



SUPPORT VECTOR MACHINE (SVM)

TEAM TEACHING



○ Support Vector Machine

- Konsep Klasifikasi dengan Support Vector Machine (SVM) adalah mencari hyperplane terbaik yang berfungsi sebagai pemisah dua kelas data.
- Ide sederhana dari SVM adalah memaksimalkan margin, yang merupakan jarak pemisah antara kelas data.
- SVM mampu bekerja pada dataset yang berdimensi tinggi dengan menggunakan kernel trik.
- SVM hanya menggunakan beberapa titik data terpilih yang berkontribusi (Support Vector) untuk membentuk model yang akan digunakan dalam proses klasifikasi.



○ Support Vector Machine

- Macam-Macam Training untuk SVM :
 - Chunking (Quadratic Programming).
 - Osuna (Dekomposisi).
 - Sequential Minimum Optimation (SMO).
 - Least Square (LS) dan lainnya.





Model SVM

- Titik data : $x_i = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^n$
- Kelas data : $y_i \in \{-1, +1\}$
- Pasangan data dan kelas :
 $\{(x_i, y_i)\}_{i=1}^N$
- Memaksimalkan fungsi berikut :

$$Ld = \sum_{i=1}^N \alpha_i - \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j) \text{ syarat : } 0 \leq \alpha_i \leq C \text{ dan } \sum_{i=1}^N \alpha_i y_i = 0$$





Model SVM

- Menghitung nilai w dan b :

$$w = \sum_{i=1}^N \alpha_i y_i x_i \quad b = -\frac{1}{2} (w \cdot x^+ + w \cdot x^-)$$

- Fungsi keputusan klasifikasi $\text{sign}(f(x))$:

$$f(x) = w \cdot x + b \quad \text{atau} \quad f(x) = \sum_{i=1}^m \alpha_i y_i K(x, x_i) + b$$

- Keterangan :

- N (banyaknya data), n (dimensi data atau banyaknya fitur), L_d (Dualitas Lagrange Multiplier), α_i (nilai bobot setiap titik data), C (nilai konstanta), m (jumlah support vector/titik data yang memiliki $\alpha_i > 0$), $K(x, x_i)$ (fungsi kernel).





Model SVM

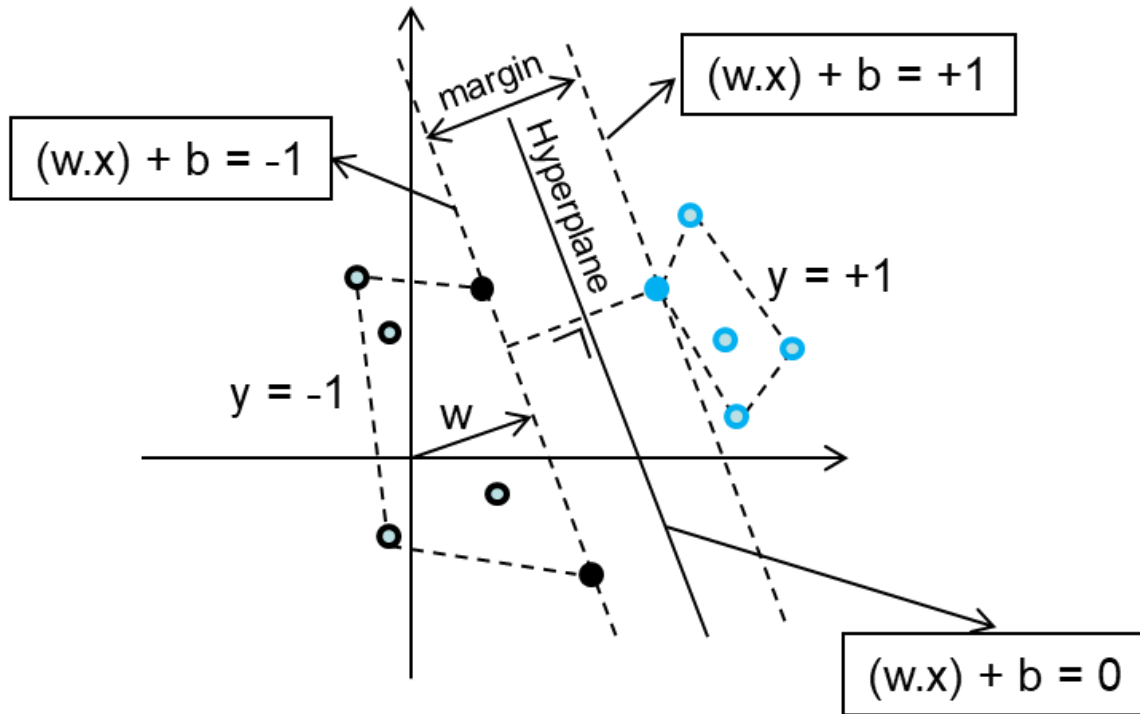
- Beberapa Macam Fungsi Kernel Support Vector Machine (SVM) :

No	Nama Kernel	Definisi Fungsi
1	Linier	$K(x,y) = x.y$
2	Polinomial of degree d	$K(x,y) = (x.y)^d$
3	Polinomial of degree up to d	$K(x,y) = (x.y + c)^d$
4	Gaussian RBF	$K(x,y) = \exp\left(\frac{-\ x-y\ ^2}{2\sigma^2}\right)$
5	Sigmoid (Tangen Hiperbolik)	$K(x,y) = \tanh(\sigma(x.y)+c)$
6	Invers Multi Kuadratik	$K(x,y) = \frac{1}{\sqrt{\ x-y\ ^2+c^2}}$
7	Additive	$K(x,y) = \sum_{i=1}^n K_i(x_i, y_i)$

- Kernel Linier digunakan ketika data yang akan diklasifikasi dapat terpisah dengan sebuah garis/hyperplane.
- Kernel non-Linier digunakan ketika data hanya dapat dipisahkan dengan garis lengkung atau sebuah bidang pada ruang dimensi tinggi (Kernel Trik, No.2 sampai 6).

Visualisasi SVM

- Linier Kernel :



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

- Support Vector kelas -1
- Support Vector kelas +1

● Jarak titik (x_i) ke Hyperplane :

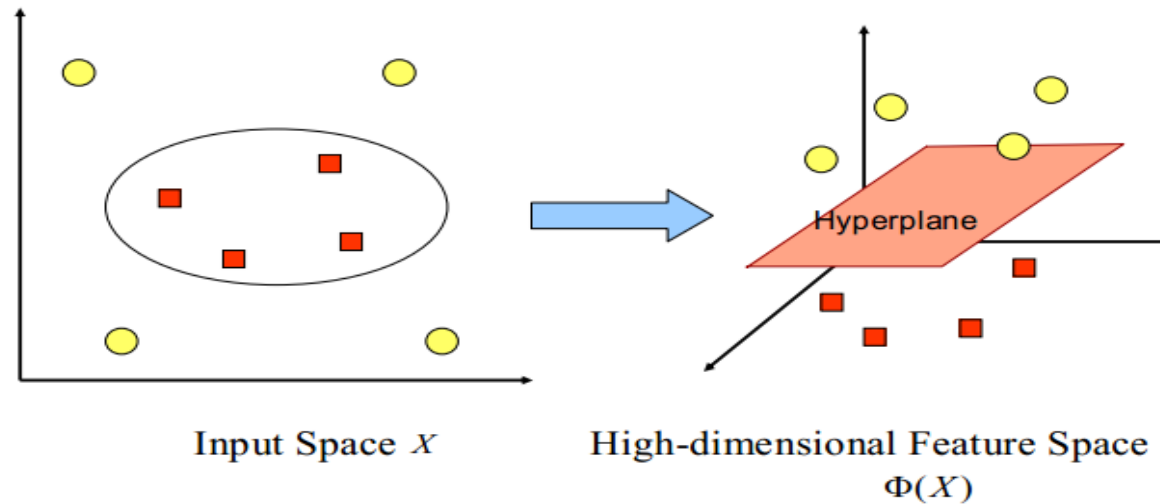
$$d((w, b), x_i) = \frac{y_i(x_i \cdot w + b)}{\|w\|} \geq \frac{1}{\|w\|}$$

$$d((w, b), x_i) = \frac{y_i f(x_i)}{\|w\|} \geq \frac{1}{\|w\|}$$



Visualisasi SVM

- Non-Linear Kernel :





Karakteristik SVM

- SVM memerlukan proses pelatihan dengan menyimpan hasil support vektor yang didapatkan untuk digunakan kembali pada saat proses prediksi/testing.
- SVM selalu memberikan model yang sama dan solusi yang sama dengan margin maksimal.
- SVM dapat memisahkan data yang distribusi kelasnya bersifat linier maupun non linier.
- SVM tidak dipengaruhi oleh dimensi data yang tinggi, sehingga tidak ada proses reduksi dimensi didalamnya.
- Memori yang digunakan dalam SVM dipengaruhi oleh banyaknya data, bukan besarnya dimensi data.



○ Kelebihan dan Kekurangan

Kelebihan

- Training relatif mudah
 - tidak mengalami local optimal, tidak seperti pada neural networks
- Relatif baik untuk data berdimensi tinggi
- Tradeoff antara kompleksitas classifier dan kesalahan dapat dikontrol secara eksplisit
- Data non-tradisional seperti string dan tree dapat digunakan sebagai input ke SVM, bukan vektor fitur

