

《第5章 鸢尾花分类项目》实验报告

实验要求

一、测试数据文件

- 1、将原测试文件data.txt作为测试数据，运行程序。
- 2、将原测试文件data.txt中所有的记录随机打乱，生成新的测试数据文件，再运行程序

二、输出结果文件

- 1、根据以上两个测试数据文件，分别运行，生成输出结果文件。
- 2、输出结果文件中每条记录的格式为：

记录号：数据1，数据2，数据3，数据4，分类名称，聚类编号

例如：

```
132:4.80,3.40,1.60,0.20,Iris-setosa, 1
147:5.00,3.50,1.60,0.60,Iris-setosa, 1
1:6.80,3.20,5.90,2.30,Iris-virginica, 2
12:7.70,2.80,6.70,2.00,Iris-virginica, 2
14:6.70,3.30,5.70,2.10,Iris-virginica, 2
```

- 3、在输出结果文件的末尾，输出每个分类的正确率，和分类的总正确率。

三、程序功能要求

- 1、原作业所要求的功能。
- 2、不输入记录数，由读文件结束判断来读取测试数据文件。
- 3、增加对原测试数据文件记录进行随机打乱，生成新测试数据文件的功能。
- 4、增加对各分类正确率和分类总正确率的计算和输出。

四、

1、程序运行截图：顺序数据：程序输出

```
sxing@MacBook-Pro-XX 5# % ls
CMakeLists.txt      data.txt             randata.txt
a.out               head.h              result.txt
cmake-build-debug   main.c              实验报告.pages
sxing@MacBook-Pro-XX 5# % clang ./main.c
sxing@MacBook-Pro-XX 5# % ./a.out data.txt result.txt

Rounds: 1
cluster 0 point cnt: 22
cluster 0 center: length_of_sepal: 7.12, width_of_sepal: 3.11, length_of_petal: 6.03, width_of_petal: 2.13
cluster 1 point cnt: 42
cluster 1 center: length_of_sepal: 5.79, width_of_sepal: 2.75, length_of_petal: 4.74, width_of_petal: 1.70
cluster 2 point cnt: 87
cluster 2 center: length_of_sepal: 5.48, width_of_sepal: 3.15, length_of_petal: 2.67, width_of_petal: 0.71

Rounds: 2
cluster 0 point cnt: 32
cluster 0 center: length_of_sepal: 6.91, width_of_sepal: 3.10, length_of_petal: 5.85, width_of_petal: 2.13
cluster 1 point cnt: 61
cluster 1 center: length_of_sepal: 6.04, width_of_sepal: 2.80, length_of_petal: 4.58, width_of_petal: 1.51
cluster 2 point cnt: 58
cluster 2 center: length_of_sepal: 4.95, width_of_sepal: 3.24, length_of_petal: 1.57, width_of_petal: 0.34

Rounds: 3
cluster 0 point cnt: 35
cluster 0 center: length_of_sepal: 6.87, width_of_sepal: 3.09, length_of_petal: 5.79, width_of_petal: 2.12
cluster 1 point cnt: 64
cluster 1 center: length_of_sepal: 5.95, width_of_sepal: 2.76, length_of_petal: 4.45, width_of_petal: 1.44
cluster 2 point cnt: 52
cluster 2 center: length_of_sepal: 4.91, width_of_sepal: 3.33, length_of_petal: 1.47, width_of_petal: 0.26

Rounds: 4
cluster 0 point cnt: 37
cluster 0 center: length_of_sepal: 6.87, width_of_sepal: 3.09, length_of_petal: 5.75, width_of_petal: 2.09
cluster 1 point cnt: 63
cluster 1 center: length_of_sepal: 5.90, width_of_sepal: 2.75, length_of_petal: 4.41, width_of_petal: 1.43
cluster 2 point cnt: 51
cluster 2 center: length_of_sepal: 4.91, width_of_sepal: 3.35, length_of_petal: 1.44, width_of_petal: 0.24

Rounds: 5
cluster 0 point cnt: 38
cluster 0 center: length_of_sepal: 6.85, width_of_sepal: 3.07, length_of_petal: 5.74, width_of_petal: 2.07
cluster 1 point cnt: 62
cluster 1 center: length_of_sepal: 5.90, width_of_sepal: 2.75, length_of_petal: 4.39, width_of_petal: 1.43
cluster 2 point cnt: 51
cluster 2 center: length_of_sepal: 4.91, width_of_sepal: 3.35, length_of_petal: 1.44, width_of_petal: 0.24

Rounds: 6
cluster 0 point cnt: 38
cluster 0 center: length_of_sepal: 6.85, width_of_sepal: 3.07, length_of_petal: 5.74, width_of_petal: 2.07
cluster 1 point cnt: 62
cluster 1 center: length_of_sepal: 5.90, width_of_sepal: 2.75, length_of_petal: 4.39, width_of_petal: 1.43
cluster 2 point cnt: 51
cluster 2 center: length_of_sepal: 4.91, width_of_sepal: 3.35, length_of_petal: 1.44, width_of_petal: 0.24

Clustering is completed after 6 times of clustering
Result saved in result.txt
Total accuracy: 89.19%
From Database: data.txt
sxing@MacBook-Pro-XX 5# %
```

result.txt:

