



# **Session 32**

## **Classification II**



# Table of Content

## Apa yang Akan Kita Pelajari Hari Ini?

1. SVM
2. Multiclass classification



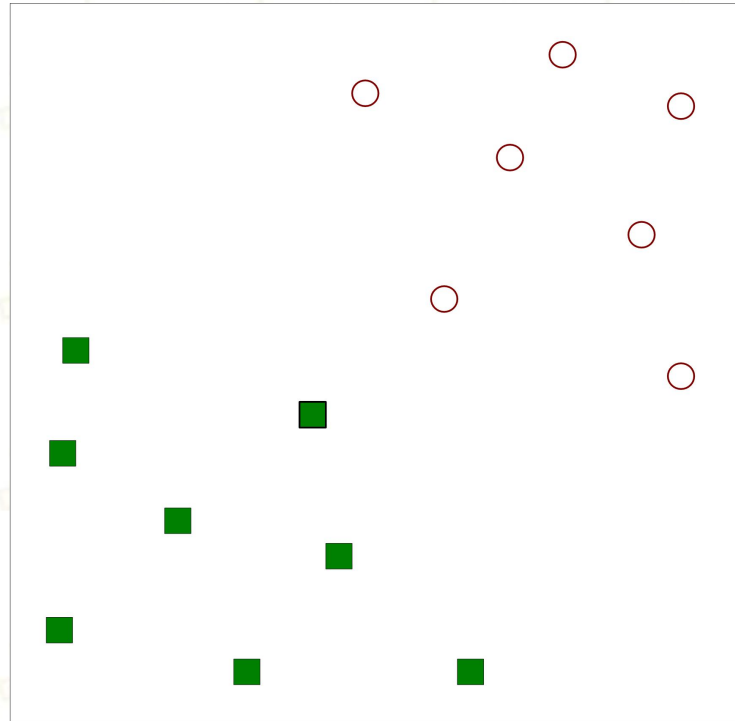


# SVM





# Support Vector Machines

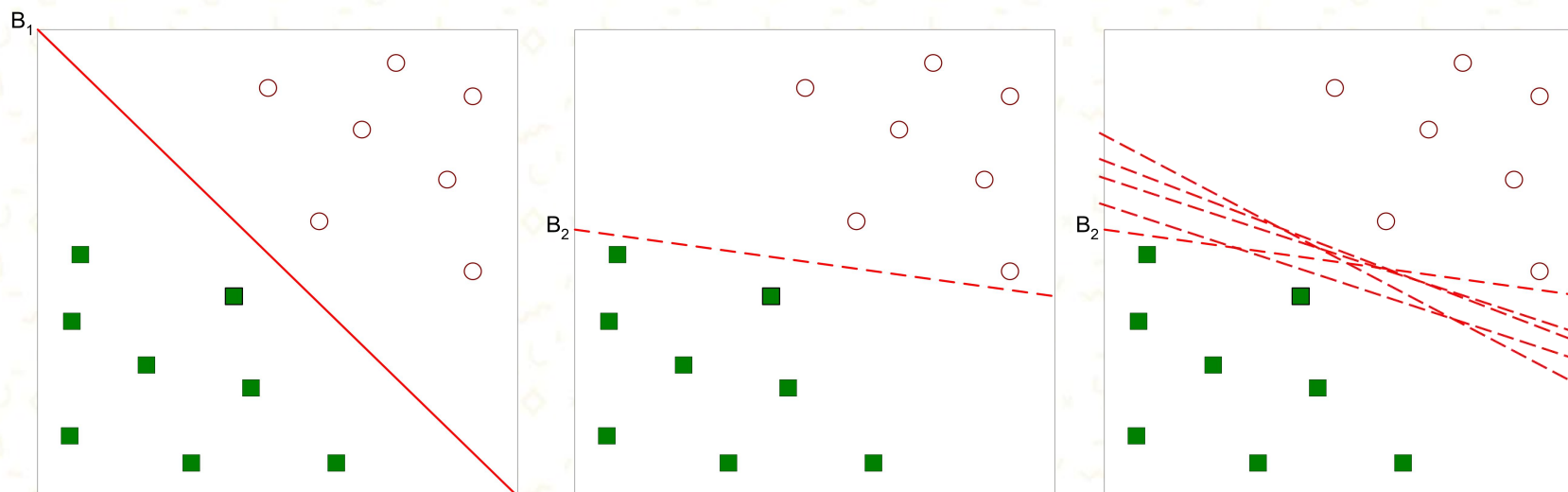


- Temukan sebuah *linear hyperplane* (decision boundary) yang dapat memisahkan data