Kekurangan

- Membutuhkan fungsi kernel yang bagus



○ SVM vs Neural Networks

SVM

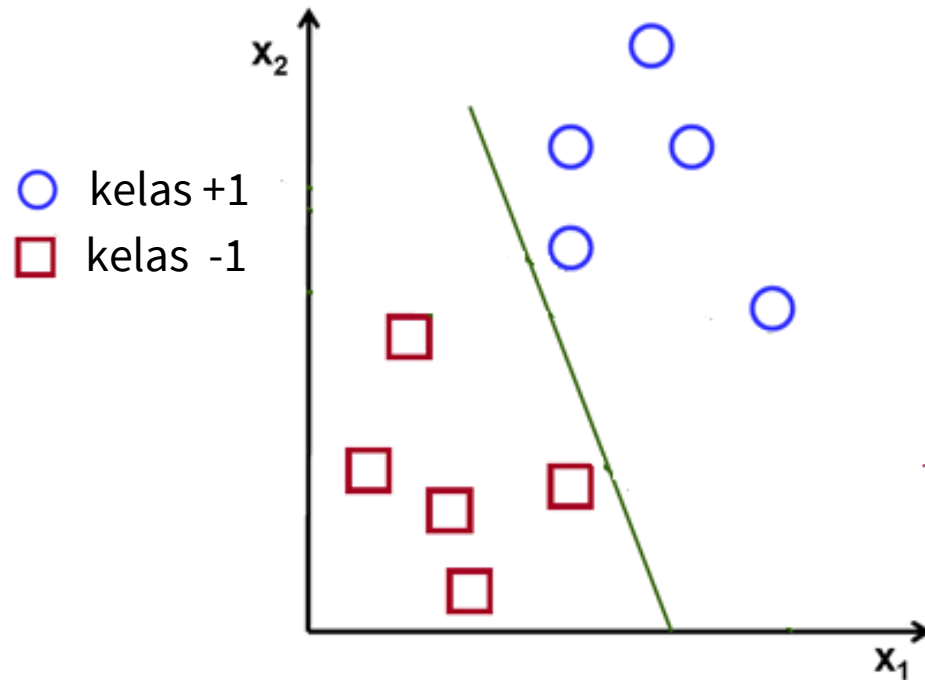
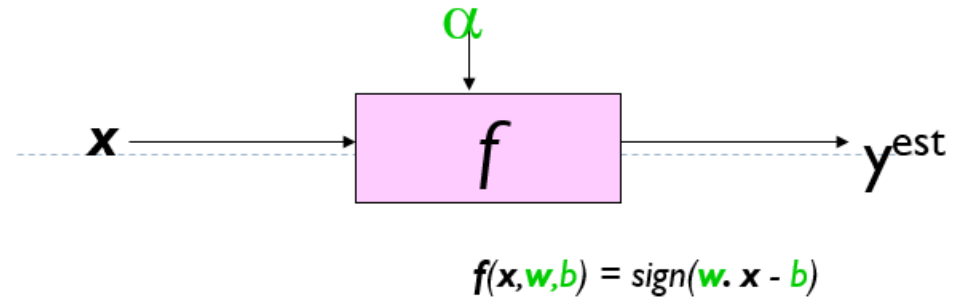
- Konsep yang relatif baru
- Sifat generalisasi yang bagus
- Sulit untuk dipelajari → teknik QP (Quadratic Programming)
- Menggunakan kernel, dapat mempelajari fungsi yang sangat kompleks

ANN

- Men-generalisasi dengan baik tetapi tidak memiliki dasar matematika
- Dapat dengan mudah dipelajari secara bertahap
- Untuk mempelajari fungsi kompleks → gunakan struktur multi layer yang kompleks.



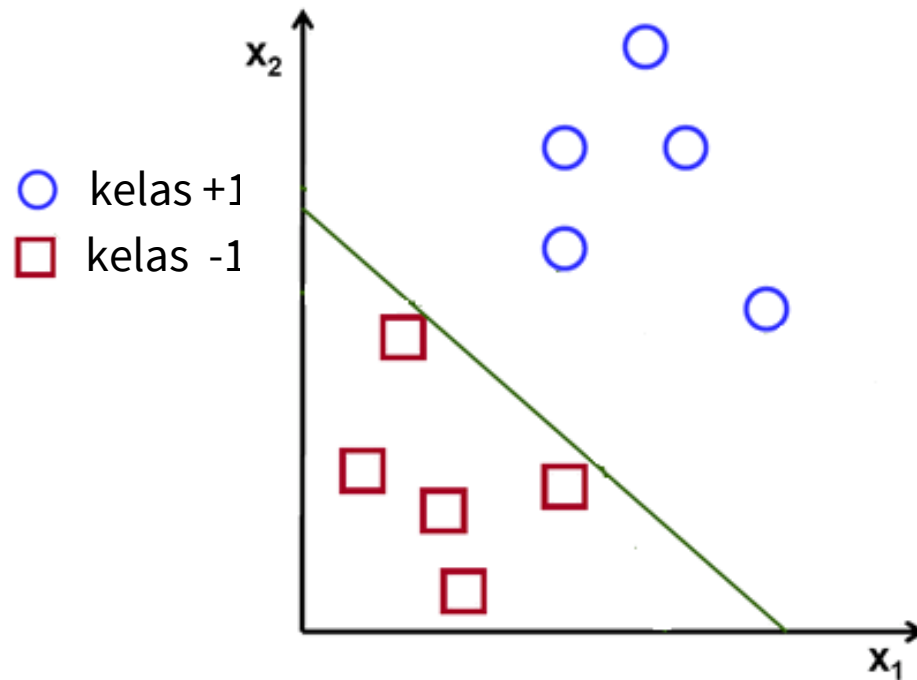
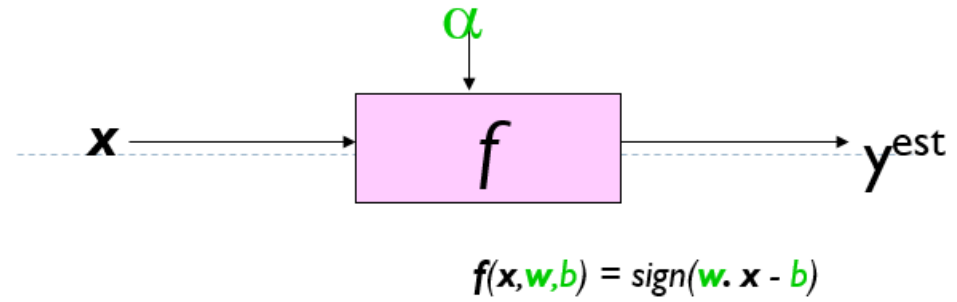
Linear Classifiers



Bagaimana
mengklasifikasikan data ini?



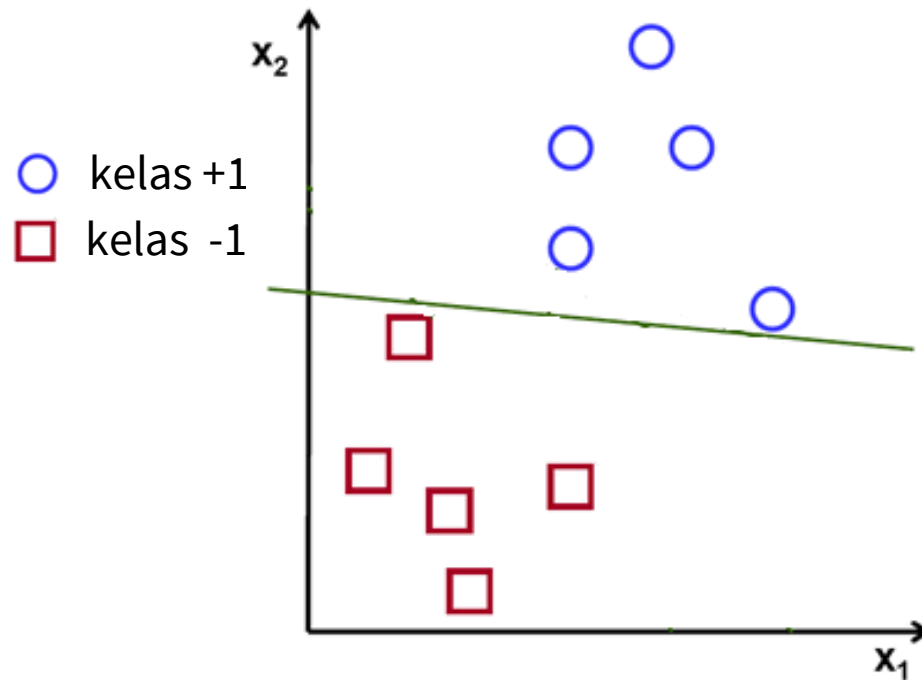
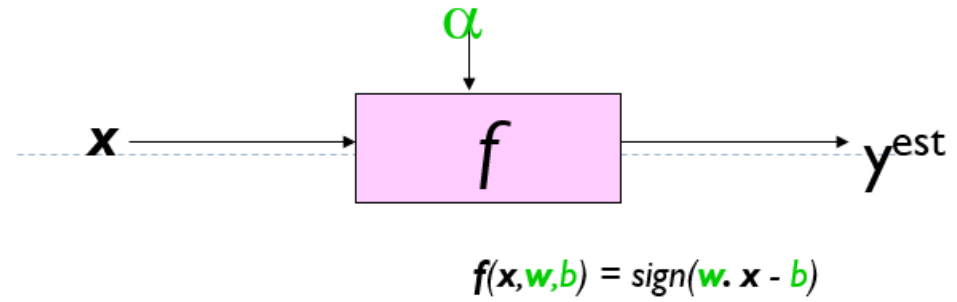
Linear Classifiers



Bagaimana
mengklasifikasikan data ini?



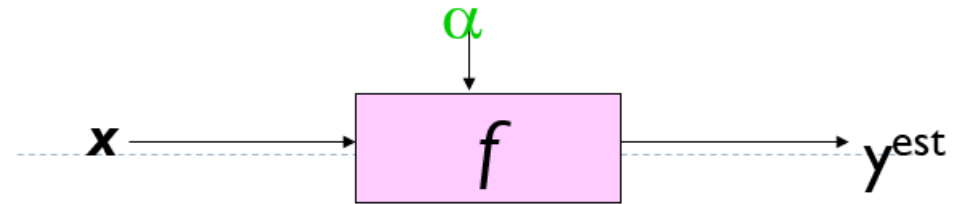
Linear Classifiers



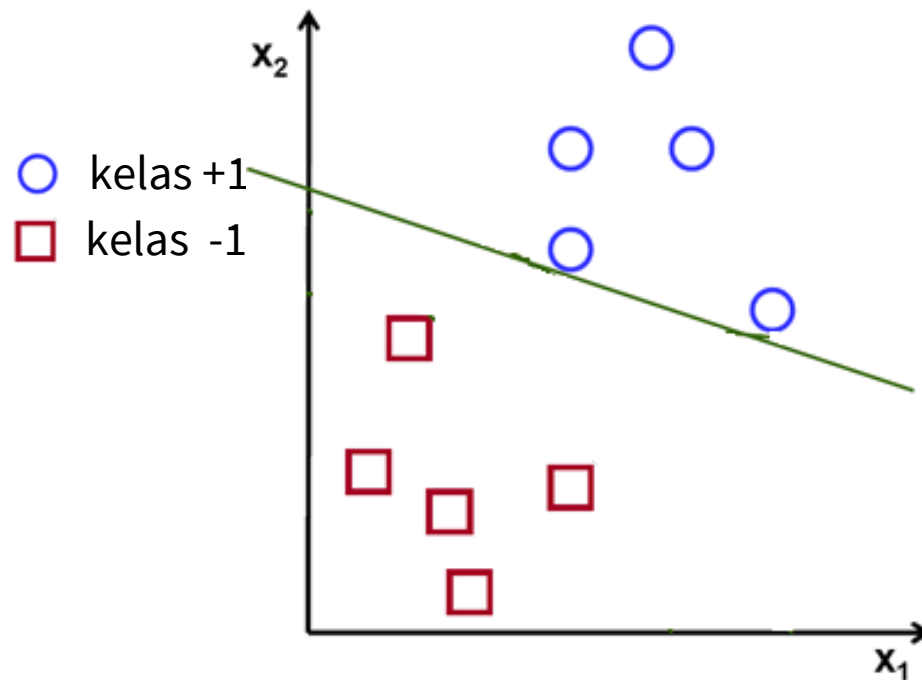
Bagaimana
mengklasifikasikan data ini?



