MLPA W-1

1) $w = \begin{bmatrix}1\\2\\3\end{bmatrix}$ $b = -4$ $x = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$ $w^T x + b = 0$

Sol^n :- ~~x_1~~ $x_1 + 2x_2 + 3x_3 = -4$

$\Rightarrow$ $\left[\begin{array}{ccc|c} \underset{P}{1} & \underbrace{2 \quad 3}_{F.V} & & -4 \end{array}\right]$

$\Rightarrow$ $x_1 = -2x_2 - 3x_3 + 4$

$x_2 = x_2$

$x_3 = x_3$

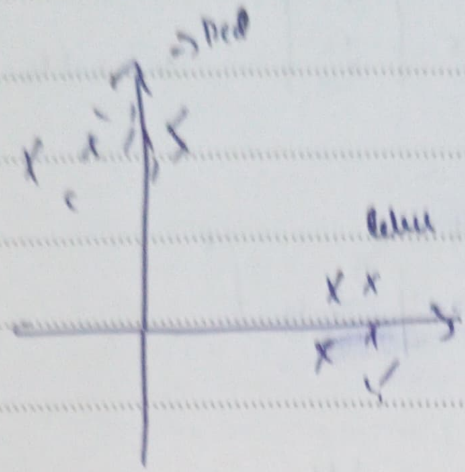
$$\Rightarrow X = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = x_2\begin{bmatrix}-2\\1\\0\end{bmatrix} + x_3\begin{bmatrix}-3\\0\\1\end{bmatrix} + \begin{bmatrix}4\\0\\0\end{bmatrix}$$

$\boxed{\text{Scaler projection} = \dfrac{v_1^T w}{\|w\|}}$ =

Date: / /

<u>W-2</u>

Hyper plane



→ Find w & b

→ Find hyperplane that maximally separates the red (−1) and blue (+1)

Find w & b s.t $w^T x^{(i)} + b \geq 1$ for positive sample

$w^T x^{(i)} + b \leq -1$ for negative

If $y^{(i)}$ is sample +1 or −1

$$y^{(i)} (w^T x^{(i)} + b) \geq 1$$

$$\frac{w^T x^{(i)} + b}{\|w\|}$$

Maximize the minimum among $\frac{|w^T x^{(i)} + b|}{\|w\|}$

s.t $y^{(i)} \left(w^T x^{(i)} + b \right) \geq 1$

Find w and b maximise $\frac{1}{\|w\|}$

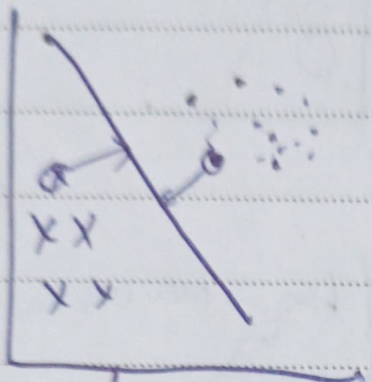
S.t $y^{(i)} (w^T x^{(i)} + b) \geq 1$

$\frac{||w||^2}{2}$

Find w & b s.t minimize $\frac{||w||^2}{2}$

s.t $y^{(i)}(w^T x^{(i)} + b) \geq 1$

The hyperplanes out of solving this optimization for hard margin SVM



maximum margin hyper pl.

