COM S 474/574
Introduction to Machine Learning

Spring 2022

Homework 8

Yuichi Hamamoto

1 Directions:

- Due: Thursday April 14, 2022 at 9pm. Late submissions will be accepted for 24 hours with a 15% penalty. (the enforcement is strict, beginning at 9:01pm, except for extreme situations; having a poor wifi connection or minor computer problems is not sufficient for the penalty to be waived.)
- Upload the homework to Canvas as a single pdf file.
- If the graders cannot easily read your submission (writing is illegible, image is too dark,

  or if the contrast is too low) then you might receive a zero or only partial credit.

- Any non-administrative questions must be asked in office hours or (if a brief response is sufficient) Piazza.

2 Problems

Problem 1. [15 points; 3 each] Consider the following data set. There are two classes for Y (mapped to $\{-1, +1\}$). There is one feature X that can be used for prediction.

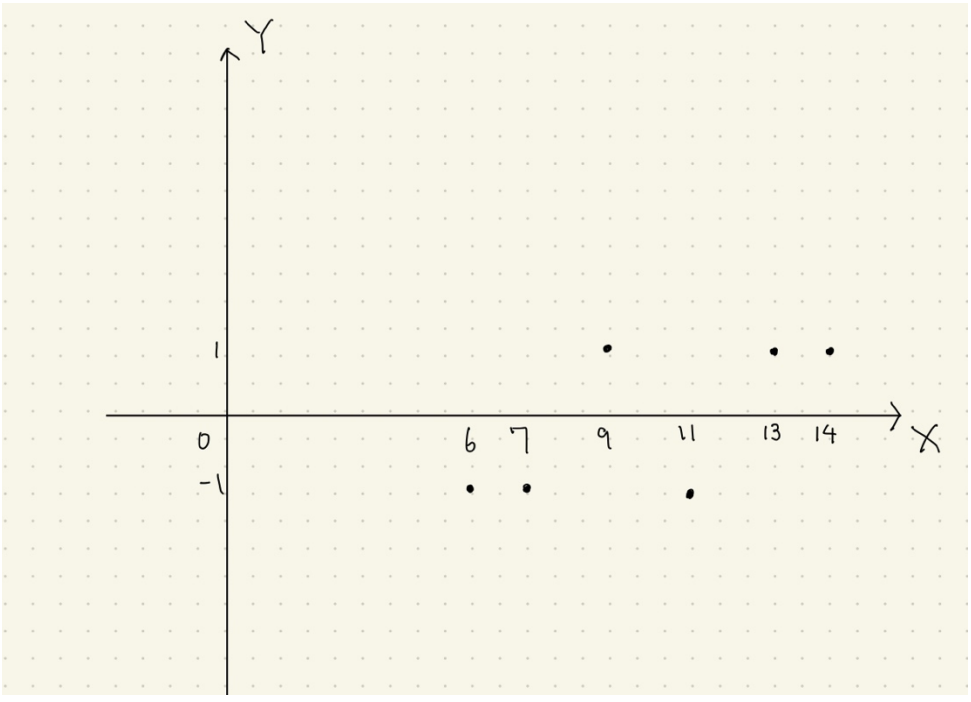
Y	X
-1	6
-1	7
-1	11
+1	9
+1	13
+1	14

a) Draw a scatter plot of the data (by hand).

