Cs434 Assignment 1 solution

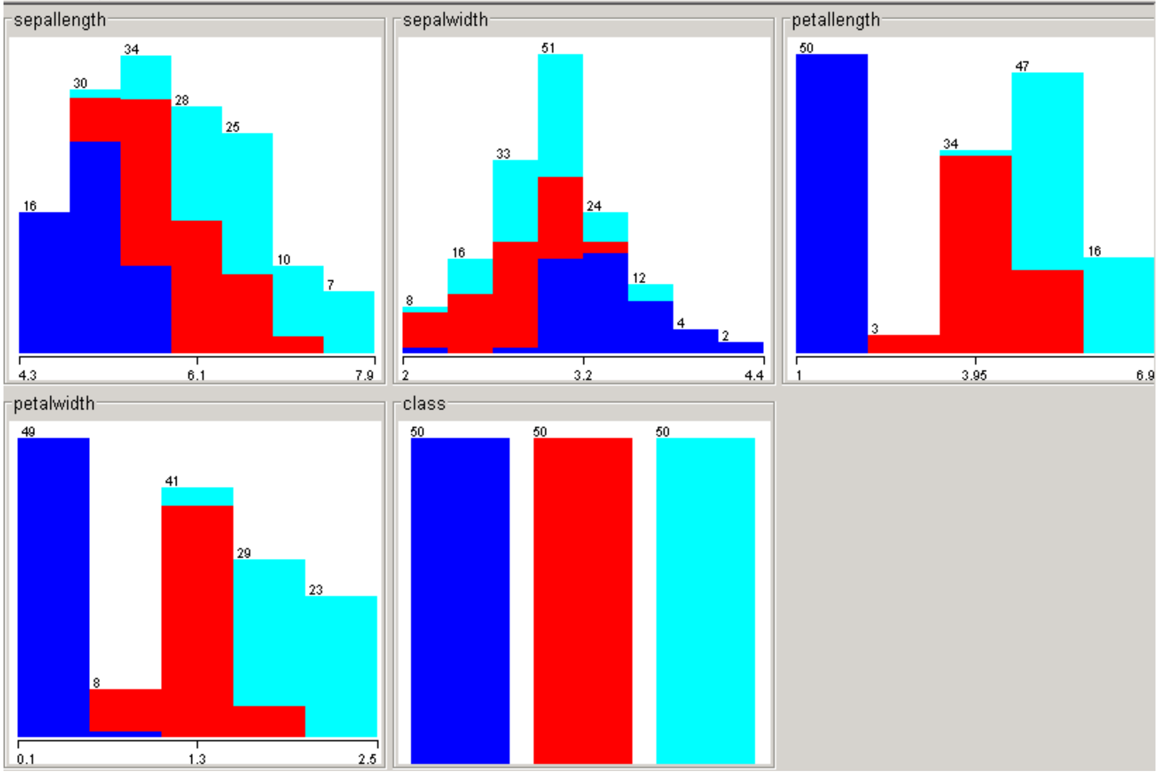
Part I

1. How many classes there are in this data set?
   Here we are assuming the “class” attribute is the target class label to predict. There are three classes: “Setosa”, “Versicolor” and “Virginica”

2. How many attributes there are?
   Excluding the class attribute, we have four numerical attributes: “sepal length”, “sepal width”, “petal length” and “pedal width”

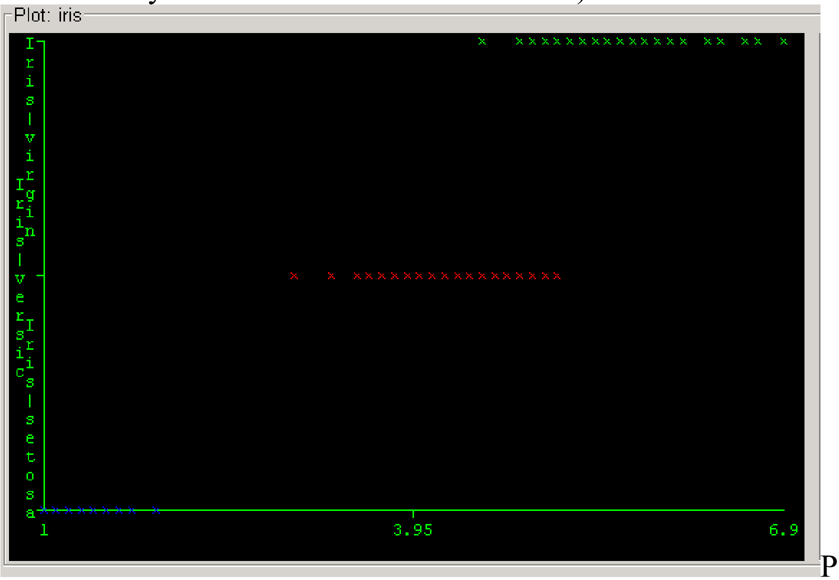
3. What are the mean and standard deviation for each attribute?
   SL: 5.843 (0.828)
   SW: 3.054 (0.434)
   PL: 3.759 (1.764)
   PW: 1.199 (0.763)