5.10,3.50,1.40,0.20,Iris-setosa	,1
4.90,3.00,1.40,0.20,Iris-setosa	,2
4.70,3.20,1.30,0.20,Iris-setosa	,2
4.60,3.10,1.50,0.20,Iris-setosa	,2
5.00,3.60,1.40,0.20,Iris-setosa	,2
5.40,3.90,1.70,0.40,Iris-setosa	,2
4.60,3.40,1.40,0.30,Iris-setosa	,2
5.00,3.40,1.50,0.20,Iris-setosa	,2
4.40,2.90,1.40,0.20,Iris-setosa	,2
4.90,3.10,1.50,0.10,Iris-setosa	,2
5.40,3.70,1.50,0.20,Iris-setosa	,2
4.80,3.40,1.60,0.20,Iris-setosa	,2
4.80,3.00,1.40,0.10,Iris-setosa	,2
4.30,3.00,1.10,0.10,Iris-setosa	,2
5.80,4.00,1.20,0.20,Iris-setosa	,2
5.70,4.40,1.50,0.40,Iris-setosa	,2
5.40,3.90,1.30,0.40,Iris-setosa	,2
5.10,3.50,1.40,0.30,Iris-setosa	,2
5.70,3.80,1.70,0.30,Iris-setosa	,2
5.10,3.80,1.50,0.30,Iris-setosa	,2
5.40,3.40,1.70,0.20,Iris-setosa	,2
5.10,3.70,1.50,0.40,Iris-setosa	,2
4.60,3.60,1.00,0.20,Iris-setosa	,2
5.10,3.30,1.70,0.50,Iris-setosa	,2
4.80,3.40,1.90,0.20,Iris-setosa	,2
5.00,3.00,1.60,0.20,Iris-setosa	,2
5.00,3.40,1.60,0.40,Iris-setosa	,2
5.20,3.50,1.50,0.20,Iris-setosa	,2
5.20,3.40,1.40,0.20,Iris-setosa	,2
4.70,3.20,1.60,0.20,Iris-setosa	,2
4.80,3.10,1.60,0.20,Iris-setosa	,2
5.40,3.40,1.50,0.40,Iris-setosa	,2
5.20,4.10,1.50,0.10,Iris-setosa	,2
5.50,4.20,1.40,0.20,Iris-setosa	,2
4.90,3.10,1.50,0.10,Iris-setosa	,2
5.00,3.20,1.20,0.20,Iris-setosa	,2
5.50,3.50,1.30,0.20,Iris-setosa	,2
4.90,3.10,1.50,0.10,Iris-setosa	,2
4.40,3.00,1.30,0.20,Iris-setosa	,2
5.10,3.40,1.50,0.20,Iris-setosa	,2
5.00,3.50,1.30,0.30,Iris-setosa	,2
4.50,2.30,1.30,0.30,Iris-setosa	,2
4.40,3.20,1.30,0.20,Iris-setosa	,2
5.00,3.50,1.60,0.60,Iris-setosa	,2
5.10,3.80,1.90,0.40,Iris-setosa	,2
4.80,3.00,1.40,0.30,Iris-setosa	,2
5.10,3.80,1.60,0.20,Iris-setosa	,2

