K Nearst Neighbours - Exercise

Today we will implement the k nearest neighbours algorithm on the famous iris data set.
We will do this in six small steps.

1. **Handle the data:** write a function that will open the dataset and split it to training and testing. You can either:
   A. use the function *sklearn.datasets.load_iris()* which provides an object where *.data* and *.target* are the data and the labels
   B*. Download the data from the following link:
   https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data
   store it in a text file and use numpy's *genfromtext* function to read it and then sort the data and the labels.
2. **Distance function:** write a function that can calculate the distance between two instances.
3. **Nearest neighbours:** write a function that searches the whole dataset for the k nearest neighbours.
   Hint: you can sort the disntances and find the k first elements using the function
   https://docs.scipy.org/doc/numpy/reference/generated/numpy.sort.html
   Predict from k nearst neighbours: now that we have the k nearest neighbours we can calculate an average of them to predict the result, or for categorical data we can do voting i.e. finding the mode (השכיח) between these k nearest points.
4. **Calculate the accuracy on the test data:** calcualte the prediction on every element of the test data and compare to the expected values. Calculate the percentage of the data sets that we predicted accurately.
5. **Main function:** write a main function that contains everything and calls all the functions that we have written.

If you have questions you can check online on the following link along which we followed:
http://machinelearningmastery.com/tutorial-to-implement-k-nearest-neighbors-in-python-from-scratch/