4. If you were to choose only one attribute to build your classifier, which attribute should you choose?

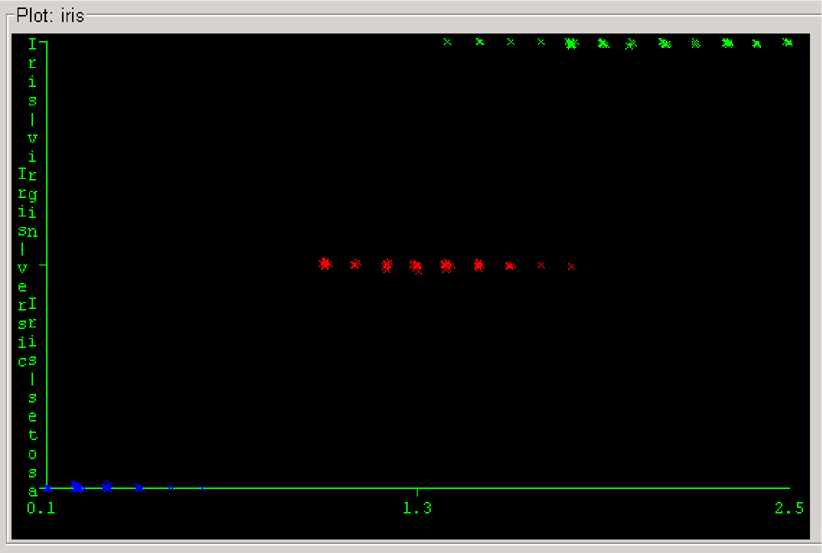


Above figure contains the histograms of the instances (each color represents one class).

Arguably one would choose either petallength or petalwidth. Both attributes provide a good separation between the setosa and the other two classes. This can be seen more clearly through the figures below, the first figure is the class label vs petallength, and the second is the class label vs petal width. Note that the points are jittered slightly in the petalwidth figure so that the points don't overlap significantly (weka provides this functionality as one of the visualization tools).



