

## 9181040G0818 黄海浪 机器学习第 7 次作业

Svmlib

介绍: <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>

下载: <https://github.com/cjlin1/libsvm>

通过 python 直接安装 `pip install libsvm`

核心代码:

```
# 输入数据
problem = svm_problem(dataY, dataX, isKernel=True)
# 采用多项式核函数-t 1,5 倍交叉验证模式-v 5
options = svm_parameter('-s 0 -t 1 -v 5')
# 训练
model = svm_train(problem, options)
```

得到的结果:

准确率: **79.7468%**

输出:

```
*
optimization finished, #iter = 25
nu = 0.637976
obj = -34.494488, rho = 1.912918
nSV = 42, nBSV = 39
Total nSV = 42
*
optimization finished, #iter = 29
nu = 0.708004
obj = -39.696073, rho = 1.842492
nSV = 46, nBSV = 44
Total nSV = 46
*
optimization finished, #iter = 21
nu = 0.634921
obj = -34.694327, rho = 1.856937
nSV = 40, nBSV = 40
Total nSV = 40
*
optimization finished, #iter = 26
nu = 0.643034
obj = -35.503476, rho = 1.873430
nSV = 42, nBSV = 40
Total nSV = 42
*
```

```
optimization finished, #iter = 24  
nu = 0.700000  
obj = -39.339908, rho = 1.847530  
nSV = 46, nBSV = 44  
Total nSV = 46  
Cross Validation Accuracy = 79.7468%
```

附录:

```
#!/usr/bin/python3
# -*- coding: utf-8 -*-
# @time    : 2020/11/14 19:56
# @author  : lerogo
# @fileName: main.py

# 导入包
from libsvm.svm import svm_problem, svm_parameter
from libsvm.svmutil import svm_train
import numpy as np
import pandas as pd

# 从 data 读取数据
def readData():
    data_x = pd.read_table("./data/ex4x.dat", sep=' ').values
    data_y = pd.read_table("./data/ex4y.dat", sep=' ').values
    x1 = []
    x2 = []
    x0 = []
    y = []
    for i in range(len(data_y[:, 3])):
        x0.append(1.0)
        x1.append(float(data_x[:, 3][i]))
        x2.append(float(data_x[:, 6][i]))
        y.append(int(data_y[:, 3][i]))
    dataX = np.array([x0, x1, x2])
    dataY = np.array(y)
    return dataX.T, dataY.T

# 归一化
def normalized(para):
    para = (para - para.min()) / (para.max() - para.min())
    return para

if __name__ == '__main__':
    dataX, dataY = readData() # 初始化训练数据
    # 归一化
    dataX[:, 1] = normalized(dataX[:, 1])
    dataX[:, 2] = normalized(dataX[:, 2])

    # 输入数据
```

```
problem = svm_problem(dataY, dataX, isKernel=True)
# 采用多项式核函数-t 1,5 倍交叉验证模式-v 5
options = svm_parameter('-s 0 -t 1 -v 5')
# 训练
model = svm_train(problem, options)
```