CSCE 633: Machine Learning Homework-3

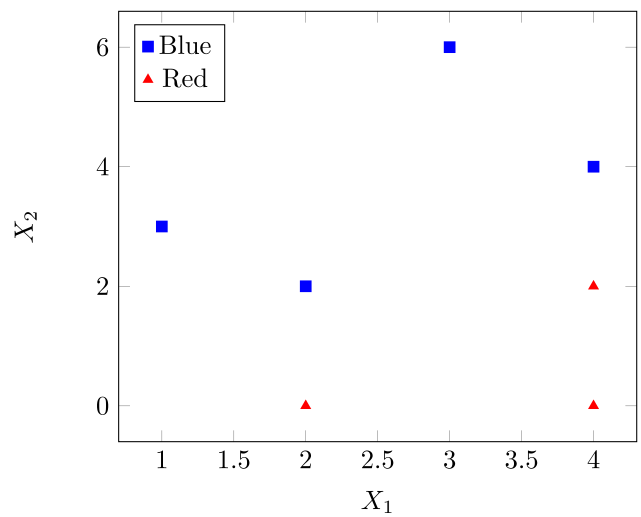
Arunachalam Venkatachalam

29th March 2024

1 Hyperplane Fitting

Index	X1	X2	Y
1	3	6	Blue
2	2	2	Blue
3	4	4	Blue
4	1	3	Blue
5	2	0	Red
6	4	2	Red
7	4	0	Red

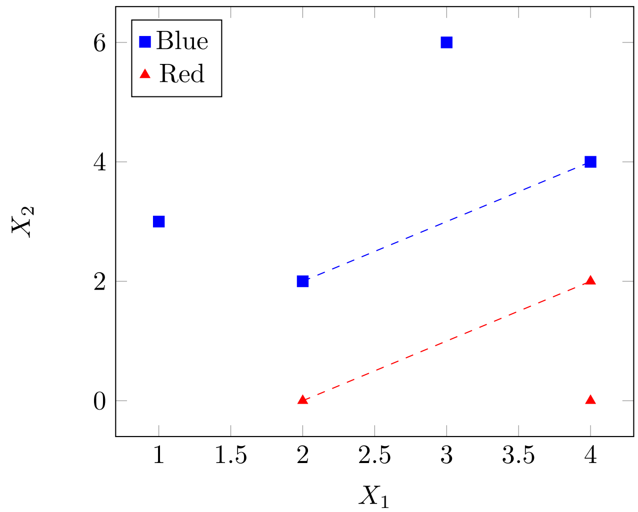
Plotting the data on a 2D plane



1. Support vectors:

   - Blue: $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$

- Red: $\mathbf{v}_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$, $\mathbf{v}_4 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$



I'm essentially trying to find a plane/line such that it maximizes the margin, which is the minimum perpendicular distance from the hyperplane to the support vectors.

Assuming equation of the separating line is

$$Ax + By + C$$

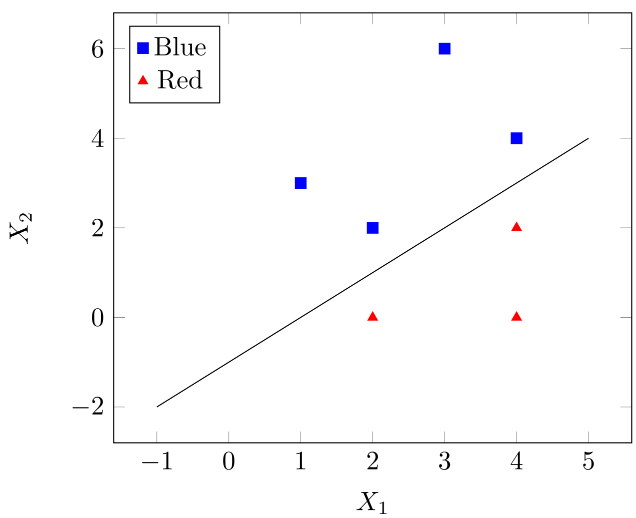
Perpendicular distance of a point from a line is

$$d = \frac{|Ax_1 + By_1 + C|}{\sqrt{A^2 + B^2}}$$

Following the above method, I end with a hyperplane/line equation:

$$X_1 - X_2 = 1$$

Here is a scatter plot with the separating line:


Blue
Red
6
4
2
0
−2
X_2
−1
0
1
2
3
4
5
X_1

2. If

$$-1 + X_1 - X_2 >= 0$$

classify as 1(RED) or if

$$-1 + X_1 - X_2 < 0$$

then classify it as -1 (BLUE).

