实验报告

1. **源码运行环境**

Win10系统下Python3.5，需安装tensorflow,matplotlib,numpy,pandas,sklearn

等库。

**2．任务定义**

波士顿房价 线性回归（20分）

波士顿房价 SVM回归（30分）

手写数字识别 前馈神经网络（30分）

手写数字识别 卷积神经网络（30分）

**3.波士顿房价问题**

输入输出：输入数据包括506个样本，每个样本包括12个特征变量和该地区的平均房价。输出的是在某特征下的对房价的预测。

**3.1 线性回归**

**3.1.1方法描述**

由于房价和多个特征变量相关，所以使用多元线性回归建模。其公式如下：



房价预测结果由不同特征的输入值和对应的权重相乘求和，加上偏置项计算求解。

具体步骤：

(a)读取数据

(b)数据预处理：归一化，打乱数据顺序

考虑到不同特征值取值范围大小不同，会有影响，所以要对数据进行归一化。记特征值为x,最小特征值为，最大特征值为，归一化后的结果为，则归一化公式为：



打乱数据顺序是为了防止过拟合。

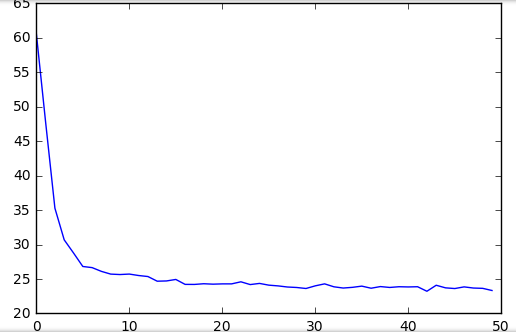
(c)定义模型：包括定义输入输出，模型结构与命名空间等。

(d)模型训练：这里设置训练轮次为50，学习率为0.01，定义均方差损失函数，选择梯度下降优化器。

(e)模型应用：可视化损失函数，对房价进行预测。

**3.1.2结果分析**

训练过程中损失函数图像如下所示：



随机测试一个数据，其结果如下：



随训练的进行，损失不断减小，最后收敛至25左右。

**3.2 SVM回归**

**3.2.1方法描述**

（1）数据预处理

先用train\_test\_split切割出70%的训练集和30%的测试集。由于该数据集各样本取值范围差异很大，直接将数据输入到SVM中的话，学习将会变得困难且容易被噪声干扰。解决方法是对每个特征做标准化或归一化或正则化，本次采用sklearn库自带的z\_值标准化。

（2）模型训练和评估

分别使用线性核，高斯核，sigmod核，多项式核进行训练。

**3.2.2结果分析**

线性核函数：

训练集评分： 0.748439055371

测试集评分： 0.602628182275

测试集均方差： 0.351696191687

测试集R2分： 0.602628182275

高斯核函数：

训练集评分： 0.968482433896

测试集评分： 0.845455687936

测试集均方差： 0.136780324058

测试集R2分： 0.845455687936

sigmoid核函数：

训练集评分： -57.7870855639

测试集评分： -62.8185590217

测试集均方差： 56.48297933

测试集R2分： -62.8185590217

多项式核函数：

训练集评分： 0.91701495815

测试集评分： 0.790786837991

测试集均方差： 0.18516530123

测试集R2分： 0.790786837991

从结果上看，模型评分：高斯> 多项式>线性>sigmoid

**4.Mnist手写数字识别**

输入输出：输入的是一张图片，利用模型进行预测，输出的是图片上的数字。

**4.1前馈神经网络(FNN)**

**4.1.1方法描述**

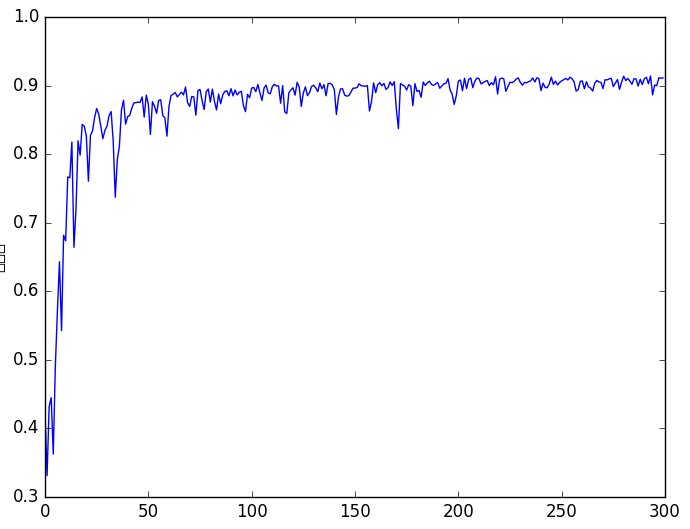
在网络中，输出层上节点的值（输出值）通过输入值乘以权重值直接得到。取出其中一个元进行讨论，其输入到输出的变换关系为



对单张图片而言，这里x是[1,784]的数组，w是[784,10]的数组，b是[1,10]的数组，预测值的输出是[1,10]的数组，用one-hot编码表示，即index为对应数字时该值为1，其余值为0.如9的one-hot编码为[0, 0, 0, 0, 0, 0, 0, 0, 0,1].

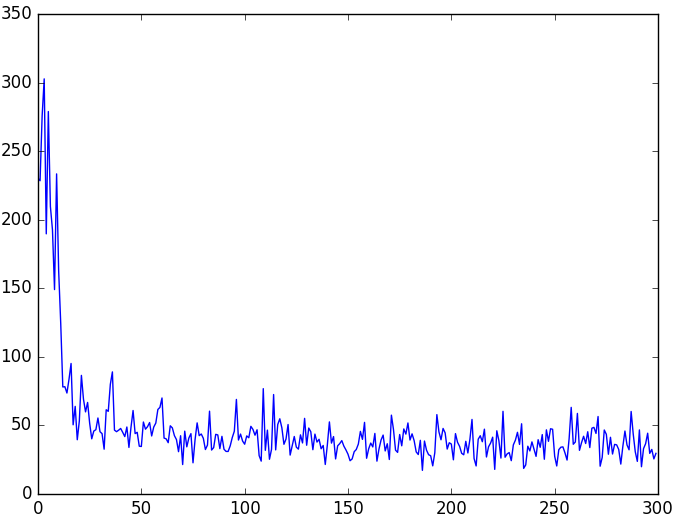
用梯度下降法对网络进行训练，是交叉熵损失函数达到最小，通过反向传播更新网络参数。

**4.1.2结果分析**



网络训练过程中准确率如上图所示，可以看到，准确率最终收敛到0.9左右.

网络训练过程中交叉熵损失函数如下图所示，可以看到，损失最终收敛到40左右.



**4.2卷积神经网络(FNN)**

**4.1.1方法描述**

卷积神经网络模型采用两个卷积层和全连接层构成，第一卷积层为32个大小为5\*5\*1的卷积核，然后通过2\*2的最大值池化下采样。第二个卷积层为64个5\*5\*32的卷积核，2\*2最大值池化下采样。最后连接到全连接层，通过softmax输出分类。

为了避免模型过拟合，加入dropout。dropout给隐藏层的神经元加上概率为keep\_prob的失活率，从而同时训练出指数规模个共享权值的子网络用于分类，增加模型的鲁棒性。

采用随机批梯度下降法使损失函数达到最小，取每100个训练数据的均值来迭代更新一次权值，有利于模型更好的收敛到最小值。

**4.2.2结果分析**



最终在测试集识别准确率达到0.977，效果较好。