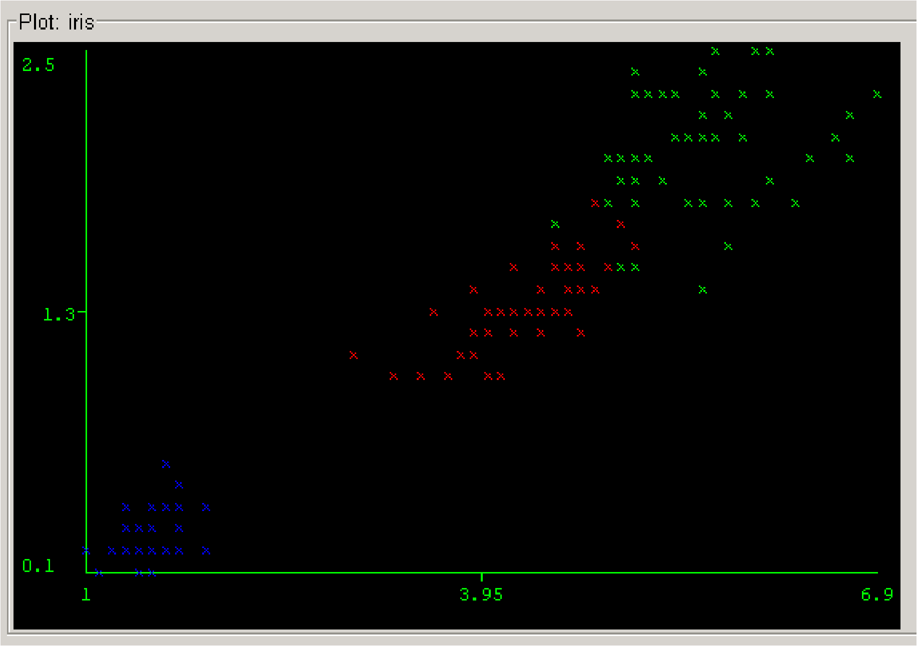
Petallength



Petalwidth

Because petallength values have a wider spread (larger std), it may appear to you that petallength should be better. But overall, it would be fair to say these attributes are highly similar in their distinguishing power.

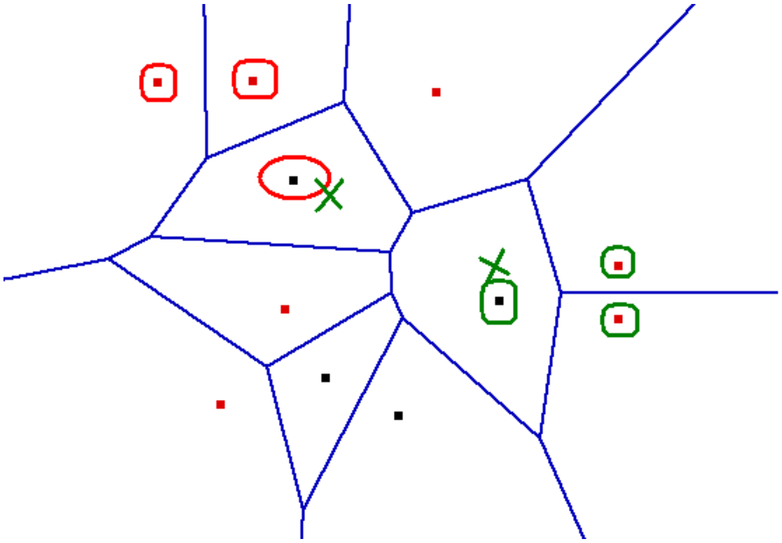
5. Which pair of attributes provides the best discrimination among classes? (Suggestion: use the visualization tool to look for good separations among classes.)
   Clearly PW and PL are the best. These two attributes combined provide a pretty reasonable separation among classes as shown in the figure (we call this figure a scatter plot of the training set) below, although it is arguable whether using two features together is adding much discriminative power because these two features are clearly highly correlated with one another.


Plot: iris
2.5
1.3
0.1
1
3.95
6.9

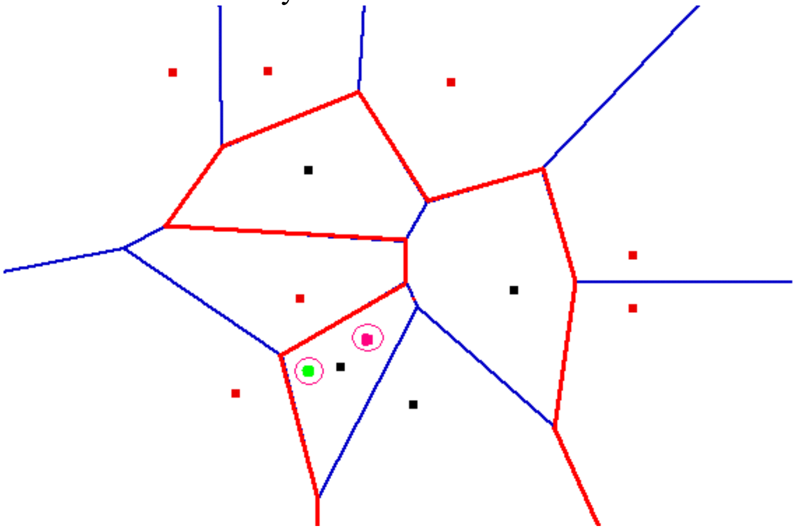
Part II

a) Training error of 1-nn is always 0, since the nearest neighbor of a data point is always itself (because itself is contained in the training set).

b) Training error of 3-nn is 2. The two points marked out with "X" are the errors. The circles mark out their 3 nearest neighbors. Note that one of the 3-nearest neighbors is always itself (again because itself is contained in the training set).



c) The decision boundary of 1-nn is marked out with red color.



