Live Interaction #7:

25th February 2024

E-masters Next Generation Wireless Technologies

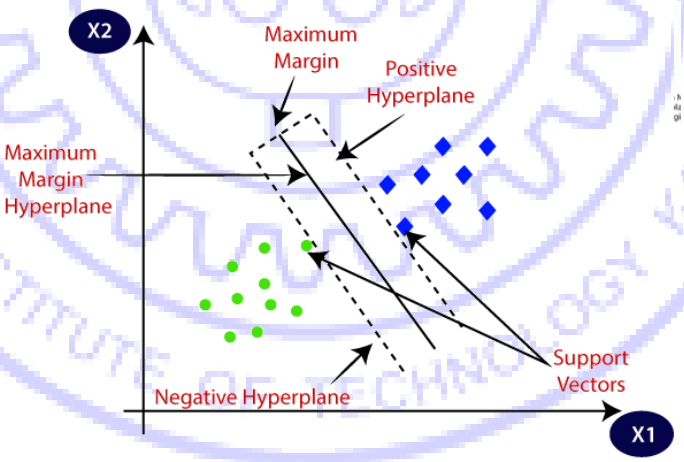
EE902 Advanced ML Techniques for Wireless Technology

- **SVM:**

$$\max \frac{2}{\|\bar{\mathbf{a}}\|_2} \equiv \min \|\bar{\mathbf{a}}\|$$

$$\mathcal{C}_0 : \bar{\mathbf{a}}^T \bar{\mathbf{x}}_i + b \geq 1,\ 1 \leq i \leq M$$

$$\mathcal{C}_1 : \bar{\mathbf{a}}^T \bar{\mathbf{x}}_i + b \leq -1,\ M+1 \leq i \leq 2M$$



- We introduce a new variable y_i

$$\mathcal{C}_0 : y_i = 1,\ 1 \leq i \leq M$$

$$\mathcal{C}_1 : y_i = -1,\ M+1 \leq i \leq 2M$$

- SVM

$$\max \frac{2}{\|\bar{\mathbf{a}}\|_2} \equiv \min \|\bar{\mathbf{a}}\|$$

$$\mathcal{C}_0: y_i(\bar{\mathbf{a}}^T\bar{\mathbf{x}}_i + b) \geq 1,\ 1 \leq i \leq M$$
$$\mathcal{C}_1: y_i(\bar{\mathbf{a}}^T\bar{\mathbf{x}}_i + b) \geq 1,\ M+1 \leq i \leq 2M$$

- Therefore, the equivalent problem is

$$\min \|\bar{\mathbf{a}}\|$$
$$y_i(\bar{\mathbf{a}}^T\bar{\mathbf{x}}_i + b) \geq 1$$
$$\min \frac{1}{2}\|\bar{\mathbf{a}}\|^2$$
$$-(y_i(\bar{\mathbf{a}}^T\bar{\mathbf{x}}_i + b) - 1) \leq 0$$

- What is the next step?

$$\frac{1}{2}\|\bar{\mathbf{a}}\|^2 - \sum_{i=1}^{2M} \lambda_i(y_i(\bar{\mathbf{a}}^T\bar{\mathbf{x}}_i + b) - 1)$$

- Setting gradient with respect to $\bar{\mathbf{a}}$ and set it equal to 0.

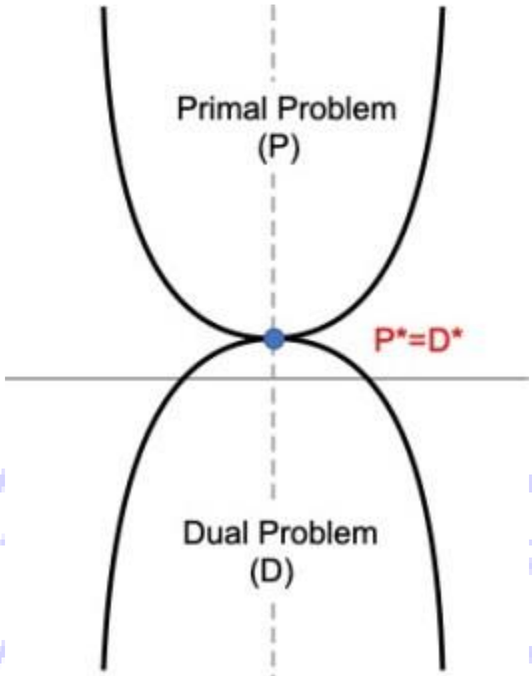
$$\bar{\mathbf{a}} = \sum_{i=1}^{2M} \lambda_i y_i \bar{\mathbf{x}}_i$$