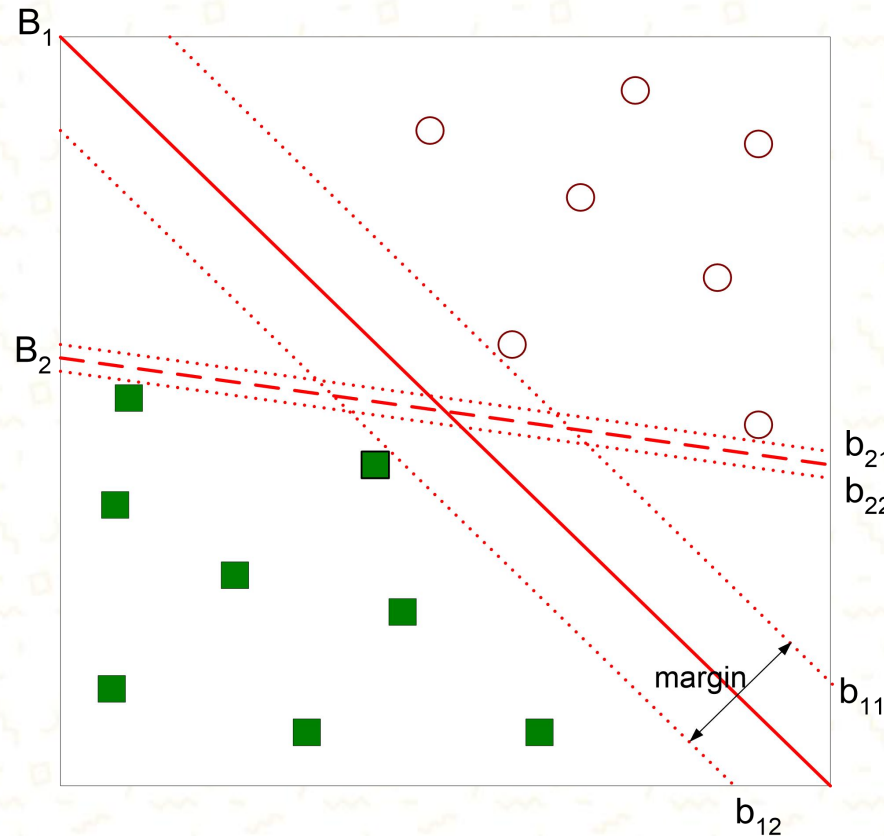
# Support Vector Machines



- Ada banyak kemungkinan solusi



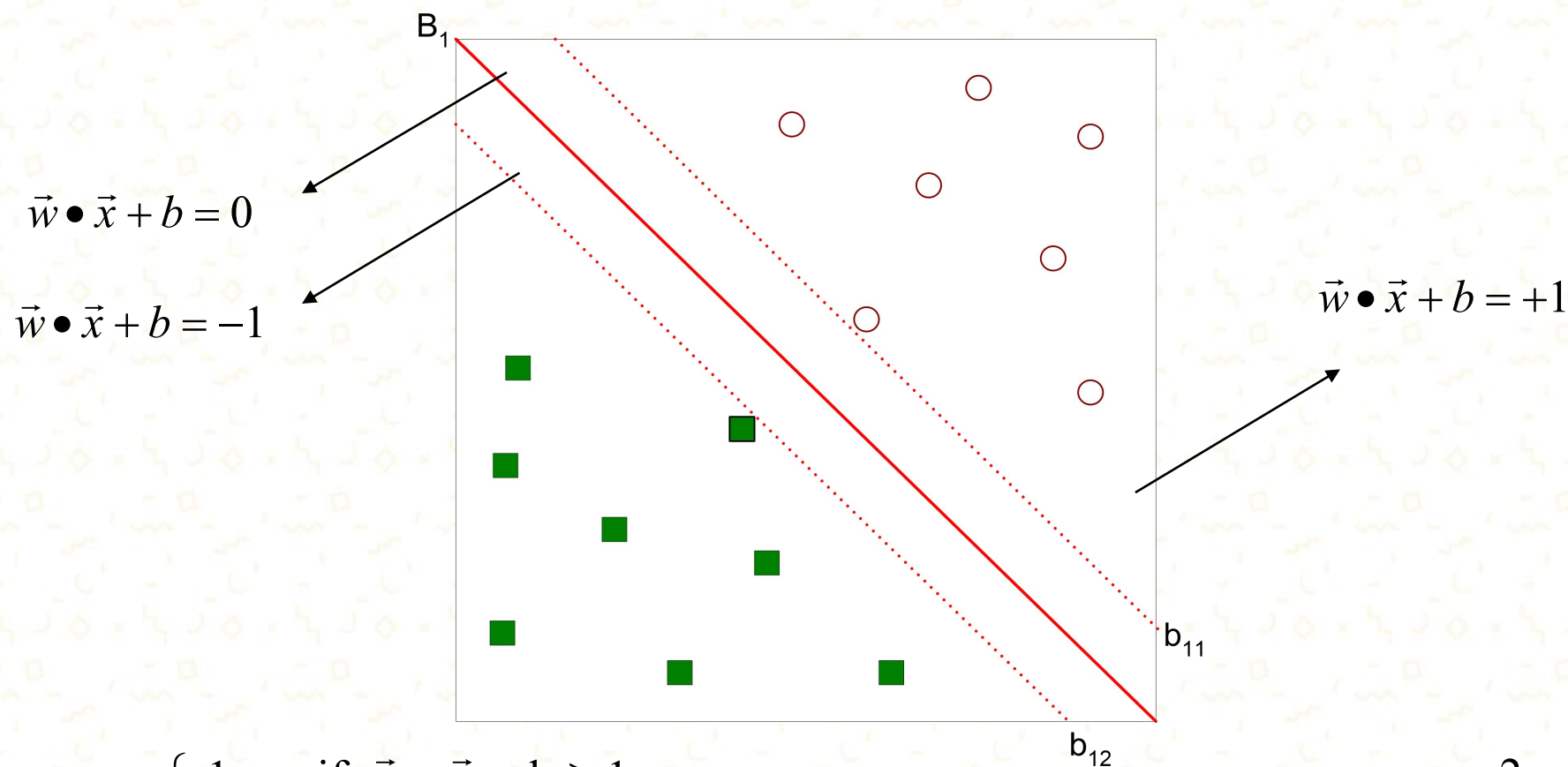
# Support Vector Machines



- Temukan *hyperplane* yang memaksimalkan the margin
