**averages.py**

import sys

import copy

filename = 'iris\_data.txt' if len(sys.argv) < 2 else sys.argv[1]

num\_read = 50 if len(sys.argv) < 3 else int(sys.argv[2])

file = open(filename, 'r')

data = []

next\_ele = file.readline()

next\_ele = file.readline()

data\_list = [0,0,0,0]

i = 0

while next\_ele != '' and i < num\_read:

flower\_data = next\_ele.split(',')

species = str(flower\_data.pop())

for j in range(len(flower\_data)):

data\_list[j] = float(flower\_data[j])

data.append(copy.deepcopy(data\_list))

next\_ele = file.readline()

i = i + 1

sums = [0.0,0.0,0.0,0.0]

averages = [0.0,0.0,0.0,0.0]

for flower in data:

for i in range(len(flower)):

sums[i] = sums[i] + flower[i]

for i in range(len(sums)):

averages[i] = sums[i] / len(data)

print averages

**compare.py**

import sys

#AVERAGES TAKEN FROM OUTPUT OF FIRST 80% OF EACH CLASS

setosa\_avg = [5.0375, 3.44, 1.4625, 0.2325]

versi\_avg = [6.01, 2.78, 4.3175, 1.35]

virgin\_avg = [6.6225, 2.96, 5.6075, 1.99]

averages = [setosa\_avg, versi\_avg, virgin\_avg]

filename = 'iris\_data.txt' if len(sys.argv) < 2 else sys.argv[1]

num\_start = 0 if len(sys.argv) < 3 else int(sys.argv[2])

num\_stop = -1 if len(sys.argv) < 4 else int(sys.argv[3])

file = open(filename, 'r')

#data = []

next\_ele = file.readline()

next\_ele = file.readline()

data\_list = [0,0,0,0]

i = num\_start

for x in range(i):

file.readline()

print "0 -> setosa\n1 -> versicolor\n2 -> virginica"

diff\_list = [0,0,0,0]

while next\_ele != '' and i != num\_stop:

flower\_data = next\_ele.split(',')

species = str(flower\_data.pop())[5:-1]

dist = [0.0,0.0,0.0]

for j in range(len(flower\_data)):

data\_list[j] = float(flower\_data[j])

for j in range(len(averages)):

for k in range(len(data\_list)):

dist[j] = dist[j] + ((data\_list[k] - averages[j][k]) \*\* 2) #EUCLID

#dist[j] = dist[j] + abs((data\_list[k] - averages[j][k])) #MANHATTAN

#diff\_list[k] = data\_list[k] - averages[j][k] #CHEBYCHEV

#dist[j] = max(diff\_list) # CHEBYCHEV

print dist.index(min(dist)), species

#data.append(copy.deepcopy(data\_list))

next\_ele = file.readline()

i = i + 1

#UNCOMMENT THE SPECIFIED SECTIONS TO CHANGE DISTANCE

**Results (Euclidean):** 3.33% Error

0 -> setosa

1 -> versicolor

2 -> virginica

0 setosa

0 setosa

0 setosa

0 setosa

0 setosa

0 setosa

0 setosa

0 setosa

0 setosa

0 setosa

1 versicolor

1 versicolor

1 versicolor

1 versicolor

1 versicolor

1 versicolor

1 versicolor

1 versicolor

1 versicolor

1 versicolor

2 virginica

2 virginica

1 virginica

2 virginica

2 virginica

2 virginica

2 virginica

2 virginica

2 virginica

2 virginica

**Results (Manhattan):** 3.33% Error

0 setosa

0 setosa

0 setosa

0 setosa

0 setosa

0 setosa

0 setosa

0 setosa

0 setosa

0 setosa

1 versicolor

1 versicolor

1 versicolor

1 versicolor

1 versicolor

1 versicolor

1 versicolor

1 versicolor

1 versicolor

1 versicolor

2 virginica

2 virginica

1 virginica

2 virginica

2 virginica

2 virginica

2 virginica

2 virginica

2 virginica

2 virginica

**Results (Cheyshev):** 40.0% Error

0 setosa

2 setosa

0 setosa

0 setosa

0 setosa

2 setosa

0 setosa

0 setosa

0 setosa

0 setosa

2 versicolor

2 versicolor

2 versicolor

2 versicolor

2 versicolor

2 versicolor

2 versicolor

2 versicolor

2 versicolor

2 versicolor

2 virginica

2 virginica

2 virginica

2 virginica

2 virginica

2 virginica

2 virginica

2 virginica

2 virginica

2 virginica