4.60,3.20,1.40,0.20,Iris-setosa	,2
5.30,3.70,1.50,0.20,Iris-setosa	,2
5.00,3.30,1.40,0.20,Iris-setosa	,2
7.00,3.20,4.70,1.40,Iris-versicolor	,1
6.40,3.20,4.50,1.50,Iris-versicolor	,1
6.90,3.10,4.90,1.50,Iris-versicolor	,0
5.50,2.30,4.00,1.30,Iris-versicolor	,1
6.50,2.80,4.60,1.50,Iris-versicolor	,1
5.70,2.80,4.50,1.30,Iris-versicolor	,1
6.30,3.30,4.70,1.60,Iris-versicolor	,1
4.90,2.40,3.30,1.00,Iris-versicolor	,1
6.60,2.90,4.60,1.30,Iris-versicolor	,1
5.20,2.70,3.90,1.40,Iris-versicolor	,1
5.00,2.00,3.50,1.00,Iris-versicolor	,1
5.90,3.00,4.20,1.50,Iris-versicolor	,1
6.00,2.20,4.00,1.00,Iris-versicolor	,1
6.10,2.90,4.70,1.40,Iris-versicolor	,1
5.60,2.90,3.60,1.30,Iris-versicolor	,1
6.70,3.10,4.40,1.40,Iris-versicolor	,1
5.60,3.00,4.50,1.50,Iris-versicolor	,1
5.80,2.70,4.10,1.00,Iris-versicolor	,1
6.20,2.20,4.50,1.50,Iris-versicolor	,1
5.60,2.50,3.90,1.10,Iris-versicolor	,1
5.90,3.20,4.80,1.80,Iris-versicolor	,1
6.10,2.80,4.00,1.30,Iris-versicolor	,1
6.30,2.50,4.90,1.50,Iris-versicolor	,1
6.10,2.80,4.70,1.20,Iris-versicolor	,1
6.40,2.90,4.30,1.30,Iris-versicolor	,1
6.60,3.00,4.40,1.40,Iris-versicolor	,1
6.80,2.80,4.80,1.40,Iris-versicolor	,1
6.70,3.00,5.00,1.70,Iris-versicolor	,0
6.00,2.90,4.50,1.50,Iris-versicolor	,1
5.70,2.60,3.50,1.00,Iris-versicolor	,1
5.50,2.40,3.80,1.10,Iris-versicolor	,1
5.50,2.40,3.70,1.00,Iris-versicolor	,1
5.80,2.70,3.90,1.20,Iris-versicolor	,1
6.00,2.70,5.10,1.60,Iris-versicolor	,1
5.40,3.00,4.50,1.50,Iris-versicolor	,1
6.00,3.40,4.50,1.60,Iris-versicolor	,1
6.70,3.10,4.70,1.50,Iris-versicolor	,1
6.30,2.30,4.40,1.30,Iris-versicolor	,1
5.60,3.00,4.10,1.30,Iris-versicolor	,1
5.50,2.50,4.00,1.30,Iris-versicolor	,1
5.50,2.60,4.40,1.20,Iris-versicolor	,1
6.10,3.00,4.60,1.40,Iris-versicolor	,1
5.80,2.60,4.00,1.20,Iris-versicolor	,1
5.00,2.30,3.30,1.00,Iris-versicolor	,1
5.60,2.70,4.20,1.30,Iris-versicolor	,1

5.70,3.00,4.20,1.20,Iris-versicolor	,1
5.70,2.90,4.20,1.30,Iris-versicolor	,1
6.20,2.90,4.30,1.30,Iris-versicolor	,1
5.10,2.50,3.00,1.10,Iris-versicolor	,1
5.70,2.80,4.10,1.30,Iris-versicolor	,1
6.30,3.30,6.00,2.50,Iris-virginica	,0
5.80,2.70,5.10,1.90,Iris-virginica	,1
7.10,3.00,5.90,2.10,Iris-virginica	,0
6.30,2.90,5.60,1.80,Iris-virginica	,0
6.50,3.00,5.80,2.20,Iris-virginica	,0
7.60,3.00,6.60,2.10,Iris-virginica	,0
4.90,2.50,4.50,1.70,Iris-virginica	,1
7.30,2.90,6.30,1.80,Iris-virginica	,0
6.70,2.50,5.80,1.80,Iris-virginica	,0
7.20,3.60,6.10,2.50,Iris-virginica	,0
6.50,3.20,5.10,2.00,Iris-virginica	,0
6.40,2.70,5.30,1.90,Iris-virginica	,0
6.80,3.00,5.50,2.10,Iris-virginica	,0
5.70,2.50,5.00,2.00,Iris-virginica	,1
5.80,2.80,5.10,2.40,Iris-virginica	,1
6.40,3.20,5.30,2.30,Iris-virginica	,0
6.50,3.00,5.50,1.80,Iris-virginica	,0
7.70,3.80,6.70,2.20,Iris-virginica	,0
7.70,2.60,6.90,2.30,Iris-virginica	,0
6.00,2.20,5.00,1.50,Iris-virginica	,1
6.90,3.20,5.70,2.30,Iris-virginica	,0
5.60,2.80,4.90,2.00,Iris-virginica	,1
7.70,2.80,6.70,2.00,Iris-virginica	,0
6.30,2.70,4.90,1.80,Iris-virginica	,1
6.70,3.30,5.70,2.10,Iris-virginica	,0
7.20,3.20,6.00,1.80,Iris-virginica	,0
6.20,2.80,4.80,1.80,Iris-virginica	,1
6.10,3.00,4.90,1.80,Iris-virginica	,1
6.40,2.80,5.60,2.10,Iris-virginica	,0
7.20,3.00,5.80,1.60,Iris-virginica	,0
7.40,2.80,6.10,1.90,Iris-virginica	,0
7.90,3.80,6.40,2.00,Iris-virginica	,0
6.40,2.80,5.60,2.20,Iris-virginica	,0
6.30,2.80,5.10,1.50,Iris-virginica	,1
6.10,2.60,5.60,1.40,Iris-virginica	,0
7.70,3.00,6.10,2.30,Iris-virginica	,0
6.30,3.40,5.60,2.40,Iris-virginica	,0
6.40,3.10,5.50,1.80,Iris-virginica	,0
6.00,3.00,4.80,1.80,Iris-virginica	,1
6.90,3.10,5.40,2.10,Iris-virginica	,0
6.70,3.10,5.60,2.40,Iris-virginica	,0
6.90,3.10,5.10,2.30,Iris-virginica	,0
5.80,2.70,5.10,1.90,Iris-virginica	,1

```

6.80,3.20,5.90,2.30,Iris-virginica      ,0
6.70,3.30,5.70,2.50,Iris-virginica      ,0
6.70,3.00,5.20,2.30,Iris-virginica      ,0
6.30,2.50,5.00,1.90,Iris-virginica      ,1
6.50,3.00,5.20,2.00,Iris-virginica      ,0
6.20,3.40,5.40,2.30,Iris-virginica      ,0
5.90,3.00,5.10,1.80,Iris-virginica      ,1
0.00,0.00,0.00,0.00,                    ,2

