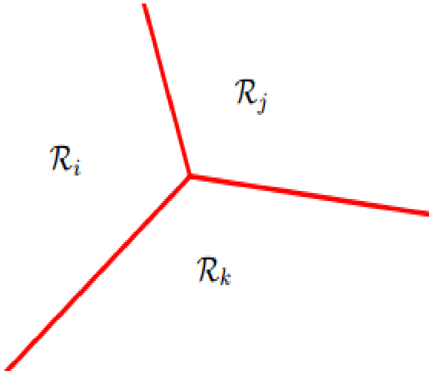
SVM

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 :

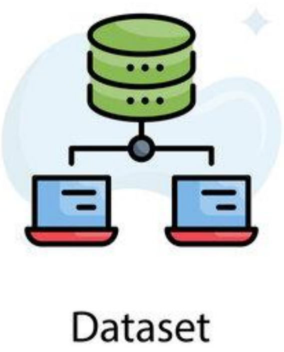
1. Rechercher un **hyperplan** qui sépare le mieux l'espace des exemples en différentes **régions de décisions**.

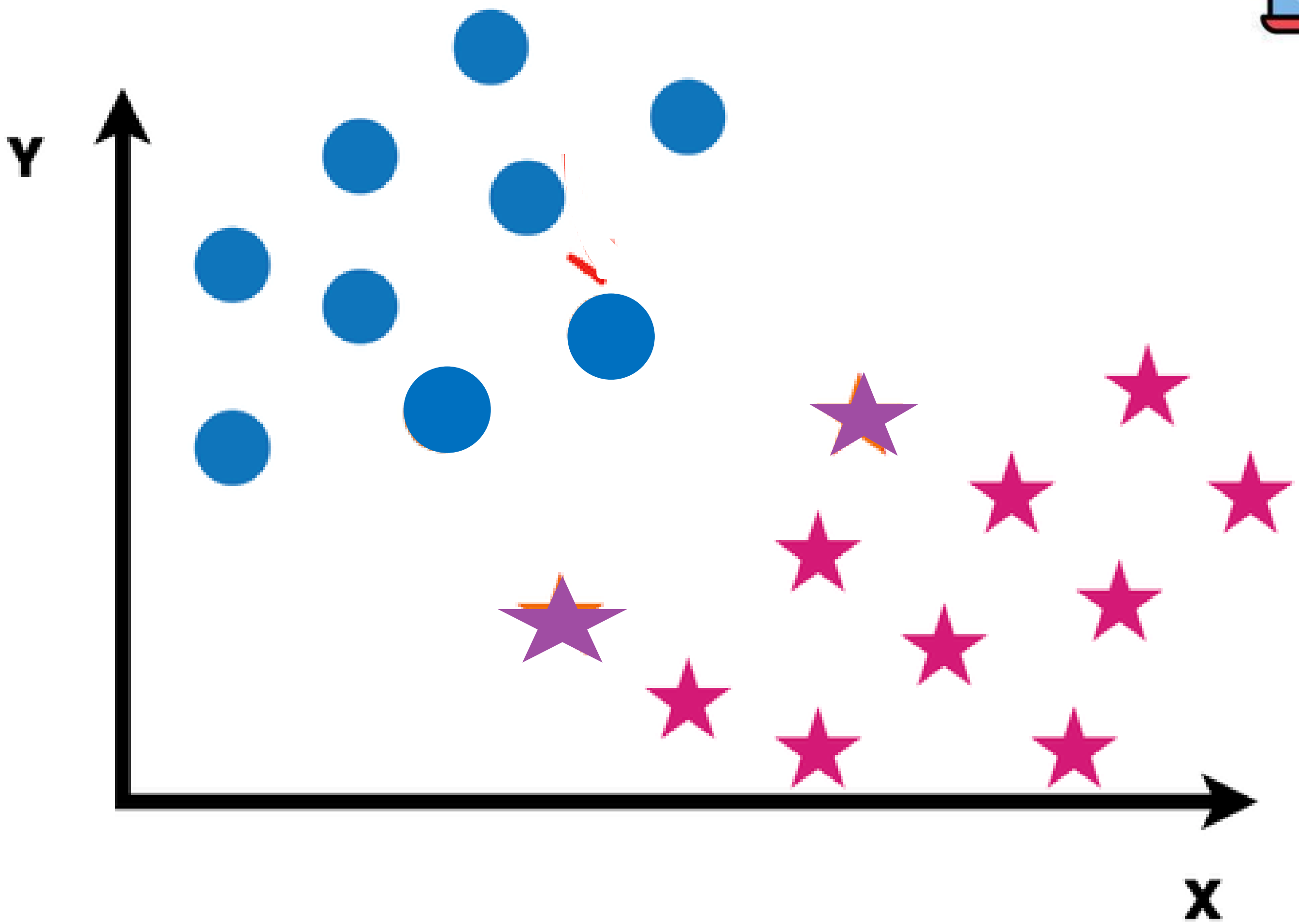
2. Une **frontière de décision** séparant les exemples d'une classe à une autre. Chaque région est associée à une classe.

3. Une recherche qui s'appuie uniquement sur quelques exemples de la base d'entrainement, appelés **Vecteurs Support**, ainsi que la **marge** entre eux.

Machines à Vecteurs de Support

Principe de base:

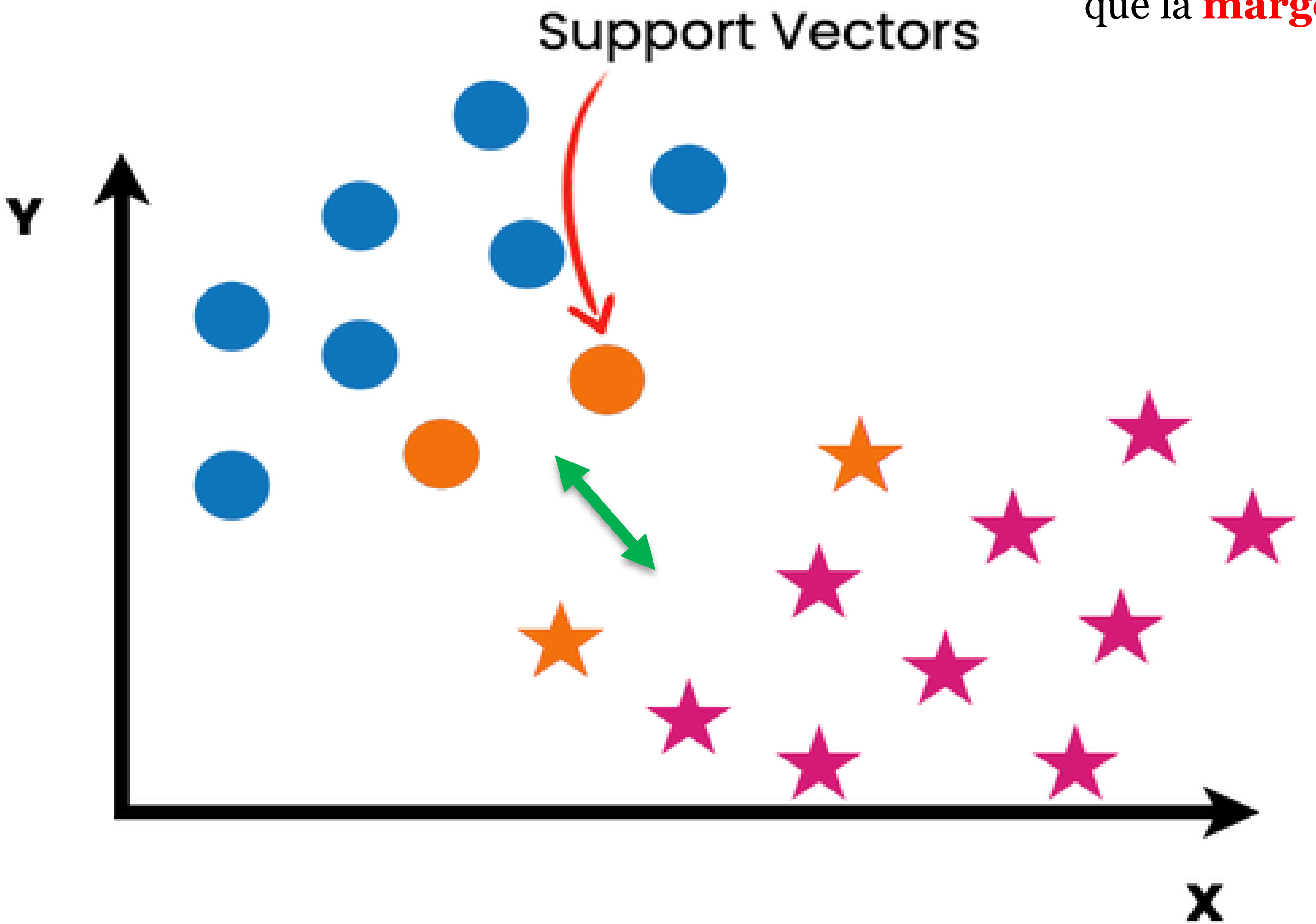




Machines à Vecteurs de Support

- **Principe de base:** Comment ?

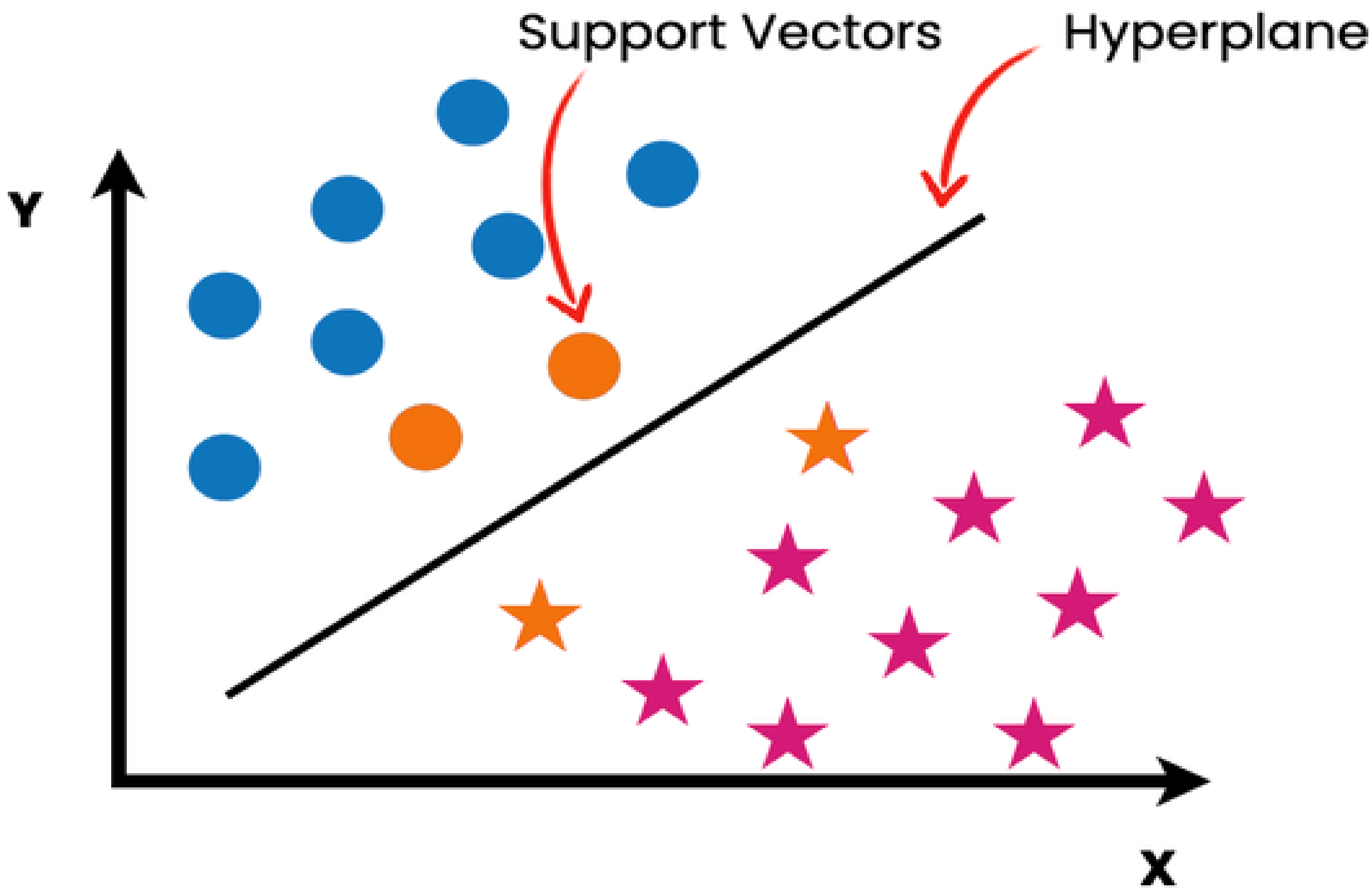
S'appuie sur quelques exemples de la base d'entrainement, appelés **Vecteurs Support**, ainsi que la **marge** entre eux.



Machines à Vecteurs de Support

- **Principe de base:** Pourquoi ?

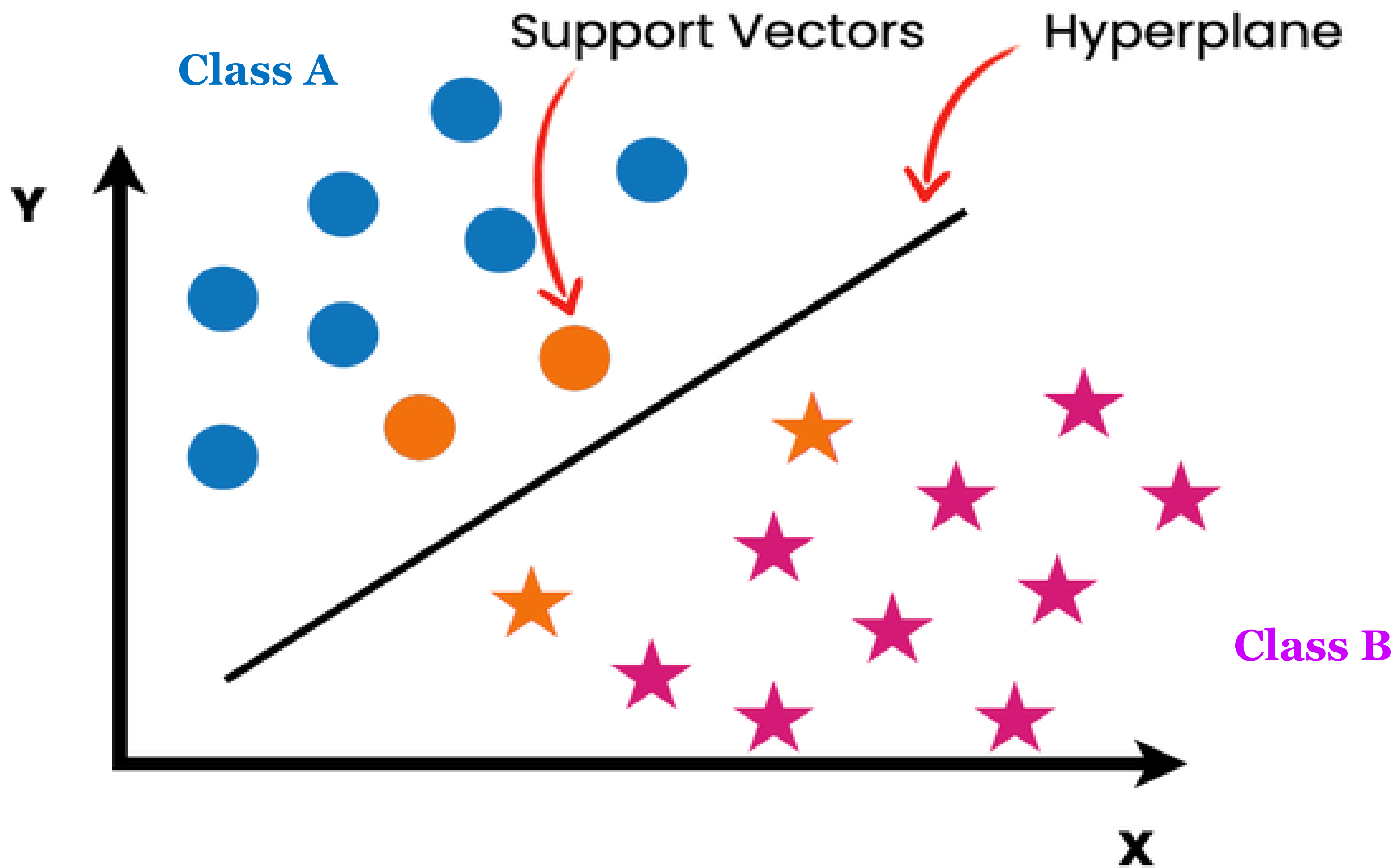
Rechercher un **hyperplan** qui sépare le mieux l'espace des exemples en différentes **régions de décisions**.



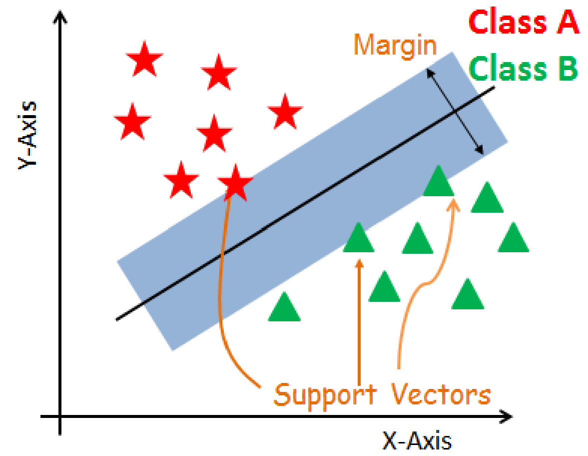
Machines à Vecteurs de Support

- **Principe de base:** Pourquoi ?

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



Machines à Vecteurs de Support



- **Concepts de base en SVM :**

- **Marge** - La marge est l'écart entre l'hyperplan et les vecteurs de support.