- $\Rightarrow$  B1 lebih baik dari B2



# Support Vector Machines



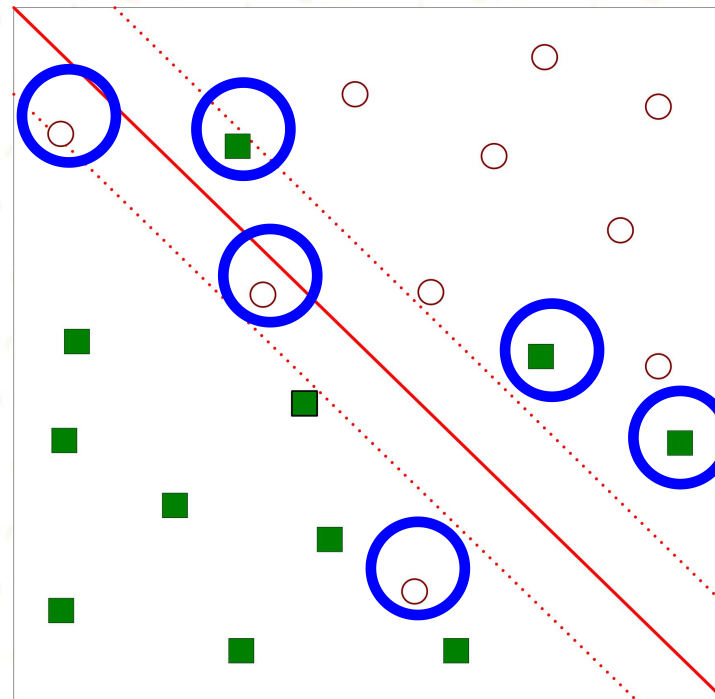
$$f(\vec{x}) = \begin{cases} 1 & \text{if } \vec{w} \bullet \vec{x} + b \geq 1 \\ -1 & \text{if } \vec{w} \bullet \vec{x} + b \leq -1 \end{cases}$$