```

```

Category: Iris-setosa      ,Accuracy: 100.00%
Category: Iris-versicolor ,Accuracy: 96.00%
Category: ris-virginica    ,Accuracy: 72.00%

```

2、程序运行截图：随机数据

```

[sxing@MacBook-Pro-XX 5# % ./a.out randata.txt result.txt

Rounds: 1
cluster 0 point cnt: 58
cluster 0 center: length_of_sepal: 5.85, width_of_sepal: 2.73, length_of_petal: 4.36, width_of_petal: 1.43
cluster 1 point cnt: 42
cluster 1 center: length_of_sepal: 6.84, width_of_sepal: 3.06, length_of_petal: 5.65, width_of_petal: 2.01
cluster 2 point cnt: 50
cluster 2 center: length_of_sepal: 5.01, width_of_sepal: 3.42, length_of_petal: 1.46, width_of_petal: 0.24

Rounds: 2
cluster 0 point cnt: 61
cluster 0 center: length_of_sepal: 5.88, width_of_sepal: 2.74, length_of_petal: 4.39, width_of_petal: 1.43
cluster 1 point cnt: 39
cluster 1 center: length_of_sepal: 6.85, width_of_sepal: 3.08, length_of_petal: 5.72, width_of_petal: 2.05
cluster 2 point cnt: 50
cluster 2 center: length_of_sepal: 5.01, width_of_sepal: 3.42, length_of_petal: 1.46, width_of_petal: 0.24

Rounds: 3
cluster 0 point cnt: 61
cluster 0 center: length_of_sepal: 5.88, width_of_sepal: 2.74, length_of_petal: 4.39, width_of_petal: 1.43
cluster 1 point cnt: 39
cluster 1 center: length_of_sepal: 6.85, width_of_sepal: 3.08, length_of_petal: 5.72, width_of_petal: 2.05
cluster 2 point cnt: 50
cluster 2 center: length_of_sepal: 5.01, width_of_sepal: 3.42, length_of_petal: 1.46, width_of_petal: 0.24

Clustering is completed after 3 times of clustering
Result saved in result.txt
Total accuracy: 88.59%
From Database: randata.txt_

```

result.txt

```

5.80,2.70,3.90,1.20,Iris-versicolor      ,-1
4.60,3.10,1.50,0.20,Iris-setosa          ,2
6.10,3.00,4.90,1.80,Iris-virginica      ,0
6.50,3.00,5.20,2.00,Iris-virginica      ,1
4.70,3.20,1.30,0.20,Iris-setosa          ,2
5.90,3.20,4.80,1.80,Iris-versicolor      ,0
4.90,3.10,1.50,0.10,Iris-setosa          ,2
6.10,2.80,4.00,1.30,Iris-versicolor      ,0
4.40,3.20,1.30,0.20,Iris-setosa          ,2
6.70,2.50,5.80,1.80,Iris-virginica      ,1