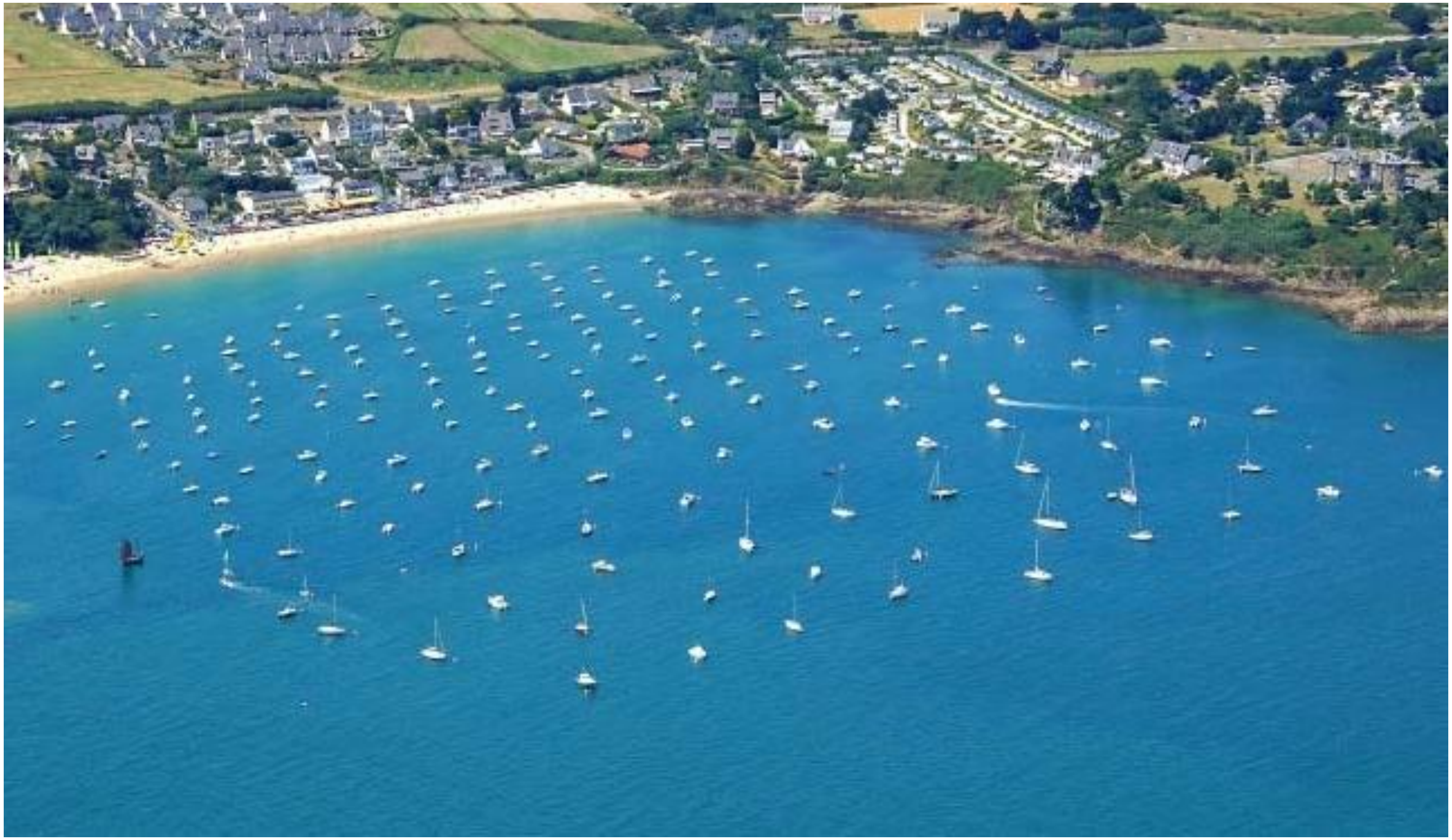
- **Hyperplan** - Les hyperplans sont des frontières de décision qui aident à classer les points de données.