$$\text{Margin} = \frac{2}{\|\vec{w}\|^2}$$



# Support Vector Machines

- Bagaimana jika tidak bisa dipisahkan secara linear (*not linearly separable*)?







# Support Vector Machines

- Bagaimana jika tidak bisa dipisahkan secara linear (*not linearly separable*)?

- *Introduce slack variables*

- *Need to minimize:*

$$L(w) = \frac{\|\vec{w}\|^2}{2} + C \left( \sum_{i=1}^N \xi_i \right)$$

- *Subject to:*

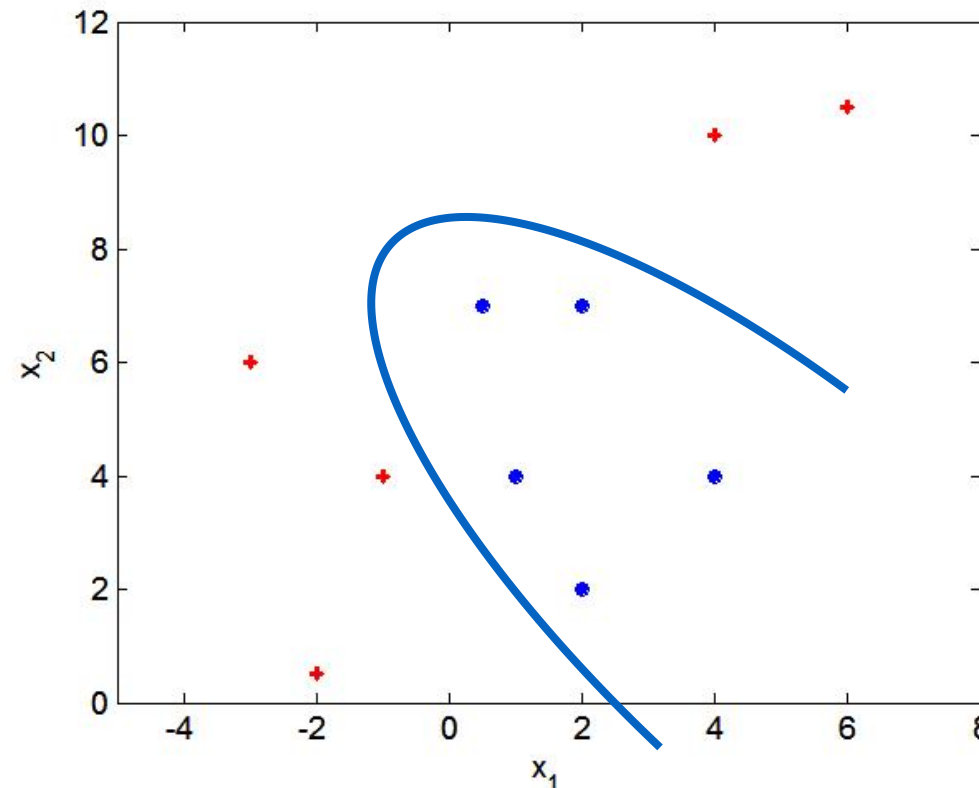
$$f(\vec{x}_i) = \begin{cases} 1 & \text{if } \vec{w} \bullet \vec{x}_i + b \geq 1 - \xi_i \\ -1 & \text{if } \vec{w} \bullet \vec{x}_i + b \leq -1 + \xi_i \end{cases}$$





