

## 实验 3: softmax 回归

### 一、实验目的:

- 1、理解 softmax 回归和交叉熵损失函数的概念。
- 2、掌握利用神经网络解决多分类问题。
- 3、观察拟合后得到的模型并观察预测的结果，并计算准确率。

### 二、实验内容:

#### 1、聚类 and 分类

通过 sklearn 库提供的聚类算法生成 K 类数据，以这些数据作为数据集训练神经网络，利用 softmax 层和交叉熵损失函数对数据进行分类。聚类参数要求  $k>3$ , 数据样本不少于 1000, 其它参数参考课件。对聚类后的数据按 9:1 的原则划分训练集和测试集，利用在训练集上训练得到的模型对测试集上的数据进行验证，要求模型准确率不低于 99%。

#### 2、鸢尾花分类

iris 数据集包含 150 个样本，对应数据集的每行数据。每行数据包含每个样本的四个特征和样本的类别信息，iris 数据集是用来给鸢尾花做分类的数据集，每个样本包含了花萼长度、花萼宽度、花瓣长度、花瓣宽度四个特征，请用神经网络训练一个分类器，分类器可以通过样本的四个特征来判断样本属于山鸢尾、变色鸢尾还是维吉尼亚鸢尾。数据集文件 iris.csv。要求模型准确率不低于 99%。



### 三、实验准备

- 1、下载和安装 Anaconda, Pytorch 和 cuda 库
- 2、下载和安装 Jupyter Notebook 或 Pycharm 或 Vs code

#### 四、实验指导

参考课堂资料和课件