50.021 Artificial Intelligence

Quiz 1

Student Name: **Student ID:**

[Question1]. True or False (Give one sentence explanation) 6 points

a Use a evaluation/testing dataset T_n disjoint from the training data D_n.
Solution: True. Those sets have to be disjoint to test the performance of the algorithm.

b When the hypothesis space is richer, over fitting is more likely.
Solution: True. Too many dimensions will cause over-fitting because it might introduce more meaningless regularity in the data.

c Lets consider a linear classifier which has the form $f(x) = \omega x + b$. The decision boundary are all those vectors x which are parallel to the vector ω.
Solution: False. Decision boundary is composed of vectors which are perpendicular to ω.

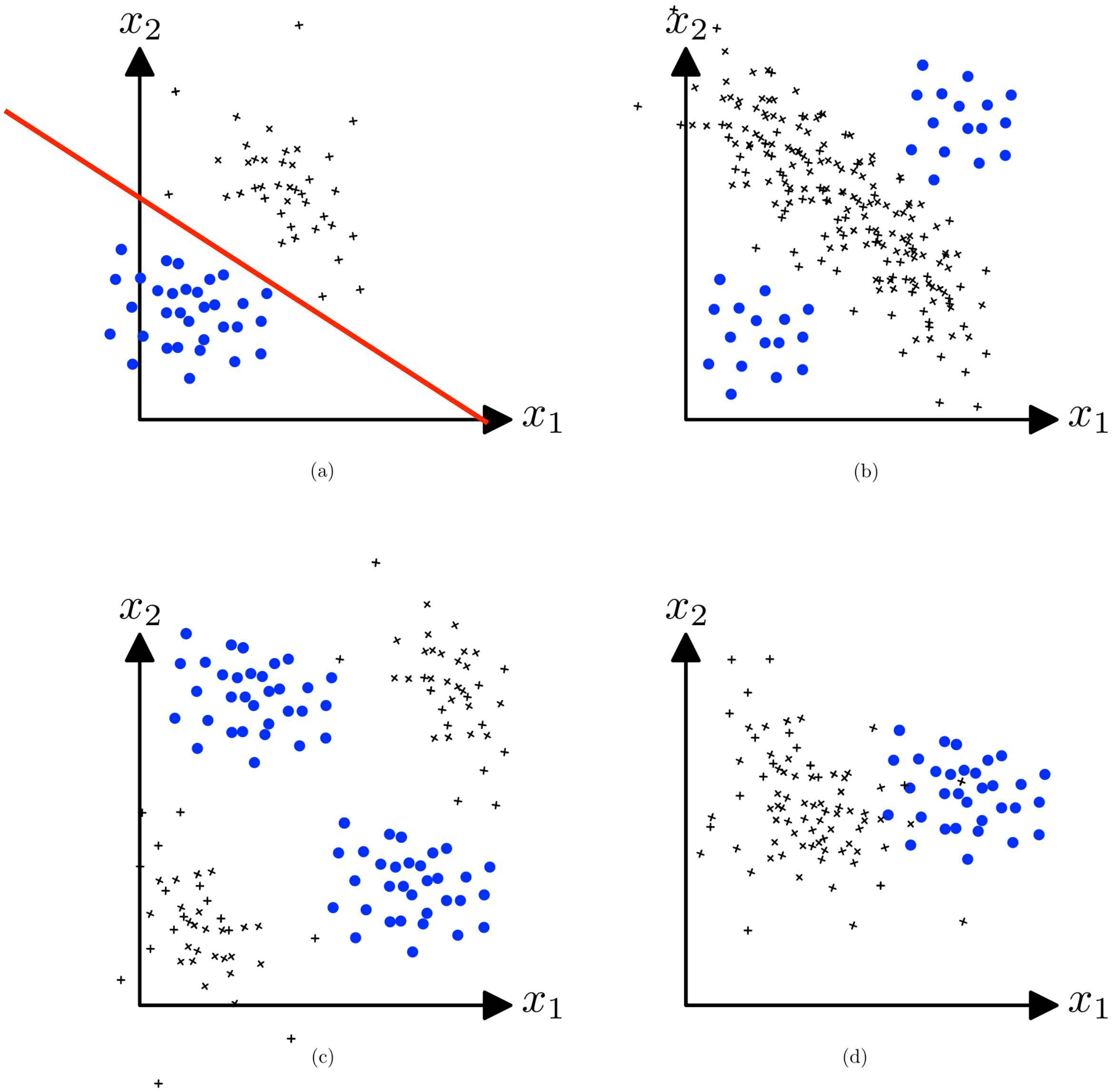
d Loss function can measure a loss $L(f(x), y)$ for how wrong our prediction f (x) is relative to its ground truth y. For example, regression function: $L(f(x), y) = (f(x) - y)^3$ is a good loss function for regression
Solution: False. $L(f(x), y) = (f(x) - y)^3$ is a bad loss function for regression because it might cancel out negative and positive errors.

e The long term goal for the prediction is to get a model to have high loss L on new, unseen data.
Solution: False. The goal is always to have minimum loss (thats why its called "loss", we'd like to always minimize that).

f Lets consider a linear classifier which has the form $f(x) = \omega x + b$. We know that for $b = 0$, the decision boundary are all those vectors x which are orthogonal to vector w. The bias b shifts the decision boundary corresponding to $\omega x + b = 0$ parallel to the direction of w. Positive $b > 0$ shift the hyperplane to the positive side and vice versa.
Solution: False. It should be the other way around. Positive b shift the hyperplane to the negative side and vice versa.

[Question2]. Each of the plots in figure 1 show training samples in $\mathbb{R}^2$, with two classes : x and o. For each of the four plots, answer the following question,

1. Can we run linear regression with this data? - answer please for all 4 plots (1 points)
   Solution: Yes, to all 4. Linear regression can be done on **any** data, even if its a classification and not regression problem (just that its not going to give meaningful results).
2. Is the data linearly separable? (yes or no will suffice) - answer please for all 4 plots (2 points)
   textbfSolution: Only plot (a) is linearly separable. Plot (b), (c), and (d) are not.
3. If yes, sketch the decision boundary on the plots below (approximate drawing is fine)- answer please for all 4 plots (2 points).
   Solution: Sketch any line that completely separates the two data in figure (a).



(a) (b) (c) (d)

Figure 1: Training sample plots