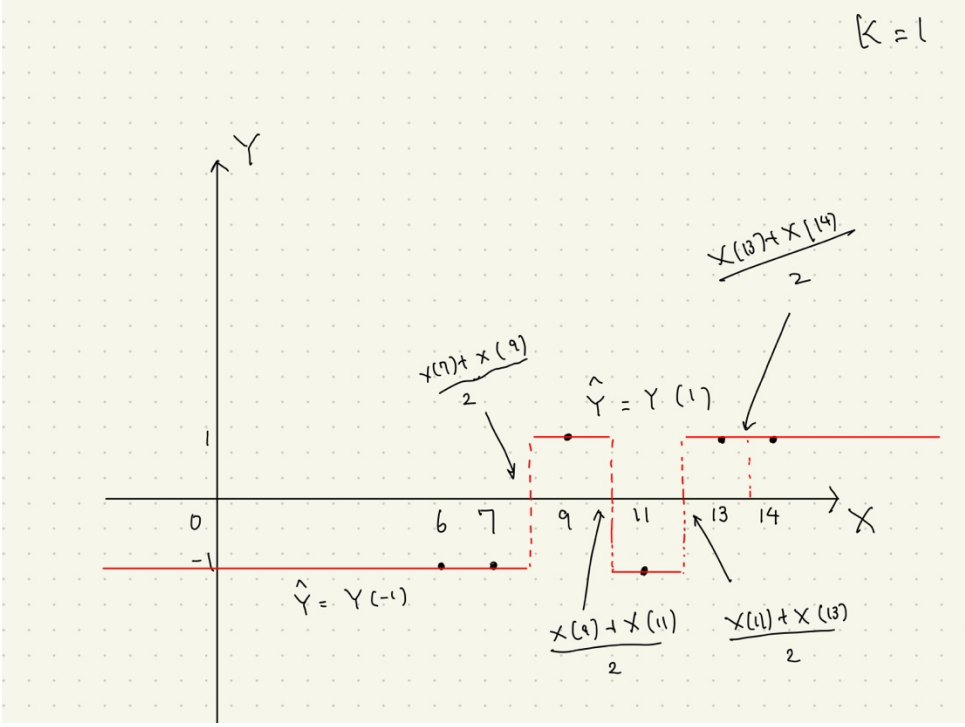
- Horizontal axis for X
- Vertical axis for Y
- The marker size(s), color(s), shape(s) do not matter as long as the scatter plot is clear.

In the following, you will draw prediction functions Y(x) for different classifiers.
• You may draw them all on a single scatter plot as long as your drawing is clear. • You may also draw them on separate scatter plots.

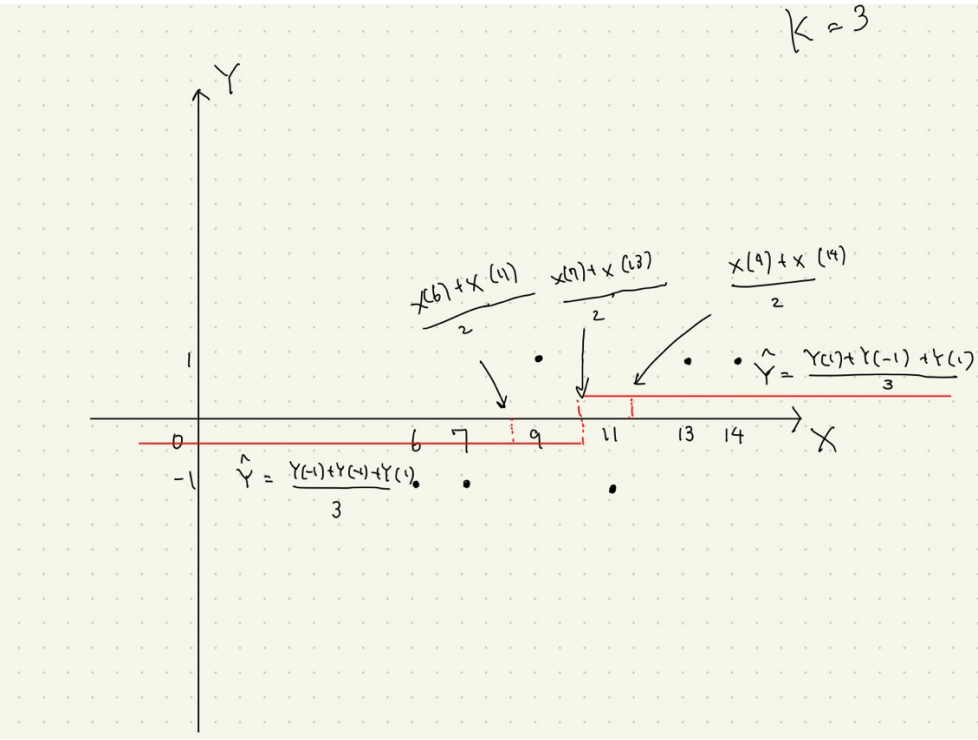
• If different classifiers have the same prediction functions Y(x), it is ok to explain that and only draw the function once.