→ $w^T x + b > 0$ +ve

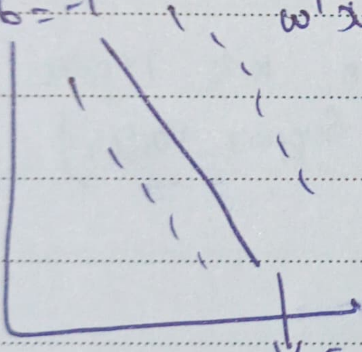
↓ w lies on this side

↓

$w^T x + b < 0$ -ve

w lies on this side

$w^T + b = -1$, $w^T x + b = 1$



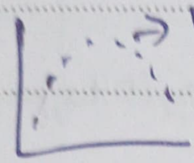
↓ $w^T x + b = 0$

Find hyper plan $\frac{2}{||w||}$

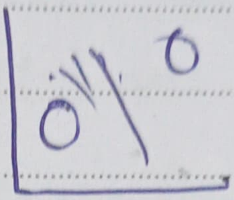
Date: / /

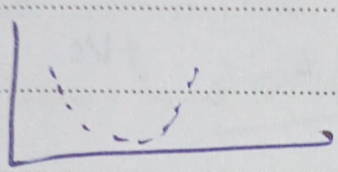
$$\max = \frac{2}{\|w\|}$$

$$\min \frac{\|w\|}{2} \longrightarrow \min \frac{\|w\|^2}{2}$$

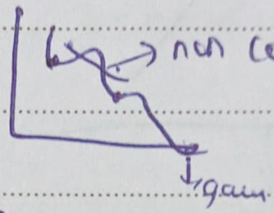


→ margin is most max

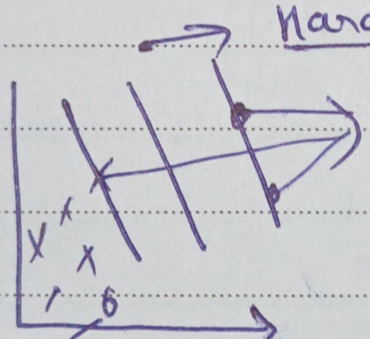




min



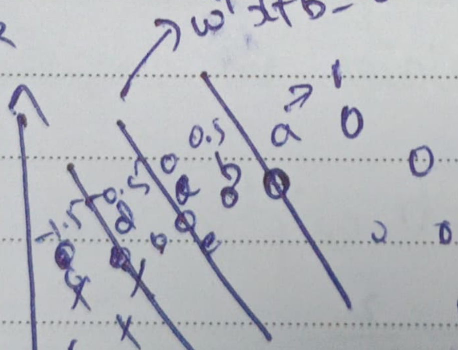
→ Hard margin (Equal agi separate irodu)



→ Support vectors

(margin mele irodu Support vector)

slack



Soft margin

$C = 0.0001$

$C = 1000.0$ (C less idea

(low training error) leads to — regularisation overfit

$$\min = \frac{\|w\|^2}{2} + C \sum_{n=1}^{n} \xi_n$$

Regularisation.

loss%

classification error ↓

margin ↑

C ↑ R ↓

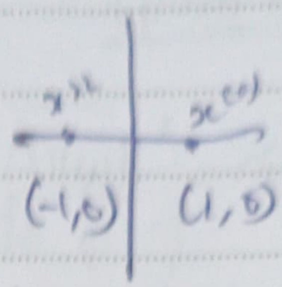
C ↓ R ↑

Date: / /

W-3

$x^{(1)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad y^{(1)} = 1$

$x^{(2)} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \quad y^{(2)} = -1$



linear seperable

Solve the dual problem for $\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$ and then calculate the solution for the primal problem as

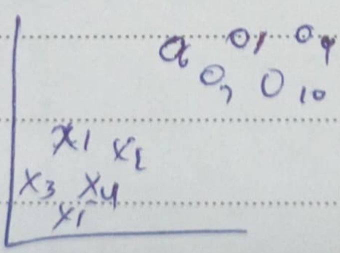
$$w = \alpha_1 y^{(1)} x^{(1)} + \alpha_2 y^{(2)} x^{(2)}$$

Primal problem :-

$$\min_{w,b} \frac{\|w\|^2}{2}$$

Dual problem

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \underbrace{x_i x_j}_{\text{Dot product}}$$



$x_1 \cdot x_2$

$x_1 \cdot x_3$

$\vdots$

$x_1 \quad o_{10}$

Date: / /

Polynomial kernal

$$K(x_1, x_2) = (x_i^T x_j + c)^d$$

$a \quad b$

$\begin{bmatrix}1\\0\end{bmatrix} \begin{bmatrix}0\\1\end{bmatrix}$

$\begin{bmatrix}1\\1\end{bmatrix} \begin{bmatrix}3\\4\end{bmatrix}$

$\begin{bmatrix}-2\\-4\end{bmatrix}$

$\begin{bmatrix}-3\\4\end{bmatrix}$

RBF (Radias Basis function)

$$K_{RBF}(x_i, x_j) = \exp\left(\frac{-\|x_1 - x_2\|^2}{2\sigma^2}\right)$$

$$= e\left(\frac{\text{distance}}{\sigma^2}\right)$$

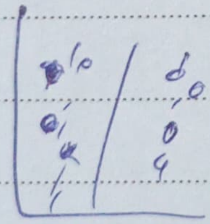
more distance — less similarity

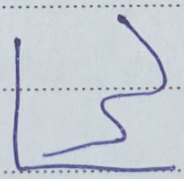
less — " — more — " —

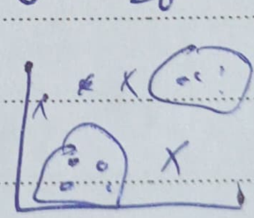
$\alpha \geq 0$ | Support vech ≥ 0

Non support v $\alpha = 0$

Linear	Polynomial	RBF
$x_i x_j$	$(x_i x_j + c)^d$	$e\left(\frac{-\|x_i - x_j\|^2}{2\sigma}\right)$







