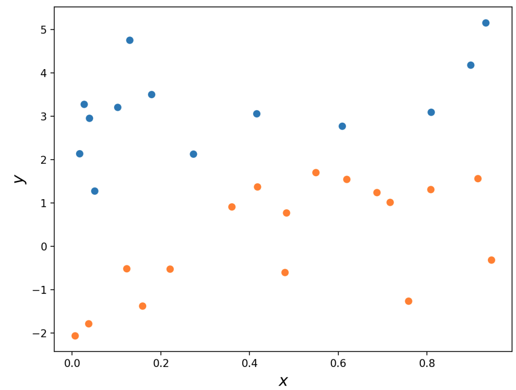
Quantitative Problems

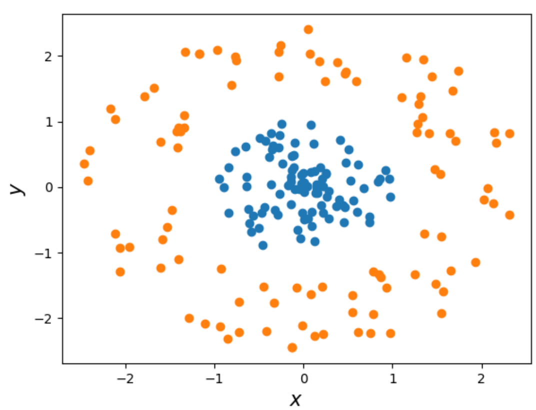
1. Let $X \times Y$ be the data set shown in Figure 1 and assume that $Y = \{-1, 1\}$ where the label -1 corresponds to the orange points and 1 corresponds to the blue points.



Data Set

   (a) Draw a good decision linear boundary for this data set.
   (b) Define the linear classifier that you drew in part (a).
   (c) What would be a better classifier? How does this classifier differ from the classifier from parts (a),(b)?

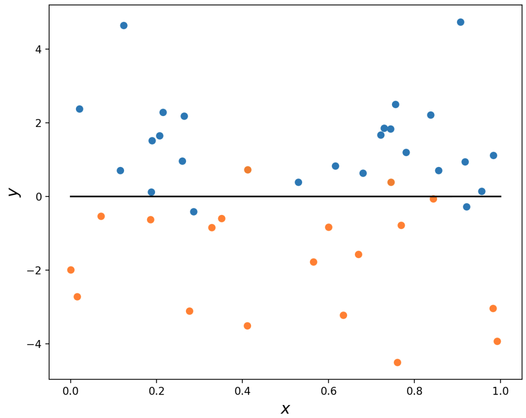
2. Define a nonlinear classifier for the following data set and draw the corresponding decision boundary.



Data Set

3. Let $X \times Y = \{(x_1, y_1) \ldots, (x_n, y_n)\}$ be the data set shown in Figure 3 with the sample space being $X = \mathbb{R}$ and label space $Y = \{-1, 1\}$. Define the function $f : X \to Y$ be the classifier defined by

$$f(x) = \begin{cases} 1 & y > 0 \\ -1 & y \leq 0 \end{cases}$$



Data Set

(a) Let $\mathcal{L}_1 : Y \times Y \to \mathbb{R}^+$ be the loss function

$$\mathcal{L}_1(X) = \sum_i [f(x_i) \neq y_i]$$

and compute the loss corresponding to the classifier f.

(b) Let $\mathcal{L}_2 : Y \times Y \to \mathbb{R}^+$ be the loss function

$$\mathcal{L}_2(X) = \sum_i [f(x_i) \cdot y_i < 0]$$

and compute the loss corresponding to the classifier f.

(c) Is there any diffrence between these losses?

4. (Challenge) Let X be the set of linear separators. Then the set X is said to *shatter* n points if there exists a classifier $f \in X$ that separates any labelling of a collection of n points. What is the largest number of points that X can shatter?

Conceptual Problems

1. Are there any tasks that you don't think that a computer learn how to do and why?