- **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:**

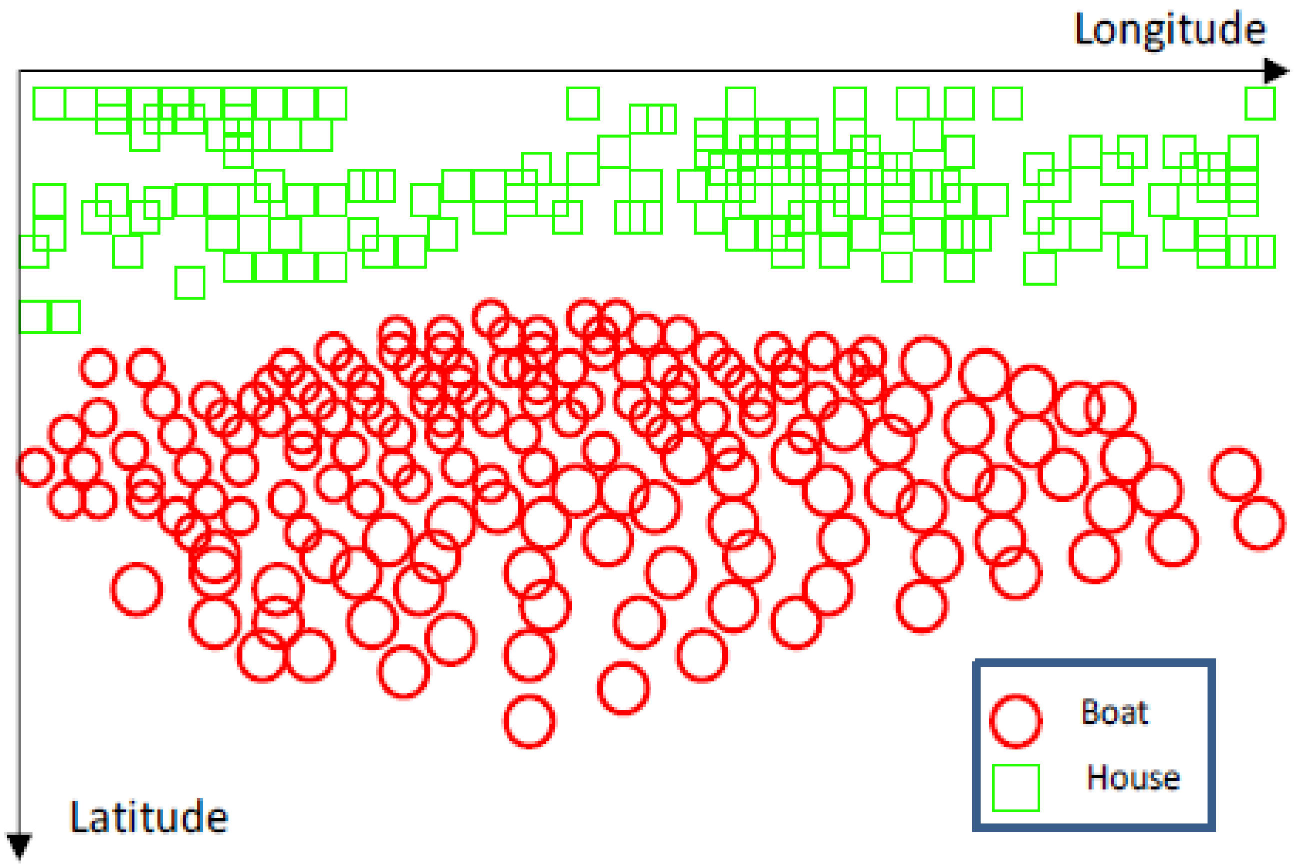


Machines à Vecteurs de Support

- **Principe de base:**

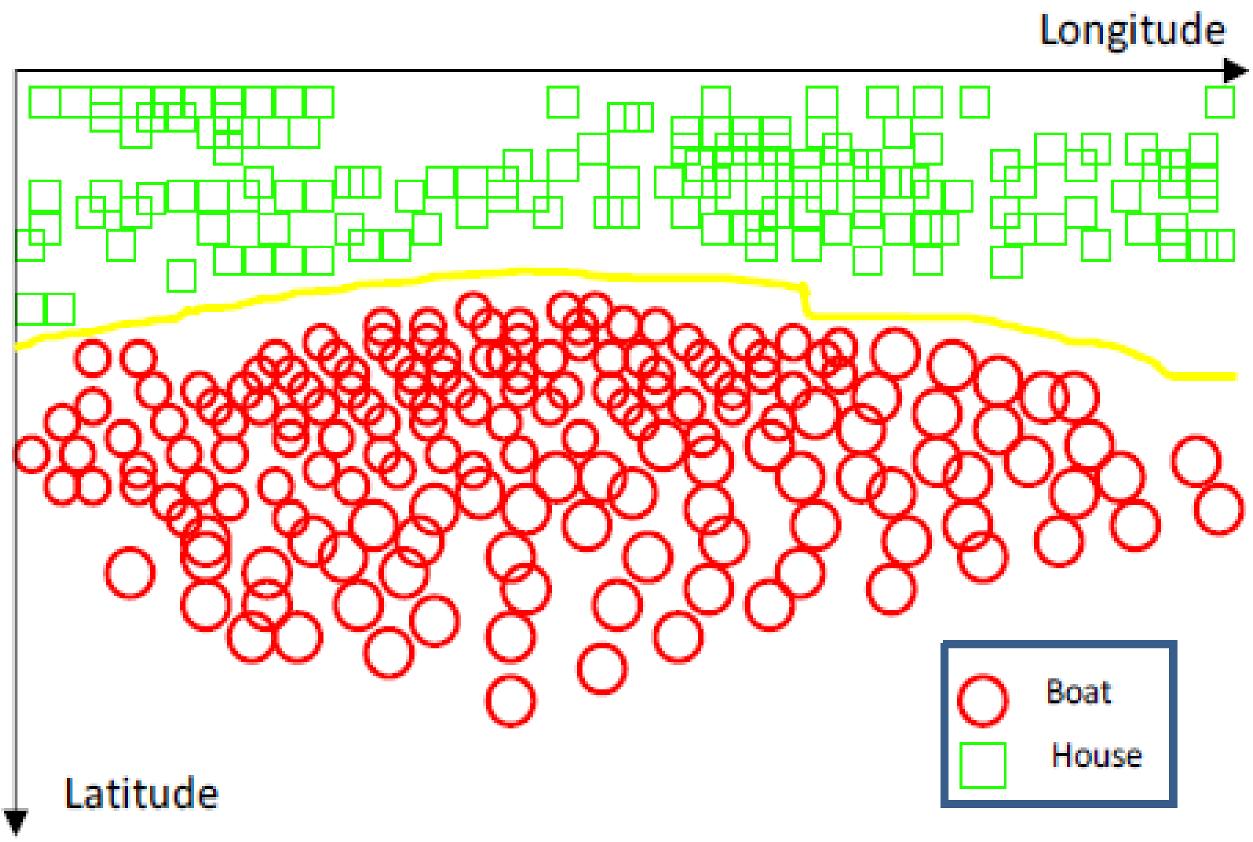
Machines à Vecteurs de Support

- **Principe de base:**



Machines à Vecteurs de Support

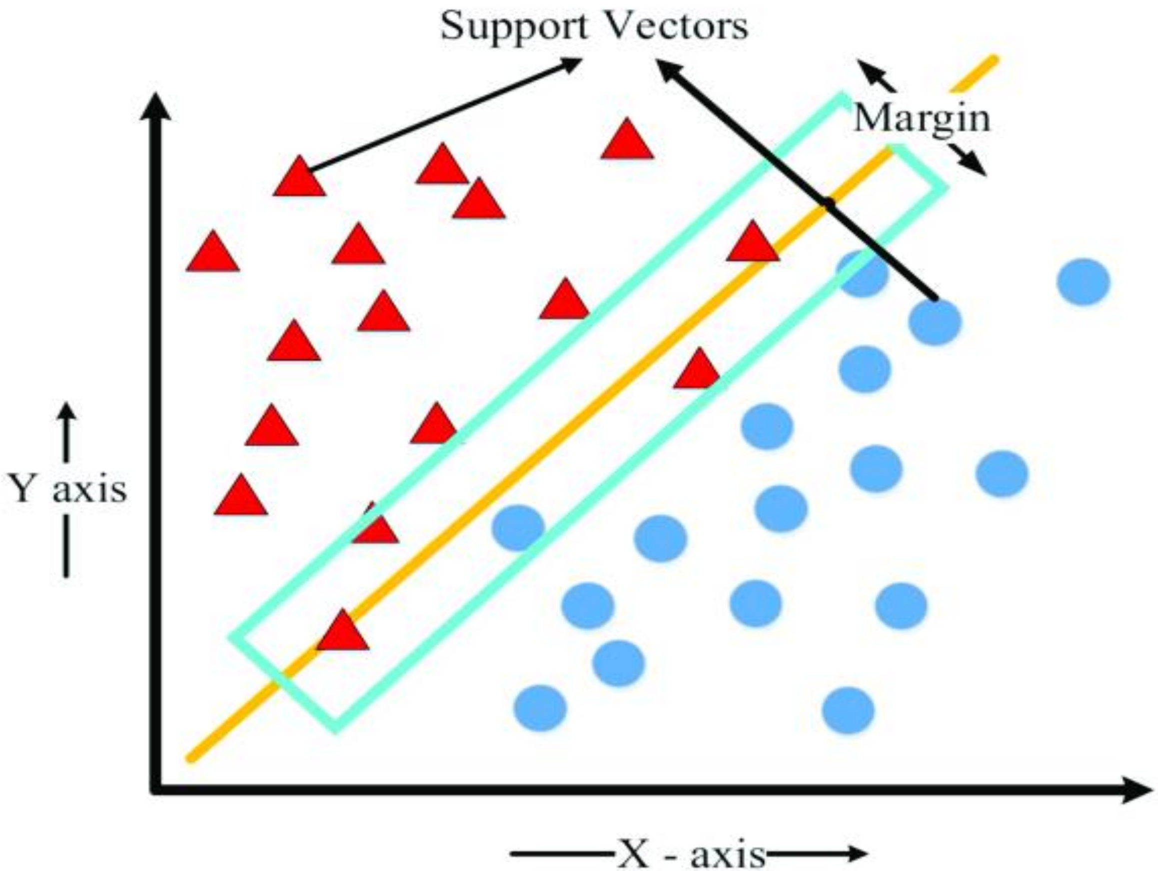
- **Principe de base:**



frontière de décision

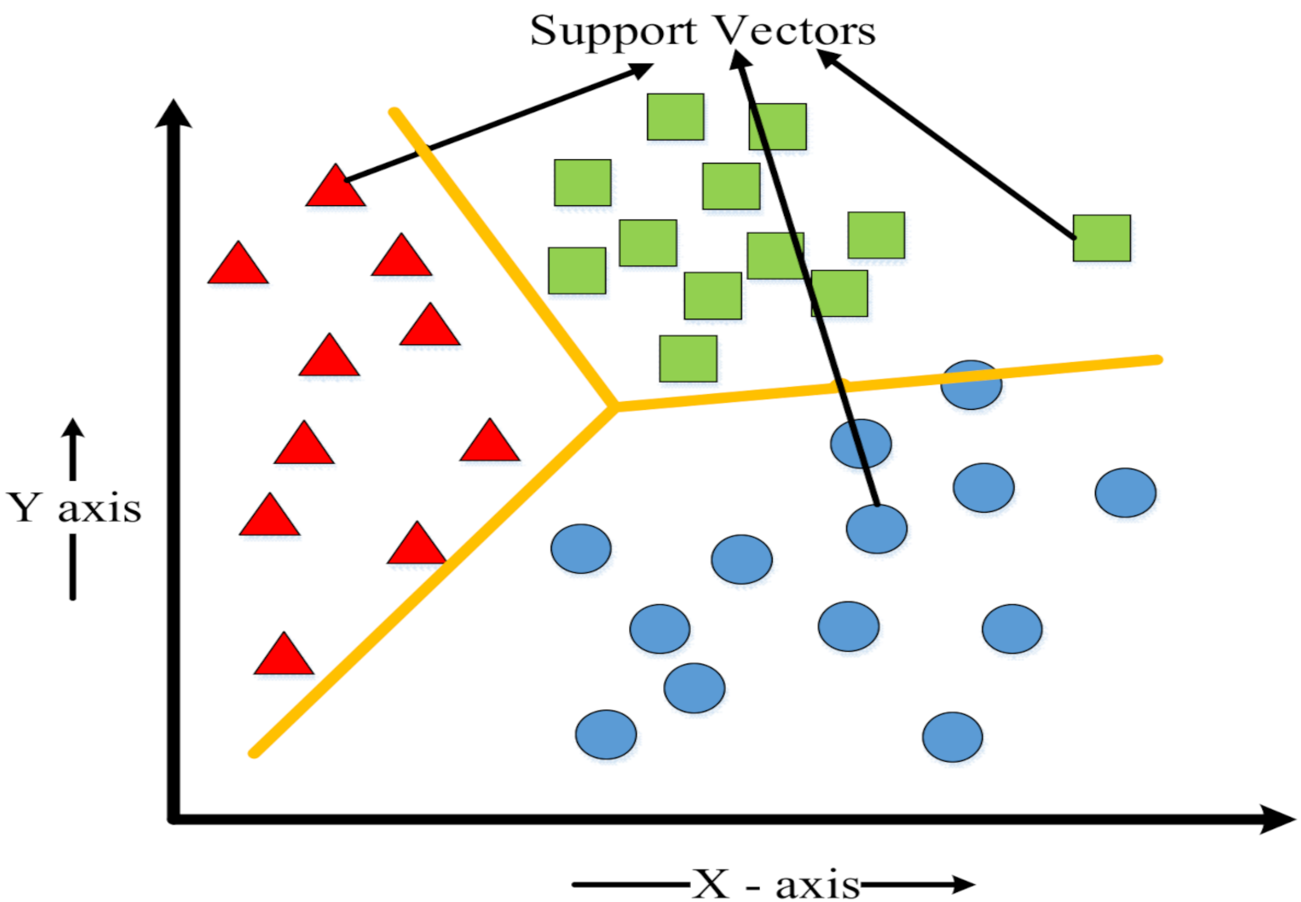
Machines à Vecteurs de Support

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



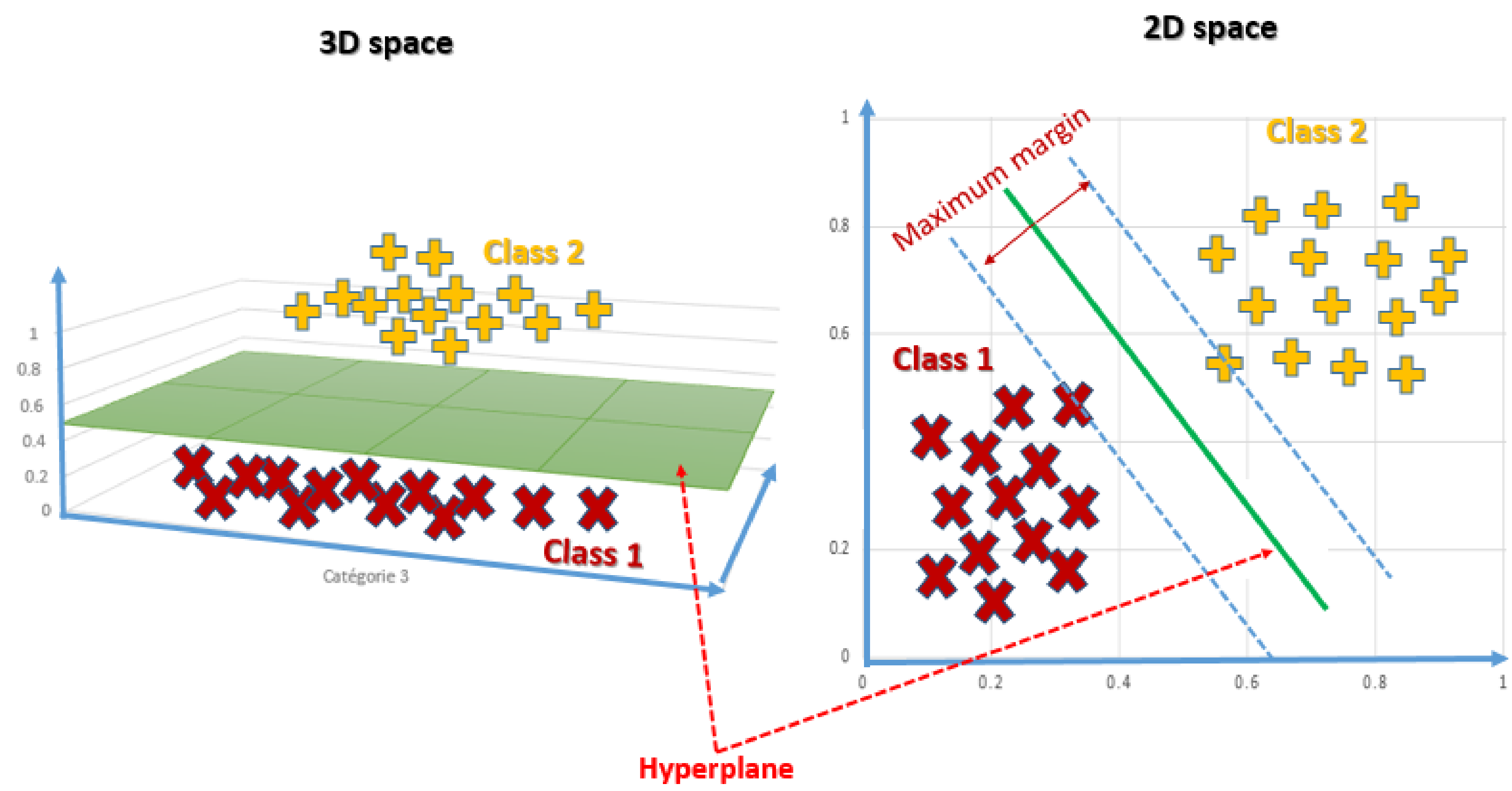
Machines à Vecteurs de Support

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



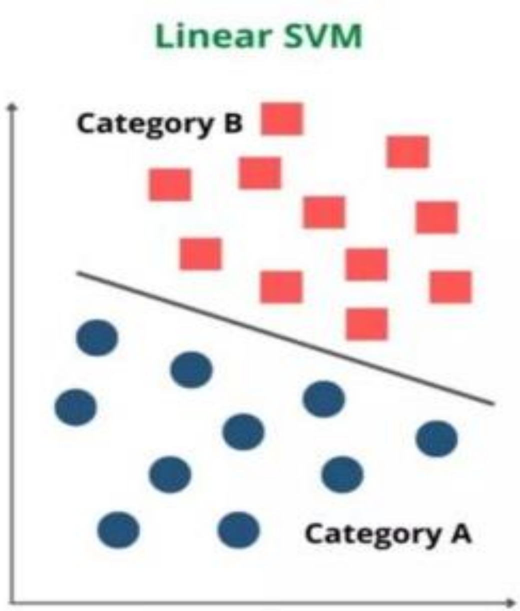
Machines à Vecteurs de Support

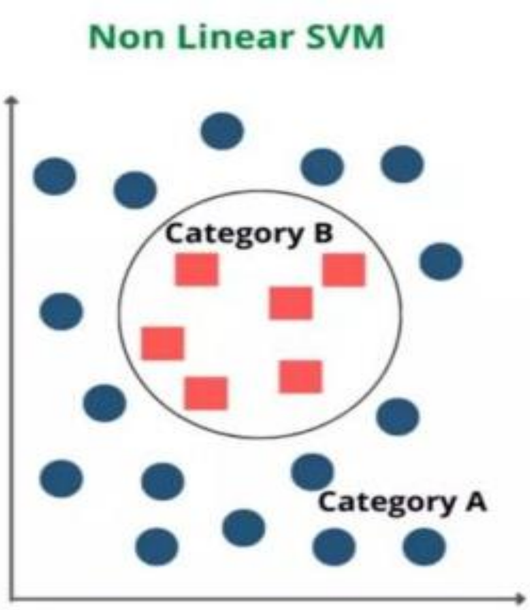
- **L'hyperplan:** Cas d'une classification **binaire** à deux attributs => **Plan 3D, 2 régions, séparateur =** **plan**. **> 3D : hyperplan,**



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.



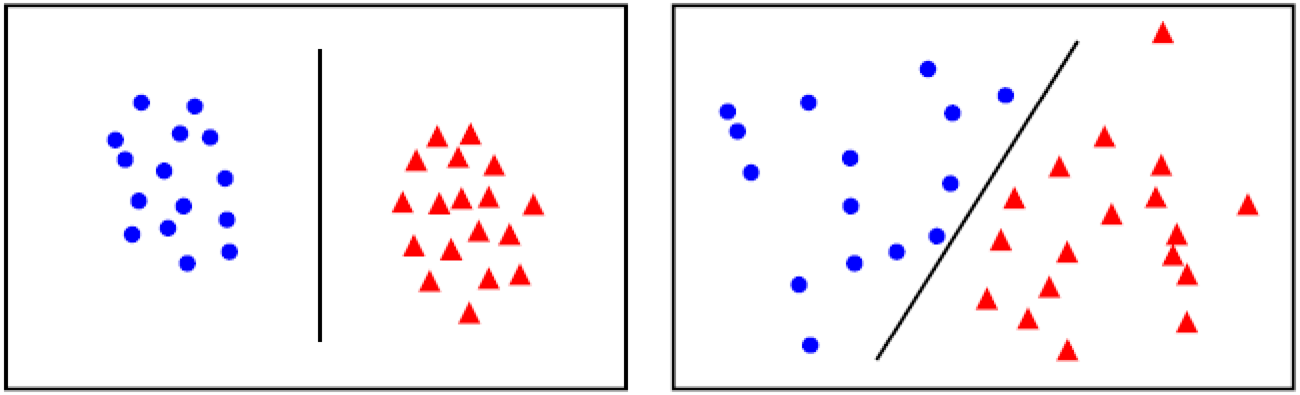


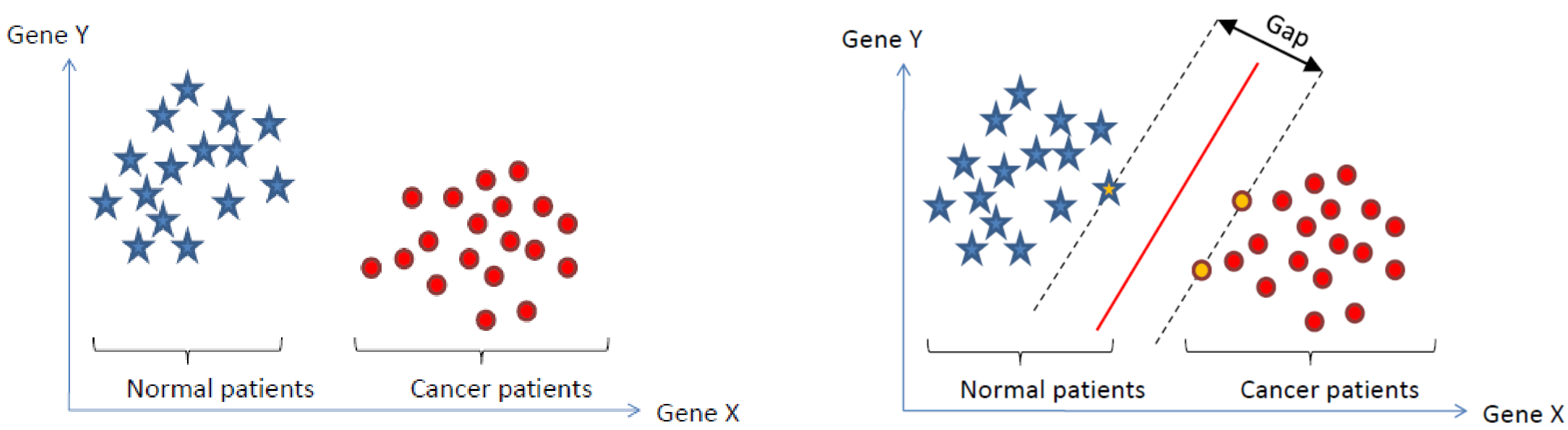
Machines à Vecteurs de Support

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.



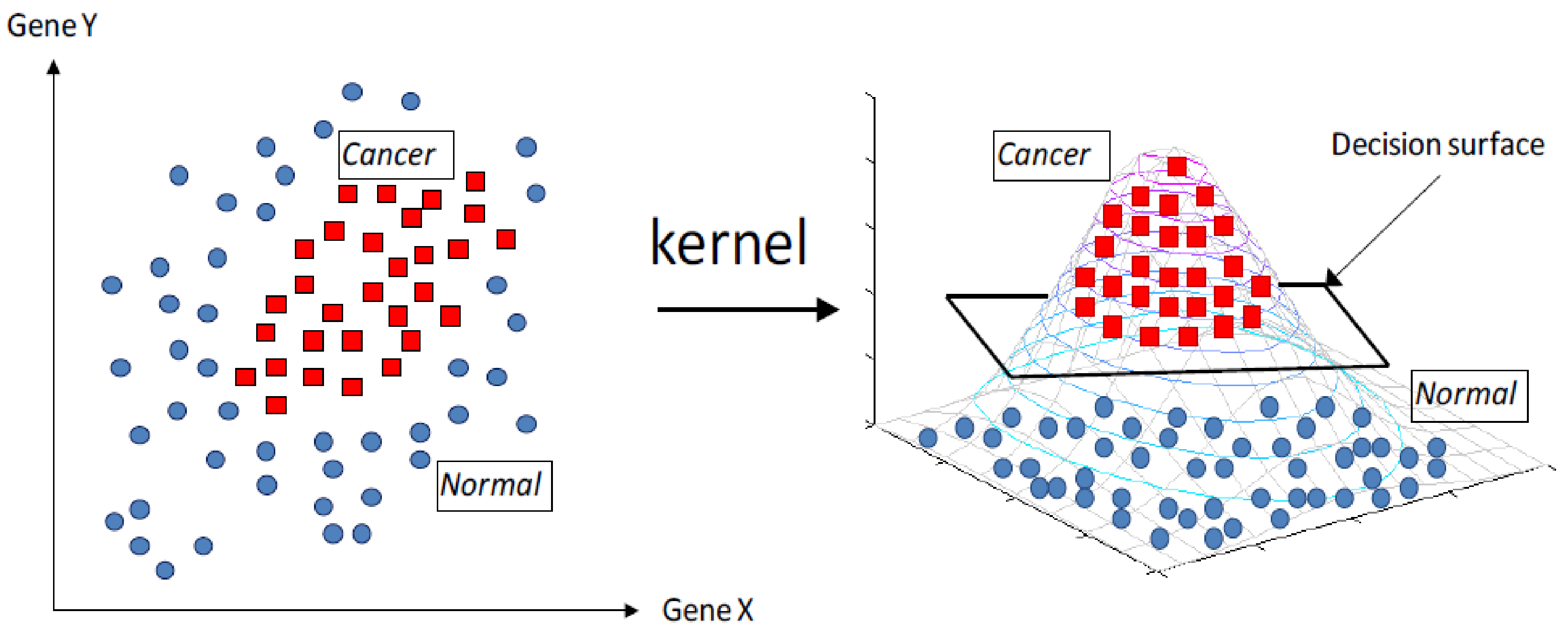


Machines à Vecteurs de Support

Deux cas du problème posé :

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

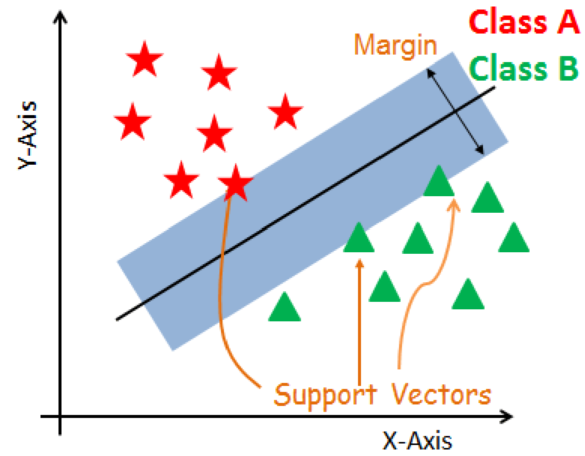
On dit qu'un problème est linéairement séparable si une surface linéaire permet de classifier parfaitement.



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.

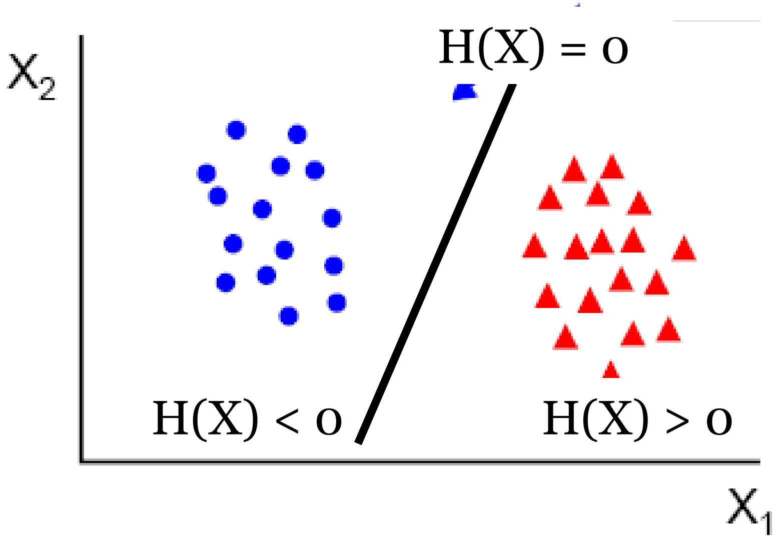
SVM Linéaire

Cas de séparabilité linéaire : Linear SVM

- Supposons une classification binaire, deux classes :
  - Yi = **{-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

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

Cas de séparabilité linéaire

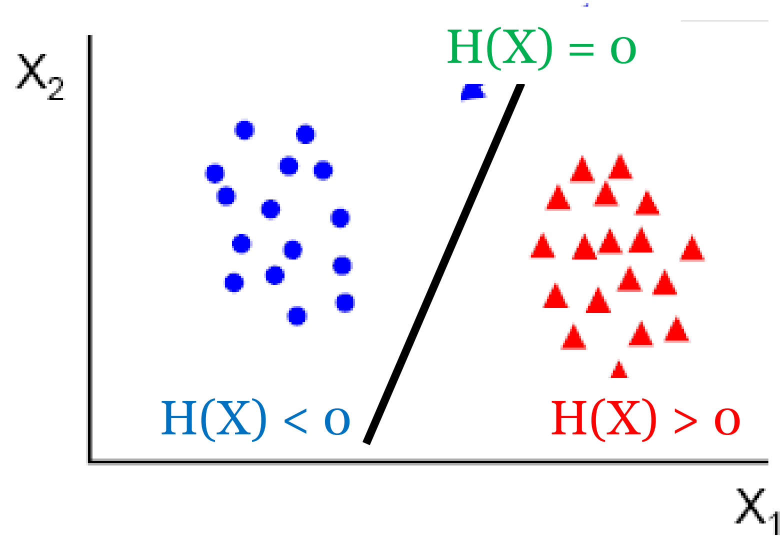
- Supposons deux classes :
  - Yi = **{-1, +1}**

