Lecture9作业

李星毅 U201712072 自实1701

1，PPT中的作业题

（1） ③

（2） ④

（3） ④

2，对于一幅300\*300大小的彩色（RGB）图像，（1）如果输入端与有100个神经元的第一层隐含层用全链接方式（Fully Connected neural Network）连接时，请问这一层会包含多少参数？（2）如果用100个5\*5大小的滤波器作卷积操作，那么这一层的参数为多少？如果滤波器移动步长（stride=1）为1，经过卷积计算后的输出端神经元个数有多少？

解：这里需要讨论每一层是否包含偏置。

（1）300\*300大小的彩色（RGB）图像的size是300\*300\*3，因此输入层神经元个数为300\*300\*3+1=270000+1=270001，则这一层会包含的参数个数为270001\*100=27000100。

（2）记输入图像尺寸为，输出图像尺寸为，步长（Stride）为，滤波器大小为，滤波器个数为，填充（Padding）为，则有



这里假设Padding=0。

100个5\*5大小的滤波器对一幅300\*300大小的彩色（RGB）图像作卷积操作，输出图像尺寸为，但这一层包含的参数个数为(5\*5\*3+1)\*100=7600。而经过卷积计算后的输出端神经元个数有296\*296\*100=8761600。

3，有训练样本集为： 假设某神经网络结构为第一层有两个神经元，第二层有三个神经元，第三层有一个神经元，前两层每个神经元的激活函数为ReLU（即，这里代表第l层第d个神经元的输入，代表该神经元的输出），第三层为线性输出，即。误差函数为：，学习率为0.01。假设初始权系数矩阵定义如下：

，，

其中w的下标0代表迭代次数为0（即初始状态），上标数字分别代表第1、2、3层。要求将上述训练样本集的样本用反向传播法按顺序进行一轮训练，写出每一次迭代时各层的权系数矩阵，即：t=1时，进入样本，得到；t=2时，进入样本，得到；t=3时，进入样本，得到；t=4时，进入样本，得到

解：首先推导单个样本反向传播算法的矩阵形式，以便书写，可轻易推广到mini-batch的矩阵形式。其中记表示对向量或矩阵增广，表示对向量或矩阵去增广（考虑到偏置的影响），表示Hadamard Product，即对应元素相乘。

损失函数为：



则有：



综上所述，神经网络算法总结如下：

| **Algorithm** **1** Backpropagation using ReLU with Stochastic Gradient Descent |
| --- |
| Initialize  For   1. Forward Propagation      1. Backpropagation   Compute    Update by    until  or enough iterations  return |

下面将所推到的算法应用于所给的数据集中：

（1） 时，，，先进行前向传播：



下面进行反向传播更新权重：



（2）时，，，先进行前向传播：



下面进行反向传播更新权重：



（3）时，，，先进行前向传播：



下面进行反向传播更新权重：



（4）时，，，先进行前向传播：

下面进行反向传播更新权重：



4，在下表中填写前四周学过的算法：

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | 线性分类 | 非线性分类 |
| 监督模式识别 | 二分类 | 1. 感知器算法 2. 线性回归 3. Logistic回归 4. Fisher线性判别 5. 支撑向量机 6. 神经网络、卷积神经网络 | 1. 利用非线性变换 2. 支撑向量机 3. 神经网络、卷积神经网络 |
| 多分类 | 1. OVA & OVO 2. Softmax回归 3. 神经网络、卷积神经网络 | |
| 无监督模式识别 | | 1. K-Means聚类 2. 层次聚类（HAC） 3. 主成分分析（PCA） 4. 自动编码器（Autoencoder） | |