KNN

Consider the problem where we want to predict the gender of a person from a set of input parameters, namely height, weight, and age. Assume our training data is given as follows:

D = { ((170, 57, 32), W),

((190, 95, 28), M),

((150, 45, 35), W),

((168, 65, 29), M),

((175, 78, 26), M),

((185, 90, 32), M),

((171, 65, 28), W),

((155, 48, 31), W),

((165, 60, 27), W),

((182, 80, 30), M),

((175, 69, 28), W),

((178, 80, 27), M),

((160, 50, 31), W),

((170, 72, 30), M), }

a) Using Cartesian distance as the similarity measurements show the results of the gender prediction for the following data items for values of K of 1, 3, and 5. Include the intermedia steps (i.e. distance calculation, neighbor selection, prediction).

(162, 53, 28), (168, 75, 32), (175, 70, 30), (180, 85, 29)

b) Implement the KNN algorithm for this problem. Your implementation should work with different training data sets and allow to input a data point for the prediction.

c) Repeat the prediction using KNN when the age data is removed (i.e. when only the first two features in the input data are available). Try to determine (using multiple target values) which data gives you better predictions. Show your intermediate results.