```

4.80,3.40,1.60,0.20,Iris-setosa	,2
4.60,3.60,1.00,0.20,Iris-setosa	,2
5.70,3.00,4.20,1.20,Iris-versicolor	,0
5.50,4.20,1.40,0.20,Iris-setosa	,2
7.60,3.00,6.60,2.10,Iris-virginica	,1
6.80,3.20,5.90,2.30,Iris-virginica	,1
6.00,2.20,4.00,1.00,Iris-versicolor	,0
4.60,3.40,1.40,0.30,Iris-setosa	,2
4.80,3.10,1.60,0.20,Iris-setosa	,2
5.20,2.70,3.90,1.40,Iris-versicolor	,0
6.60,3.00,4.40,1.40,Iris-versicolor	,0
6.50,3.20,5.10,2.00,Iris-virginica	,1
5.10,3.70,1.50,0.40,Iris-setosa	,2
6.70,3.10,4.40,1.40,Iris-versicolor	,0
6.00,2.20,5.00,1.50,Iris-virginica	,0
5.50,3.50,1.30,0.20,Iris-setosa	,2
5.00,3.50,1.30,0.30,Iris-setosa	,2
4.90,3.10,1.50,0.10,Iris-setosa	,2
6.00,3.00,4.80,1.80,Iris-virginica	,0
6.80,2.80,4.80,1.40,Iris-versicolor	,0
6.30,3.40,5.60,2.40,Iris-virginica	,1
5.40,3.00,4.50,1.50,Iris-versicolor	,0
5.40,3.70,1.50,0.20,Iris-setosa	,2
5.50,2.40,3.70,1.00,Iris-versicolor	,0
6.40,3.20,5.30,2.30,Iris-virginica	,1
5.80,2.70,4.10,1.00,Iris-versicolor	,0
6.90,3.10,5.10,2.30,Iris-virginica	,1
7.70,2.80,6.70,2.00,Iris-virginica	,1
7.00,3.20,4.70,1.40,Iris-versicolor	,1
5.90,3.00,4.20,1.50,Iris-versicolor	,0
5.70,2.60,3.50,1.00,Iris-versicolor	,0
6.70,3.10,4.70,1.50,Iris-versicolor	,0
7.70,2.60,6.90,2.30,Iris-virginica	,1
7.40,2.80,6.10,1.90,Iris-virginica	,1
5.00,3.60,1.40,0.20,Iris-setosa	,2
6.10,3.00,4.60,1.40,Iris-versicolor	,0
4.60,3.20,1.40,0.20,Iris-setosa	,2
6.90,3.10,4.90,1.50,Iris-versicolor	,1
5.40,3.90,1.30,0.40,Iris-setosa	,2
6.30,3.30,4.70,1.60,Iris-versicolor	,0
6.40,3.20,4.50,1.50,Iris-versicolor	,0
5.20,3.50,1.50,0.20,Iris-setosa	,2
5.40,3.90,1.70,0.40,Iris-setosa	,2
5.60,2.50,3.90,1.10,Iris-versicolor	,0
6.70,3.30,5.70,2.50,Iris-virginica	,1
4.80,3.00,1.40,0.10,Iris-setosa	,2
6.40,3.10,5.50,1.80,Iris-virginica	,1
6.20,2.90,4.30,1.30,Iris-versicolor	,0

7.10,3.00,5.90,2.10,Iris-virginica	,1
6.70,3.00,5.00,1.70,Iris-versicolor	,1
6.10,2.60,5.60,1.40,Iris-virginica	,1
5.00,2.30,3.30,1.00,Iris-versicolor	,0
5.20,3.40,1.40,0.20,Iris-setosa	,2
6.30,3.30,6.00,2.50,Iris-virginica	,1
6.40,2.90,4.30,1.30,Iris-versicolor	,0
4.80,3.40,1.90,0.20,Iris-setosa	,2
5.50,2.50,4.00,1.30,Iris-versicolor	,0
5.20,4.10,1.50,0.10,Iris-setosa	,2
6.10,2.90,4.70,1.40,Iris-versicolor	,0
6.70,3.00,5.20,2.30,Iris-virginica	,1
6.30,2.90,5.60,1.80,Iris-virginica	,1
6.60,2.90,4.60,1.30,Iris-versicolor	,0
5.00,3.40,1.50,0.20,Iris-setosa	,2
5.70,2.80,4.50,1.30,Iris-versicolor	,0
5.10,2.50,3.00,1.10,Iris-versicolor	,0
5.60,3.00,4.10,1.30,Iris-versicolor	,0
6.40,2.80,5.60,2.20,Iris-virginica	,1
5.10,3.40,1.50,0.20,Iris-setosa	,2
5.70,3.80,1.70,0.30,Iris-setosa	,2
5.10,3.80,1.50,0.30,Iris-setosa	,2
7.70,3.80,6.70,2.20,Iris-virginica	,1
4.30,3.00,1.10,0.10,Iris-setosa	,2
7.20,3.60,6.10,2.50,Iris-virginica	,1
5.00,3.50,1.60,0.60,Iris-setosa	,2
6.00,2.90,4.50,1.50,Iris-versicolor	,0
4.40,2.90,1.40,0.20,Iris-setosa	,2
4.70,3.20,1.60,0.20,Iris-setosa	,2
5.50,2.30,4.00,1.30,Iris-versicolor	,0
6.50,3.00,5.80,2.20,Iris-virginica	,1
5.60,2.80,4.90,2.00,Iris-virginica	,0
5.10,3.80,1.60,0.20,Iris-setosa	,2
5.70,2.90,4.20,1.30,Iris-versicolor	,0
6.10,2.80,4.70,1.20,Iris-versicolor	,0
7.30,2.90,6.30,1.80,Iris-virginica	,1
6.20,3.40,5.40,2.30,Iris-virginica	,1
7.70,3.00,6.10,2.30,Iris-virginica	,1
5.70,2.50,5.00,2.00,Iris-virginica	,0
5.90,3.00,5.10,1.80,Iris-virginica	,0
4.80,3.00,1.40,0.30,Iris-setosa	,2
5.00,2.00,3.50,1.00,Iris-versicolor	,0
5.50,2.60,4.40,1.20,Iris-versicolor	,0
5.80,2.60,4.00,1.20,Iris-versicolor	,0
7.20,3.20,6.00,1.80,Iris-virginica	,1
5.60,2.70,4.20,1.30,Iris-versicolor	,0
6.00,3.40,4.50,1.60,Iris-versicolor	,0
5.10,3.50,1.40,0.20,Iris-setosa	,2