Linear Classifiers



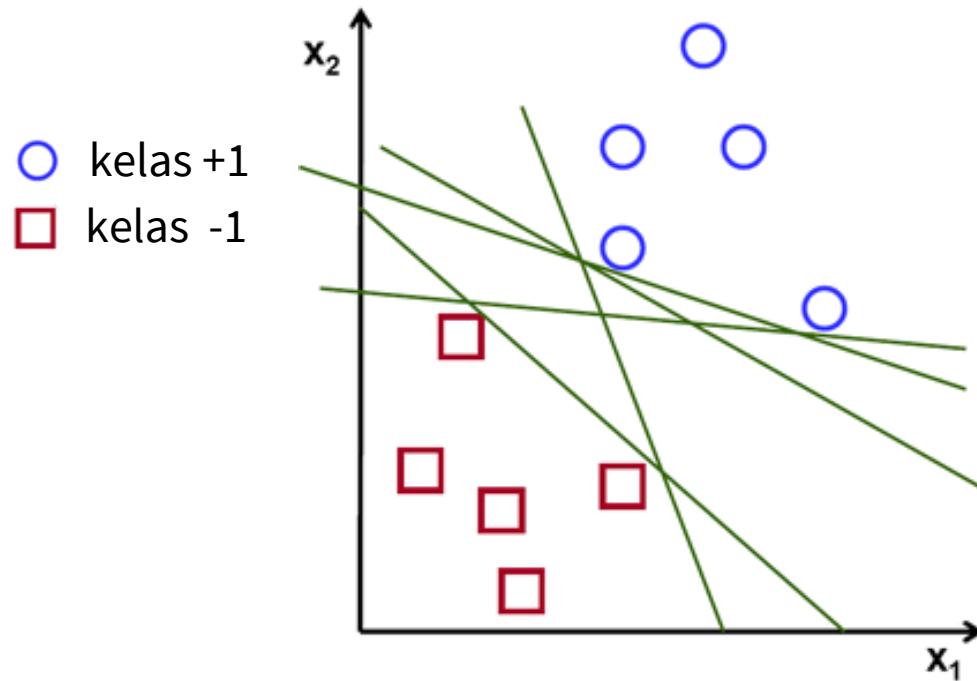
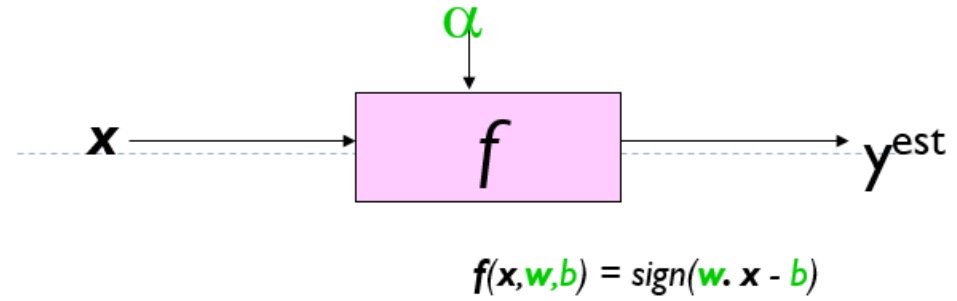
$$f(\mathbf{x}, \mathbf{w}, b) = \text{sign}(\mathbf{w} \cdot \mathbf{x} - b)$$



Bagaimana
mengklasifikasikan data ini?



Linear Classifiers



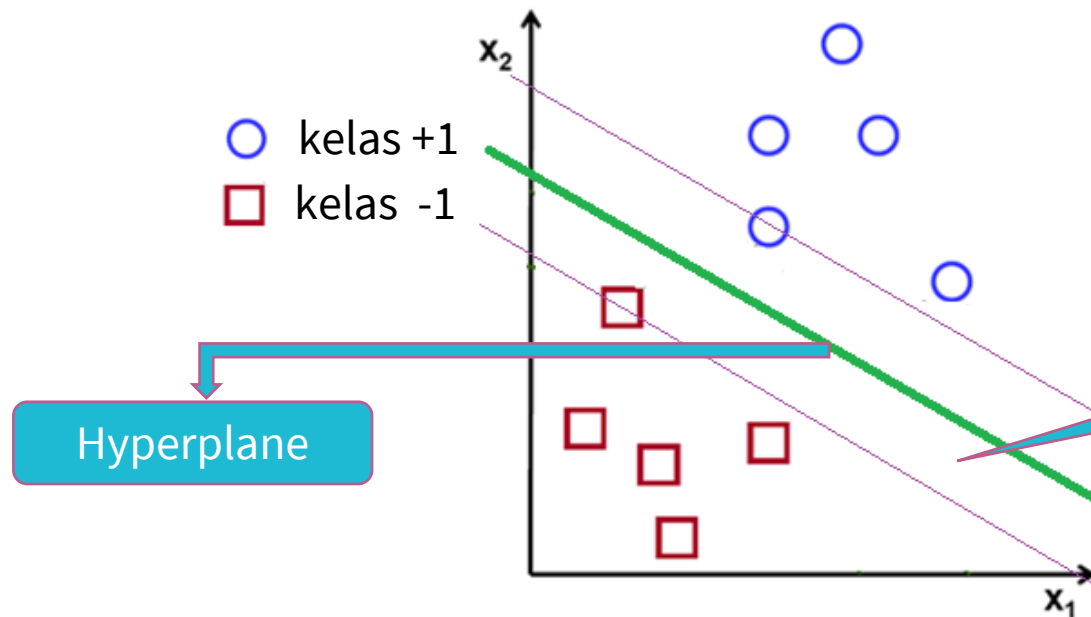
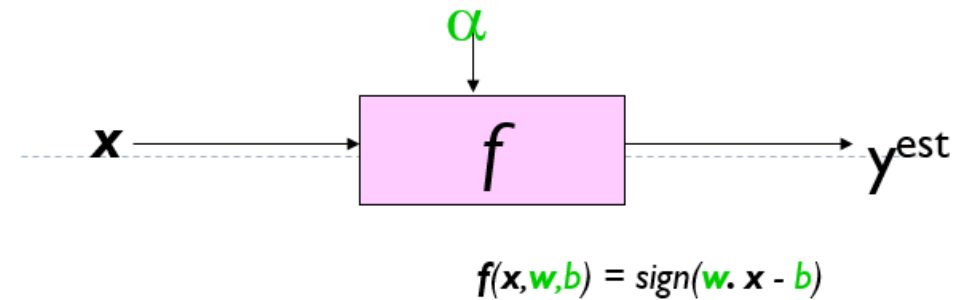
Yang mana yang terbaik?
(Hyperplane terbaik)?





Classifiers Margin

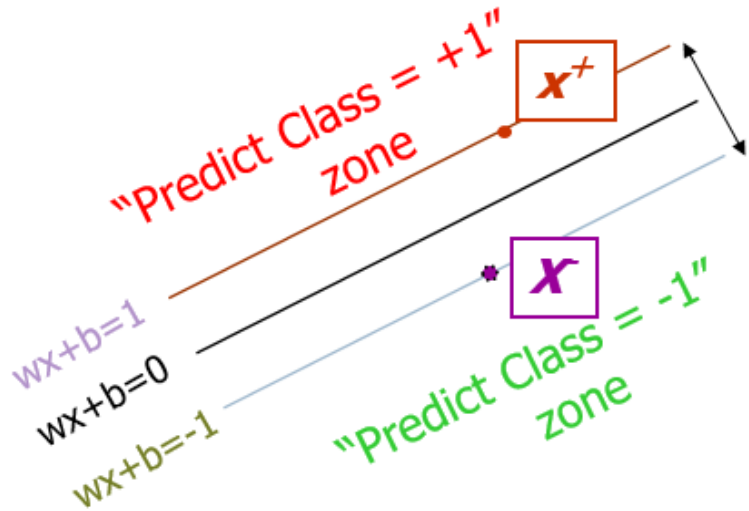
- **Ide dasar SVM:** memaksimalkan distance (jarak) antara hyperplane dan titik sampel terdekat



Definisikan **margin/batas** linear classifier sebagai lebarnya dimana batas tersebut dapat dimaksimalkan sebelum mencapai titik data



Linear SVM secara Matematis



- ▶ $w \cdot x^+ + b = +1$
- ▶ $w \cdot x^- + b = -1$
- ▶ $w \cdot (x^+ - x^-) = 2$

$$M = \frac{(x^+ - x^-) \cdot w}{|w|} = \frac{2}{|w|}$$

Dimana,

M : Margin

w : bobot

x^+ : support vector kelas +1

x^- : support vector kelas -1





Problem Optimasi

- $\{x_1, \dots, x_n\}$ adalah data set dan $y_i \in \{1, -1\}$ adalah kelas label dari x_i
- Batas keputusan harus dapat mengklasifikasi semua titik dengan benar

$$y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1, \quad \forall i$$

- Problem optimasi terbatas

$$\text{meminimalkan } \frac{1}{2} \|\mathbf{w}\|^2$$

$$\text{dimana } y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 \quad \forall i$$





Problem Optimasi

- Kita dapat mengubah problem menjadi bentuk dual

$$\max. W(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1, j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j$$

dimana $\alpha_i \geq 0, \sum_{i=1}^n \alpha_i y_i = 0$

- Problem *quadratic programming* (QP)
 - Global maximum pada α_i dapat selalu ditemukan
- $\mathbf{w}$ dapat diperbaiki menjadi $\mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i$