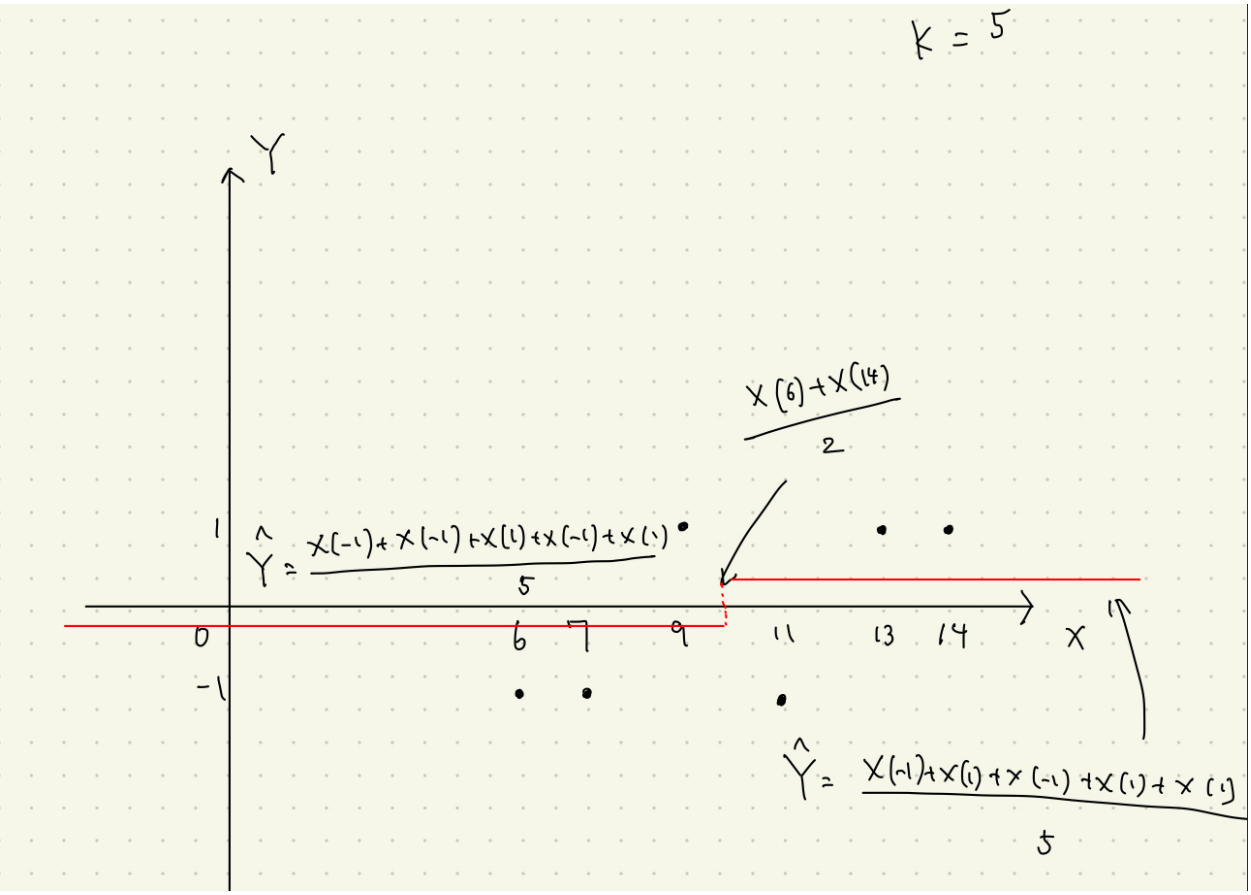
b) Sketch the prediction function Y(x) for the k = 1 nearest neighbor classifier on the scatter plot. Briefly explain your process.



c) Sketch the prediction function Y(x) for the k = 3 nearest neighbor classifier on the scatter plot. Briefly explain your process.



d) Sketch the prediction function Y(x) for the k = 5 nearest neighbor classifier on the scatter plot. Briefly explain your process.



e) Describe what would happen if you tried to use the max margin classifier for this data set.