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

- 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



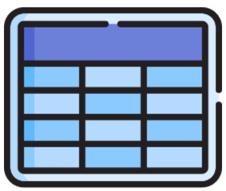
SVM Linéaire

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

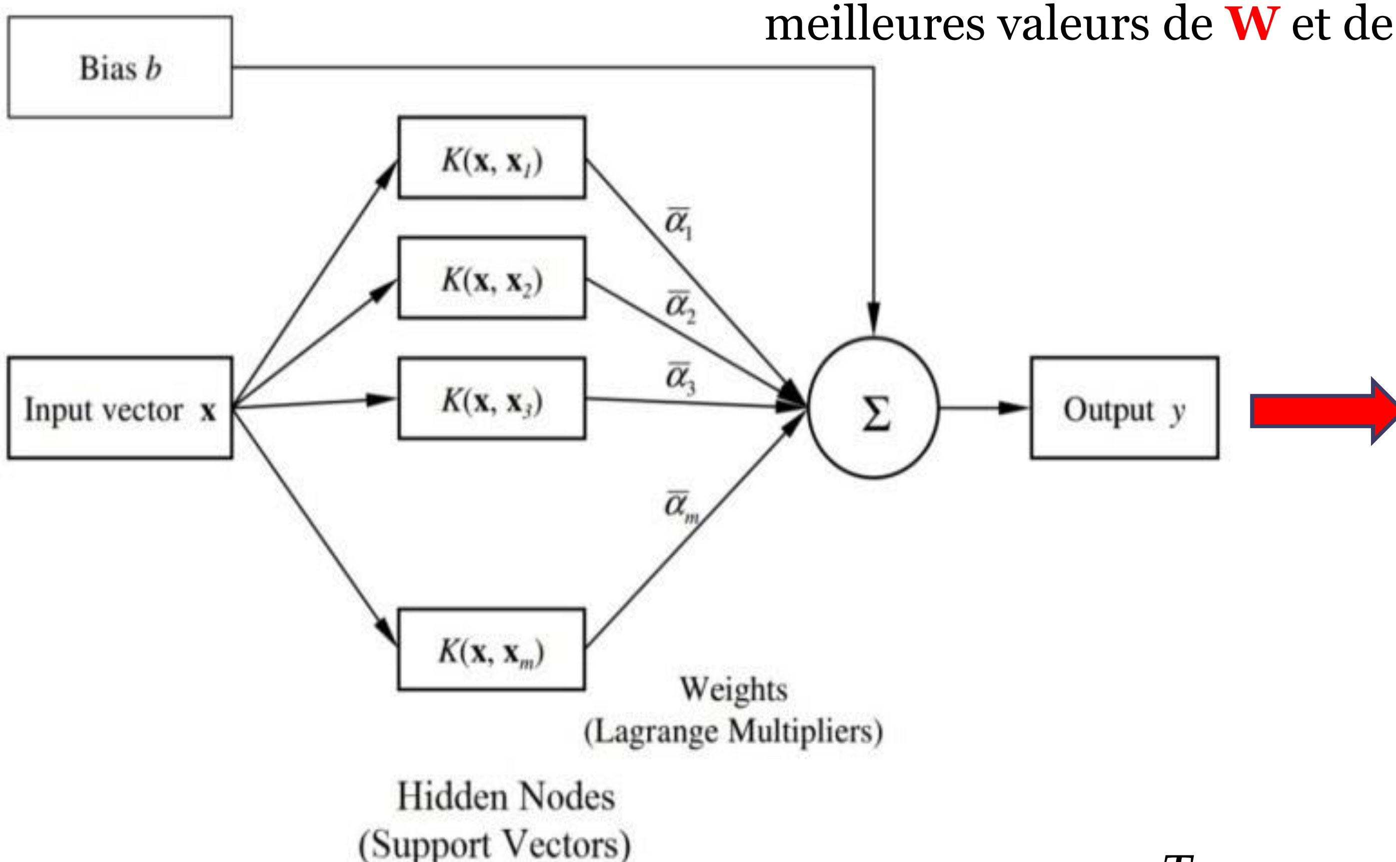
<u>Cas de séparabilité linéaire</u>

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

X



Training Dataset



Predict True Class of the Input Vector

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

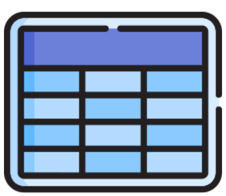
SVM Linéaire

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

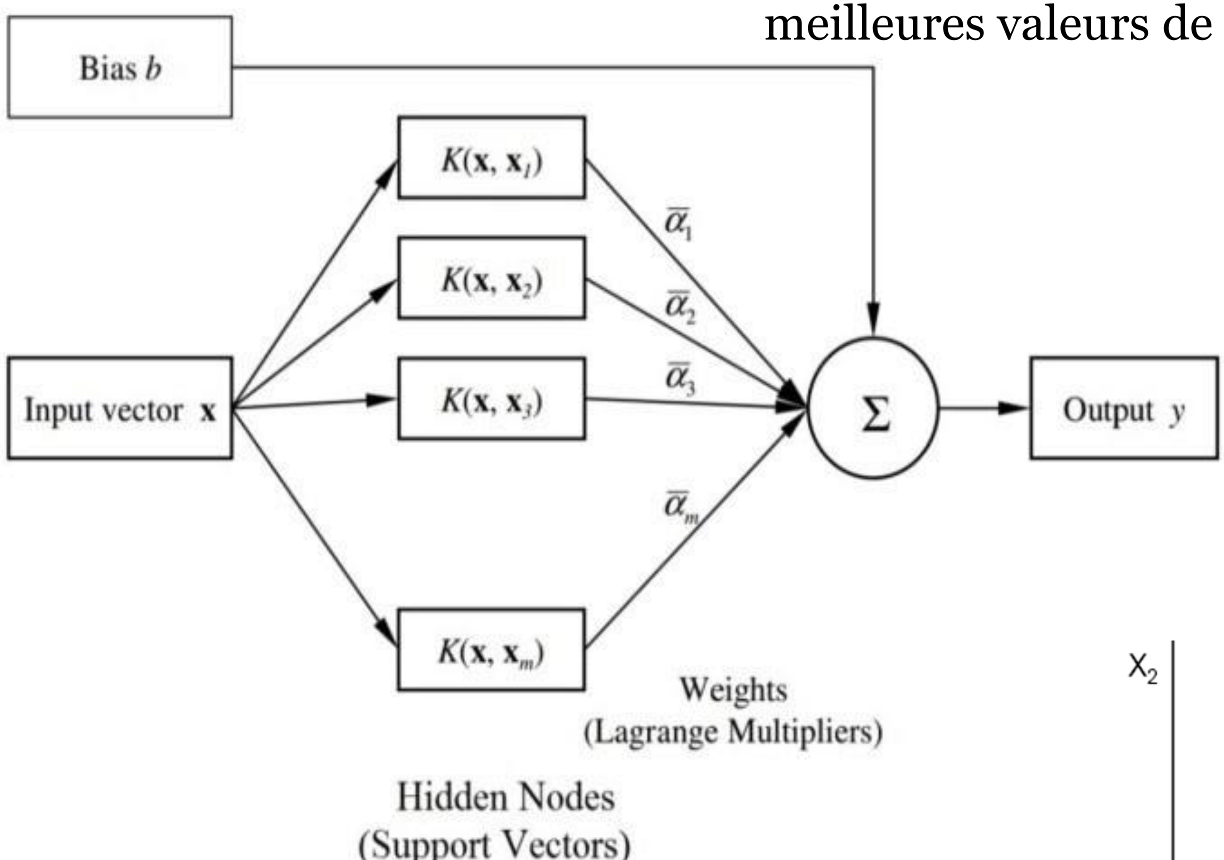
<u>Cas de séparabilité linéaire</u>

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

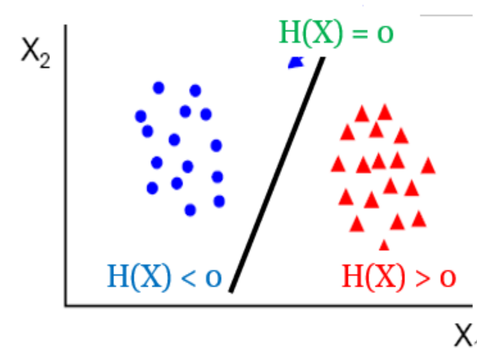
X



Training Dataset



Based on the Sign of H(x)



SVM Linéaire

<u>Cas de séparabilité linéaire</u>

➢ Supposons deux classes :

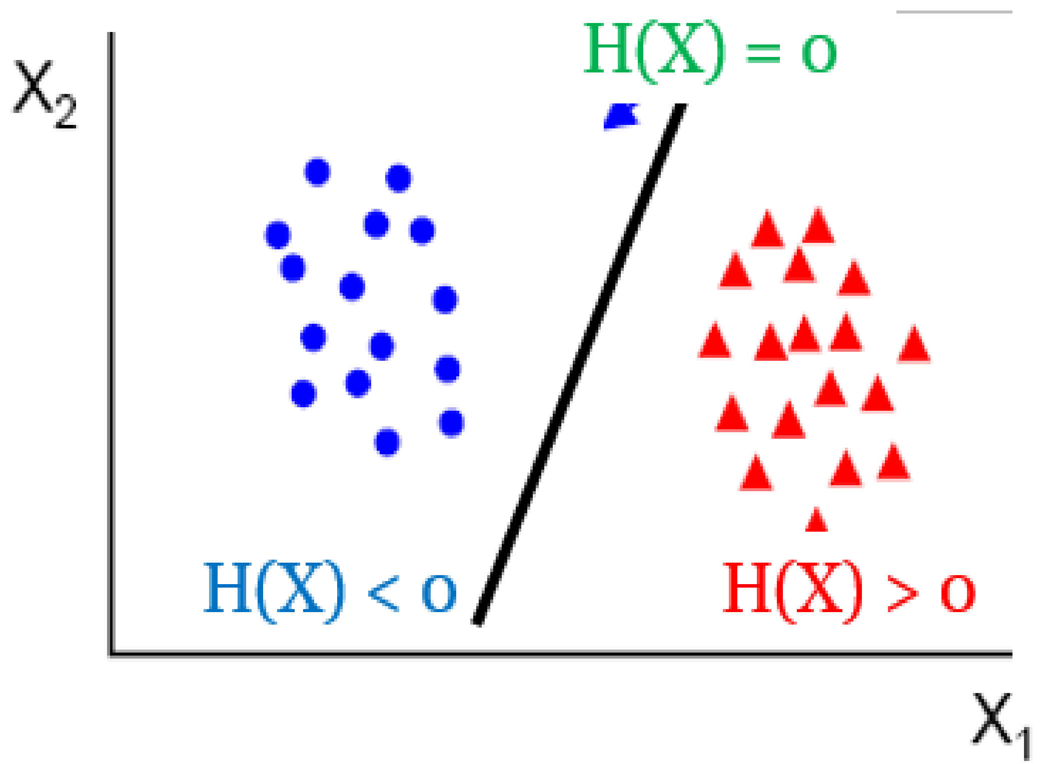
- Yi = **{-1, +1}**

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

W: Vecteur de poids
b : biais

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

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



SVM Linéaire

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

Cas de séparabilité linéaire

$$\begin{cases} Ck = +1 \quad \text{Si } H(X) > 0 \quad : \quad H+ \\ Ck = -1 \quad \text{Si } H(X) < 0 \quad : \quad H- \end{cases}$$

Déterminer la classe selon le **signe (H(x))**

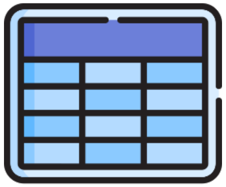
- Aucun exemple ne se situe sur l'hyperplan, donc :

$$\begin{cases} \mathbf{Ck = +1 \quad Si\ H(X) >= 1} \\ \mathbf{Ck = -1 \quad Si\ H(X) <= -1} \end{cases}$$

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

➔ SVM doit satisfaire : - i.e. doit correctement classifié tous les **Xi** (les datapoints dans le dataset) pour être correctement entrainé :

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

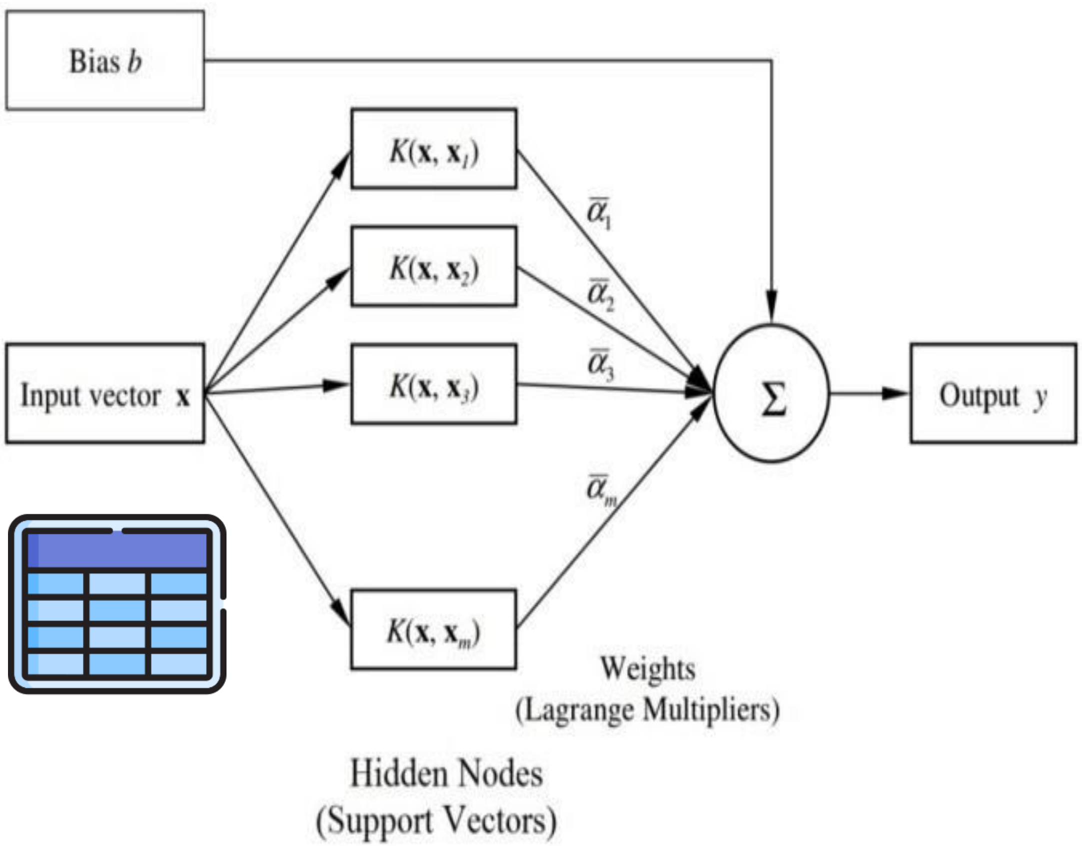


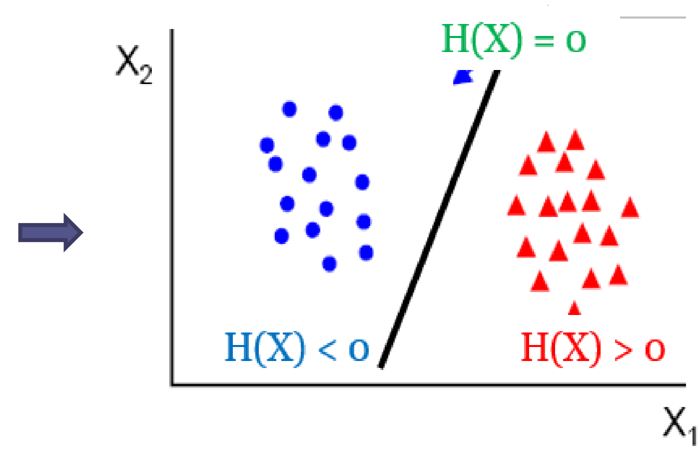
SVM Linéaire

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

<u>Cas de séparabilité linéaire</u>

- SVM doit satisfaire : - i.e. doit correctement classifié tous les **Xi** (les datapoints dans le dataset) pour être correctement entrainé :





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

SVM Linéaire

Cas de séparabilité linéaire

- Le but d'un SVM lors de l'entrainement est donc de trouver :

- Un vecteur de poids **W** et un biais **b** tel que :

- Pour tout xi ayant comme classe yi appartenant aux données d'entrainement :

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

- Etapes SVM:

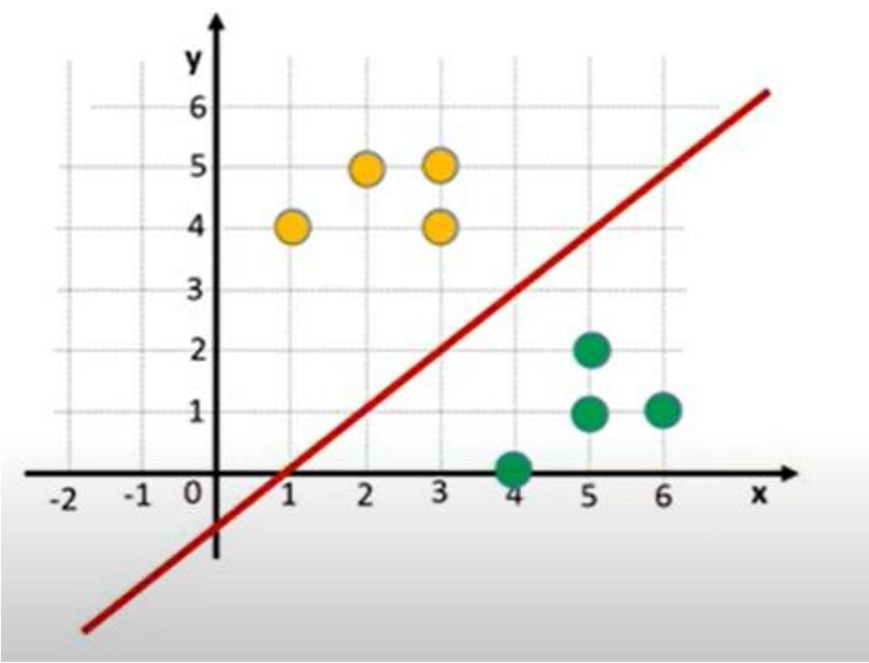
1 - Générer et tester des hyperplans qui séparent les classes, en utilisant et testant différentes valeurs de **W** et de **b.**