Betas for RED

$$\beta_0, \beta_1, \beta_2 = -1, 1, -1$$

Betas for BLUE

$$\beta_0, \beta_1, \beta_2 = -1, 1, -1$$

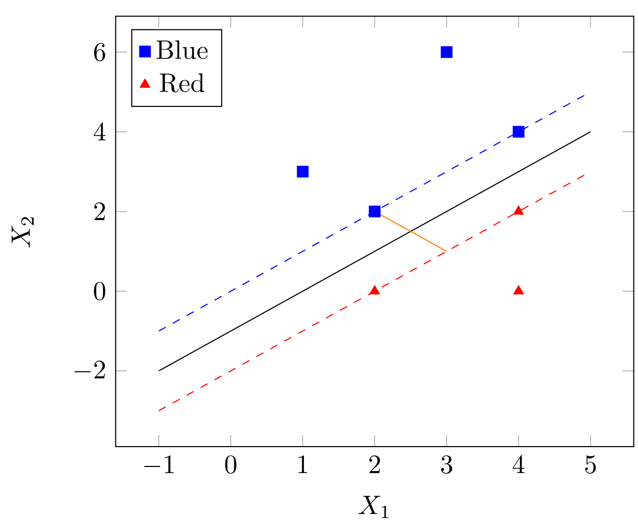
3. Maximal Margin Line

We could use the below equation to find the point at which perpendicular from (2,2) would cut the line X2 = X1 - 2

$$k = \frac{(y_2 - y_1)(x_3 - x_1) - (x_2 - x_1)(y_3 - y_1)}{(y_2 - y_1)^2 + (x_2 - x_1)^2}$$

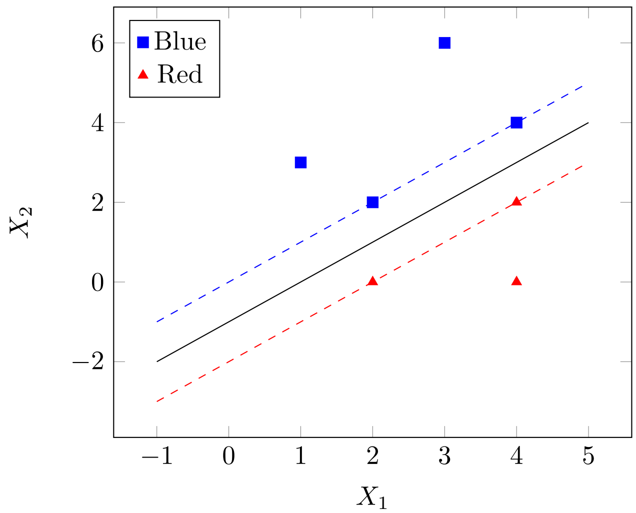
$$x_4 = x_3 - k(y_2 - y_1)$$

$$y_4 = y_3 + k(x_2 - x_1)$$



The orange line represents the maximal margin.

4. Support Vectors for the maximal margin hyperplane denoted as red and blue dashed lines.

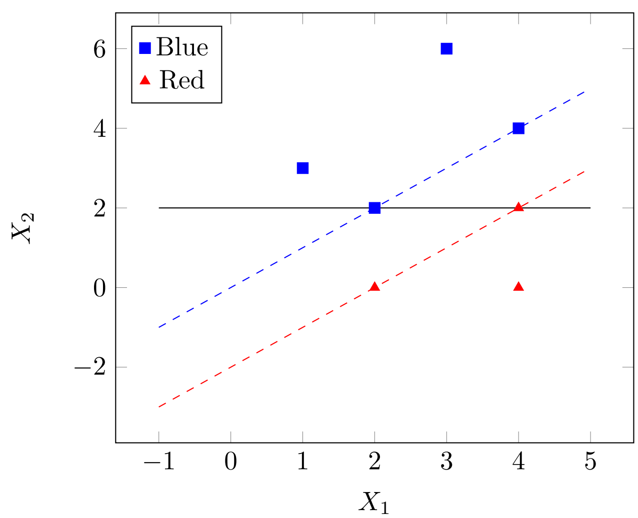


5. Slight movement of the seventh observation

   Slight movement of the seventh observation (4,0) does not affect the alignment of the hyperplane unless it becomes a support vector, meaning unless it moves closer to the hyperplane no changes would be there.

