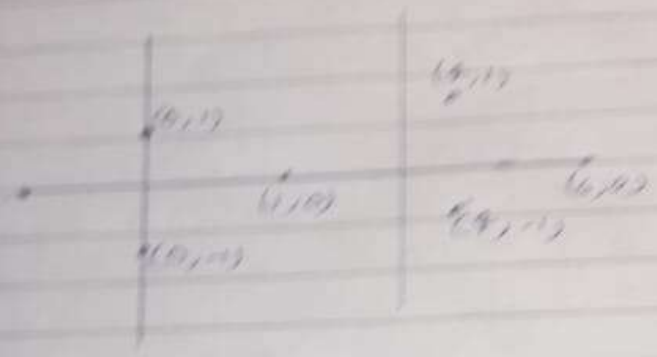
Sol 6

Points $(4,1)$, $(4,-1)$ & $(6,0)$ belong to positive class. Points $(1,0)$, $(0,1)$ and $(0,-1)$ belong to negative class. Draw an optimal hyperplane.



$s_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $s_2 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$ $s_3 = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$

$\bar{s}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\bar{s}_2 = \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}$ $\bar{s}_3 = \begin{pmatrix} 4 \\ -1 \\ 1 \end{pmatrix}$

$\alpha_1 \bar{s}_1 \bar{s}_1 + \alpha_2 \bar{s}_2 \bar{s}_1 + \alpha_3 \bar{s}_3 \bar{s}_1 = 0$
$\alpha_1 \bar{s}_1 \bar{s}_2 + \alpha_2 \bar{s}_2 \bar{s}_2 + \alpha_3 \bar{s}_3 \bar{s}_2 = 0$
$\alpha_1 \bar{s}_1 \bar{s}_3 + \alpha_2 \bar{s}_2 \bar{s}_3 + \alpha_3 \bar{s}_3 \bar{s}_3 = 0$

On solving, $\alpha_1 = -\frac{22}{9}$, $\alpha_2 = \frac{7}{18}$, $\alpha_3 = \frac{7}{18}$

$w = \sum \alpha_i \bar{s}_i$

$= \begin{pmatrix} 2/3 \\ 0 \\ -5/3 \end{pmatrix}$

$\Sigma = 0 \quad \frac{2}{3} x_1 + 0 x_2 + \frac{-5}{3} = 0$

$2x_1 = 5$

$x_1 = 2.5$

boundary at $x_1 = 2.5$

For Iris, CSV dataset, what is accuracy of classifier using linear and RBF kernel? which of the two has better performance?

Linear kernel = 100%

RBF kernel = 100%

Generally RBF performs better but since we had small dataset, we get equal performance.

RBF is more suitable as iris dataset is non linear.

For "letter recognition, CSV" dataset, present & interpret the confusion matrix. Are there any specific letters frequently confused with others? How does the SVM model perform compared to Iris dataset.

The confusion matrix shows how many times each letter was predicted

Rows => actual

Columns => predicted

diagonal => correct predictions