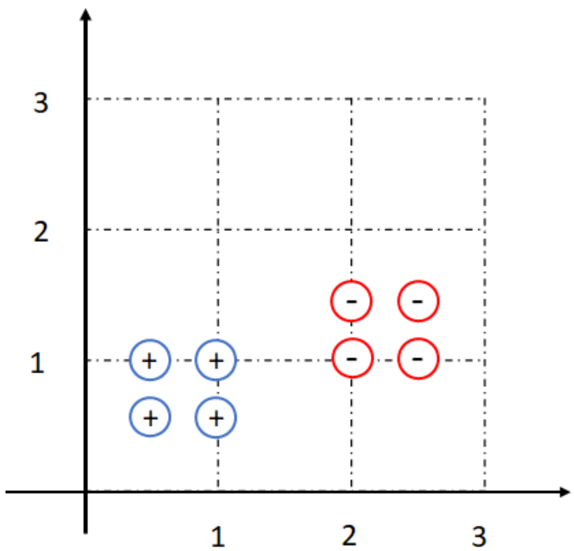
AI534 — Written Homework Assignment 3 (40 pts) — Due Nov 12th, 2021

1. Apply linear SVM without soft margin to the following problem.



   a. (2pts) Please mark out the support vectors, the decision boundary ($\mathbf{w}^T\mathbf{x}+b=0$) and $\mathbf{w}^T\mathbf{x}+b=1$ and $\mathbf{w}^T\mathbf{x}+b=-1$. Note that you don't need to solve the optimization problem for this, just eyeballing the solution should be sufficient.
   Your answer here.

   b. (6 pts) Please solve for $\mathbf{w}$ and b based on the support vectors and decision boundary you identified in (a).
   Your answer here.

2. L_2 SVM
   Given a set of training examples $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$, where $y_i \in \{1, -1\}$ for all i. The following is the primal formulation of L_2 SVM, a variant of the standard SVM obtained by squaring the hinge loss:

$$\min_{\mathbf{w},b,\xi} \; \mathbf{w}^T\mathbf{w} + \lambda \sum_{i=1}^{N} \xi_i^2$$
$$\text{s.t. } y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \xi_i, \; i \in \{1, \cdots, N\}$$
$$\xi_i \geq 0, \; i \in \{1, \cdots, N\}$$

   a. (5pts) Show that removing the second constraint $\xi_i \geq 0$ will not change the solution to the problem. In other words, let $(\mathbf{w}^*, b^*, \xi^*)$ be the optimal solution to the problem without this set of constraints, show that $\xi_i^* \geq 0$, $\forall i \in \{1, \cdots, N\}$. (Hint: use proof by contradiction.)
   Your answer here.

   b. (2 pts) After removing the second set of constraints, we have a simpler problem with only one set of constraints. Now provide the lagrangian of this new problem.
   Your answer here.

   c. (5pts) Derive the dual of this problem. How is it different from the standard SVM with hinge loss? Which formulation is more sensitive to outliers?
   Your answer here.

3. (Naive Bayes Classifier) Consider the following training set:

A	B	C	Y
0	1	1	0
1	1	1	0
0	0	0	0
1	1	0	1
0	1	0	1
1	0	1	1

(a) (5 pts) Learn a Naive Bayes classifier by estimating all necessary probabilities (there should be 7 probabilities in total).

Your answer here.

(b) (5 pts) Compute the probability $P(y = 1|A = 1, B = 0, C = 0)$.

Your answer here.

(c) (2 pts) Suppose we know that A, B and C are independent random variables, can we say that the Naive Bayes assumption is valid? (Note that the particular data set is irrelevant for this question). If your answer is yes, please explain why; if you answer is no please give an counter example.

Your answer here.

4. (Naive Bayes learns linear decision boundary.) (8 pts) Consider a naive Bayes binary classifier with a set of binary features $x_1, x_2, ..., x_d$. Show that the Naive Bayes classifier learns a linear decision boundary $w_0 + w_1x_1 + w_2x_2 + ... + w_dx_d = 0$. Express the weights using the Naive Bayes parameters. Hint: consider the decision rule of predicting $y = 1$ if $P(y = 1|\mathbf{x}) > P(y = 1|\mathbf{x})$. This is equivalent to having a decision boundary defined by $\log \frac{P(y=1|\mathbf{x})}{P(y=0|\mathbf{x})} = 0$.

Your answer here.