# Nonlinear Support Vector Machines

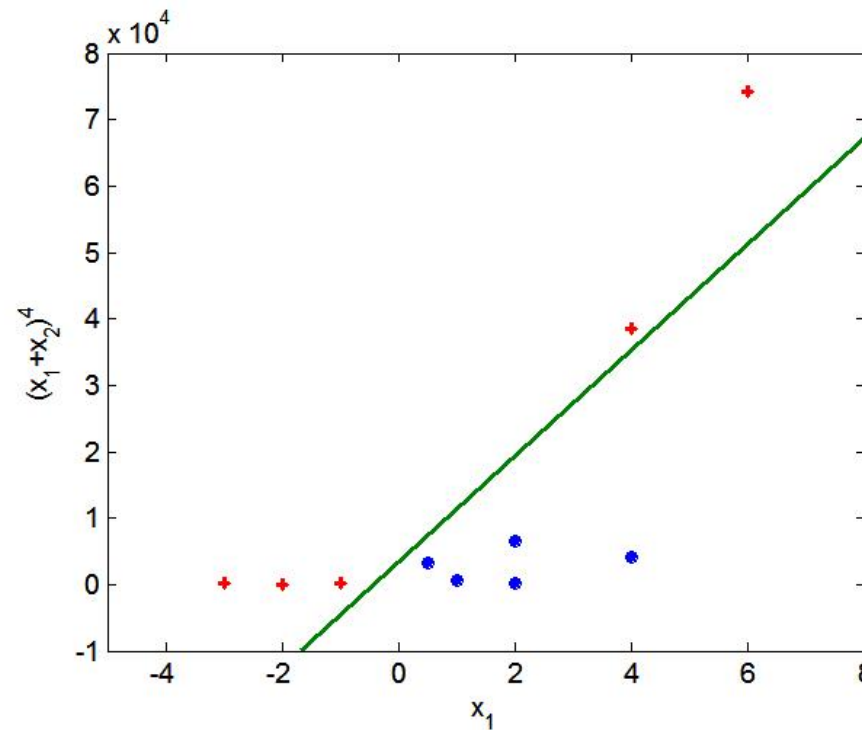
- Bagaimana jika batas keputusan tidak linier (*decision boundary is not linear*)?





# Nonlinear Support Vector Machines

- Ubah data ke ruang dimensi yang lebih tinggi (*higher dimensional space*)





# Multiclass







# Basic Classification in ML

Input  
 $\mathbf{x} \in \mathcal{X}$

Output  
 $y \in \mathcal{Y}$

Cancer  
detection

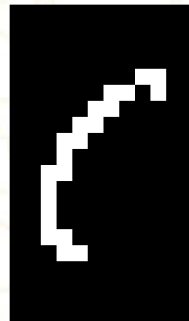
```
breast-cancer - 복사본.csv - 메모장
파일(F)  편집(E)  서식(O)  보기(V)  도구들(T)
age,menopause,tumor-size,inv-nodes,node-caps,deg-
malig,breast,breast-quad,irradiat
30-39,premeno,30-34,0-2,no,3,left,left_low,no
40-49,premeno,20-24,0-2,no,2,right,right_up,no
40-49,premeno,20-24,0-2,no,2,left,left_low,no
60-69,ge40,15-19,0-2,no,2,right,left_up,no
40-49,premeno,0-4,0-2,no,2,right,right_low,no
60-69,ge40,15-19,0-2,no,2,left,left_low,no
50-59,premeno,25-29,0-2,no,2,left,left_low,no
60-69,ge40,20-24,0-2,no,1,left,left_low,no
40-49,premeno,50-54,0-2,no,2,left,left_low,no
```



**Binary**

- Cancer
- Not cancer

Character  
recognition



**Multi-Class**

- A
- B
- C
- ...




# Multiclass classification

- Multiclass classification
  - Klasifikasi yang melibatkan lebih dari dua Class
  - Setiap data hanya dapat dimiliki oleh satu Class
- Multilabel classification
  - Klasifikasi yang melibatkan lebih dari dua Class
  - Setiap data dapat dimiliki oleh beberapa Class


Three Type of Classification Tasks

**Binary Classification**




- Spam
- Not spam

**Multiclass Classification**



- Dog
- Cat
- Horse
- Fish
- Bird
- ...