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

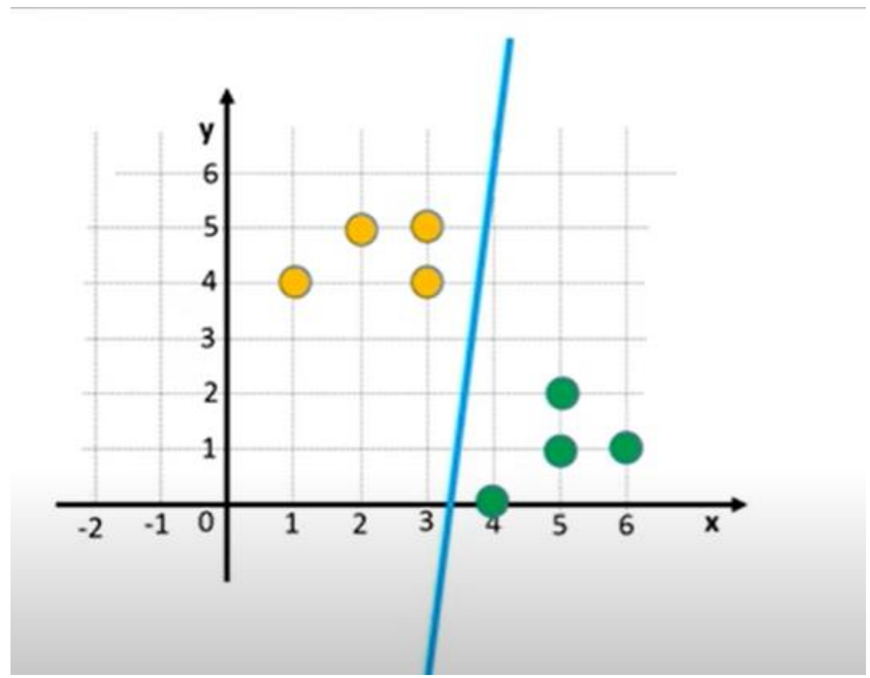
SVM Linéaire

Cas de séparabilité linéaire

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



Hyperplan 1

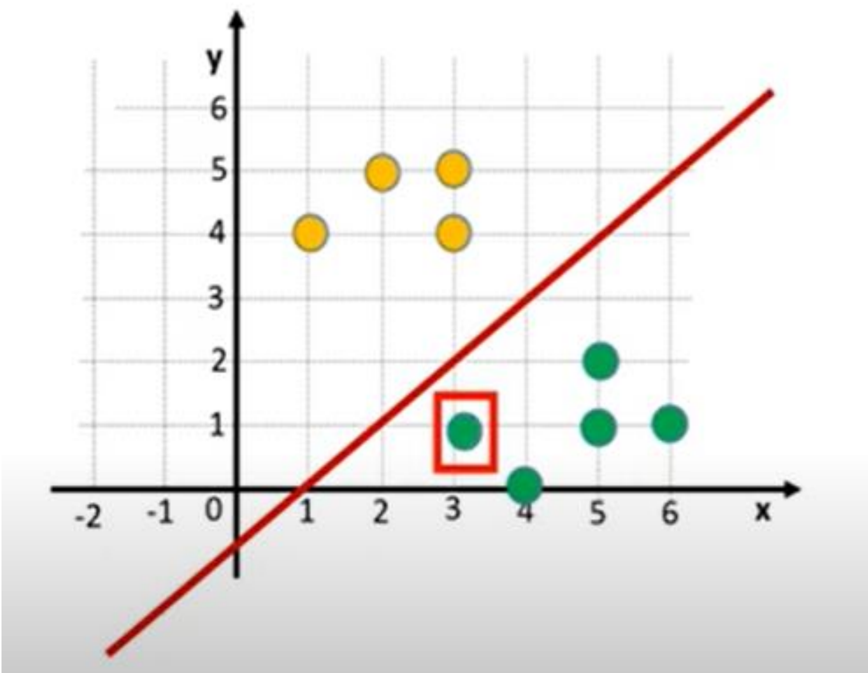


Hyperplan 2

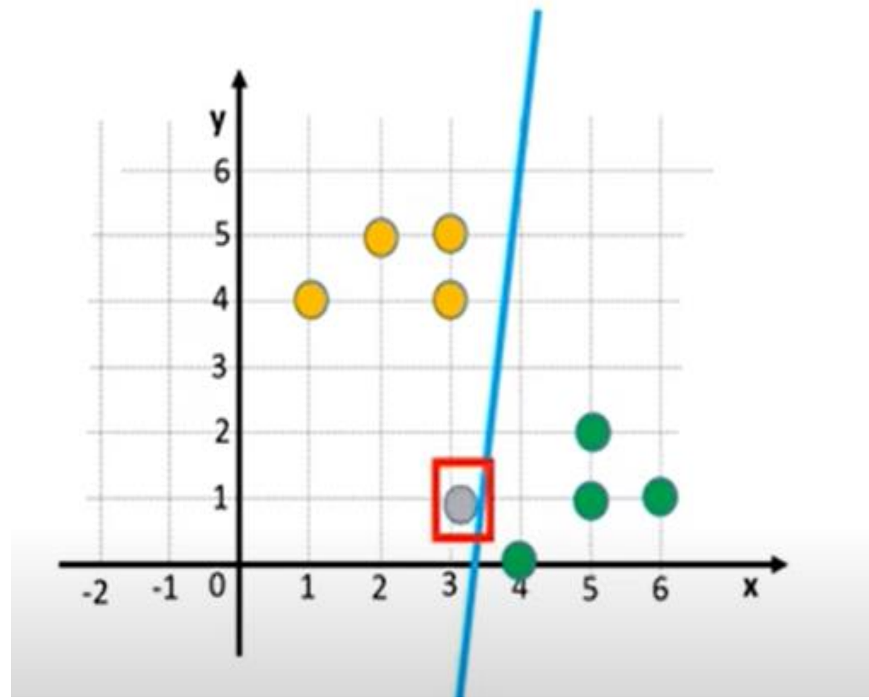
SVM Linéaire

Cas de séparabilité linéaire

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



Prédire un nouvel exemple : correct

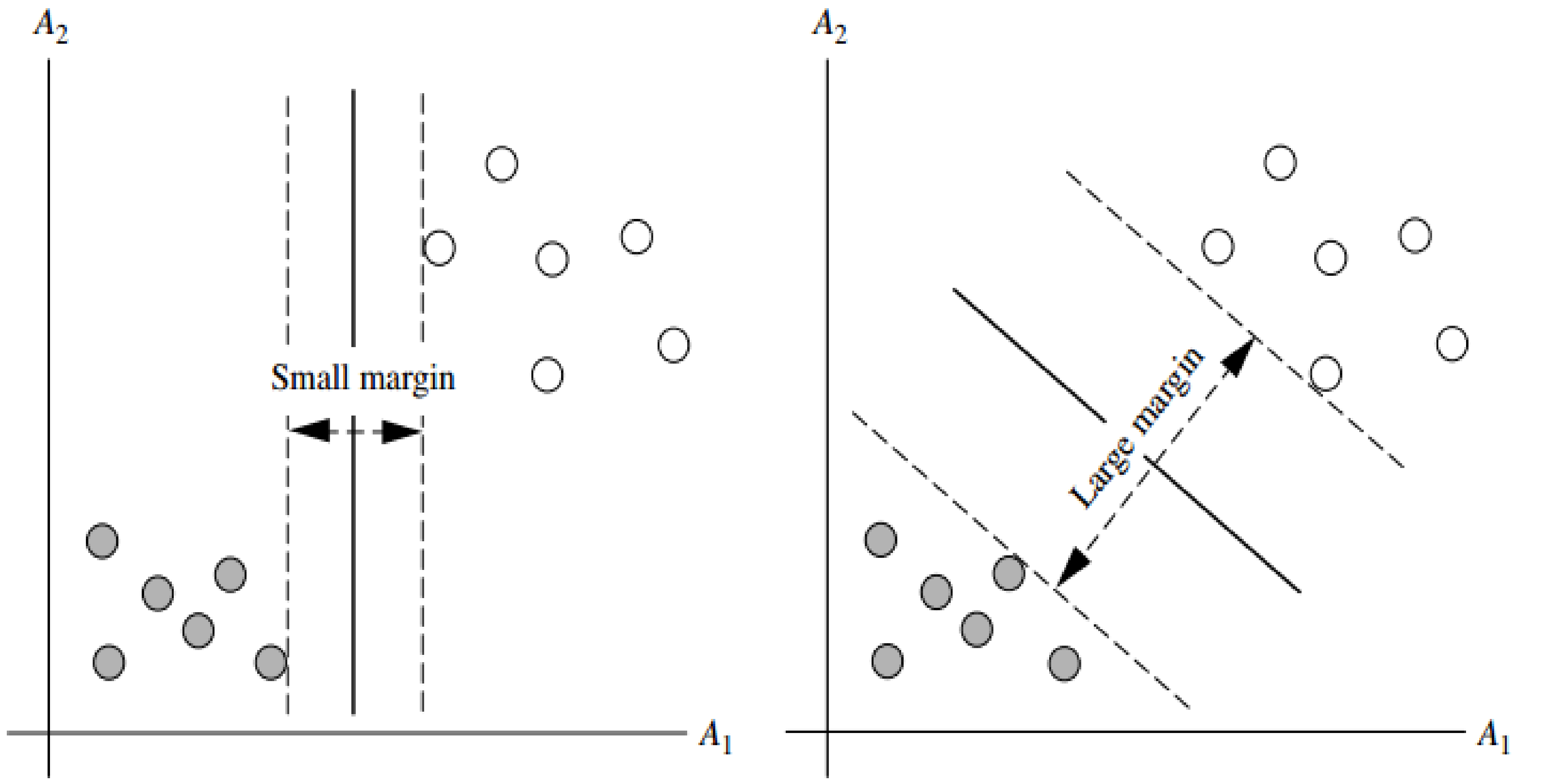


Prédire un nouvel exemple : faux

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 ?

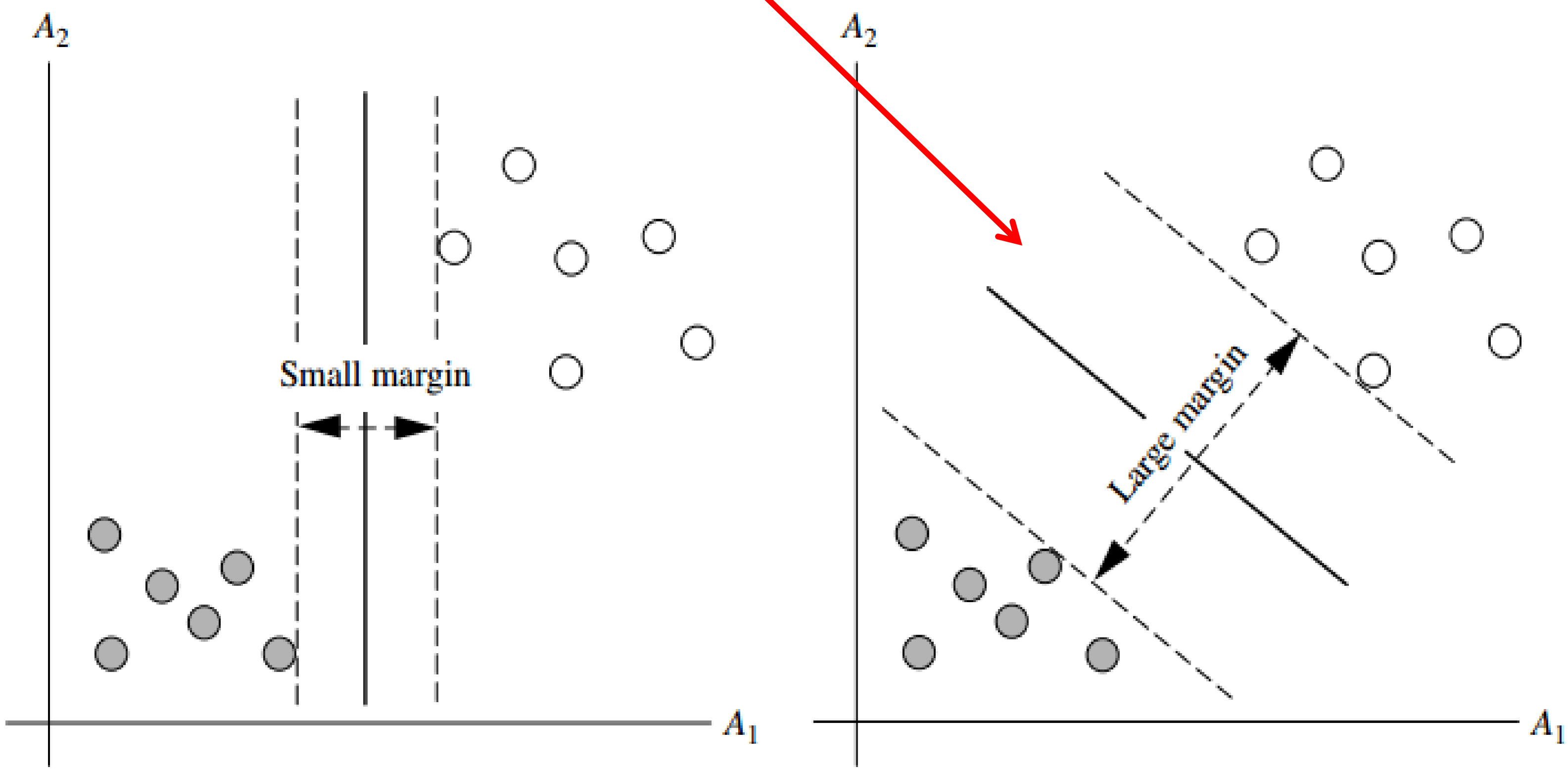


SVM Linéaire

Cas de séparabilité linéaire

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

Solution : **Maximiser la marge**

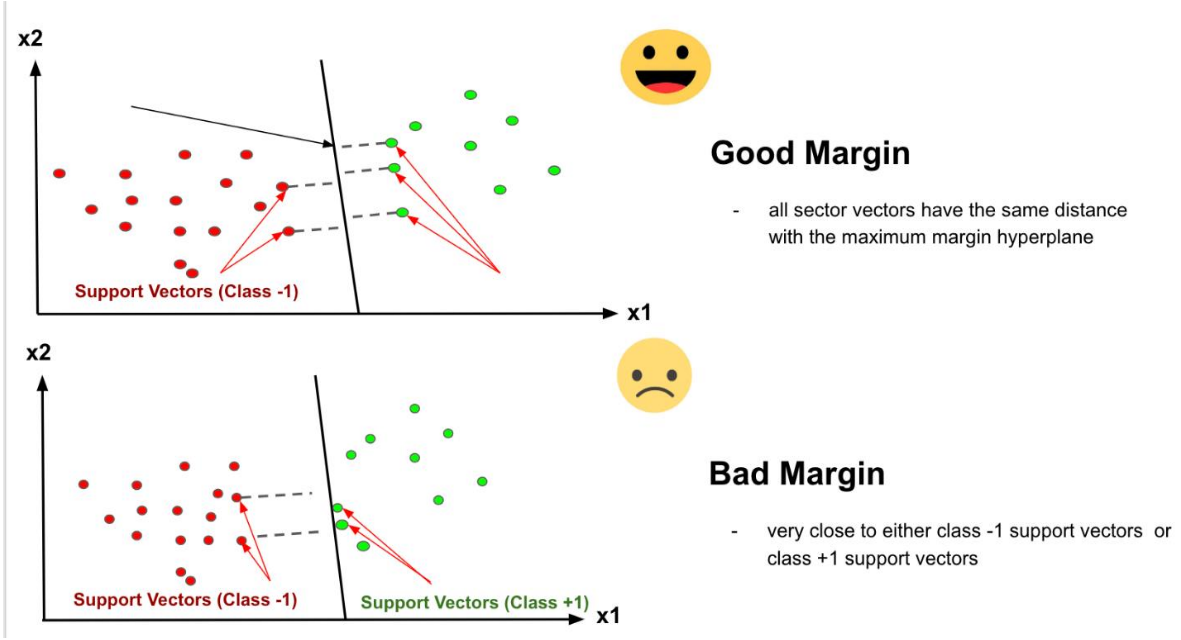


SVM Linéaire

<u>Cas de séparabilité linéaire</u>

- Etapes SVM:

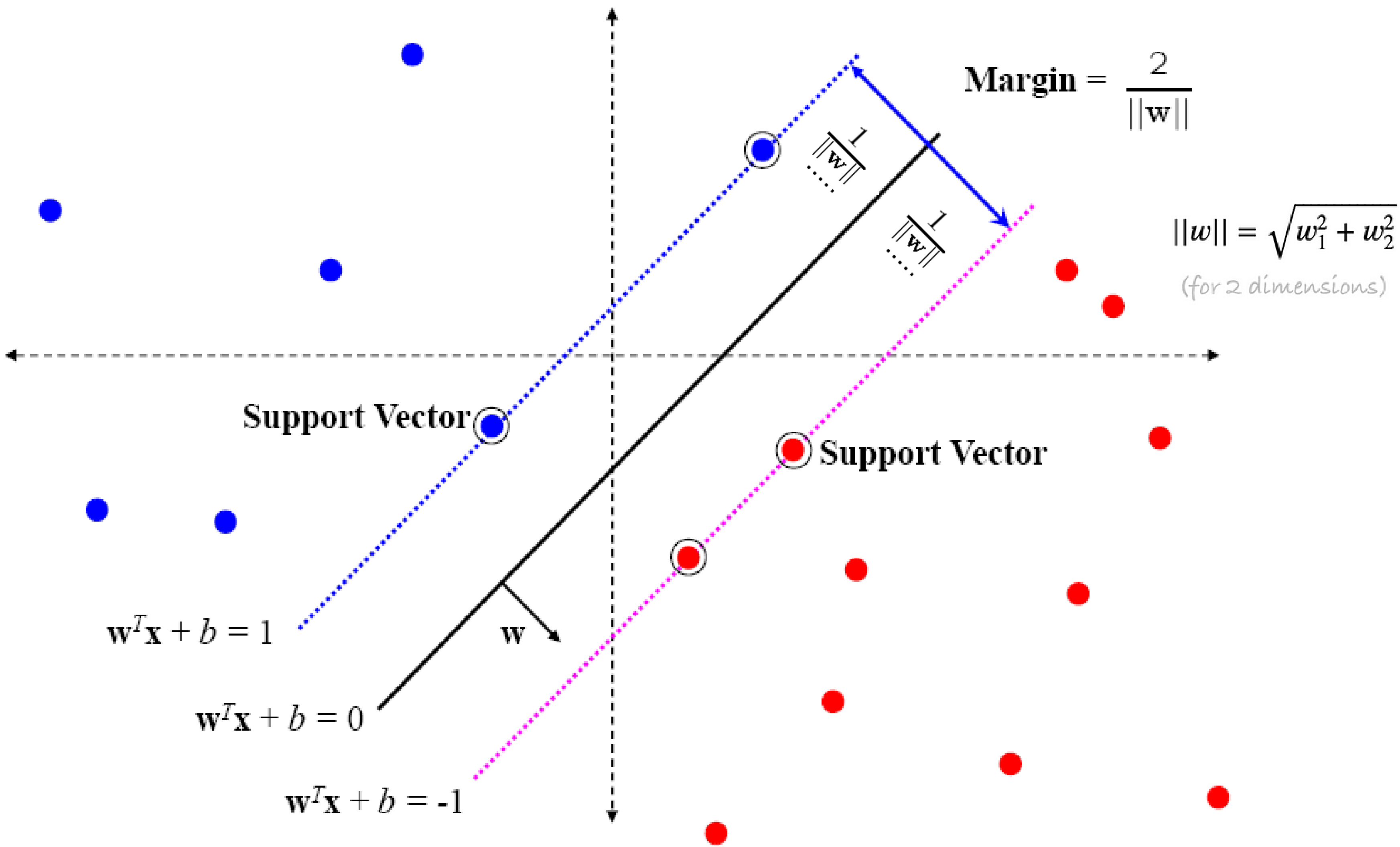
2 – **Calculer la marge** avec chaque hyperplan généré et sélectionné l'hyperplan qui a une **marge maximale** séparant le mieux les données (voire les **vecteurs supports** seulement, trouvés lors de l'entrainement).



