Introduction to Machine Learning
Homework 5: Support Vector Machine

Prof. Linda Sellie

1. Consider the following training data,

x_1	x_2	y
0	0	-1
2	2	-1
2	0	1
1	1.5	-1
3	0.5	1

   (a) Plot the training data, and the hyperplane given by: $\mathbf{w} = [12, -32]$ and $w_0 = -5$

   (b) Are the points linearly separable?

   (c) For $\mathbf{w}$ and w_0 given above:

      i. compute the *functional* margin for each training example, and show the functional margin with respect to the set of training examples

      ii. compute the *geometric* margin for each training example, and show the geometric margin with respect to set of training examples

      iii. compute the canonical weights with respect to the training examples

   (d) Identify which of the training examples are support vectors

   (e) If we add the point $x = (1, 3)$ and $y = -1$ to the training data, does the margin change? Does separating hyperplane change? Do the support vectors change?

   (f) If we remove the point $(1, 1.5)$ does the margin change? Does the separating hyperplane change?

   (g) If we remove the point $(0, 0)$ does the margin change? Does the separating hyperplane change?

   (h) Specify a *constrained optimization* to find a hyperplane that separates the training examples above where the separating hyperplane has the largest possible margin[1]

[1] You do not need to solve your constrained optimization problem.

2. Consider the following training data,

x_1	y
1	-1
2	1
3	1
4	-1

Let $\sigma = 1$. Without performing feature scaling (to make the numbers easier to compute), create new features using the Gaussian Radial Basis function.

If you ran the hard margin SVM algorithm on these data points you would receive coefficients:

Using those coefficients, what would you predict for these new points:

x_1
1.2
2.5
3.25

3. Would the following constrained optimizations create the same decision boundary? Justify your answer.

`max` γ
Subject to:
$-1(5.1w_1 + 3.5w_2 + 1.4w_3 + 0.2w_4 + w_0) \geq \gamma$
$-1(4.9w_1 + 3.0w_2 + 1.4w_3 + 0.2w_4 + w_0) \geq \gamma$
$-1(4.7w_1 + 3.2w_2 + 1.3w_3 + 0.2w_4 + w_0) \geq \gamma$
$-1(4.6w_1 + 3.1w_2 + 1.5w_3 + 0.2w_4 + w_0) \geq \gamma$
$1(7.0w_1 + 3.2w_2 + 4.7w_3 + 1.4w_4 + w_0) \geq \gamma$
$1(16.4w_1 + 3.2w_2 + 4.5w_3 + 1.5w_4 + w_0) \geq \gamma$
$1(6.9w_1 + 3.1w_2 + 4.9w_3 + 1.5w_4 + w_0) \geq \gamma$
$1(.5w_1 + 2.3w_2 + 4.0w_3 + 1.3w_4 + w_0) \geq \gamma$
$||\mathbf{w}||_2 = 1$

`min` $||\mathbf{w}||_2^2$
Subject to:
$-1(5.1w_1 + 3.5w_2 + 1.4w_3 + 0.2w_4 + w_0) \geq 1$
$-1(4.9w_1 + 3.0w_2 + 1.4w_3 + 0.2w_4 + w_0) \geq 1$
$-1(4.7w_1 + 3.2w_2 + 1.3w_3 + 0.2w_4 + w_0) \geq 1$

$$-1(4.6w_1 + 3.1w_2 + 1.5w_3 + 0.2w_4 + w_0) \geq 1$$
$$1(7.0w_1 + 3.2w_2 + 4.7w_3 + 1.4w_4 + w_0) \geq 1$$
$$1(16.4w_1 + 3.2w_2 + 4.5w_3 + 1.5w_4 + w_0) \geq 1$$
$$1(6.9w_1 + 3.1w_2 + 4.9w_3 + 1.5w_4 + w_0) \geq 1$$
$$1(.5w_1 + 2.3w_2 + 4.0w_3 + 1.3w_4 + w_0) \geq 1$$