Karakteristik Solusi

- Kita dapat mengubah problem menjadi bentuk dual

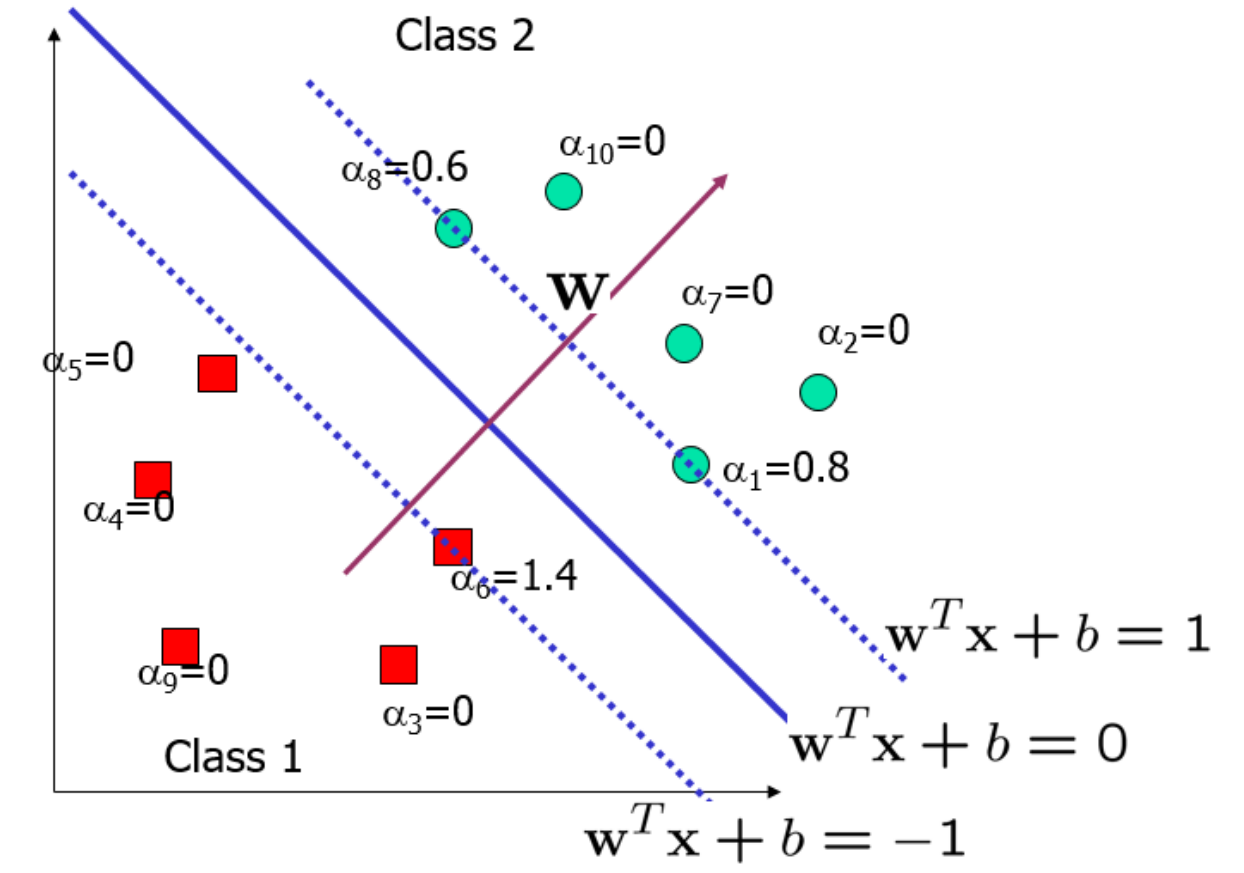
$$\max. W(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1, j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j$$

dimana $\alpha_i \geq 0, \sum_{i=1}^n \alpha_i y_i = 0$

- Problem *quadratic programming* (QP)
 - Global maximum pada α_i dapat selalu ditemukan
- $\mathbf{w}$ dapat diperbaiki menjadi $\mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i$

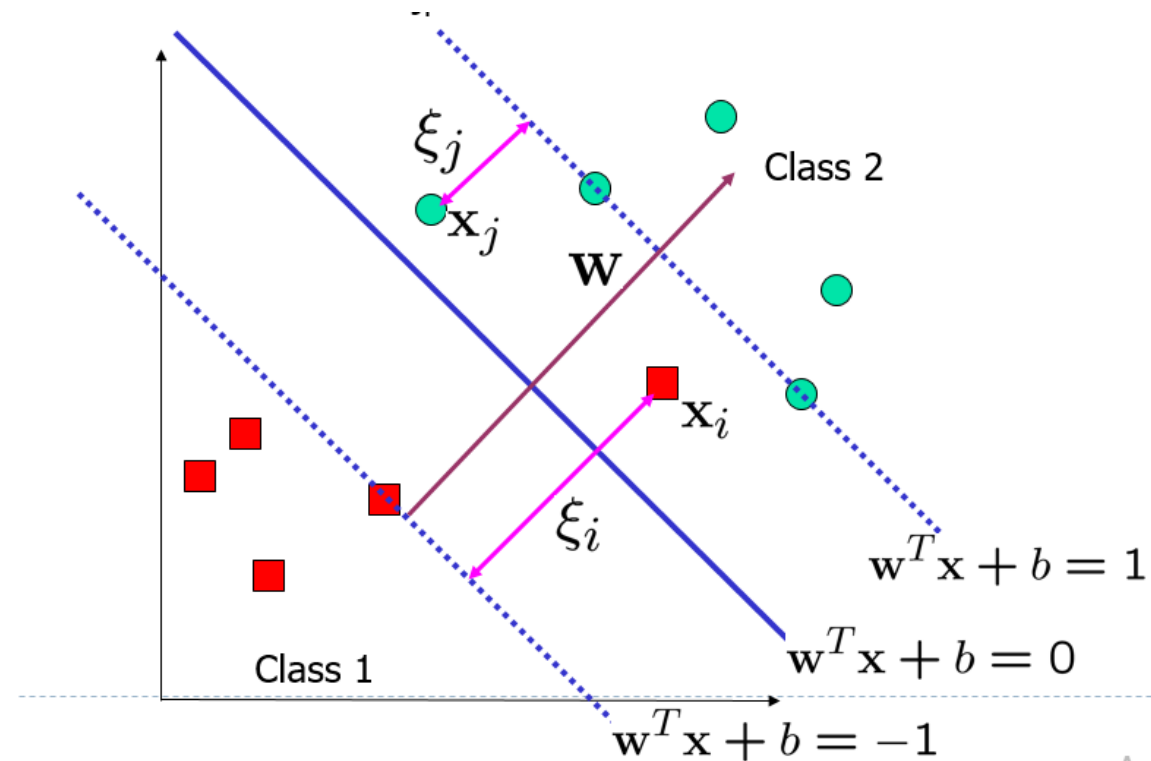


Interpretasi Geometri



● Tidak dapat Dipisahkan secara Linier?

- Memperbolehkan adanya *error* (ξ) pada klasifikasi



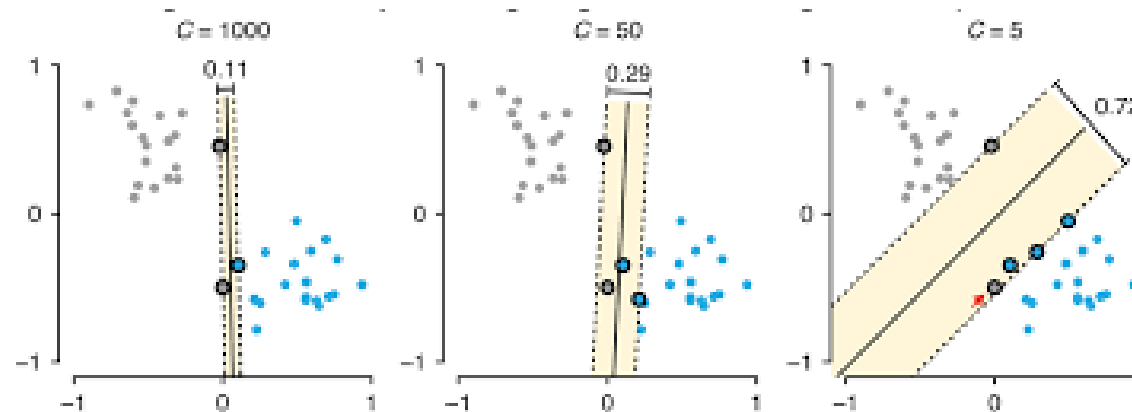
Soft Margin Hyperplane

- Definisikan $\xi_i=0$ jika tidak ada error untuk $\mathbf{x}_i$
 - ξ_i adalah variabel untuk mendefinisikan nilai error
$$\begin{cases} \mathbf{w}^T \mathbf{x}_i + b \geq 1 - \xi_i & y_i = 1 \\ \mathbf{w}^T \mathbf{x}_i + b \leq -1 + \xi_i & y_i = -1 \\ \xi_i \geq 0 & \forall i \end{cases}$$
- Minimalkan $\frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i$
 - C : tradeoff parameter antara error dan margin
- Problem optimasi menjadi
meminimalkan $\frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i$
dimana $y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0$



Soft Margin Hyperplane

- Menemukan nilai yang tepat untuk C menjadi salah satu masalah dalam SVM
- Nilai C berperan dalam mengontrol *overfitting*.
 - C besar $\rightarrow$ lebih sedikit sampel training yang berada di posisi yang tidak ideal (artinya lebih sedikit error, sehingga berdampak positif pada kinerja *classifier*) $\rightarrow$ **C terlalu besar menyebabkan *overfitting***
 - C kecil $\rightarrow$ lebih banyak sampel training yang tidak berada pada posisi ideal (artinya akan banyak error training sehingga berdampak negatif pada kinerja *classifier*) $\rightarrow$ **C terlalu kecil menyebabkan *underfitting***



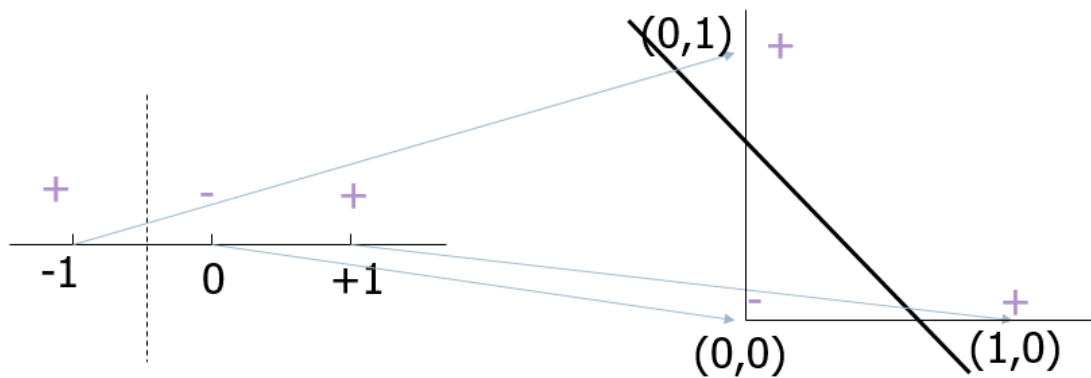
● Permasalahan Non-linear

- Ide: transformasi $\mathbf{x}_i$ ke ruang berdimensi lebih tinggi untuk memudahkan perhitungan
 - Ruang Input : ruang $\mathbf{x}_i$
 - Ruang Fitur: ruang $\phi(\mathbf{x}_i)$ setelah transformasi
 - Mengapa perlu transformasi?
 - Operasi linear pada ruang fitur ekuivalen dengan operasi non-linear pada ruang input
 - Proses klasifikasi lebih mudah dilakukan dengan transformasi.
- Contoh: XOR