SVM Linéaire

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

<u>Cas de séparabilité linéaire</u>



SVM Linéaire

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

Cas de séparabilité linéaire

- Marge : distance entre deux hyperplans parallèles
- Marge = **2/||W||,** avec ||W|| : norme du vecteur W (w1, w2)

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

(for 2 dimensions)

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

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

- Déterminer le vecteur **W** qui possède la norme euclidienne minimale ;
- Et qui vérifie la contrainte de l'équation de bonne classification des exemples d'entrainement.

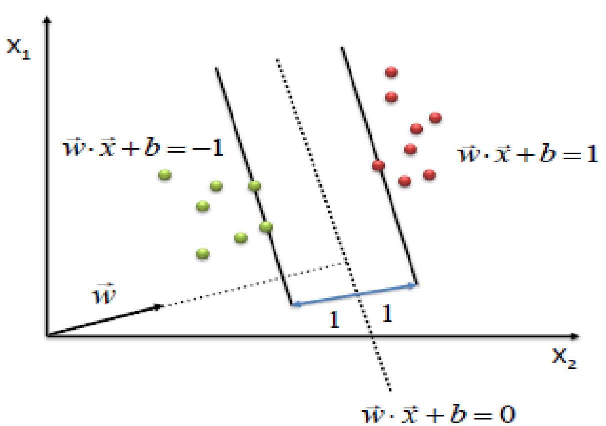
SVM Linéaire

<u>Cas de séparabilité **linéaire**</u>

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

  - **Maximiser la marge** 2/||W||
  - Sous contraintes de : $y_i(W^T x_i + b) \geq 1, i = 1..n$

    - W.Xi + b >= 1, pour classe +1
    - W.Xi + b <= -1, pour classe -1



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

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

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

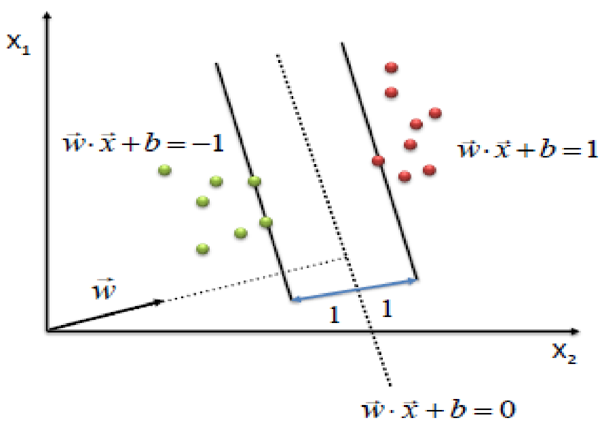
SVM Linéaire

<u>Cas de séparabilité **linéaire**</u>

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

$$\begin{cases} \text{- Maximiser la marge } 2/||W|| \\ \text{- Sous contraintes de : } y_i(W^T x_i + b) \geq 1, i = 1..n \end{cases}$$

  - W.Xi + b = 1, pour *<u>vecteurs supports</u>* de classe +1
  - W.Xi + b = -1, pour *<u>vecteurs supports</u>* de classe -1



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

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

<u>Cas de séparabilité **linéaire**</u>

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

$$\begin{cases} \text{- Maximiser la marge } 2/||W|| : \text{fonction objective} \\ \text{- Sous contraintes de : } \quad y_i(W^T x_i + b) \geq 1, i = 1..n \end{cases}$$

  - W.Xi + b = 1, pour *vecteurs supports* de classe +1
  - W.Xi + 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.

Cas de séparabilité **linéaire**

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

  - **Minimiser ||W||/2** : fonction objective
  - Sous contraintes de : $y_i(W^T x_i + b) \geq 1, i = 1..n$

    - W.Xi + b >= 1,   pour classe +1
    - W.Xi + b <= -1,   pour 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**

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

  - **Minimiser ||W||²/2** : fonction objective

  - Sous contraintes de : $y_i(W^T x_i + b) \geq 1, i = 1..n$

    - W.Xi + b >= 1, pour classe +1
    - W.Xi + b <= -1, pour 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 – Dual SVM** :

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

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

Cas de séparabilité **linéaire**

➔ **Résolution – Dual SVM** :

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

$$\text{Avec} \quad \begin{cases} W = \sum_{i=1}^{n} \alpha_i y_i x_i \\ \sum_{i=1}^{n} \alpha_i y_i = 0 \end{cases}$$

Pour les non vecteurs supports

Cas de séparabilité **linéaire**

➔ **Résolution – Dual SVM**

- 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 :

Nombre de vecteurs support

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

Exemple à tester

Les classes des vecteurs supports

Supports vecteurs

Multiplicateurs de Lagrange

Exemple à tester

SVM Linéaire

<u>Cas de séparabilité **linéaire – Dual SVM**</u>

- 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 :

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

Cas de séparabilité **linéaire**

➔ **Résolution – Primal SVM** – Soft Margin:

➢ L'hyperplan séparateur optimal peut être obtenu en minimisant le **problème primal** : regularization et variables slack

$$\begin{cases} \min_{w,b,\zeta} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \zeta_i \\ \text{subject to } y_i(w^T x_i + b) \geq 1 - \zeta_i, \\ \zeta_i \geq 0, i = 1, \ldots, n \end{cases}$$

Slack variables

- C is a regularization parameter:
  - small C allows constraints to be easily ignored $\rightarrow$ large margin
  - large C makes constraints hard to ignore $\rightarrow$ narrow margin
  - $C = \infty$ enforces all constraints: hard margin

SVM Linéaire

Cas de séparabilité **linéaire**

➔ **Résolution – Primal SVM** – Soft Margin:

- La fonction de décision est :

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

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

$$\begin{cases} \text{La classe de X} = \mathbf{+1} & \text{Si } H(X) > 0 \\ \text{La classe de X} = \mathbf{-1} & \text{Si } H(X) < 0 \end{cases} \quad = sign(W^T X + b)$$

Cas de séparabilité **linéaire**

- Les fonctions de décision des deux versions:

Primal version of classifier:

$$f(\mathbf{x}) = \mathbf{w}^\top \mathbf{x} + b$$

Dual version of classifier:

$$f(\mathbf{x}) = \sum_{i}^{N} \alpha_i y_i (\mathbf{x}_i{}^\top \mathbf{x}) + b$$

SVM Linéaire

<u>Cas de séparabilité linéaire</u>

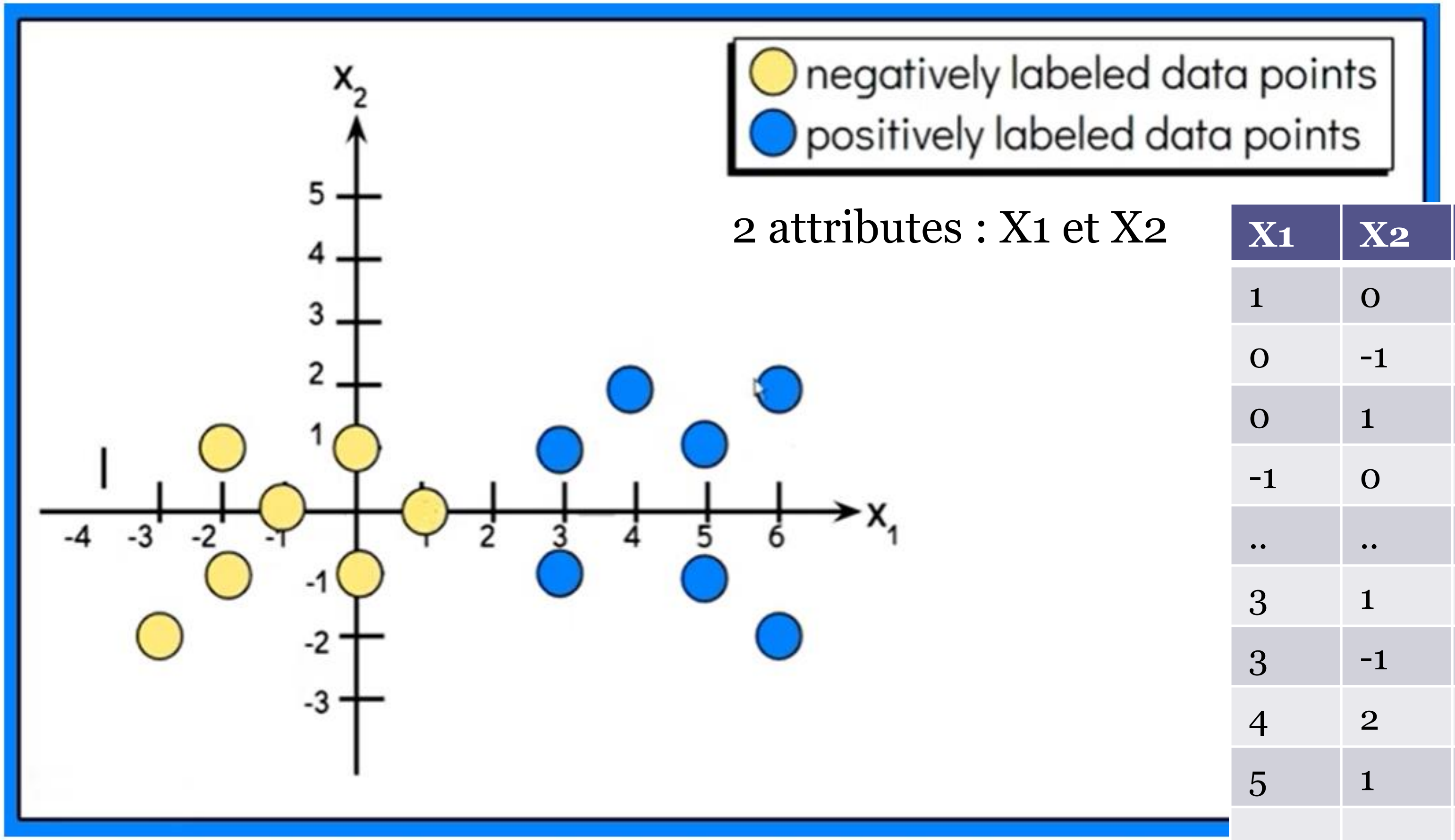
- Implémentation des SVM: pour la classification binaire consiste à la résolution du problème dual de programmation quadratique.
- Résolution ➔ déterminer les multiplicateurs Lagrangiens optimaux.
- Si nombre d'exemples est important => problème de temps et de mémoire.
- Solution : Méthode de Shnuking. Entrainement 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=TPVzIKJOcN0
https://axon.cs.byu.edu/Dan/678/miscellaneous/SVM.example.pdf

Cas de séparabilité linéaire

➢ Implémentation des SVM : Exemple - Déroulement



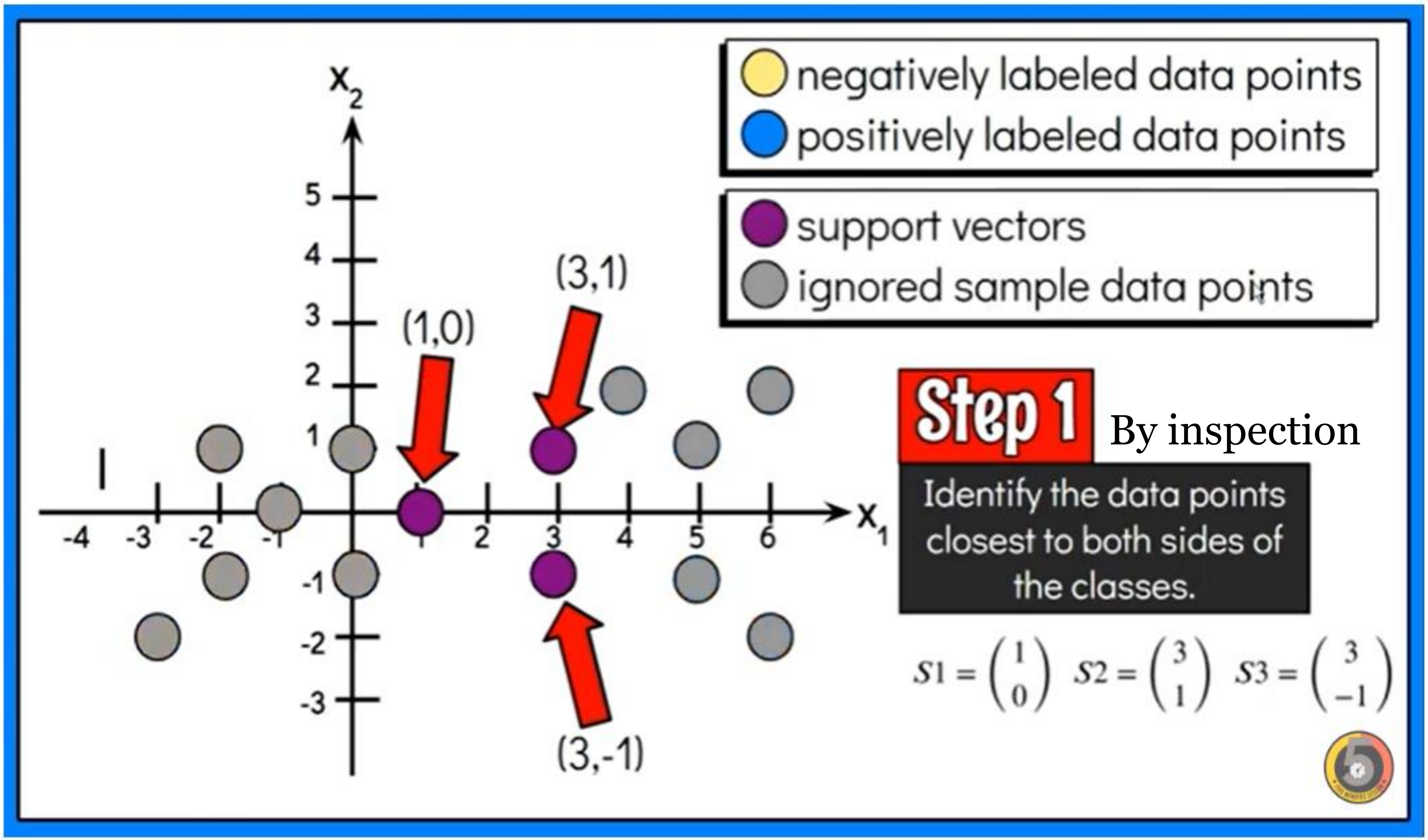
2 attributes : X1 et X2

X1	X2	Y
1	0	-1
0	-1	-1
0	1	-1
-1	0	-1
..	..	-1
3	1	+1
3	-1	+1
4	2	+1
5	1	+1
..	..	+1

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.

$$S1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad S2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad S3 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

$$\downarrow \qquad\qquad \downarrow \qquad\qquad \downarrow$$

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

SVM Linéaire

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

<u>Cas de séparabilité **linéaire**</u>

- **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.Xi + b = 1, pour <u>*vecteurs supports*</u> de classe +1
    - W.Xi + b = -1, pour <u>*vecteurs supports*</u> 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

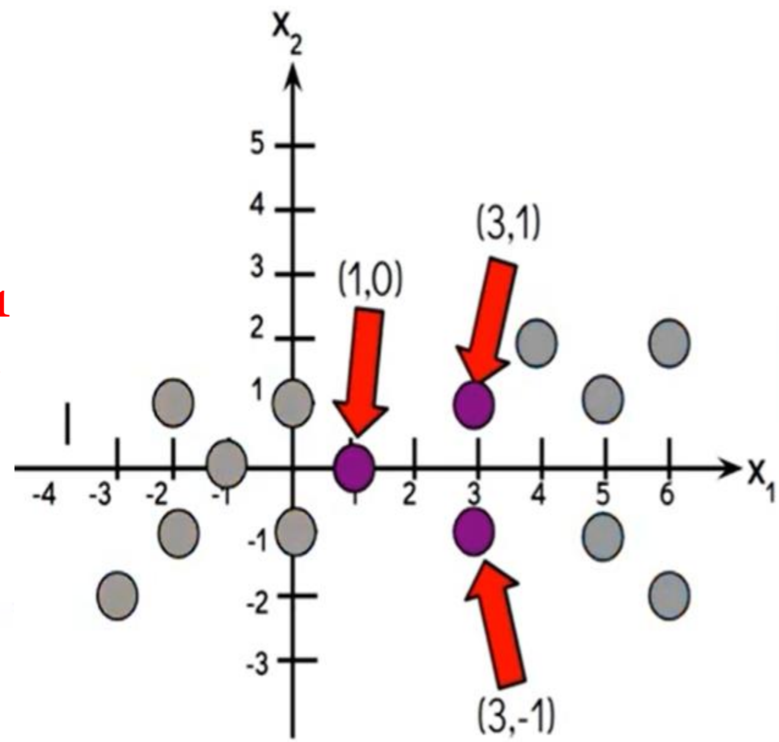
Step 3 Write the equations needed to calculate the weighted vector.