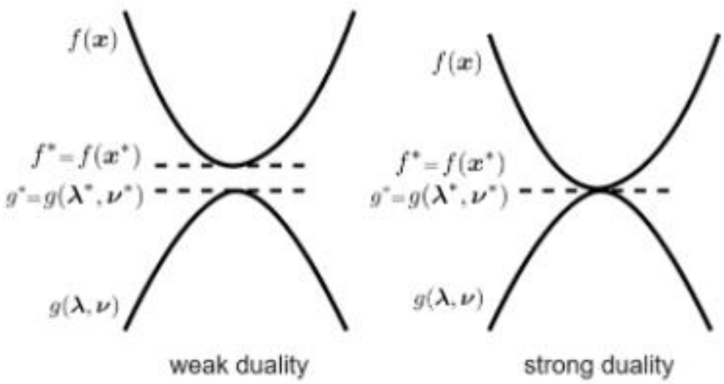
- **Complementary slackness**.
- The support vectors are those points for which $\lambda_i \neq 0$.
- **Dual Problem:**

$$\max \sum_{i=1}^{2M} \lambda_i - \frac{1}{2}\sum_{i,j=1}^{2M} \lambda_i\lambda_j y_i y_j \bar{\mathbf{x}}_i^T\bar{\mathbf{x}}_j$$

$$\text{subject to } \lambda_i \geq 0$$

$$\sum_{i=1}^{2M} \lambda_i y_i = 0$$





- What is interesting?
- **Dual Problem:**

$$\max \sum_{i=1}^{2M} \lambda_i - \frac{1}{2} \sum_{i,j=1}^{2M} \lambda_i \lambda_j y_i y_j \langle \bar{\mathbf{x}}_i, \bar{\mathbf{x}}_j \rangle$$

$$\text{subject to } \lambda_i \geq 0$$

$$\sum_{i=1}^{2M} \lambda_i y_i = 0$$

- The inner product can be replaced by a suitable **Kernel** which can be **non-linear** in nature.

$$\max \sum_{i=1}^{2M} \lambda_i - \frac{1}{2}\sum_{i,j=1}^{2M} \lambda_i \lambda_j y_i y_j K(\bar{\mathbf{x}}_i, \bar{\mathbf{x}}_j)$$

$$\text{subject to } \lambda_i \geq 0$$

$$\sum_{i=1}^{2M} \lambda_i y_i = 0$$

- This helps us consider **non-linear features**!!

$$K(\bar{\mathbf{x}}_i, \bar{\mathbf{x}}_j) = \phi^T(\bar{\mathbf{x}}_i)\phi(\bar{\mathbf{x}}_j) = (\bar{\mathbf{x}}_i^T \bar{\mathbf{x}}_j)^2$$

$$\phi(\bar{\mathbf{x}}_i) = \begin{bmatrix} x_i(1)x_i(1) \\ x_i(1)x_i(2) \\ x_i(1)x_i(3) \\ x_i(2)x_i(1) \\ x_i(2)x_i(2) \\ x_i(2)x_i(3) \\ x_i(3)x_i(1) \\ x_i(3)x_i(2) \\ x_i(3)x_i(3) \end{bmatrix} = \bar{\mathbf{x}}_i \otimes \bar{\mathbf{x}}_i$$

- Popular kernel: **Gaussian Kernel**

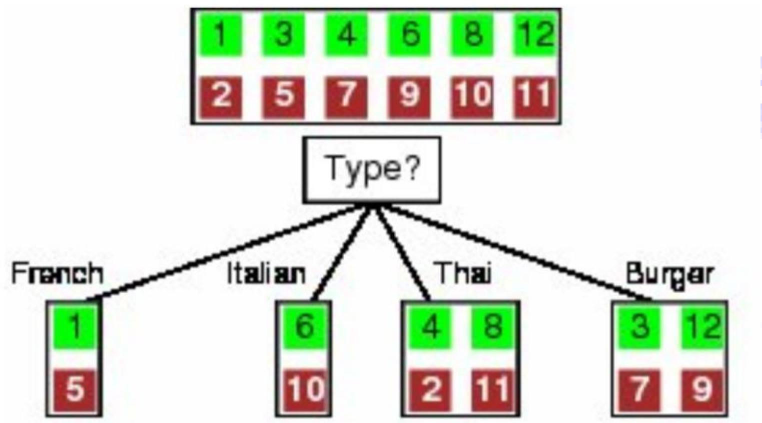
$$K(\bar{\mathbf{x}}_i, \bar{\mathbf{x}}_j) = \exp\left(-\frac{\|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j\|^2}{2\sigma^2}\right)$$

- This can be written as the inner product between two **infinite dimensional** feature vectors!!