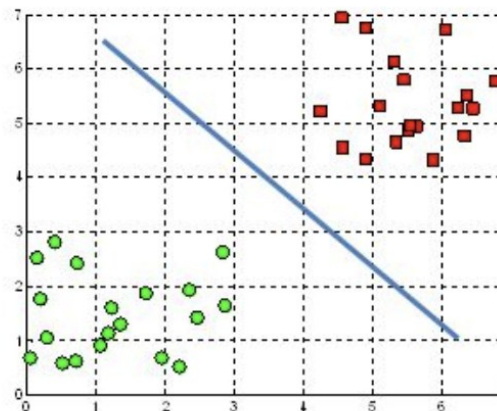
Dimensi Tinggi

- Proyeksikan data ke ruang berdimensi tinggi agar data-data tersebut dapat dipisahkan secara linear dan dapat menggunakan linear SVM – (Using Kernels)

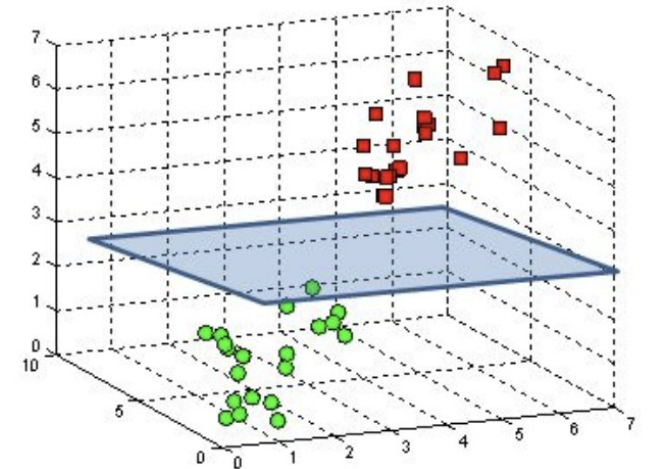


Data dari $\mathbb{R}^1$ ditransformasi ke $\mathbb{R}^2$

A hyperplane in $\mathbb{R}^2$ is a line

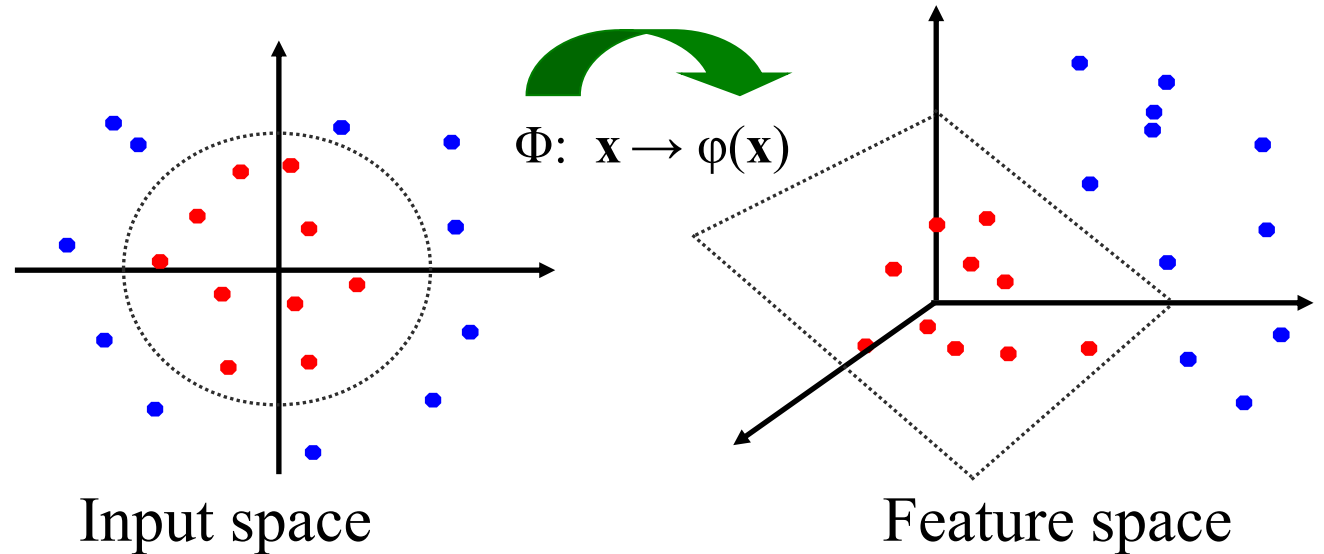


A hyperplane in $\mathbb{R}^3$ is a plane



● Permasalahan Non-linear

- Kemungkinan problem transformasi
 - Komputasi yang tinggi dan sulit memperoleh estimasi bagus
- SVM menyelesaikan masalah ini secara bersamaan
 - Kernel tricks untuk komputasi yg efisien
 - Meminimalkan $\|\mathbf{w}\|^2$ dapat menghasilkan classifier yang baik





Kernel Trick

- Hubungan antara fungsi kernel K dan mapping $\phi(\cdot)$

$$\mathbf{K}(\mathbf{x}, \mathbf{y}) = \langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle$$

- Disebut sebagai **kernel trick**
- Secara intuitif, $\mathbf{K}(\mathbf{x}, \mathbf{y})$ merepresentasikan kemiripan antara data $\mathbf{x}$ dan $\mathbf{y}$, dan ini diperoleh dari pengetahuan sebelumnya



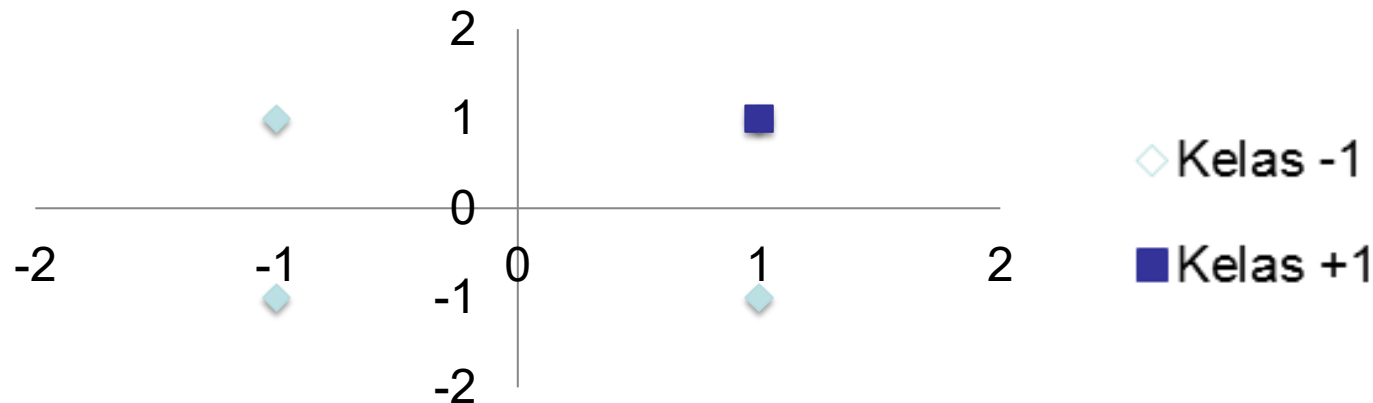
Contoh Studi Kasus 1

- Contoh SVM Linier pada dataset berikut :

Tentukan Hyperplanenya !

x_1	x_2	Kelas (y)	Support Vector (SV)
1	1	1	1
1	-1	-1	1
-1	1	-1	1
-1	-1	-1	0

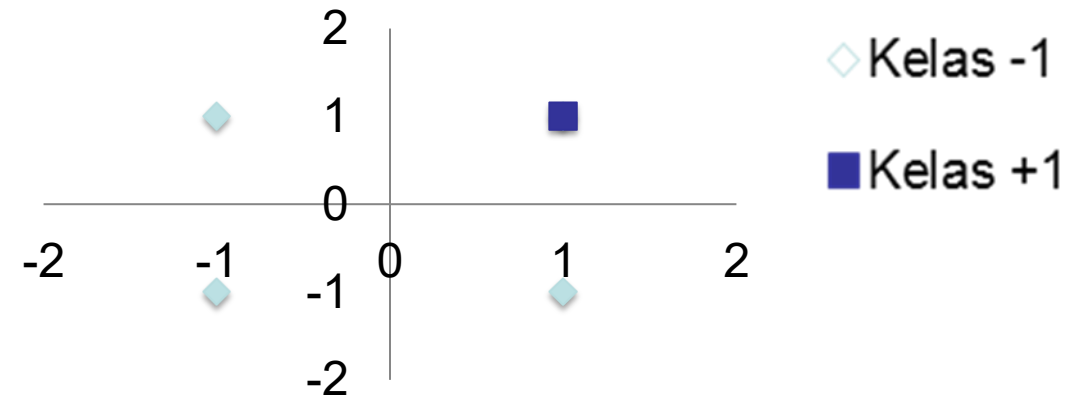
- Bentuk Visualisasi data :



Contoh Studi Kasus 1

- Contoh SVM Linier :

x_1	x_2	Kelas (y)
1	1	1
1	-1	-1
-1	1	-1
-1	-1	-1



- Karena ada dua fitur (x_1 dan x_2), maka w juga akan memiliki 2 fitur (w_1 dan w_2).
- Formulasi yang digunakan adalah sebagai berikut :
 - Meminimalkan nilai : $\frac{1}{2} \|w\|^2 = \frac{1}{2} (w_1^2 + w_2^2)$
 - Syarat : $y_i(w \cdot x_i + b) \geq 1, \quad i = 1, 2, 3, \dots, N$



● Contoh Studi Kasus 1

- Karena ada dua fitur (x_1 dan x_2), maka w juga akan memiliki 2 fitur (w_1 dan w_2).
- Formulasi yang digunakan adalah sebagai berikut :

- Meminimalkan nilai margin :

$$\frac{1}{2} \|w\|^2 = \frac{1}{2} (w_1^2 + w_2^2)$$

- Syarat :

$$y_i(w \cdot x_i + b) \geq 1, \text{ dimana } i = 1, 2, 3, \dots, N$$

$$y_i(w_1 \cdot x_1 + w_2 \cdot x_2 + b) \geq 1$$

- Sehingga didapatkan beberapa persamaan berikut :

$$1. (w_1 + w_2 + b) \geq 1, \text{ untuk } y_1 = 1, x_1 = 1, x_2 = 1$$

$$2. (-w_1 + w_2 - b) \geq 1, \text{ untuk } y_2 = -1, x_1 = 1, x_2 = -1$$

$$3. (w_1 - w_2 - b) \geq 1, \text{ untuk } y_3 = -1, x_1 = -1, x_2 = 1$$

$$4. (w_1 + w_2 - b) \geq 1, \text{ untuk } y_4 = -1, x_1 = -1, x_2 = -1$$

x_1	x_2	Kelas (y)
1	1	1
1	-1	-1
-1	1	-1
-1	-1	-1



Contoh Studi Kasus 1

- Dari 4 persamaan berikut, carilah nilai w_1 , w_2 dan b

- $(w_1 + w_2 + b) \geq 1$
- $(-w_1 + w_2 - b) \geq 1$
- $(w_1 - w_2 - b) \geq 1$
- $(w_1 + w_2 - b) \geq 1$