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

W.Xi + b = 1, pour *VS* de classe +1

W.Xi + 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

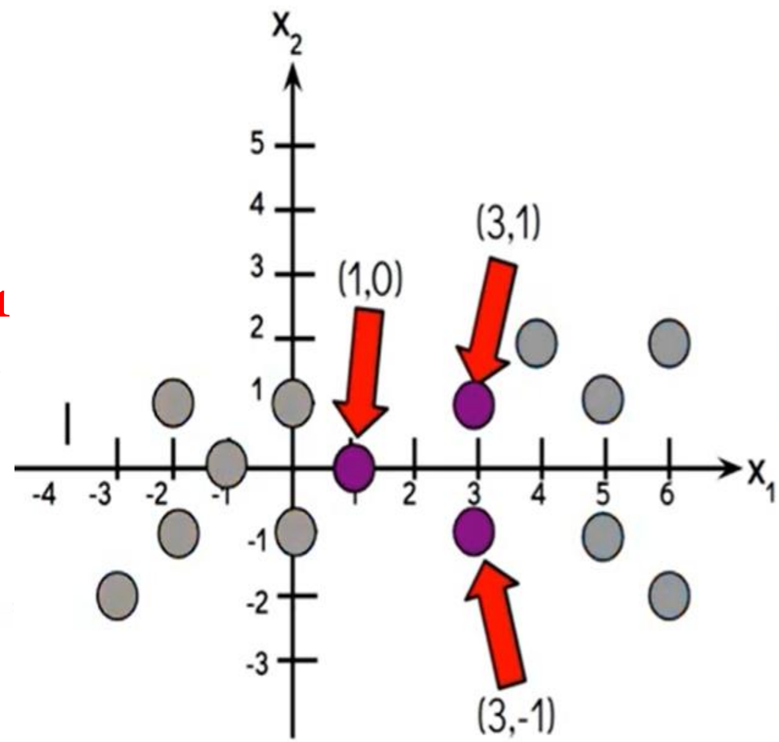
Cas de séparabilité linéaire - Exemple - Déroulement

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

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

W.Xi + b = 1, pour *VS* de classe +1
W.Xi + 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

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

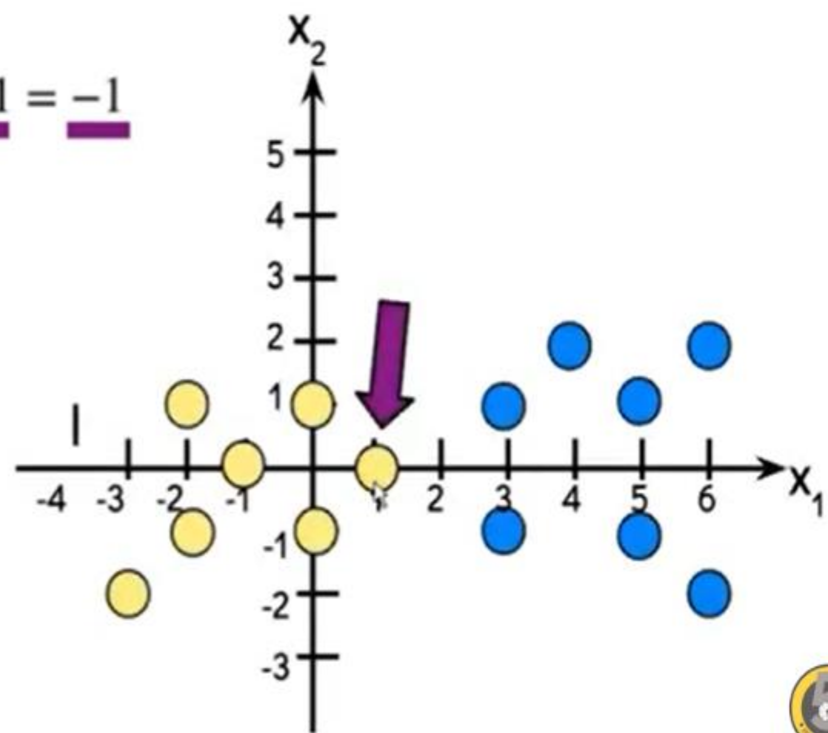
<u>Cas de séparabilité linéaire</u> - Exemple - Déroulement

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

- W.Xi + b = 1, pour *VS* de classe +1
- W.Xi + b = -1, pour *VS* de classe -1

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

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



Cas de séparabilité linéaire - Exemple - Déroulement

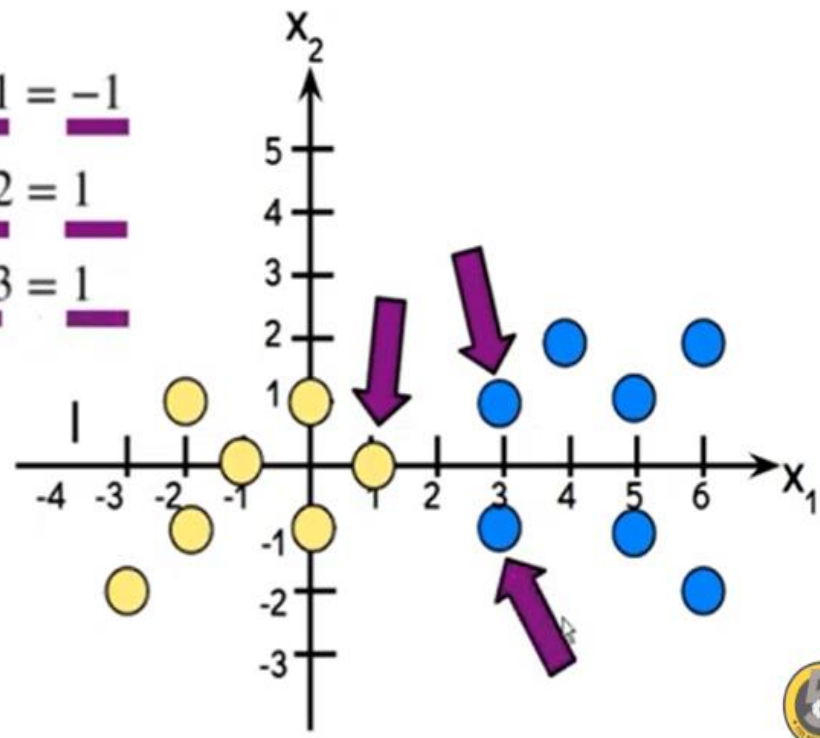
Step 3 Write the equations needed to calculate the weighted vector.

- W.Xi + b = 1, pour *VS* de classe +1
- W.Xi + b = -1, pour *VS* de classe -1

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

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

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



SVM Linéaire

Cas de séparabilité linéaire - Exemple - Déroulement

Step 4 Substitute the values into the equation.

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

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

$$\alpha 1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha 2 \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha 3 \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = -1$$

SVM Linéaire

<u>Cas de séparabilité linéaire</u> - Exemple - Déroulement

Step 4 Substitute the values into the equation.

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

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

$$\Downarrow$$

$$\alpha 1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha 2 \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha 3 \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = -1$$

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

$$\Downarrow$$

$$\alpha 1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} + \alpha 2 \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} + \alpha 3 \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} = 1$$

SVM Linéaire

Cas de séparabilité linéaire - Exemple - Déroulement

Step 4 Substitute the values into the equation.

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

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

$$\Downarrow$$

$$\alpha 1 \begin{pmatrix}1\\0\\1\end{pmatrix}\begin{pmatrix}1\\0\\1\end{pmatrix} + \alpha 2 \begin{pmatrix}3\\1\\1\end{pmatrix}\begin{pmatrix}1\\0\\1\end{pmatrix} + \alpha 3 \begin{pmatrix}3\\-1\\1\end{pmatrix}\begin{pmatrix}1\\0\\1\end{pmatrix} = -1$$

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

$$\Downarrow$$

$$\alpha 1 \begin{pmatrix}1\\0\\1\end{pmatrix}\begin{pmatrix}3\\-1\\1\end{pmatrix} + \alpha 2 \begin{pmatrix}3\\1\\1\end{pmatrix}\begin{pmatrix}3\\-1\\1\end{pmatrix} + \alpha 3 \begin{pmatrix}3\\-1\\1\end{pmatrix}\begin{pmatrix}3\\-1\\1\end{pmatrix} = 1$$

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

$$\Downarrow$$

$$\alpha 1 \begin{pmatrix}1\\0\\1\end{pmatrix}\begin{pmatrix}3\\1\\1\end{pmatrix} + \alpha 2 \begin{pmatrix}3\\1\\1\end{pmatrix}\begin{pmatrix}3\\1\\1\end{pmatrix} + \alpha 3 \begin{pmatrix}3\\-1\\1\end{pmatrix}\begin{pmatrix}3\\1\\1\end{pmatrix} = 1$$

SVM Linéaire

Cas de séparabilité linéaire - Exemple - Déroulement

Step 5 Perform dot product operation between two vectors.

$$\alpha 1 \begin{pmatrix}1\\0\\1\end{pmatrix}\begin{pmatrix}1\\0\\1\end{pmatrix} + \alpha 2 \begin{pmatrix}3\\1\\1\end{pmatrix}\begin{pmatrix}1\\0\\1\end{pmatrix} + \alpha 3 \begin{pmatrix}3\\-1\\1\end{pmatrix}\begin{pmatrix}1\\0\\1\end{pmatrix} = -1$$

$$\alpha 1\,(1+0+1) + \alpha 2\,(3+0+1) + \alpha 3\,(3+0+1) = -1$$

$$2\alpha 1 + 4\alpha 2 + 4\alpha 3 = -1$$

$$\alpha 1 \begin{pmatrix}1\\0\\1\end{pmatrix}\begin{pmatrix}3\\1\\1\end{pmatrix} + \alpha 2 \begin{pmatrix}3\\1\\1\end{pmatrix}\begin{pmatrix}3\\1\\1\end{pmatrix} + \alpha 3 \begin{pmatrix}3\\-1\\1\end{pmatrix}\begin{pmatrix}3\\1\\1\end{pmatrix} = 1$$

$$\alpha 1\,(3+0+1) + \alpha 2\,(9+1+1) + \alpha 3\,(9-1+1) = 1$$

$$4\alpha 1 + 11\alpha 2 + 9\alpha 3 = 1$$

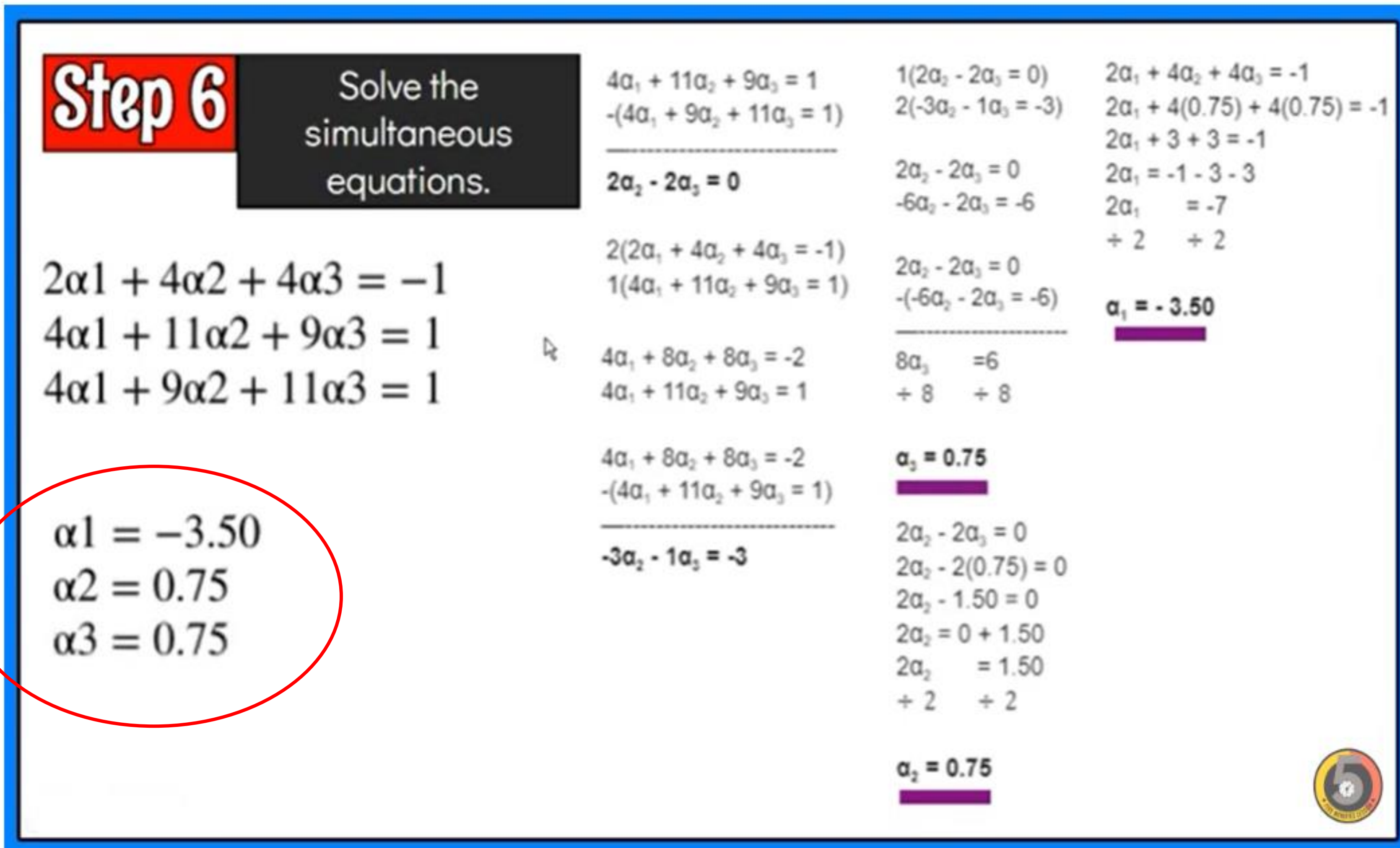
$$\alpha 1 \begin{pmatrix}1\\0\\1\end{pmatrix}\begin{pmatrix}3\\-1\\1\end{pmatrix} + \alpha 2 \begin{pmatrix}3\\1\\1\end{pmatrix}\begin{pmatrix}3\\-1\\1\end{pmatrix} + \alpha 3 \begin{pmatrix}3\\-1\\1\end{pmatrix}\begin{pmatrix}3\\-1\\1\end{pmatrix} = 1$$

$$\alpha 1\,(3+0+1) + \alpha 2\,(9-1+1) + \alpha 3\,(9+1+1) = 1$$

$$4\alpha 1 + 9\alpha 2 + 11\alpha 3 = 1$$

SVM Linéaire

Cas de séparabilité linéaire - Exemple - Déroulement



Cas de séparabilité linéaire - Exemple - Déroulement

Step 7 Calculate the weighted vector.

$$\alpha 1 = -3.50$$
$$\alpha 2 = 0.75$$
$$\alpha 3 = 0.75$$

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

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

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

Cas de séparabilité linéaire - Exemple - Déroulement

Step 8 Interpret the computed weighted vector to draw the hyperplane.

