K-Nearest Neighbors Exercise with Solution

(From Prof. Winston's book; Original file by Kimberle Koile,
original file date: before Fall 2004)

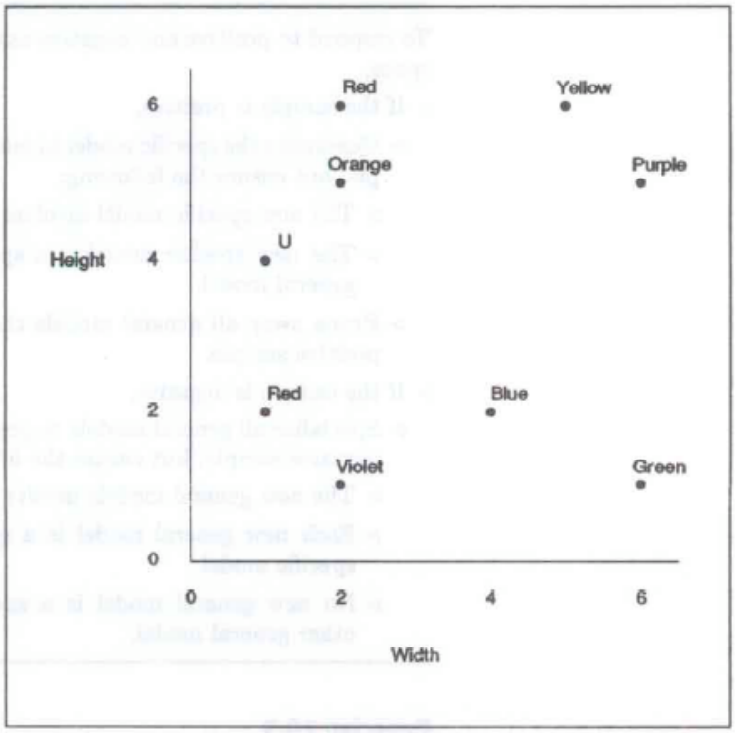

Figure E19.4 More than one nearest neighbor may be used to determine properties.

Exercise 19.6

If you wish to predict properties, it sometimes makes sense to look at several close neighbors, rather than at just one. But this approach raises the question of what to do when the close neighbors predict differing values. One solution is to let the neighbors vote with equal strength, with victory going to the value with the most votes. Another solution is to let the neighbors vote in proportion to a weight, w, that is determined by distance. Here are two possibilities:

- Each nearby neighbor votes according to the inverse square of its distance, $w = 1/d^2$.
- Each nearby neighbor votes according to a Gaussian function of the square of its distance, $w = e^{-kd^2}$.

Suppose that three nearby neighbors are used to predict the color of the unknown U in figure E19.4. Determine whether U is orange or red for inverse-square voting and for Gaussian voting with $k = 0.2$ and with $k = 0.4$.

Solution 19.6

For inverse-square voting, there are $1/2 = 0.5$ orange votes and $1/4 + 1/5 = 0.45$ red votes, so orange wins. For Gaussian voting with $k = 0.2$, there are 0.67 orange votes and 0.82 red votes, so red wins. For Gaussian voting with $k = 0.4$, voting strength falls off faster than with $k = 0.2$; there are 0.45 orange votes and 0.34 red votes, so orange wins.