Tahap 1: Jumlahkan persamaan (1) dan (2)

$$\begin{array}{r} (w_1 + w_2 + b) \geq 1 \\ (-w_1 + w_2 - b) \geq 1 \\ \hline + \\ 2w_2 = 2 \\ \text{Maka } w_2 = 1 \end{array}$$



Tahap 2: Jumlahkan persamaan (1) dan (3):

$$\begin{array}{r} (w_1 + w_2 + b) \geq 1 \\ (w_1 - w_2 - b) \geq 1 \\ \hline + \\ 2w_1 = 2 \\ \text{Maka } w_1 = 1 \end{array}$$

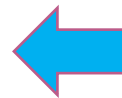


Tahap 3: Jumlahkan persamaan (1) dan (3):

$$\begin{array}{r} (-w_1 + w_2 - b) \geq 1 \\ (w_1 - w_2 - b) \geq 1 \\ \hline + \\ -2b = 2 \\ \text{Maka } b = -1 \end{array}$$

Tahap 4: didapatkan persamaan *hyperplane*

$$\begin{aligned} w_1 \cdot x_1 + w_2 \cdot x_2 + b &= 0 \\ x_1 + x_2 + 1 &= 0 \\ x_2 &= 1 - x_1 \end{aligned}$$



Contoh Studi Kasus 1

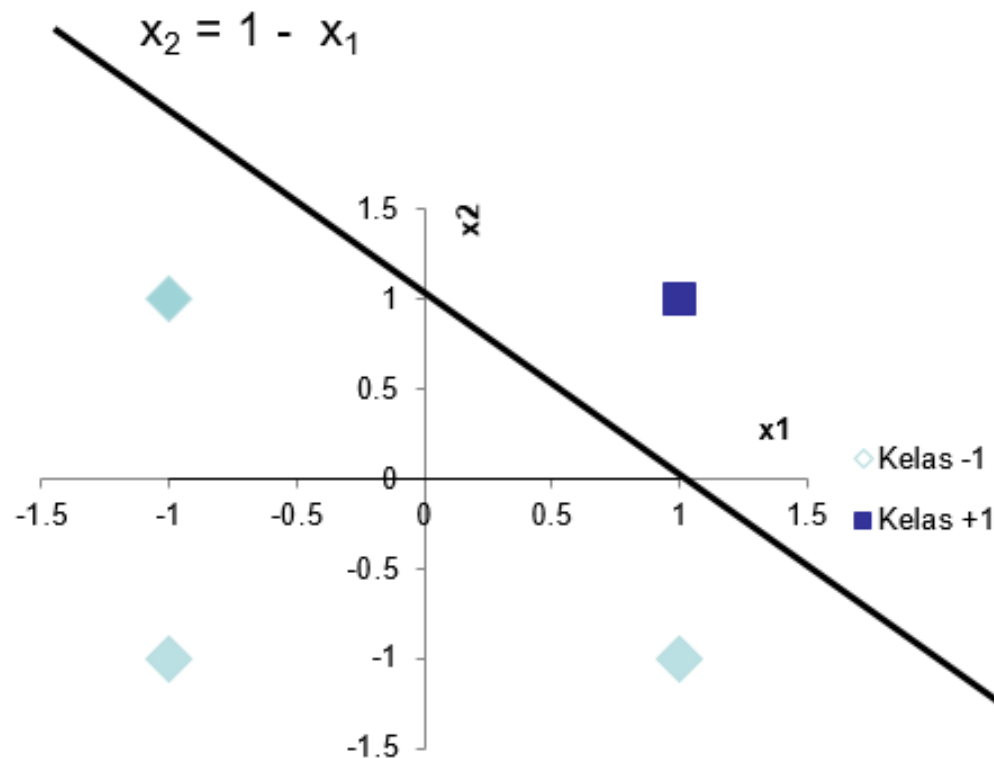
Visualisasi garis hyperplane (sebagai fungsi klasifikasi) :

$$w_1x_1 + w_2x_2 + b = 0$$

$$x_1 + x_2 - 1 = 0$$

$$x_2 = 1 - x_1$$

x_1	$x_2 = 1 - x_1$
-2	3
-1	2
0	1
1	0
2	-1



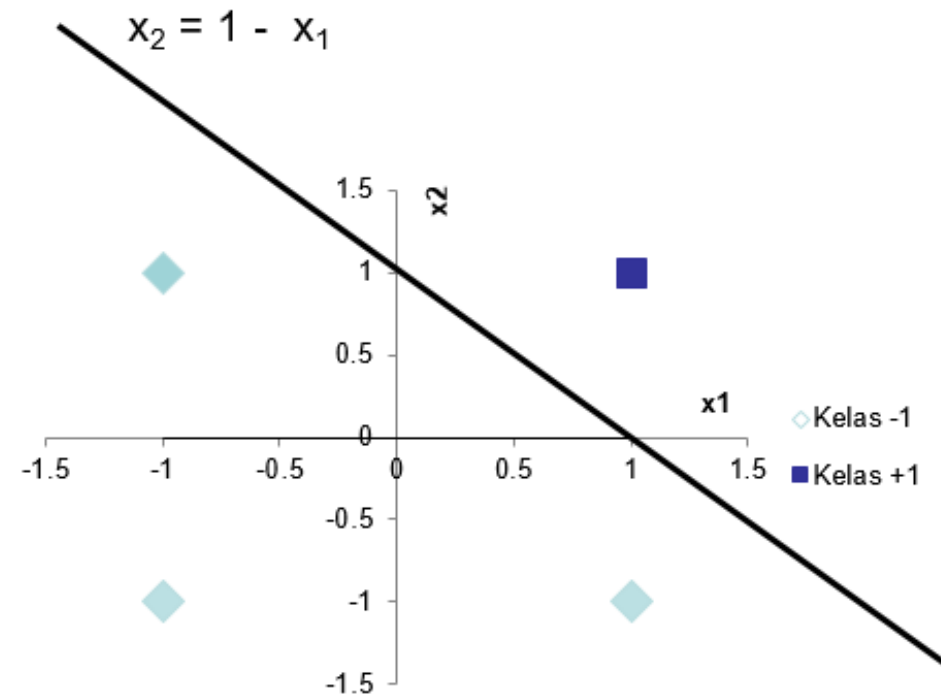
Contoh Studi Kasus 1

Misalkan diketahui data uji/ data testing berikut :

Diketahui : $f(x) = x_1 + x_2 - 1$

Kelas = $\text{sign}(f(x))$

No	Data Uji		Hasil Klasifikasi
	x_1	x_2	Kelas = $\text{sign}(x_1 + x_2 - 1)$
1	1	5	$\text{sign}(1 + 5 - 1) = +1$
2	-1	4	$\text{sign}(-1 + 4 - 1) = +1$
3	0	7	$\text{sign}(0 + 7 - 1) = +1$
4	-9	0	$\text{sign}(-9 + 0 - 1) = -1$
5	2	-2	$\text{sign}(2 - 2 - 1) = -1$



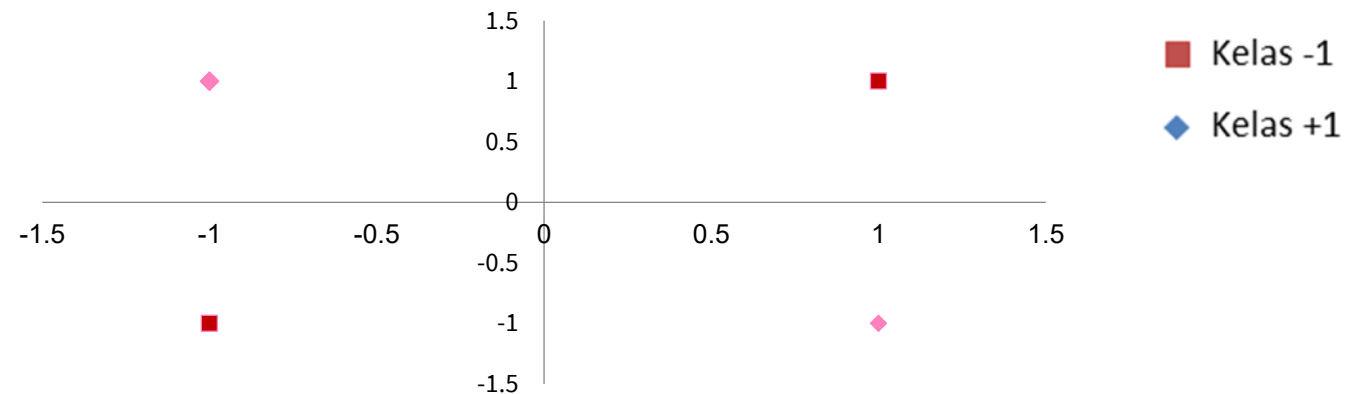


Contoh Studi Kasus 2

- Contoh SVM Non Linier pada dataset berikut :

x_1	x_2	Kelas (y)	Support Vector (SP)
1	1	-1	1
1	-1	1	1
-1	1	1	1
-1	-1	-1	1

- Bentuk Visualisasi data :