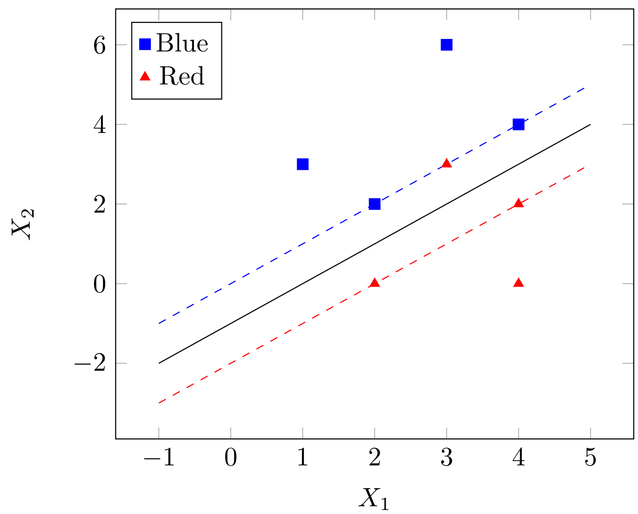
6. Alternative hyperplane that is not the optimal separating hyperplane

   Equation of alternative hyperplane that is not the optimal separating hyperplane is X2 = 2


Blue
Red
6
4
2
0
−2
−1
0
1
2
3
4
5
X_1
X_2

7. Additional observation on the plot so that the two classes are no longer separable by a hyperplane

   Adding a RED observation with the coordinates (3,3) would make the dataset not linearly separable.



2 SVM

(a) The table representing this training set is:

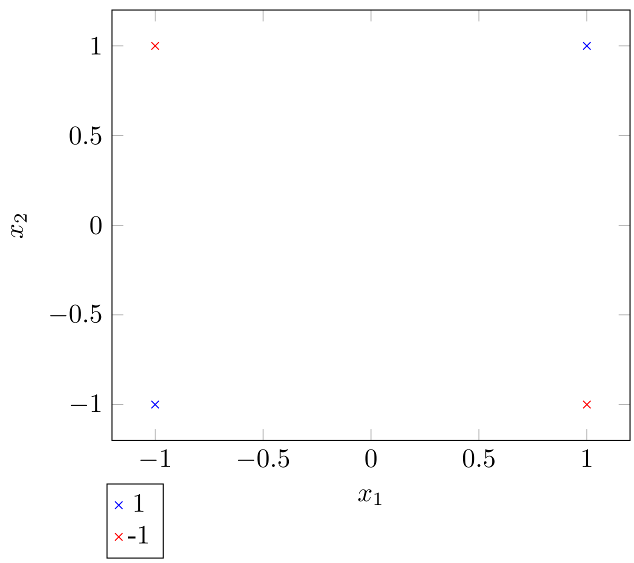
x_1	x_2	y
1	1	1
-1	-1	1
1	-1	-1
-1	1	-1

If I assume positive values to be 1 and negative values to be 0 because a truth table consists only of binary data, the table becomes

x_1	x_2	y
1	1	1
0	0	1
1	0	0
0	1	0

The shape of X is $(4, 2)$ and the shape of y is $(4, 1)$.
(a.bonus) The logic gate representing this truth table is XOR gate.

(b) These points are not linearly separable because no straight line can separate the positive samples from the negative samples.



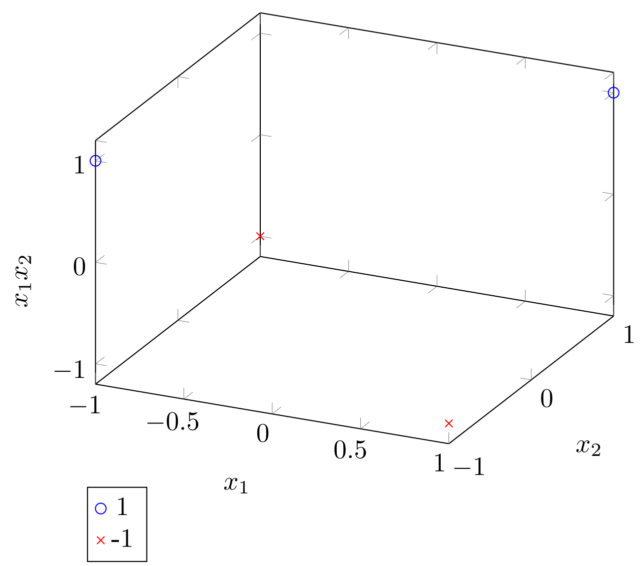
(c) The feature transformation function $\phi(x) = [x1, x2, x1 * x2]$. Applying this transformation to each data point:

$$\begin{aligned}
\phi(1, 1) &= [1, 1, 1] \\
\phi(-1, -1) &= [-1, -1, 1] \\
\phi(1, -1) &= [1, -1, -1] \\
\phi(-1, 1) &= [-1, 1, -1]
\end{aligned}$$

The transformed points are:

x_1	x_2	x_1x_2	y
1	1	1	1
−1	−1	1	1
1	−1	−1	−1
−1	1	−1	−1

Now, these transformed points are linearly separable. A plane can be drawn to separate the points.



x_1x_2
1
0
−1
−1
−0.5
0
0.5
1
x_1
−1
0
1
x_2
1
-1

(d) The margin size after the transformation is $\frac{2}{\sqrt{3}} \approx 1.155$. The support vectors are the points that lie on the margin, which are the transformed points $(1, 1, 1)$ and $(-1, -1, 1)$.