4.90,3.00,1.40,0.20,Iris-setosa	,2
6.40,2.80,5.60,2.10,Iris-virginica	,1
4.90,2.50,4.50,1.70,Iris-virginica	,0
6.90,3.20,5.70,2.30,Iris-virginica	,1
6.20,2.80,4.80,1.80,Iris-virginica	,0
6.20,2.20,4.50,1.50,Iris-versicolor	,0
5.70,2.80,4.10,1.30,Iris-versicolor	,0
5.10,3.50,1.40,0.30,Iris-setosa	,2
5.50,2.40,3.80,1.10,Iris-versicolor	,0
4.40,3.00,1.30,0.20,Iris-setosa	,2
4.90,2.40,3.30,1.00,Iris-versicolor	,0
6.30,2.50,5.00,1.90,Iris-virginica	,0
6.70,3.10,5.60,2.40,Iris-virginica	,1
4.50,2.30,1.30,0.30,Iris-setosa	,2
5.70,4.40,1.50,0.40,Iris-setosa	,2
5.60,2.90,3.60,1.30,Iris-versicolor	,0
5.40,3.40,1.50,0.40,Iris-setosa	,2
5.00,3.30,1.40,0.20,Iris-setosa	,2
7.20,3.00,5.80,1.60,Iris-virginica	,1
5.80,2.70,5.10,1.90,Iris-virginica	,0
5.80,2.70,5.10,1.90,Iris-virginica	,0
6.30,2.30,4.40,1.30,Iris-versicolor	,0
6.70,3.30,5.70,2.10,Iris-virginica	,1
5.80,4.00,1.20,0.20,Iris-setosa	,2
4.90,3.10,1.50,0.10,Iris-setosa	,2
6.40,2.70,5.30,1.90,Iris-virginica	,1
6.90,3.10,5.40,2.10,Iris-virginica	,1
6.50,2.80,4.60,1.50,Iris-versicolor	,0
5.80,2.80,5.10,2.40,Iris-virginica	,0
5.00,3.00,1.60,0.20,Iris-setosa	,2
5.00,3.20,1.20,0.20,Iris-setosa	,2
6.50,3.00,5.50,1.80,Iris-virginica	,1
6.30,2.70,4.90,1.80,Iris-virginica	,0
5.30,3.70,1.50,0.20,Iris-setosa	,2
7.90,3.80,6.40,2.00,Iris-virginica	,1
5.40,3.40,1.70,0.20,Iris-setosa	,2
6.30,2.80,5.10,1.50,Iris-virginica	,0
6.30,2.50,4.90,1.50,Iris-versicolor	,0
6.80,3.00,5.50,2.10,Iris-virginica	,1
6.00,2.70,5.10,1.60,Iris-versicolor	,0
5.10,3.80,1.90,0.40,Iris-setosa	,2
5.60,3.00,4.50,1.50,Iris-versicolor	,0
5.10,3.30,1.70,0.50,Iris-setosa	,2
5.00,3.40,1.60,0.40,Iris-setosa	,2

Category: Iris-setosa	,Accuracy: 100.00%
Category: Iris-versicolor	,Accuracy: 93.88%
Category: ris-virginica	,Accuracy: 72.00%

