ML- Assignment - 2

Divyam Anshumaan 2017147

Date:
Page No. 1

Q3. Yes. Using the kernel trick, we transform the feature space of XOR:

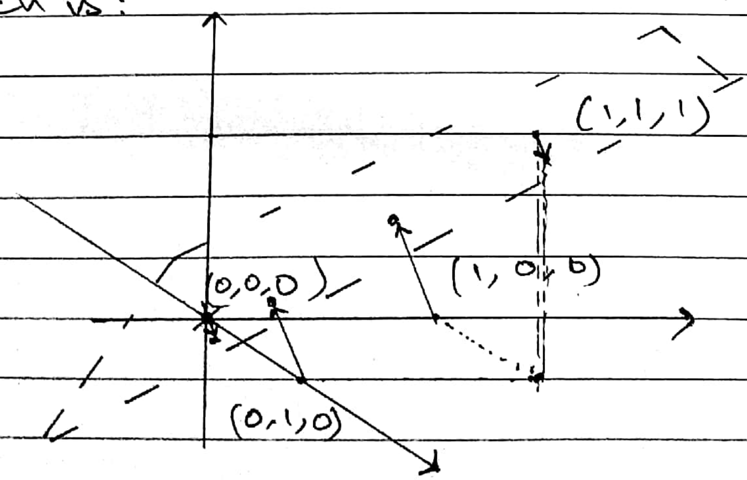
a	b	out
0	0	0
1	0	1
0	1	1
1	1	0

, $x^{(i)} = \begin{bmatrix} a \\ b \end{bmatrix}$

Using $\phi : \mathbb{R}^2 \to \mathbb{R}^3$ where $\phi\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) \to \begin{bmatrix} a \\ b \\ ab \end{bmatrix}$

From this we obtain: $(0,0) \to (0,0,0)$
$(0,1) \to (0,1,0)$
$(1,0) \to \cancel{(0,1,}(1,0,0)$
$(1,1) \to (1,1,1)$

which is:



and ~~ca~~ ~~x~~ is linearly separable in this particular feature space, using the decision boundary:

$w^T x + b = 0$, where $w = \begin{bmatrix} \\ \\ \\ \end{bmatrix}$ and $b = -0.25$

Given $K(x, x') = \langle \phi(x), \phi(x') \rangle$

$= (1 + x^T x)^2$

$= \cancel{1 + x_1^2 + x_2^2} \quad (1 + x_1 x_1' + x_2 x_2')^2$

$= 1 + (x_1 x_1')^2 + (x_2 x_2')^2 + 2 x_1 x_1' + 2 x_1 x_2 x_1' x_2' + 2 x_2 x_2'$

$= 1 \cdot 1 + (x_1^2) \cdot (x_1')^2 + (x_2^2) \cdot (x_2')^2 + (\sqrt{2} x_1)(\sqrt{2} x_1') + (\sqrt{2} x_1 x_2)(\sqrt{2} x_1' x_2') + (\sqrt{2} x_2)(\sqrt{2} x_2')$

$$= \cancel{\begin{bmatrix} \\ \\ \\ \end{bmatrix}} = \left(\begin{bmatrix} 1 \\ x_1^2 \\ x_2^2 \\ \sqrt{2} x_1 \\ \sqrt{2} x_1 x_2 \\ \sqrt{2} x_2 \end{bmatrix} \right)^T \left(\begin{bmatrix} 1 \\ x_1'^2 \\ x_2'^2 \\ \sqrt{2} x_1' \\ \sqrt{2} x_1' x_2' \\ \sqrt{2} x_2' \end{bmatrix} \right)$$

$= \langle \phi(x), \phi(x') \rangle$.

So, $\phi : \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ x_1^2 \\ x_2^2 \\ \sqrt{2}x_1 \\ \sqrt{2}x_2 \\ \sqrt{2}x_1x_2 \end{bmatrix}$

Q5. we start by calculating the weights ~~...~~ $w \in \mathbb{R}^2$ using

$$w = \sum_{i=1}^{m} \alpha_i y^{(i)} x^{(i)}$$

where the expression is obtained from the minimizing condition of the dual of the lagrangian, ie

$$\mathcal{L}(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum \alpha_i (y^{(i)}(w^T x^{(i)} + b) - 1)$$

to get $\nabla_w \mathcal{L}(\cdot) = 0$, and m = no. of training examples.

So, $w = \begin{bmatrix} 2.008 \\ 3.872 \end{bmatrix}$

- Now we know from the KKT conditions, $\alpha_i > 0$ strictly for support vectors. So

x_1, x_4, x_7, x_9 are support vectors.

