

# 机器学习

## 第7次实验：使用SVM对手写体数字图片分类

# 实验目的

---

01

会用Python创建多分类SVM神经模型

02

使用多分类SVM模型对手写体数字图片分类

03

会对分类结果进行评价

# 实验内容

## 使用sklearn.svm类对手写体数字图片进行分类

- ✓ 先对1类和2类的数据进行计算，得到1类和2类的分割超平面，然后对2类和3类的数据进行计算，以此类推，直到两两类别分别完成计算
- ✓ 在用模型做预测时，对两两类别之间的分割超平面分别进行匹配，统计有多少次判别将其误分类1类，多少次判别为2类...，判定所属类别次数最多的就是最后预测的类别

**训练数据: digits\_training.csv**

**测试数据: digits\_testing.csv**

**第1列是类别，其他列是特征属性**

# 实验内容

1

载入训练数据、分出特征属性和类别，对特征属性标准化，显示读入数据的行数

```
xTrain = trainData[:,1:NTrain]  
yTrain = trainData[:,0]
```

标准化函数

```
def normalizeData(X):  
    return (X - X.mean())/X.max()
```

# 实验内容

2

训练多分类SVM模型

```
model = svm.SVC(decision_function_shape='ovo')
```

3

保存训练模型名字为 “svm\_classifier\_model1.m”

4

载入测试数据，分出特征属性和类别，对特征属性标准化，显示读入数据的行数

5

使用模型对测试集进行预测，显示预测错误数据的数目、预测数据的准确率和模型内建正确率

# 执行结果图例

1. 载入训练数据，对数据进行标准化处理.....

训练数据： 5000 条

2. 训练模型.....

3. 保存模型...

4. 加载测试数据，对数据进行标准化...

测试数据： 500 条

5. 加载训练好的模型，进行预测...

预测错误数据： 25 条

测试数据正确率： 0.95

模型内建的正确率估计： 0.95