源代码:

```
<head.h>
//
// Created by Sxing on 04/10/2021.
//

#ifndef INC_5__HEAD_H
#define INC_5__HEAD_H

#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <math.h>
#include <time.h>

#define K 3 // num of categories of iris
#define num_of_dim // num of dimension
#define MAX_ROUNDS 100 // maximum number of iterations

typedef enum {
    True = 1,
    False = 0
} bool;

typedef struct {
    double length_of_sepal;
    double width_of_sepal;
    double length_of_petal;
    double width_of_petal;
    char name_of_iris[30];
    int clusterID;
} Iris;

typedef struct {
    char iris_name[30];
    double accuracy;
} test_result;

int is_continue; // is or not continue
Iris iris[151]; // data
//Iris iris_copy[151]; // a copy of iris for verifying
Iris cluster_center[K]; // K set of clusters.
Iris new_center[K]; // a new copy
int num_of_data;
int cluster_center_init_index[K]; // the index of the data point originally
used for each cluster center
double distance_from_center[K]; // distance from center for each point.
int data_size_per_cluster[K]; // total number of samples within each cluster
```

```

test_result testResults[K]; //
double total_accuracy;

bool inputs(char file_name[]) {
    FILE *fp;
    char name_of_iris[30];
    if ((fp = fopen(file_name, "rb")) == NULL) {
        return False;
    }
    int count = 0;
    while (!feof(fp)) {
        fscanf(fp, "%lf,%lf,%lf,%lf,%s", &iris[count].length_of_sepal,
&iris[count].width_of_sepal,
        &iris[count].length_of_petal, &iris[count].width_of_petal,
iris[count].name_of_iris);
        iris[count].clusterID = -1;
        count++;
    }
    num_of_data = count;
    fclose(fp);
    return True;
}

void init_cluster() {
    int random;
    srand((unsigned) time(NULL));
    for (int i = 0; i < K; ++i) {
        cluster_center_init_index[i] = -1;
    }
    for (int i = 0; i < K; ++i) {
        random = rand() % (num_of_data - 1);
        int j = 0;
        while (j < i) {
            if (random == cluster_center_init_index[j]) {
                random = rand() % (num_of_data);
                j = 0;
            } else {
                j++;
            }
        }
        cluster_center_init_index[i] = random;
    }
    for (int i = 0; i < K; ++i) {
        cluster_center[i].length_of_sepal =
iris[cluster_center_init_index[i]].length_of_sepal;
        cluster_center[i].length_of_petal =
iris[cluster_center_init_index[i]].length_of_petal;
    }
}

```

```

        cluster_center[i].width_of_sepal =
iris[cluster_center_init_index[i]].width_of_sepal;
        cluster_center[i].width_of_petal =
iris[cluster_center_init_index[i]].width_of_petal;
        cluster_center[i].clusterID = i;
        iris[cluster_center_init_index[i]].clusterID = i;
    }
}

void partition_4_all_point_one_cluster();

void cal_cluster_center();

void KMeans() {
    int rounds;
    for (rounds = 0; rounds < MAX_ROUNDS; ++rounds) {
        printf("\nRounds: %d\n", rounds + 1);
        partition_4_all_point_one_cluster();
        cal_cluster_center();
        if (is_continue == 0) {
            printf("\nClustering is completed after %d times of clustering\n",
rounds + 1);
            break;
        }
    }
}

void cal_distance_2_one_centers(int pointID, int centerID) {
    double x1 = pow(iris[pointID].length_of_sepal -
cluster_center[centerID].length_of_sepal, 2.0);
    double x2 = pow(iris[pointID].length_of_petal -
cluster_center[centerID].length_of_petal, 2.0);
    double x3 = pow(iris[pointID].width_of_sepal -
cluster_center[centerID].width_of_sepal, 2.0);
    double x4 = pow(iris[pointID].width_of_petal -
cluster_center[centerID].width_of_petal, 2.0);
    distance_from_center[centerID] = sqrt(x1 + x2 + x3 + x4);
}

void cal_distance_2_all_centers(int pointID) {
    for (int i = 0; i < K; ++i) {
        cal_distance_2_one_centers(pointID, i);
    }
}

void partition_4_one_point(int pointID) {
    int min_index = 0;
    double min_value = distance_from_center[0];

```

```

    for (int i = 0; i < K; ++i) {
        if (distance_from_center[i] < min_value) {
            min_value = distance_from_center[i];
            min_index = i;
        }
    }
    iris[pointID].clusterID = cluster_center[min_index].clusterID;
}

void partition_4_all_point_one_cluster() {
    for (int i = 0; i < num_of_data; ++i) {

        cal_distance_2_all_centers(i);
        partition_4_one_point(i);

    }
}

void compare_new_old_cluster_center();

void cal_cluster_center() {
    memset(new_center, 0, sizeof(new_center));
    memset(data_size_per_cluster, 0, sizeof(data_size_per_cluster));
    for (int i = 0; i < num_of_data; ++i) {
        new_center[iris[i].clusterID].length_of_sepal +=
iris[i].length_of_sepal;
        new_center[iris[i].clusterID].length_of_petal +=
iris[i].length_of_petal;
        new_center[iris[i].clusterID].width_of_sepal += iris[i].width_of_sepal;
        new_center[iris[i].clusterID].width_of_petal += iris[i].width_of_petal;
        data_size_per_cluster[iris[i].clusterID]++;
    }
    for (int i = 0; i < K; ++i) {
        if (data_size_per_cluster[i] != 0) {
            new_center[i].length_of_sepal = new_center[i].length_of_sepal /
(double) data_size_per_cluster[i];
            new_center[i].length_of_petal = new_center[i].length_of_petal /
(double) data_size_per_cluster[i];
            new_center[i].width_of_sepal = new_center[i].width_of_sepal /
(double) data_size_per_cluster[i];
            new_center[i].width_of_petal = new_center[i].width_of_petal /
(double) data_size_per_cluster[i];
            printf("cluster %d point cnt: %d\n", i, data_size_per_cluster[i]);
            printf("cluster %d center: length_of_sepal: %.2lf, width_of_sepal:
%.2lf, length_of_petal: %.2lf, width_of_petal: %.2lf\n",
                i, new_center[i].length_of_sepal,
new_center[i].width_of_sepal, new_center[i].length_of_petal,
                new_center[i].width_of_petal);