Example	Input Attributes										Goal
	Alt	*Bar*	*Fri*	*Hun*	*Pat*	*Price*	*Rain*	*Res*	*Type*	*Est*	*WillWait*
$\mathbf{x}_1$	Yes	No	No	Yes	Some	$$$	No	Yes	French	0–10	$y_1 =$ Yes
$\mathbf{x}_2$	Yes	No	No	Yes	Full	$	No	No	Thai	30–60	$y_2 =$ No
$\mathbf{x}_3$	No	Yes	No	No	Some	$	No	No	Burger	0–10	$y_3 =$ Yes
$\mathbf{x}_4$	Yes	No	Yes	Yes	Full	$	Yes	No	Thai	10–30	$y_4 =$ Yes
$\mathbf{x}_5$	Yes	No	Yes	No	Full	$$$	No	Yes	French	>60	$y_5 =$ No
$\mathbf{x}_6$	No	Yes	No	Yes	Some	$$	Yes	Yes	Italian	0–10	$y_6 =$ Yes
$\mathbf{x}_7$	No	Yes	No	No	None	$	Yes	No	Burger	0–10	$y_7 =$ No
$\mathbf{x}_8$	No	No	No	Yes	Some	$$	Yes	Yes	Thai	0–10	$y_8 =$ Yes
$\mathbf{x}_9$	No	Yes	Yes	No	Full	$	Yes	No	Burger	>60	$y_9 =$ No
$\mathbf{x}_{10}$	Yes	Yes	Yes	Yes	Full	$$$	No	Yes	Italian	10–30	$y_{10} =$ No
$\mathbf{x}_{11}$	No	No	No	No	None	$	No	No	Thai	0–10	$y_{11} =$ No
$\mathbf{x}_{12}$	Yes	Yes	Yes	Yes	Full	$	No	No	Burger	30–60	$y_{12} =$ Yes

- Choose the feature that maximizes the **information gain**.

$$H(X) - H(X|Y)$$



- What is $H(X)$?

$$H(X) = H\left(\frac{1}{2}, \frac{1}{2}\right) = 1$$

$$H(X|French) = 1$$

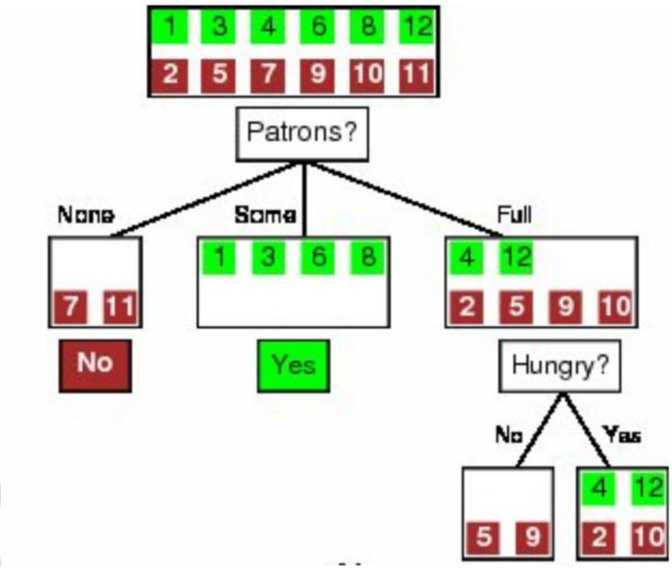
$$H(X|Italian) = 1$$

$$H(X|Thai) = 1$$

$$H(X|Burger) = 1$$

$$H(X|Type) = \frac{1}{6} \times 1 + \frac{1}{6} \times 1 + \frac{1}{3} \times 1 + \frac{1}{3} \times 1 = 1$$

$$H(X) - H(X|TYPE) = 0$$



- Assignment #7 Deadline: 1st March Friday 11:59 PM.
- Live interaction #8: 3rd March Sunday 2:00 – 3:00 PM.
- Assignment #8 Deadline: 7th March Thursday 11:59 PM.
- Assignment #7, 8 Discussion: 8th March Friday 8:00 PM – 8:30 PM.
- Quiz #4: 8th March Friday 9:00 - 9:45 PM.
- Final Exam: 10th March Sunday 9:00 AM – 12:00 PM. (Please check!!)