**Multi-label Classification**

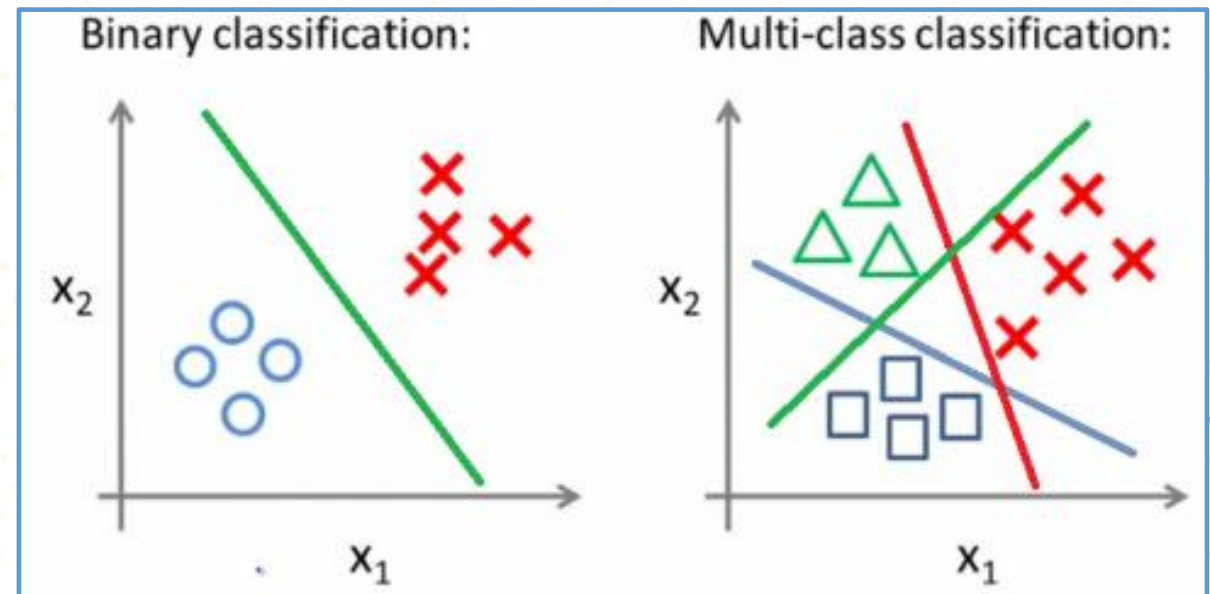


- Dog
- Cat
- Horse
- Fish
- Bird
- ...



# Binary to multiclass

- Bisakah kita menggunakan *binary classifier* untuk membangun *multiclass classifier*?
  - Mengurai prediksi menjadi *multiple binary decisions*
- Cara yang bisa dipakai?
  - *One-vs-all (One vs Rest/ OvR)*
  - *One vs One (OvO)*