The decision hypnosis would be Y = 0.

Problem 2. [20 points; 4 each A-E] For this problem you will fit classifiers to the same data sets you used in the last homework.

(a). Change the classifiers used to the following nearest neighbor classifiers:

- 1-NN
- 10-NN
- 10-NN with distance-based weights (include argument `weights='distance'`)
- Radius based neighbor classifiers for radii 3, 4, and 6 (uniform weights)
- Radius based neighbor classifiers for radii 3, 4, and 6 with distance-based weights

(b). Make the following changes (in addition to those you needed to do in HW 7):

- To make sense of distances, it will help to set

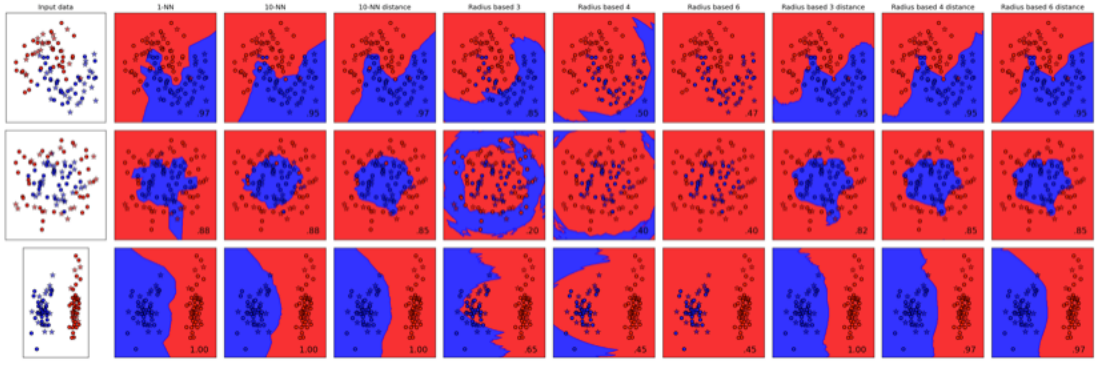
  ```
  ax.set_aspect('equal')
  ```

- Modify the code used to make the background color just be the prediction. Do so by modifying the code setting the value for `Z` (used to color the background) to be the prediction (0 or 1) and the color map to have two values:

  ```
  cm = ListedColormap(["#FF0000", "#0000FF"])
  Z = clf.predict(np.c_[xx.ravel(), yy.ravel()])
  ```

With those changes, run the script, examine the plots, and answer the following questions.

A. Include a copy of your plot.



B. Among the k-nearest neighbor classifiers, do the decision regions seem largely similar or quite different? If quite different, briefly describe the differences.

They all look very similar to each other although there is small difference among them. It is also hard to say which one is better from the samples.

C. Among the radius based neighbor classifiers, do the decision regions seem largely similar or quite different? If quite different, briefly describe the differences.

We can group them by radii 3 and 4, radii 6, and ones with weight. In the group, they share the similarities, but each group looks different. Ones with weight resulted similar to the k-nearest neighbor classifiers. Radii 3 and 4 resulted like a circle as the name suggests and seem underfitting. Radii 6 was just full of red. So, it is clearly underfitting.

D. Are the decision regions of the k-nearest neighbor classifiers and the fixed-radius nearest neighbor classifiers largely similar or quite different? If quite different, briefly describe the differences.

As I mentioned in C, radius nearest neighbor classifiers with weight resulted similar to the k-nearest neighbor classifiers. However, the rest of them look very different and the k-nearest neighbor classifiers seem working better with the given dataset.

E. Lastly, briefly discuss which, if any, of the classifiers used in HW7 had similar decision regions to the ones here across all or some of the data sets.

SVM with radial basis function kernel results similar to the k-nearest neighbor classifiers as well as radius nearest neighbor classifiers with weight.