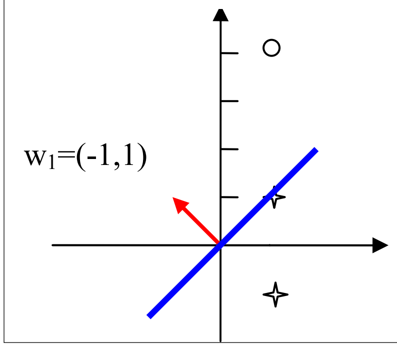
d) False. Consider the two points marked out in circles in the above figure. Using 3-nn, the green point will be classified as positive (red), and the pink point will be classified as negative (black). This indicates that there must be some boundaries within that region separating these two points, which cannot be part of the Vironoi segments. In general, deciding the decision boundaries of k-nn where k>1 is much more complicated than what Vironoi diagram can offer.

3. **Solution:**

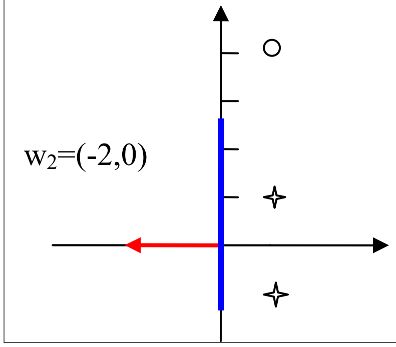
Note that you are asked to iterate until all examples are correctly classified, which requires iterating through the training examples multiple times. (This is what the perceptron algorithm is supposed to do. As a result, if the data points are not linearly separable, perceptron will cycle forever, bouncing back and force between solutions.)

Canonical representation of the data: (1, -1. -); (1, 1, -); (1,4,+)

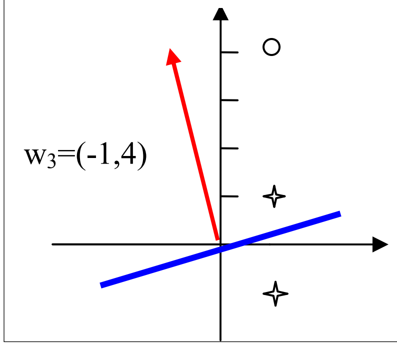
$w_0 = (0,0),\ \ y^1 w_0 \cdot x^1 = 0 \leq 0 \rightarrow$
$w_1 = w_0 + y^1 x^1 = (0,0) - (1,-1) = (-1,1)$



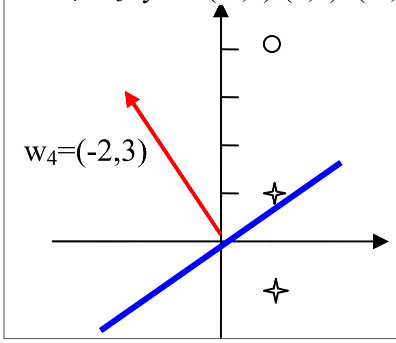
$w_1 = (-1,1),\ \ y^2 w_1 \cdot x^2 = 0 \leq 0 \rightarrow$
$w_2 = w_1 + y^2 x^2 = (-1,1) - (1,\ 1) = (-2,0)$



$w_2 = (-2,0),\ \ y^3 w_0 \cdot x^3 = -2 \leq 0 \rightarrow$
$w_3 = w_2 + y^3 x^3 = (-2,0) + (1,4) = (-1,4)$



$w_3 = (-1,4),\ \ y^1 w_3 \cdot x^1 = -5 > 0,$
$y^2 w_3 \cdot x^2 = -3 \leq 0 \ \ \rightarrow$
$w_4 = w_3 + y^2 x^2 = (-1,4) - (1,1) = (-2,3)$



$w_4 = (-2,3),\ \ y^3 w_4 \cdot x^3 = 10 > 0,$
$y^1 w_4 \cdot x^1 = -5 > 0,\ y^2 w_4 \cdot x^2 = -1 > 0,\ \ \rightarrow$
$w_5 = w_4 + y^2 x^2 = (-2,3) - (1,1) = (-3,2)$