Date: / /

Primal way

advnt

i) larger dataset

Cons

Linear boundar

Dual way

adv

1) Non linear ~~Decis~~ DB

Decision boundary

2) Own custom kernel

Cons

1) Performance will be affected when you larger data

Date: / /

Sessional Review

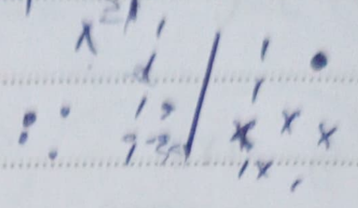
$\frac{2}{\|\omega\|}$ $\frac{1}{2}$ $\|\omega\|^2$

1) problem Wedio-1

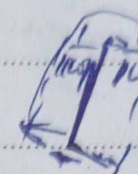
ξ slack

Hard margin $= \frac{|\omega^T x^{(i)} + b|}{\|\omega\|}$

$\frac{1}{2}\|\omega\|^2 + C \sum_{i=1}^{n} \xi^{(i)}$



$C \to 0$ low bias $\quad C^{-1}$ low variance



$\xi = 0 \Rightarrow$ right on margin $\quad \frac{2}{\|\omega\|}$

$0 < \xi < 1 \Rightarrow$ on the correct side but within margin

$\xi \geq 1 \Rightarrow$ on the incorrect side

Hinge loss

$$L_i = \sum_{j \neq y} \max\left(0, z_j^{(i)} - z_{y^{(i)}}^{(i)} + 1\right)$$

$z_j^{(i)} = (\omega x^{(i)} + b)_j$, $\quad z_{y^{(i)}}^{(i)} = (\omega x^{(i)} + b)_{y^{(i)}}$

$$= \sum_{j \neq y} \max\left(0, [\omega x^{(i)} + b]_j - [\omega x^{(i)} + b]_{y^{(i)}} + 1\right)$$

Date: / /

Primal samples = 1000

→ based on samples

→ always 1 constrains

$\min \frac{\|w\|^2}{2}$

$\max \alpha^T 1 - \frac{1}{2} \alpha^T G \alpha$

s.t $y^{(i)}(w^T x^{(i)} + b) \geq 1$

s.t $\sum_i \alpha_i y^{(i)} = 0$

$G = n \times n$

1000×1000

$\alpha \times 1000$ $= 1000000$

Date: / /

Week-4

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)} \quad \text{---} (1)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{---} (2)$$

$P(C_k/x) \Rightarrow k =$ lemon pear fan

$P(C_{lemon}/x)$ $P(C_{lem}/x)$ $P(C_{fan}/x)$

$$P(C_k/X) = \frac{P(X/C_k)\,P(C_k)}{P(X)}$$

$P(X/C_k)$

$P(C_1/X) = \frac{P(X/C_1)\,P(C_1)}{P(X)}$

$P(C_2/X)$

$\vdots$

$$\boxed{P(C_k/X) \propto P(X/C_k)\,P(C_k)}$$

Naive problem

$$P(C_k/X_i) = \prod_{i=0}^{n} (x_i - C_k)$$

Date: / /

Vedio-5

- Prior will always constant

$$P(Y / x_1, x_2, x_3 \ldots x_n) = \overset{\text{prior}}{\boxed{P(y)}} \times \{P(y/x_1) \ldots + P(y/x_n)\}$$

Naive bias on Text

i) Food is delicious - +ve
ii) Food is bad - -ve
iii) The food is awesome - +ve

Text 1	1
Text 2	0
Text -	3.

Pre prop.
- Stemming - Change Changes Change..
- Stop words - is ..

Bag of words

Food	Delicious	Bad	Awesome,	Review,
1	1	0	0	1
1	0	1	0	0,
1	0	0	1	1

} Rev.

$$P(1 / \text{Food}=1, \text{Delicious}=1, \text{Bad}=0, \text{Awesome}=0)$$

1/2 0/2 1/2

Date: 2/2

$x^{(1)}$ Food is hotel was good

$$P(y|x^{(1)}) = P(y|\text{Food}=1) \times P(y|\text{hotel}=1, \ldots$$

Multi-Naive bias

$$P(y_{(i)}) = \frac{N_y + \alpha}{N_{y(i)} + \alpha n}$$

Underfitting and overfitting

$\alpha \uparrow \rightarrow$ underfitting

$\alpha = 0 \rightarrow$ high variance overfitting

$\alpha \downarrow$	$\alpha \uparrow$
→ High variance	→ High bias
→ overfitting	→ underfitting