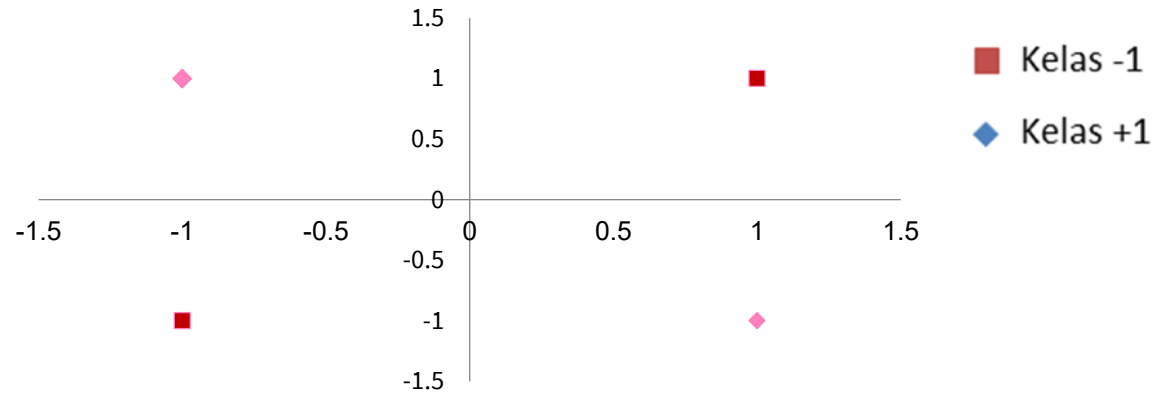




Contoh Studi Kasus 2

- Contoh SVM Non Linier :

x_1	x_2	Kelas (y)
1	1	-1
1	-1	1
-1	1	1
-1	-1	-1



- Karena ada dua fitur (x_1 dan x_2), dan kelompok datanya tidak linear, maka digunakan fungsi kernel. Misal menggunakan fungsi kernel polynomial ordo 2, yaitu :
 $\mathbf{K}(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y} + \mathbf{c})^d$ dengan $c = 1$ dan $d = 2$.
- Fungsi kernel dituliskan kembali menjadi berikut :
 $\mathbf{K}(\mathbf{x}, \mathbf{x}_i) = (\mathbf{x}^T \cdot \mathbf{x}_i + \mathbf{1})^2$ dengan
- Menghitung matrik kernel K :
 $\mathbf{K}(\mathbf{x}, \mathbf{x}_i) = \Phi(\mathbf{x}) \cdot \Phi(\mathbf{x}_i)$

$$w = \sum_{i=1..N} \alpha_i y_i \phi(x_i)$$





Contoh Studi Kasus 2

- Fungsi kernel dituliskan kembali menjadi berikut :
 $\mathbf{K}(\mathbf{x}, \mathbf{x}_i) = (\mathbf{x}^T \cdot \mathbf{x}_i + \mathbf{1})^2$ dengan $w = \sum \alpha_i y_i \phi(x_i)$
- Menghitung matrik kernel $\mathbf{K}(\mathbf{x}, \mathbf{x}_i) = \Phi(\mathbf{x}) \cdot \Phi(\mathbf{x}_i)$
- Misal, Menghitung $K(u, z)$: dengan $u=(1,1)$ dan $z=(1,-1)$

x_1	x_2	Kelas (y)
1	1	-1
1	-1	1
-1	1	1
-1	-1	-1

$$\begin{aligned}
 k(u=(1,1), z=(1,-1)) &= (((u_1 \cdot z_1) + (u_2 \cdot z_2)) + 1)^2 = ((u_1 \cdot z_1) + (u_2 \cdot z_2))^2 + 2((u_1 \cdot z_1) + (u_2 \cdot z_2)) \cdot 1 + 1^2 \\
 &= (u_1 \cdot z_1)^2 + 2(u_1 \cdot z_1)(u_2 \cdot z_2) + (u_2 \cdot z_2)^2 + 2(u_1 \cdot z_1) + 2(u_2 \cdot z_2) + 1
 \end{aligned}$$

$$= \begin{pmatrix} u_1^2 \\ \sqrt{2}u_1u_2 \\ u_2^2 \\ \sqrt{2}u_1 \\ \sqrt{2}u_2 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} z_1^2 \\ \sqrt{2}z_1z_2 \\ z_2^2 \\ \sqrt{2}z_1 \\ \sqrt{2}z_2 \\ 1 \end{pmatrix}$$

$$= \phi(u) \cdot \phi(z)$$

$$= u_1^2 z_1^2 + 2u_1 u_2 z_1 z_2 + u_2^2 z_2^2 + 2u_1 z_1 + 2u_2 z_2 + 1$$

$$= (u_1 z_1)^2 + 2(u_1 z_1)(u_2 z_2) + (u_2 z_2)^2 + 2(u_1 z_1) + 2(u_2 z_2) + 1$$





Contoh Studi Kasus 2

- Fungsi kernel dituliskan kembali menjadi berikut :
 $\mathbf{K}(\mathbf{x}, \mathbf{x}_i) = (\mathbf{x}^T \cdot \mathbf{x}_i + \mathbf{1})^2$ dengan $w = \sum \alpha_i y_i \phi(x_i)$
- Menghitung matrik kernel $\mathbf{K}(\mathbf{x}, \mathbf{x}_i) = \Phi(\mathbf{x}) \cdot \Phi(\mathbf{x}_i)$
- Misal, Menghitung $K(u, z)$: dengan $u=(1,1)$ dan $z=(1,-1)$

x_1	x_2	Kelas (y)
1	1	-1
1	-1	1
-1	1	1
-1	-1	-1

$$\begin{aligned}
 k(u=(1,1), z=(1,-1)) &= (((u_1 \cdot z_1) + (u_2 \cdot z_2)) + 1)^2 = ((u_1 \cdot z_1) + (u_2 \cdot z_2))^2 + 2((u_1 \cdot z_1) + (u_2 \cdot z_2)) \cdot 1 + 1^2 \\
 &= (u_1 \cdot z_1)^2 + 2(u_1 \cdot z_1)(u_2 \cdot z_2) + (u_2 \cdot z_2)^2 + 2(u_1 \cdot z_1) + 2(u_2 \cdot z_2) + 1
 \end{aligned}$$

$$= \begin{pmatrix} u_1^2 \\ \sqrt{2}u_1u_2 \\ u_2^2 \\ \sqrt{2}u_1 \\ \sqrt{2}u_2 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} z_1^2 \\ \sqrt{2}z_1z_2 \\ z_2^2 \\ \sqrt{2}z_1 \\ \sqrt{2}z_2 \\ 1 \end{pmatrix}$$

$$= \phi(u) \cdot \phi(z)$$

$$= u_1^2 z_1^2 + 2u_1 u_2 z_1 z_2 + u_2^2 z_2^2 + 2u_1 z_1 + 2u_2 z_2 + 1$$

$$= (u_1 z_1)^2 + 2(u_1 z_1)(u_2 z_2) + (u_2 z_2)^2 + 2(u_1 z_1) + 2(u_2 z_2) + 1$$





Contoh Studi Kasus 2

- Misal, Menghitung $K(u,z)$: dengan $u=(1,1)$ dan $z=(1,-1)$

$$\begin{aligned} k(u=(1,1),z=(1,-1)) &= (((1.1)+(1.(-1))))+1)^2 = ((1.1)+(1.(-1)))^2+2((1.1)+(1.(-1))).1 + 1^2 \\ &= (1.1)^2 + 2(1.1)(1.(-1)) + (1.(-1))^2 + 2(1.1) + 2(1.(-1)) + 1 \\ &= 1 - 2 + 1 + 2 - 2 + 1 = 1 \end{aligned}$$

$$= \begin{pmatrix} u_1^2 \\ \sqrt{2}u_1u_2 \\ u_2^2 \\ \sqrt{2}u_1 \\ \sqrt{2}u_2 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} z_1^2 \\ \sqrt{2}z_1z_2 \\ z_2^2 \\ \sqrt{2}z_1 \\ \sqrt{2}z_2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1^2 \\ \sqrt{2}.1.1 \\ 1^2 \\ \sqrt{2}.1 \\ \sqrt{2}.1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1^2 \\ \sqrt{2}.1.(-1) \\ (-1)^2 \\ \sqrt{2}.1 \\ \sqrt{2}.(-1) \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \\ \sqrt{2} \\ \sqrt{2} \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \\ \sqrt{2} \\ -\sqrt{2} \\ 1 \end{pmatrix} = 1 - 2 + 1 + 2 - 2 + 1 = 1$$

$$= \phi(u) \cdot \phi(z)$$

$$= u_1^2 z_1^2 + 2u_1 u_2 z_1 z_2 + u_2^2 z_2^2 + 2u_1 z_1 + 2u_2 z_2 + 1$$

$$= (u_1 z_1)^2 + 2(u_1 z_1)(u_2 z_2) + (u_2 z_2)^2 + 2(u_1 z_1) + 2(u_2 z_2) + 1$$





Contoh Studi Kasus 2

- Menghitung matrik kernel $K(\mathbf{x}, \mathbf{x}_i) = \Phi(\mathbf{x}) \cdot \Phi(\mathbf{x}_i)$

x_1	x_1	$K(1,1) = (x_1 \cdot x_1 + 1)^2 = (1 \cdot 1 + 1 \cdot 1 + 1)^2 = 3^2 = 9$
	x_2	$K(1,2) = (x_1 \cdot x_2 + 1)^2 = (1 \cdot 1 + 1 \cdot (-1) + 1)^2 = 1^2 = 1$
	x_3	$K(1,3) = (x_1 \cdot x_3 + 1)^2 = (1 \cdot (-1) + 1 \cdot 1 + 1)^2 = 1^2 = 1$
	x_4	$K(1,4) = (x_1 \cdot x_4 + 1)^2 = (1 \cdot (-1) + 1 \cdot (-1) + 1)^2 = (-1)^2 = 1$
x_2	x_1	$K(2,1) = (x_2 \cdot x_1 + 1)^2 = (1 \cdot 1 + (-1) \cdot 1 + 1)^2 = 1^2 = 1$
	x_2	$K(2,2) = (x_2 \cdot x_2 + 1)^2 = (1 \cdot 1 + (-1) \cdot (-1) + 1)^2 = 3^2 = 9$
	x_3	$K(2,3) = (x_2 \cdot x_3 + 1)^2 = (1 \cdot (-1) + (-1) \cdot 1 + 1)^2 = 1^2 = 1$
	x_4	$K(2,4) = (x_2 \cdot x_4 + 1)^2 = (1 \cdot (-1) + (-1) \cdot (-1) + 1)^2 = 1^2 = 1$
x_3	x_1	$K(3,1) = (x_3 \cdot x_1 + 1)^2 = ((-1) \cdot 1 + 1 \cdot 1 + 1)^2 = 1^2 = 1$
	x_2	$K(3,2) = (x_3 \cdot x_2 + 1)^2 = ((-1) \cdot 1 + 1 \cdot (-1) + 1)^2 = 1^2 = 1$
	x_3	$K(3,3) = (x_3 \cdot x_3 + 1)^2 = ((-1) \cdot (-1) + 1 \cdot 1 + 1)^2 = 3^2 = 9$
	x_4	$K(3,4) = (x_3 \cdot x_4 + 1)^2 = ((-1) \cdot (-1) + 1 \cdot (-1) + 1)^2 = 1^2 = 1$
x_4	x_1	$K(4,1) = (x_4 \cdot x_1 + 1)^2 = ((-1) \cdot 1 + (-1) \cdot 1 + 1)^2 = 1^2 = 1$
	x_2	$K(4,2) = (x_4 \cdot x_2 + 1)^2 = ((-1) \cdot 1 + (-1) \cdot (-1) + 1)^2 = 1^2 = 1$
	x_3	$K(4,3) = (x_4 \cdot x_3 + 1)^2 = ((-1) \cdot (-1) + (-1) \cdot 1 + 1)^2 = 1^2 = 1$
	x_4	$K(4,4) = (x_4 \cdot x_4 + 1)^2 = ((-1) \cdot (-1) + (-1) \cdot (-1) + 1)^2 = 3^2 = 9$

x_1	x_2	Kelas (y)
1	1	-1
1	-1	1
-1	1	1
-1	-1	-1





Contoh Studi Kasus 2

- Didapatkan matrik kernel K dengan ukuran N x N :

$K(x, x_i) =$

9	1	1	1
1	9	1	1
1	1	9	1
1	1	1	9

x_1	x_2	Kelas (y)
1	1	-1
1	-1	1
-1	1	1
-1	-1	-1

- Setiap elemen matrik kernel $K(x, x_i)$ digunakan untuk menggantikan dot-product $x_i \cdot x_j$ dalam persamaan dualitas Lagrange Multiplier.