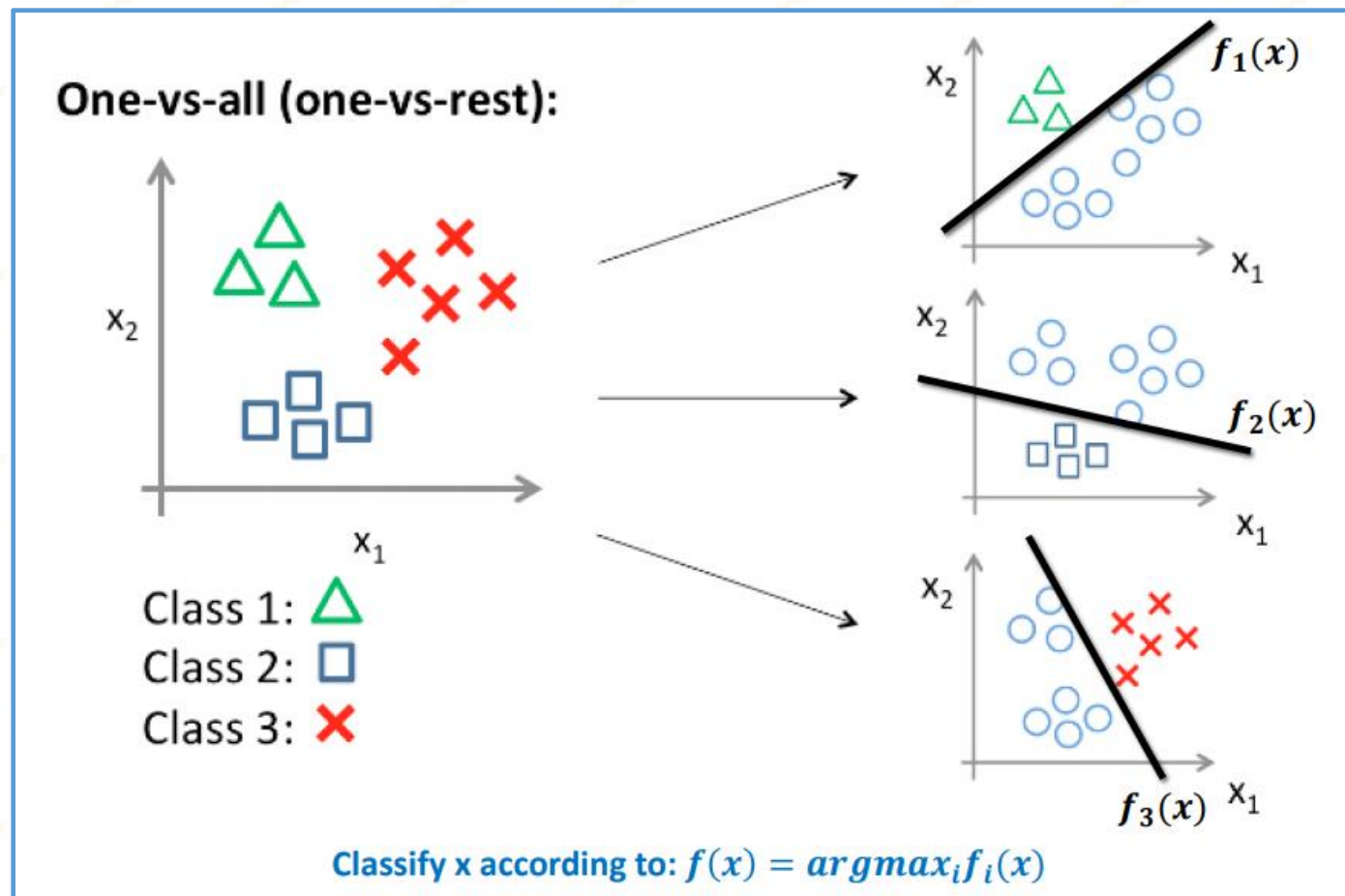
# One vs Rest (OvR)

- Buat classifier satu persatu
- Diberikan  $m$  classes, latih  $m$  classifiers: satu classifier untuk setiap class
- Classifier  $j$ : jadikan data di class  $j$  sebagai positive & sisanya sebagai negative
- Untuk memprediksi data baru  $X$ , pilih classifier dengan nilai yang maksimal





# Visualizing one vs rest





# One vs Rest learning algorithm

- Learning: Diberikan dataset  $D = \{(x_i, y_i)\}$   
 $x_i \in R^n, y_i \in \{1, 2, 3, \dots, K\}$
- Memecah menjadi K binary classification (K=jumlah class)
  - Biarkan model dengan jumlah K belajar dari training set:  
 $w_1, w_2, w_3, \dots, w_K$
  - Untuk setiap class k, buat sebuah binary classification dengan tugas mempelajari:
    - Data Positive : Data dari dataset D dengan label k
    - Data Negative: Data lain (bukan label k) di dataset D
  - Binary classification dapat diselesaikan dengan semua algorithm.