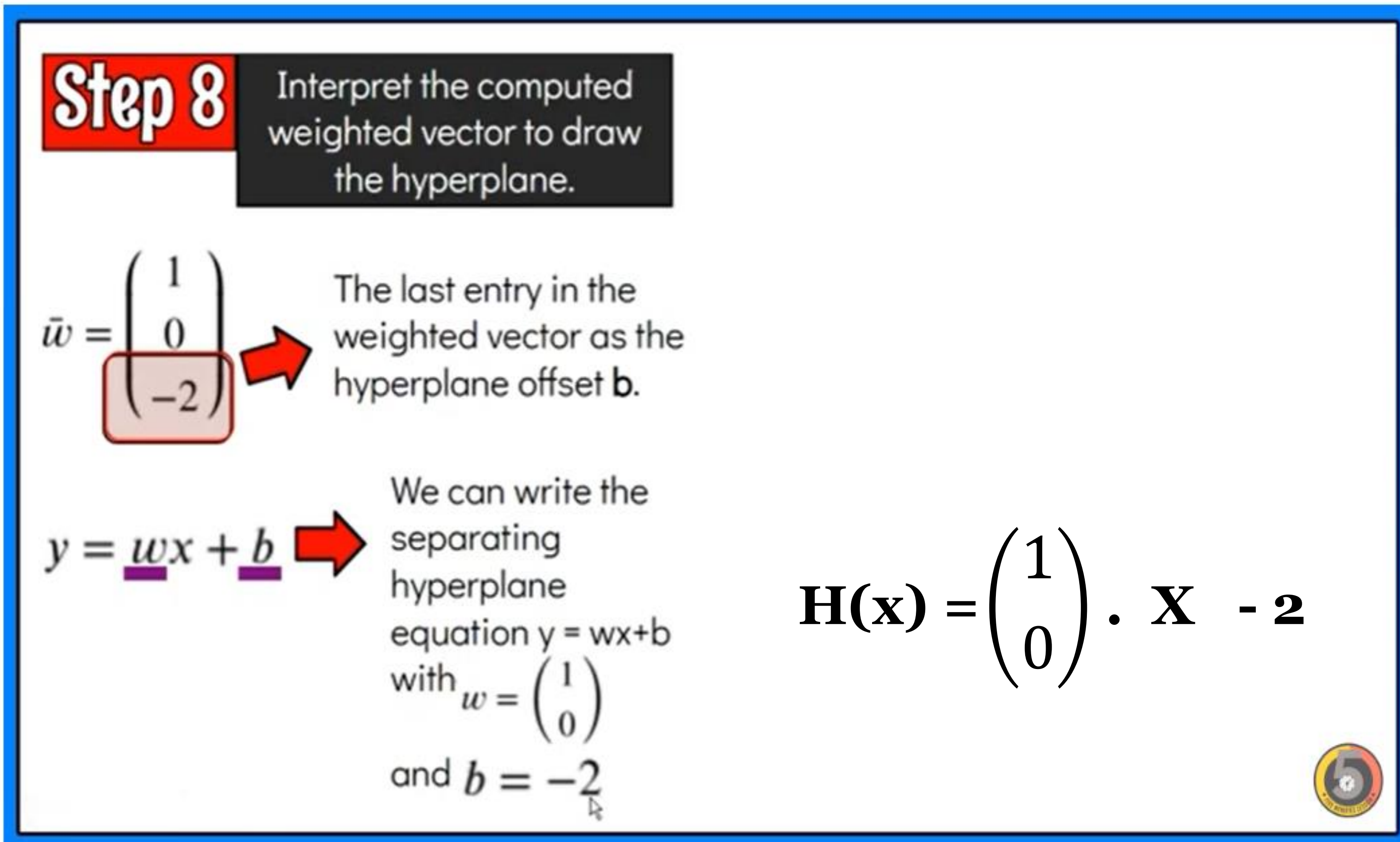
$$\bar{w} = \begin{pmatrix} 1 \\ 0 \\ \boxed{-2} \end{pmatrix}$$

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

$$= (w1, w2, b)$$

SVM Linéaire

Cas de séparabilité linéaire - Exemple - Déroulement



SVM Linéaire

Cas de séparabilité linéaire - Exemple - Déroulement

Step 8 Interpret the computed weighted vector to draw the hyperplane.

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

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

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

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

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

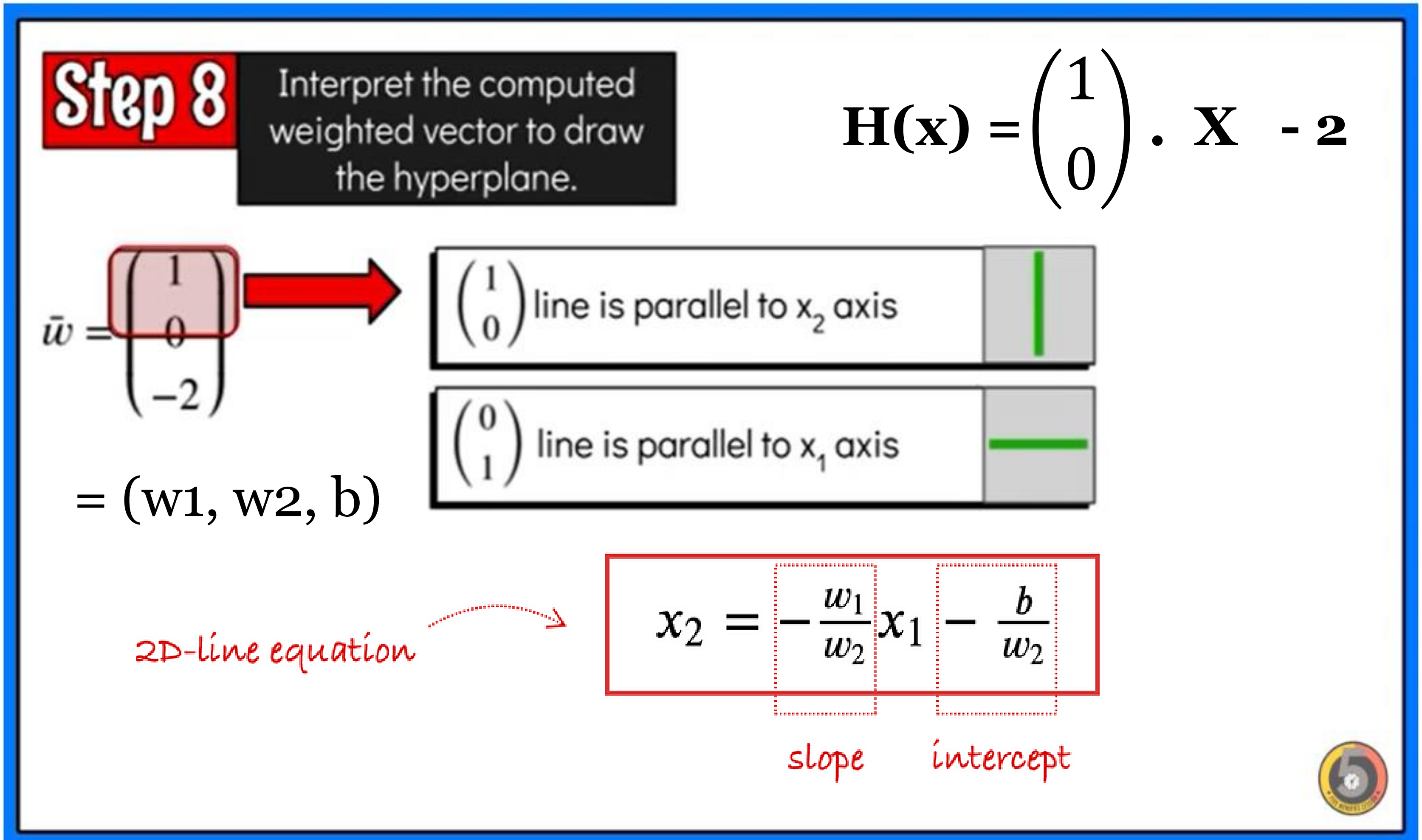
We can write the separating hyperplane equation y = wx+b with $w = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $b = -2$

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

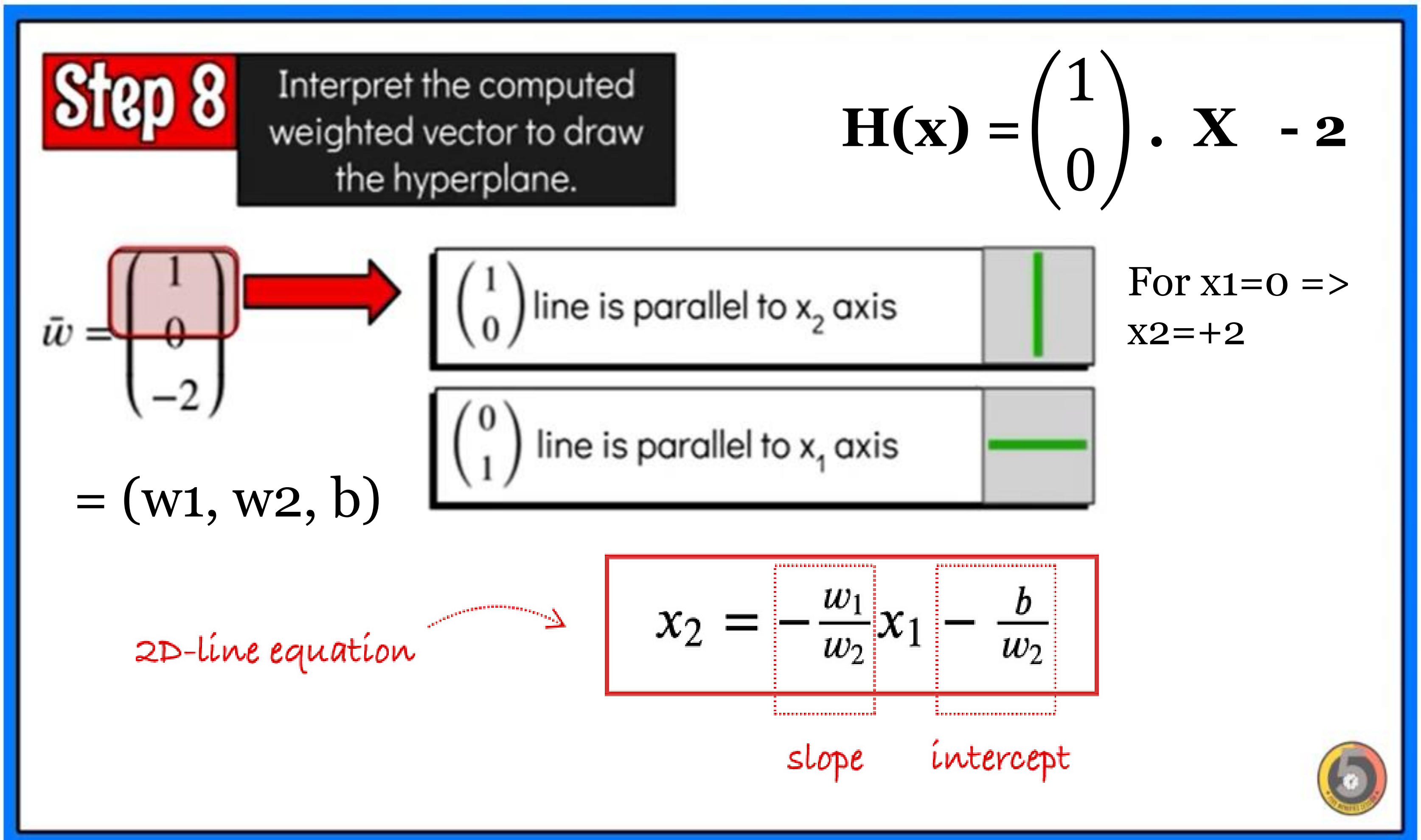
$||W|| = \mathrm{sqrt}(1^2 + 0^2)$

SVM Linéaire

Cas de séparabilité linéaire - Exemple - Déroulement



Cas de séparabilité linéaire - Exemple - Déroulement



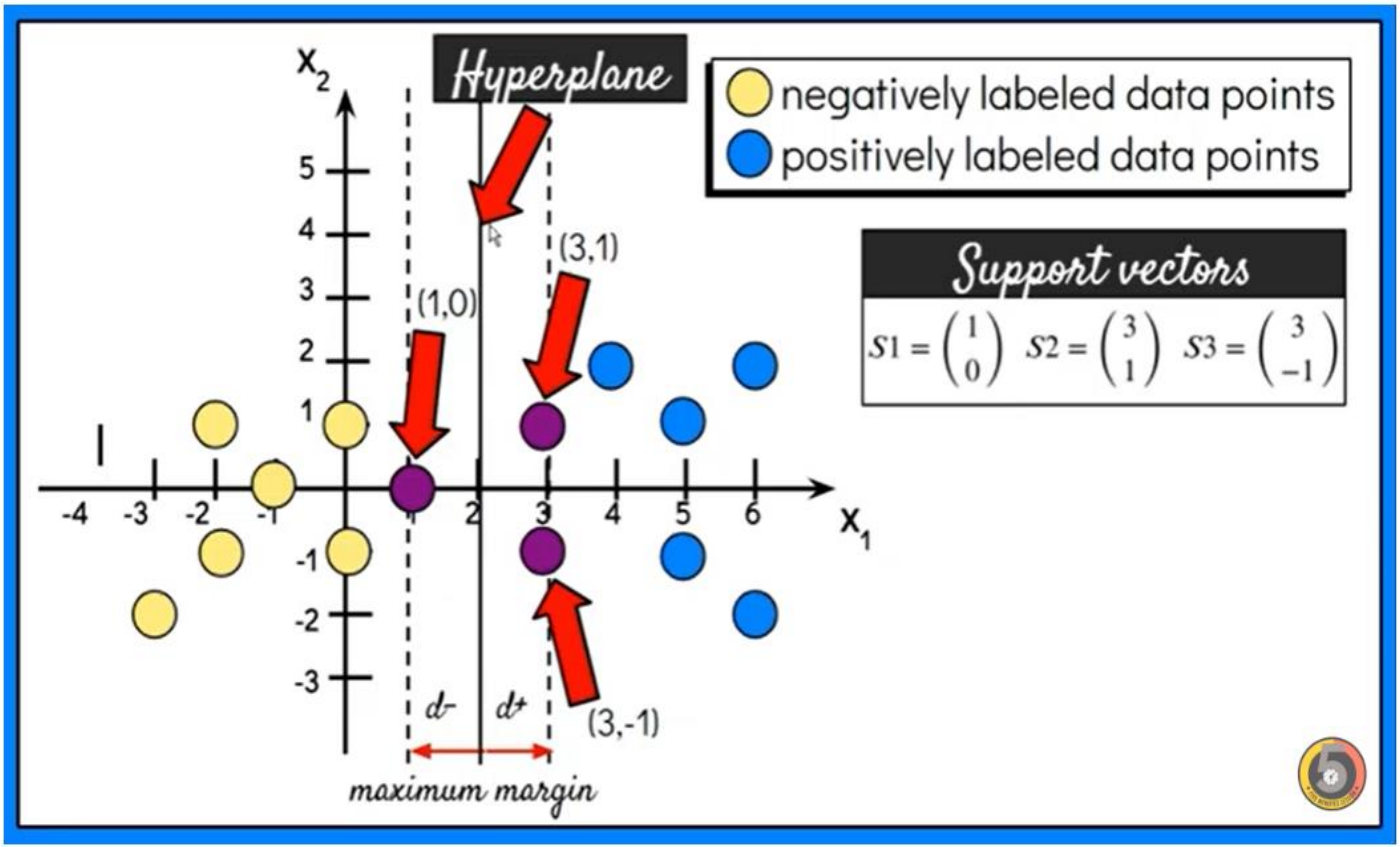
$$\mathbf{H(x)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \mathbf{X} - \mathbf{2}$$

$= (w1, w2, b)$

For x1=0 => x2=+2

SVM Linéaire

<u>Cas de séparabilité linéaire</u> - Exemple - Déroulement



SVM Linéaire

Cas de séparabilité linéaire - Exemple - Déroulement

Step 8 Make Predictions for X = (x1, x2).

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

Support vectors

$$S1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad S2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad S3 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

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

$$\mathbf{H(x)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \ \mathbf{dot} \ \begin{pmatrix} x1 \\ x2 \end{pmatrix} \mathbf{- 2}$$

$\alpha 1 = -3.50$
$\alpha 2 = 0.75$
$\alpha 3 = 0.75$

SVM Linéaire

Cas de séparabilité linéaire - Exemple - Déroulement

Step 8 Make Predictions for X = (x1, x2).

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

Support vectors

$S1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $S2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ $S3 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$

$\alpha 1 = -3.50$
$\alpha 2 = 0.75$
$\alpha 3 = 0.75$

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

$$\mathbf{H(x)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \ \mathbf{dot} \ \begin{pmatrix} x1 \\ x2 \end{pmatrix} - \mathbf{2}$$

Ex: X = (4, 2) => H(X) = (1*4 + 0*2) - 2 = 2 > 0 => X class is +1

Cas de séparabilité linéaire - Exemple - Déroulement

Step 8

Make Predictions for
X = (x1, x2).

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

Support vectors

$S1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $S2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ $S3 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$

$\alpha 1 = -3.50$
$\alpha 2 = 0.75$
$\alpha 3 = 0.75$

$$\begin{aligned} H(x) = \quad & -3{,}5 * (-1) * <1, 0> * X \\ + \quad & 0{,}75 * (+1) * <3, 1> * X \\ + \quad & 0{,}75 * (+1) * <3, -1> * X \\ - \quad & 2 \end{aligned}$$

Cas de séparabilité linéaire - Exemple - Déroulement

Step 8 Make Predictions for X = (x1, x2).

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

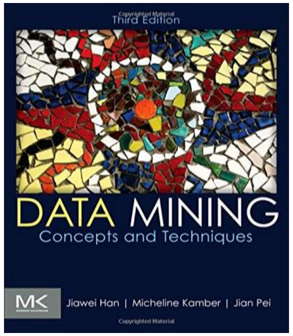
Support vectors

$S1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $S2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ $S3 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$

$$H(X(x1, x2)) = -3{,}5 * (-1) * (x1)^2 + 0{,}75 * (+1) * (3x1 + x2)^2 + 0{,}75 * (+1) * (3x1 - x2)^2 - 2$$

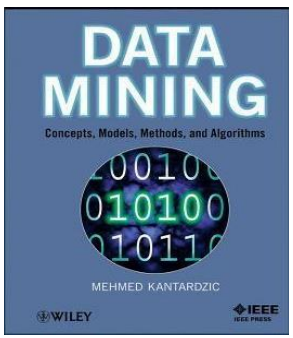
$\alpha 1 = -3.50$
$\alpha 2 = 0.75$
$\alpha 3 = 0.75$

Reférences



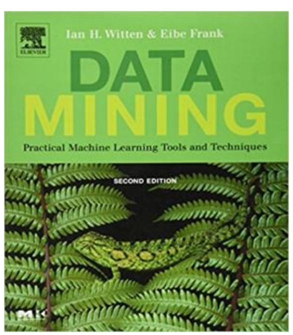
Data Mining : concepts and techniques, 3rd Edition

- ✓ 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

Reférences

Cours – Abdelhamid DJEFFAL – Fouille de données avancée

- ✓ www.abdelhamid-djeffal.net

WekaMOOC – Ian Witten – Data Mining with Weka

- ✓ https://www.youtube.com/user/WekaMOOC/featured

https://www.robots.ox.ac.uk/~az/lectures/ml/lect3.pdf

Cours - Phugo Larochelle – Apprentissage Automatique

✓https://www.youtube.com/playlist?list=PL6Xpj9I5qXYFD_rc1tttugXLfE2TcKyiO

Chris McCormick - SVM Example

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