- Maksimalkan Ld :

$$Ld = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j)$$

$$Ld = (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4) - \frac{1}{2} \begin{pmatrix} 9\alpha_1\alpha_1 - 1\alpha_1\alpha_2 - 1\alpha_1\alpha_3 + 1\alpha_1\alpha_4 - \\ 1\alpha_2\alpha_1 + 9\alpha_2\alpha_2 + 1\alpha_2\alpha_3 - 1\alpha_2\alpha_4 - \\ 1\alpha_3\alpha_1 + 1\alpha_3\alpha_2 + 9\alpha_3\alpha_3 - 1\alpha_3\alpha_4 + \\ 1\alpha_4\alpha_1 - 1\alpha_4\alpha_2 - 1\alpha_4\alpha_3 + 9\alpha_4\alpha_4 \end{pmatrix}$$

$$Ld = (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4) - \frac{1}{2} [9\alpha_1\alpha_1 + 2\alpha_1\alpha_2 - 2\alpha_1\alpha_3 + \dots]$$





Contoh Studi Kasus 2

- Setiap elemen matrik kernel $K(x, x_i)$ digunakan untuk menggantikan dot-product $\mathbf{x}_i \cdot \mathbf{x}_j$ dalam persamaan dualitas Lagrange Multiplier.
 - Maksimalkan L_d :

x_1	x_2	Kelas (y)
1	1	-1
1	-1	1
-1	1	1
-1	-1	-1

$$L_d = (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4) - \frac{1}{2} \begin{pmatrix} 9\alpha_1^2 - 2\alpha_1\alpha_2 - 2\alpha_1\alpha_3 + 2\alpha_1\alpha_4 + \\ 9\alpha_2^2 + 2\alpha_2\alpha_3 - 2\alpha_2\alpha_4 + \\ 9\alpha_3^2 - 2\alpha_3\alpha_4 + 9\alpha_4^2 \end{pmatrix}$$

$$\text{Syarat} : 0 \leq \alpha_i \leq C \text{ dan } \sum_{i=1}^N \alpha_i y_i = 0$$

$$\text{Syarat 1} : \alpha_1, \alpha_2, \alpha_3, \alpha_4 \geq 0$$

$$\text{Syarat 2} : -\alpha_1 + \alpha_2 + \alpha_3 - \alpha_4 = 0$$

(Note : Untuk hasil optimal L_d , gunakan koding)

- Misalkan didapatkan nilai Max L_d dengan $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.125$. Sehingga nilai $L_d = 0.25$.

- Hitung nilai w dan b : $X_i = \begin{pmatrix} x_1^i \\ x_2^i \end{pmatrix}$, jika X_1 adalah data ke-1,

$$X_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \phi(X_i) = \begin{pmatrix} x_1^{i^2} & \sqrt{2}x_1^i x_2^i & x_2^{i^2} & \sqrt{2}x_1^i & \sqrt{2}x_2^i & 1 \end{pmatrix}^T$$

dimana x_1^i adalah nilai pada dimensi ke-1 pada data ke-i.





Contoh Studi Kasus 2

- Hitung nilai w dan b : $X_i = \begin{pmatrix} x_1^i \\ x_2^i \end{pmatrix}$, jika X_1 adalah data ke-1,

$$X_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \phi(X_i) = \begin{pmatrix} x_1^{i^2} & \sqrt{2}x_1^i x_2^i & x_2^{i^2} & \sqrt{2}x_1^i & \sqrt{2}x_2^i & 1 \end{pmatrix}^T$$

dimana x_1^i adalah nilai pada dimensi ke-1 pada data ke-i.

x_1	x_2	Kelas (y)
1	1	-1
1	-1	1
-1	1	1
-1	-1	-1

$$w = \sum_{i=1}^N \alpha_i y_i \phi(X_i)$$

$$w = \sum_{i=1}^N \alpha_i y_i \phi(X_i) = \sum_{i=1}^4 \alpha_i y_i \phi(X_i) = \alpha_1 y_1 \phi(X_1) + \alpha_2 y_2 \phi(X_2) + \alpha_3 y_3 \phi(X_3) + \alpha_4 y_4 \phi(X_4)$$

$$w = -0.125 \begin{pmatrix} x_1^{1^2} = 1^2 = 1 \\ \sqrt{2}x_1^1 x_2^1 = \sqrt{2}(1)(1) = \sqrt{2} \\ x_2^{1^2} = 1^2 = 1 \\ \sqrt{2}x_1^1 = \sqrt{2}(1) = \sqrt{2} \\ \sqrt{2}x_2^1 = \sqrt{2}(1) = \sqrt{2} \\ 1 \end{pmatrix} + 0.125 \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \\ \sqrt{2} \\ -\sqrt{2} \\ 1 \end{pmatrix} + 0.125 \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \\ -\sqrt{2} \\ \sqrt{2} \\ 1 \end{pmatrix} - 0.125 \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \\ -\sqrt{2} \\ -\sqrt{2} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -0.71 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$





Contoh Studi Kasus 2

- Misalkan didapatkan nilai Max Ld dengan $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.125$.
Sehingga nilai Ld = 0.25.

- Hitung nilai w dan b :

$$w = \sum_{i=1}^N \alpha_i y_i \phi(X_i) = \begin{pmatrix} 0 \\ -0.71 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

x_1	x_2	Kelas (y)
1	1	-1
1	-1	1
-1	1	1
-1	-1	-1

Pilih salah satu Support Vector dari Kelas “+1” dan “-1” untuk menghitung nilai b.

$$b = -\frac{1}{2} (w \cdot x^+ + w \cdot x^-) = -\frac{1}{2} \left(\begin{pmatrix} 0 \\ -0.71 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \\ \sqrt{2} \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ -0.71 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \\ \sqrt{2} \\ 1 \end{pmatrix} \right)$$





Contoh Studi Kasus 2

- Misalkan didapatkan nilai Max Ld dengan $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.125$.
Sehingga nilai Ld = 0.25.
- Hitung nilai w dan b :

x_1	x_2	Kelas (y)
1	1	-1
1	-1	1
-1	1	1
-1	-1	-1

Pilih salah satu Support Vector dari Kelas “+1” dan “-1” untuk menghitung nilai b.

$$b = -\frac{1}{2} (w.x^+ + w.x^-) = -\frac{1}{2} \left(\begin{pmatrix} 0 \\ -0.71 \\ 0 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -\sqrt{2} \\ \sqrt{2} \\ -\sqrt{2} \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \sqrt{2} \\ \sqrt{2} \\ \sqrt{2} \\ 1 \end{pmatrix} \right)$$

$$b = -\frac{1}{2} ((-0.71)(-\sqrt{2}) + (-0.71)(\sqrt{2})) = -\frac{1}{2} ((-0.71)(-\sqrt{2}) + (-0.71)(\sqrt{2})) = 0$$





Contoh Studi Kasus 2

- Setelah didapatkan nilai w dan b :

$$w = \begin{pmatrix} 0 \\ -0.71 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad b = 0$$

$$\begin{aligned} w \cdot \phi(x_t) + b &= w \cdot \phi(x_t) + 0 \\ &= \sum_{i=1}^N \alpha_i y_i \phi(x_i) \cdot \phi(x_t) = \sum_{i=1}^4 \alpha_i y_i K(x_i, x_t) \end{aligned}$$

x_1	x_2	Kelas (y)
1	1	-1
1	-1	1
-1	1	1
-1	-1	-1

Maka model SVM siap digunakan untuk proses klasifikasi.

$$f(\phi(x)) = \text{sign}(w \cdot \phi(x) + b) = \text{sign}\left(\sum_{i=1}^N \alpha_i y_i \phi(x_i) \cdot \phi(x) + b\right)$$

Misalkan data uji/ data test $x_t = (1,5)$ maka $K(x_i, x_t) = \Phi(x_i) \cdot \Phi(x_t)$

$x_t=(1,5)$	x_1	$K(1,t) = (x_1 \cdot x_t + 1)^2 = (1 \cdot 1 + 5 \cdot 1 + 1)^2 = 7^2 = 49$	$(-0.125)(49)$
	x_2	$K(2,t) = (x_2 \cdot x_t + 1)^2 = (1 \cdot 1 + 5 \cdot (-1) + 1)^2 = 3^2 = 9$	$(0.125)(9)$
	x_3	$K(3,t) = (x_3 \cdot x_t + 1)^2 = (1 \cdot (-1) + 5 \cdot 1 + 1)^2 = 5^2 = 25$	$(0.125)(25)$
	x_4	$K(4,t) = (x_4 \cdot x_t + 1)^2 = (1 \cdot (-1) + 5 \cdot (-1) + 1)^2 = (-5)^2 = 25$	$(-0.125)(25)$

-5

+

$$f(\phi(x_t)) = \text{sign}(w \cdot \phi((1,5)) + b) = \text{sign}(-6.125 + 1.125 + 3.125 - 3.125 + 0) = \text{sign}(-5) \Rightarrow -1$$

Jadi data $x_t = (1,5)$ tersebut masuk ke kelas negatif.





Latihan Individu

1. Perhatikan dataset SVM Linier berikut :

x_1	x_2	Kelas (y)	Support Vector (SV)
1	1	+1	1
1	-1	-1	1
0	2	-1	1
-1	-1	-1	0

dataset 1

x_1	x_2	Kelas (y)	SV1	SV2
2	3	-1	1	0
3	4	-1	1	1
5	2	+1	1	1
6	3	+1	0	0

dataset 2

Tentukan Visualisasi Hyperplane masing-masing dataset di atas !





Latihan Individu

2. Perhatikan dataset SVM non-Linier berikut :

x_1	x_2	Kelas (y)	SV
0.5	0	-1	1
1	-1	+1	1
-1	1	+1	1
-0.5	-0.5	-1	1

Tentukan persamaan Hyperplanenya, lalu uji kelas data $x_t = (1,1)$!





TUGAS 2

- Tuliskan di buku catatan kemudian scan dan simpan dalam 1 file pdf.
- Tugas maksimal dikumpulkan Kamis tanggal 30 Oktober 2026 Pukul 23.59.
- Link Pengumpulan pada Ilmu2
- Aturan penamaan folder:

NPM_Nama_T2





TERIMA KASIH