- ~~We also know that for support vectors.~~

~~$y^{(i)}(w^T x^{(i)} + b) = 1$~~

~~taking (2.5, 1) $\quad 1(2.5 \times 2.008 + 3.872 \times 1 + b) = 1$~~

$= -9.$

- We also know that $b^* = -\left(\dfrac{\max_{i: y^{(i)} = -1} w^T x^{(i)} + \min_{i: y^{(i)} = +1} w^T x^{(i)}}{2}\right)$

and we have $= \dfrac{\max_{i: y^{(i)} = -1} w^T x^{(i)} + \min_{i: y^{(i)} = +1} w^T (x^{(i)})}{2}$

~~$= -\left(\dfrac{19.2608 + 12.1472}{2}\right)$~~

~~$= 15.0764 \quad = 17.32$~~

- ~~Classifying (3, 3)?~~

$= -\left(\dfrac{19.2608 \quad 12.1472}{\cancel{12.147} + \cancel{8.892}}\right)$

$= -15.704.$

- Classifying (3, 3) $\quad w^T x + b = 17.64 - 15.704 > 0$

$\Rightarrow y^{(i)} = +1.$

Scanned by CamScanner

Q6. Simplifying the expression $K(u, v)$

$$= \exp(-\gamma \cdot \|u - v\|^2) = \exp(2uv - u^2 - v^2)$$

where γ is taken as 1 and u, v are given ~~to~~ $\in \mathbb{R}$.

Also, $K(u, v) = \langle \phi(u), \phi(v) \rangle$

Now using the taylor series expansion,

$\exp(2uv - u^2 - v^2) =$

$$\exp(-u^2 - v^2) \cdot \Big(1 + 2uv + 2u^2v^2 + \frac{4u^3v^3}{3} + \frac{2u^4v^4}{3} + \dots$$

$$\dots + \frac{8u^7v^7}{315} + \dots + \frac{4u^9v^9}{2835} + \frac{4u^{10}v^{10}}{14175} + \dots\Big) \quad \text{— (1)}$$

(using $f(x) = f(0) + x \cdot f'(0) + \frac{x^2}{2!} f''(0) + \dots + \frac{x^n}{n!} f^n(0) \dots$)

(- the taylor series)

~~[illegible]~~ So, coefficient of the n^{th} term is given by $\frac{f^n(0)}{n!}$

Continuing from (1):

$= \cancel{e^{-u^2}e^{-}}$

$$= e^{-u^2}(1) \cdot e^{-v^2}(1) + \left(e^{-u^2}\sqrt{2}\,u\right)\left(e^{-v^2}\sqrt{2}\,v\right) + \left(e^{-u^2}\sqrt{2}\,u^2\right)\left(e^{-v^2}\sqrt{2}\,v^2\right) + \left(e^{-u^2} \cdot \frac{2}{\sqrt{3}} \cdot u^3\right)\left(e^{-v^2} \cdot \frac{2}{\sqrt{3}} \cdot v^3\right) + \cdots$$

$$\cdots + \left(e^{-u^2} \cdot \frac{2}{\sqrt{14175}} \cdot u^{10}\right)\left(e^{-v^2} \cdot \frac{2}{\sqrt{14175}} \cdot v^{10}\right) \cdots$$

→ Basically, we split the coefficients of each term in the expansion as $\sqrt{c} \cdot \sqrt{c}$ and conviniently we have powers of (uv) in each term which can also be split as $u^n \cdot v^n$. (from the taylor series)

→ Further more, $e^{-u^2-v^2} = e^{-u^2} \cdot e^{-v^2}$, which is multiplied by every term in the expansion and we already know from the taylor series that each term can be split into powers of $u^n \cdot v^n$.

→ Hence every term looks like :

$$\left(e^{-u^2} \cdot \sqrt{c} \cdot u^n\right)\left(e^{-v^2}\sqrt{c} \cdot v^n\right)$$

where c is taylor series coefficient.

this can be expressed as :

$$\left(\begin{bmatrix} e^{-v^2}\sqrt{c_1} \\ e^{-v^2} v\sqrt{c_2} \\ e^{-v^2} v^2\sqrt{c_3} \\ \vdots \\ e^{-v^2} v^n\sqrt{c_{n+1}} \\ \vdots \\ \vdots \end{bmatrix}\right)^T \left(\begin{bmatrix} e^{-v^2}\sqrt{c_1} \\ e^{-v^2} v\sqrt{c_2} \\ e^{-v^2} v^2\sqrt{c_3} \\ \vdots \\ e^{-v^2} \cdot v^n\sqrt{c_{n+1}} \\ \vdots \\ \vdots \end{bmatrix}\right)$$

So, $\phi : \begin{bmatrix} v \end{bmatrix} \rightarrow \begin{bmatrix} e^{-v^2}\sqrt{c_1} \\ e^{-v^2} v\sqrt{c_2} \\ \vdots \\ e^{-v^2} \cdot v^n\sqrt{c_{n+1}} \\ \vdots \\ \vdots \end{bmatrix}$

→ where c_i's are coefficients of terms of degree $i-1$ in the taylor series expansion of $\exp(2uv)$

- As there is factorial $(n!)$ in the denominator for terms of degree n, coefficients of higher order terms ≈ 0.

→ As a side note, $n!$ becomes larger than e^n @ $n \geq 10^6$