Date: / /

$x^{(1)}$ / Food is hotel was good

$P(y/x^{(1)}) = P(y \mid Food = 1) \times P(y / hotel = 1) \ldots$

<u>Multi-Naive bias</u>

$$P_{y(i)} = \frac{N_y + \alpha}{N_{y(i)} + \alpha n}$$

<u>Underfitting and overfitting</u>

$\alpha \uparrow \rightarrow$ underfitting

$\alpha = 0 \rightarrow$ high Variance overfitting

$\alpha \downarrow$	$\alpha \uparrow$
→ High Varian → overfitting	→ High bias → underfitting

i) Gaussian:

$$P(x_i/y) = \frac{1}{\sqrt{2\pi\sigma_\pi^2}} \exp\left(-\frac{(x_i-\mu_y)^2}{2\sigma_\pi^2}\right)$$

σ_y & μ_y maximum likelihood.

ii)

Multinomial

$$\phi_{y(i)} = \frac{N_{y(i)} + \alpha}{N_y + \alpha n}$$

iii) Complement Naive Bias

iv) Bernoulli Naive Bayes

$$P(x_i/y) = P(x_i = 1/y)x_i + (1 - P(x_i = 1/y))(1 - x_i)$$

v) out-of core Naive Bayes model

/ / fitting

Peace madi data madodu

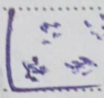
U-6

1) Clustering Algorithm ⇒ data fit madaban
2) Generative Algorith [Extra data add madodu
3) Discrimative Algorithm → Separate madodu

less data bnd mate data madabaudu

LLM → generative model

K-means



- Random Initilize centroid
- a) Calculate label for all points (while (converge)
  b) Calculat mean of new label

1) Random Init - does not generate global soln
2) K = ? K is unknown
3) Linear boundaries
4) circle
5) Not Probalistic

for avoid this we will use gaussian mixture model

GMM :-

Vedio-1

Review

Check

$$= P(c_k) \prod_{i=1}^{n} p(x_i \mid c_k)$$

Confident ($\rightarrow$ yes)	Studied ($\rightarrow$ yes)	Sick ($\rightarrow$ NO)	Result
Y	N	N	Fail
Y	N	Y	Pass
N	Y	Y	Pass
N	Y	N	Pass
Y	Y	Y	Fail

$P(\text{pass}) = 3/5$ $\quad$ $P(\text{fail}) = 2/5$

$P(\text{confident yes} / \text{pass}) = 1/3$ $\quad$ $P(\text{Yes} / \text{fail}) = 2/2$

$P(\text{studied yes} / \text{Pass}) = 2/3$ $\quad$ $P(y / \text{fail})\ 1/2$

$P(\text{sick yes} / \text{pass}) = 1/3$ $\quad$ $P(\text{Yes} / \text{fail}) = 1/2$

$P(\text{pass}) = 1/3 \times 2/3 \times 1/3 = \text{---}$

$P(\text{fail}) = 2/2 \times 1/2 \times 1/2 = \text{---}$

Date:

~~categorical~~ Numerical → For numerical Gaussian

5.1
5.2
5.3
5.2

Gaussian =

$$P(x / y_{(i)}) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}\right)$$

Multinomial :- Laplace smoothing

$$\phi_{ij} = \frac{N_{yi} + \alpha}{N + n\alpha}$$

probability 0 barabar du anta α cu madthive

por suppor

S
R
W

C → kelidru full zero agute.

$\alpha = 1$

$n = 2$

0 agodil

understanding NB

Types of NB → G & mult.

Laplaces

review.

Vedio - 2

GMM :- Gaussian mixture model

Overlap

Swalpa - Sperial →

little - diag → → some sperical/some circl

half - tied → → some sperical/some than

3/4 - full → → All that.

Covariance types

param:-

- n - component
- Covariance types
- tollerance
- EM; maximum iterati.

Full :- - Each component having there own general covariance matrix

tied :- - All components having same gener. cv.

dig :- - Each components own diagonal

Spherica - Each Component own single variance.

How to handle

K-means   Centroid

1) Randomly initilize
2) a) assign lable
   b) Calculate centers } keep doing again & again

K

GMM :- Gaussian estimators.

1) Randomly initilize [mean of these points]
2) a) Find the lables for the.
   b) Calculate new centers.

Mean