```

```

        } else {
            printf(" cluster %d count is zero\n", i);
        }
    }
    compare_new_old_cluster_center();
    for (int i = 0; i < K; ++i) {
        cluster_center[i].length_of_sepal = new_center[i].length_of_sepal;
        cluster_center[i].width_of_sepal = new_center[i].width_of_sepal;
        cluster_center[i].length_of_petal = new_center[i].length_of_petal;
        cluster_center[i].width_of_petal = new_center[i].width_of_petal;
        cluster_center[i].clusterID = i;
    }
    for (int i = 0; i < num_of_data; ++i) {
        iris->clusterID = -1;
    }
}

void compare_new_old_cluster_center() {
    is_continue = 0;
    for (int i = 0; i < K; ++i) {
        if (cluster_center[i].length_of_sepal != new_center[i].length_of_sepal
||
            cluster_center[i].width_of_sepal != new_center[i].width_of_sepal ||
            cluster_center[i].length_of_petal != new_center[i].length_of_petal
||
            cluster_center[i].width_of_petal != new_center[i].width_of_petal) {
            is_continue = 1;
            break;
        }
    }
}

void output_accuracy() {

    int category[K][K + 3] = {0}; // row as the hits of one species, 3 species
in total.
    // index k is the associated cluster_ID, index k+1 is the associated sample
count, index k+2 is the total count.
    for (int j = 0; j < num_of_data; ++j) {
        if (strcmp(iris[j].name_of_iris, "Iris-setosa") == 0) {
            category[0][iris[j].clusterID]++;
        } else if (strcmp(iris[j].name_of_iris, "Iris-versicolor") == 0) {
            category[1][iris[j].clusterID]++;
        } else if (strcmp(iris[j].name_of_iris, "Iris-virginica") == 0) {
            category[2][iris[j].clusterID]++;
        }
    }
    for (int i = 0; i < K; ++i) {

```

```

        int max = 0;
        int index_max = 0;
        int sum = 0;
        for (int j = 0; j < K; ++j) {
            sum += category[i][j];
            if (max < category[i][j]) {
                max = category[i][j];
                index_max = j;
            }
        }
        category[i][K] = index_max;
        category[i][K + 1] = max;
        category[i][K + 2] = sum;
    }

    strcpy(testResults[0].iris_name, "Iris-setosa");
    strcpy(testResults[1].iris_name, "Iris-versicolor");
    strcpy(testResults[2].iris_name, "Iris-virginica");
    int total_sum = 0, total_correct = 0;
    for (int i = 0; i < K; ++i) {
        testResults[i].accuracy = (double) category[i][K + 1] / category[i][K +
2];
        total_sum += category[i][K + 2];
        total_correct += category[i][K + 1];
    }
    total_accuracy = (double) total_correct / total_sum;
}

void verify(char *input, char *output) {
    FILE *fp;
    char name_of_iris[30];
    fp = fopen(output, "w");
    int count = 0;
    while (count < num_of_data) {
        fprintf(fp, "%-4.2lf,%-4.2lf,%-4.2lf,%-4.2lf,%-20s,%d\n",
iris[count].length_of_sepal,
            iris[count].width_of_sepal,
            iris[count].length_of_petal, iris[count].width_of_petal,
iris[count].name_of_iris,
            iris[count].clusterID);
        count++;
    }
    output_accuracy();
    fprintf(fp, "\n");
    for (int i = 0; i < K; ++i) {
        fprintf(fp, "Category: %-20s,Accuracy: %.2lf%%\n",
testResults[i].iris_name, testResults[i].accuracy * 100);
    }
}

```

```

    }
    printf("Result saved in %s\n", output);
    printf("Total accuracy: %.2lf%%\n", total_accuracy * 100);
    printf("From Database: %s\n", input);

}

#endif //INC_5__HEAD_H

<main.c>
#include "head.h"

int main(int args, char *argv[]) {
    setbuf(stdout, NULL);

    if (inputs(argv[1]) == False) {
        return -1;
    }
    init_cluster();
    KMeans();
    verify(argv[1],argv[2]);

    return 0;
}

```