# One vs Rest inference algorithm

- Inference: “Winner takes all”
  - $\hat{y} = \operatorname{argmax}_{y \in \{1, 2, \dots, K\}} w_y^T x$

For example:  $y = \operatorname{argmax}(w_{red}^T x, w_{blue}^T x, w_{green}^T x)$



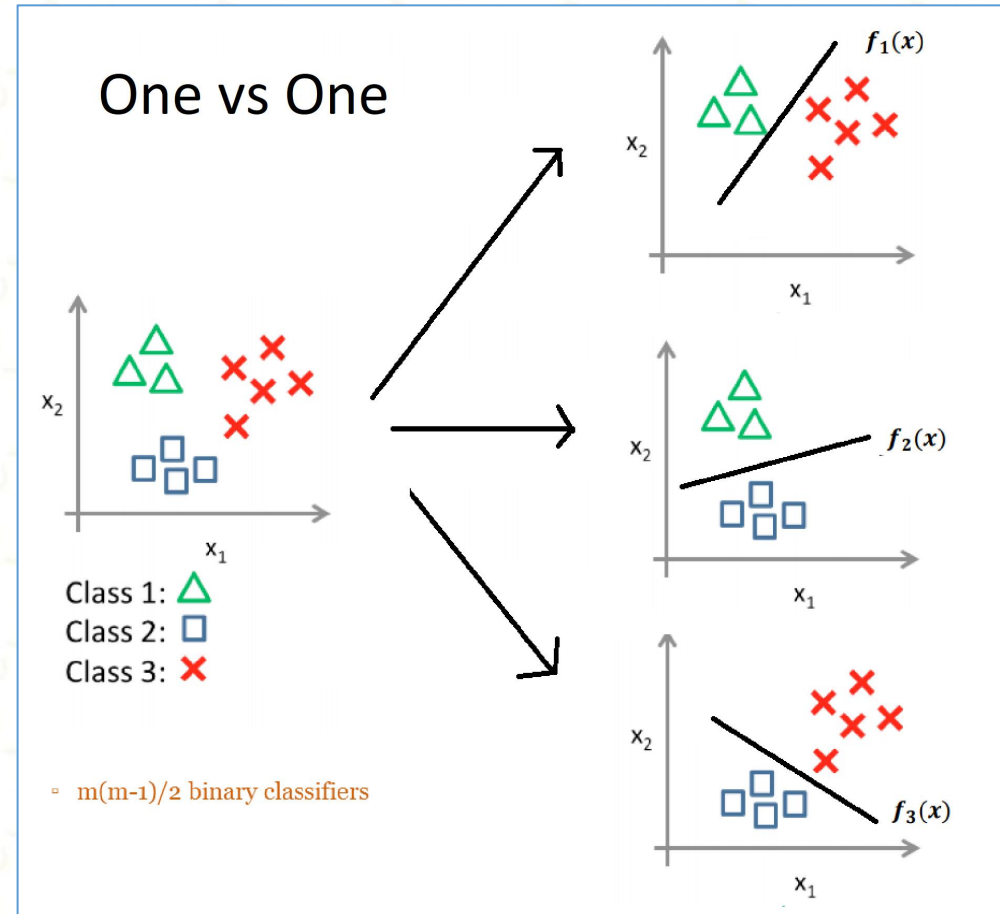
# One vs One (OvO)

- Buat sebuah classifier untuk setiap pasangan classes
- Diberikan  $m$  classes, buat binary classifiers berjumlah  $m(m-1)/2$
- Setiap classifier mempelajari data dari 2 classes (2 pasangan yang berbeda)
- Untuk memprediksi data baru  $X$ , setiap classifier melakukan voting.
- Data baru  $X$  diberikan class dengan voting suara terbanyak





# One vs One (OvO)



Classify  $x$  according to majority voting



# One vs One learning algorithm

- Learning: Diberikan dataset  $D = \{(x_i, y_i)\}$   
 $x_i \in R^n, y_i \in \{1, 2, 3, \dots, K\}$
- Memecah menjadi binary classification dengan jumlah  $K(K-1)/2$ . (K=jumlah class)
  - Biarkan model dengan jumlah  $K(K-1)/2$  belajar dari training set:  
 $w_1, w_2, w_3, \dots, w_{K*(K-1)/2}$
  - Untuk setiap pasangan class (i,j), buat sebuah binary classification dengan tugas mempelajari:
    - Data Positive : Data dari dataset D dengan label i
    - Data Negative: Data dari dataset D dengan label j
- Binary classification dapat diselesaikan dengan semua algoritma





# One vs One Inference algorithm

- Prediction:
  - Majority (mayoritas): *Pick the label with maximum votes* (pilih label dengan jumlah voting suara terbanyak)



**Lets Practice!**





Thank  
YOU