



人工智能基础与进阶

聚类与寻优算法实验

上海交通大学

第七节

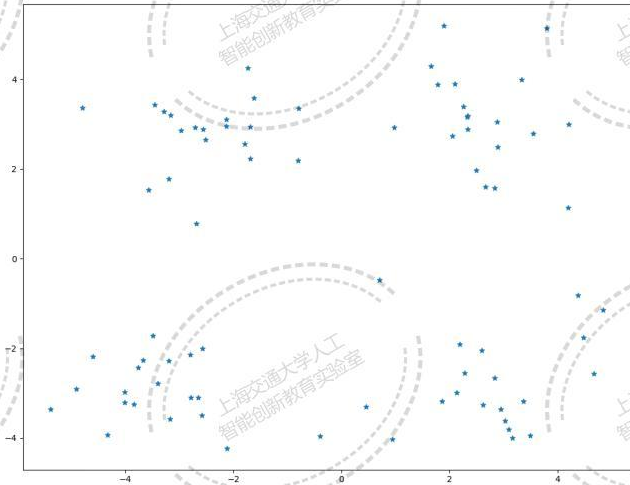
机器学习相关代码



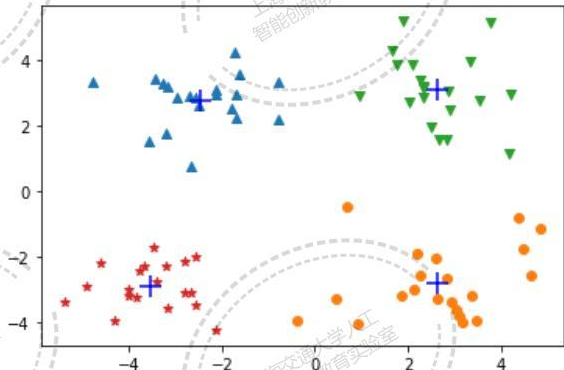
K-means 聚类—代码

```
04-K-means > 04-Kmeans.py > kMeans > dataSet
1 import numpy as np
2 import matplotlib.pyplot as plt
3
4 #读取数据
5 > def loadDataSet(fileName): #general function to parse tab -delimited floats...
15
16 #计算欧几里得距离
17 > def distEclud(vecA, vecB): ...
18
19 #计算曼哈顿距离
20 > def distMhadun(vecA, vecB): ...
21
22 # 初始化簇中心, 取k个簇中心
23 > def randCent(dataSet, k): ...
24
25 #实现k-均值聚类算法
26 > def kMeans(dataSet, k, distMeas=distEclud, createCent=randCent): ...
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53 K = 4 #设置聚类中心数
54 dataMat = np.mat(loadDataSet('testSet.txt')) #读取数据
55 markers=['^', 'o', 'v', '*', 'x']
56 centroids, clusterAssment = kMeans(dataMat, K, distMeas=distMhadun, createCent=randCent) # 调用k-均值聚类
57
58 for k in range(0,K):
59     x = []
60     y = []
61     for i in range(0, dataMat.shape[0]):
62         if int(clusterAssment[i, 0])==k:
63             x.append(dataMat[i, 0])
64             y.append(dataMat[i, 1])
65     x = np.array(x)
66     y = np.array(y)
67     plt.scatter(x,y,marker=markers[k])
68     plt.scatter(centroids[k,0],centroids[k,1], marker='+', s=72, c='b')
69
70 plt.show()
```

K-means 聚类—原始数据分布



K-means 聚类—运行结果



寻优算法

梯度下降算法

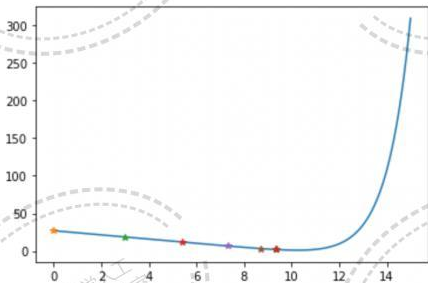
```
20 def tidu(err = 1e-5):
21     x = 0
22     cnt = 0
23     ys = [f(x) for x in xs]
24     plt.figure("梯度下降")
25     # plt.show()
26     plot(xs,ys)
27     while(np.abs(df(x)) > err):
28
29         plt.plot(x,f(x),"*")
30         x -= df(x)
31         cnt+=1
32         plt.pause(0.5) # 暂停0.5秒
33         plt.ioff() # 关闭画图的窗口
34     print(cnt)
```

牛顿法

```
35 def niudun(err = 1e-5):
36     x = 0
37     cnt = 0
38     ys = [f(x) for x in xs]
39     plt.figure("牛顿法")
40     # plt.show()
41     plot(xs,ys)
42     while(np.abs(df(x)) > 1e-5):
43         # plot(xs,ys)
44         plt.plot(x,f(x),"*")
45         x -= df(x)/ddf(x)
46         cnt+=1
47
48         plt.pause(0.5) # 暂停0.5秒
49         plt.ioff() # 关闭画图的窗口
50     print(cnt)
```

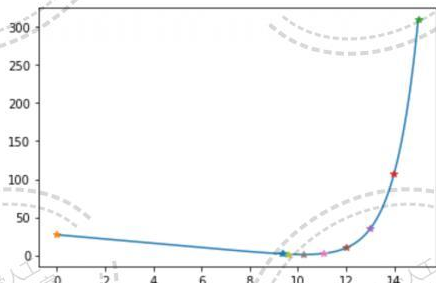
寻优算法——运行结果

梯度下降算法



误差: $1e-05$
迭代次数: 13

牛顿法



误差: $1e-05$
迭代次数: 10



上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

谢谢聆听

THANKS FOR YOUR ATTENTION

上海交通大学人工
智能创新教育实验室