

# 《数据库应用》心得与体会

BY 方朔

13120031

2014.11.25

## 目录

1 课程心得	2
2 神经网络简介	1
3 神经网络的基本结构	?
4 神经网络的训练	?

## 1 课程心得

第一次上覃老师的《数据 用》这门 程时，本以 会主要是以IBM DB2数据 用 主，但是覃老师却提出我 分配多一点时间 我 数据挖掘相 的知 ，十分高兴。因 在之前学校的 程中学 的大多是非常基 的知 而像数据挖掘这种处在前沿的技术没有办法从学校获得 在上完《数据 用》这门 之后 我也是获益匪浅。

在上这门 之前，我就了解过一些 于机器学 方面的知 ，并且在Coursera上学 了斯坦福大学吴恩达教授的机器学 程，并对于 上学到的很多算法十分感兴趣，尤其是神经网络 。而由于机器学 与数据挖掘在 多方面有着共同点，所以我的这篇心得体会主要叙述一下 于神经网络 的内容，并且有一个手写 别的样例作 神经网络 的用。

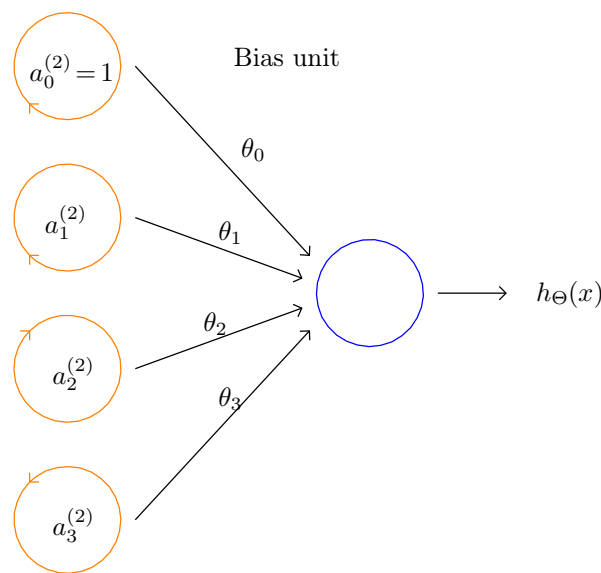
## 2 神经网络简介

人工神经网络（Artificial Neural Network, ANN），称神经网络，是一种模仿生物神经网络 的 构和功能的数学模型或 算模型。神经网络 由大量的人工神经元联 行 算。大多数情况下人工神经网络 能在外界信息的基 上改变内部 构，是一种自适 系统。 代神经网络 是一种非 性统 性数据建模工具，常用 对输入和输出间复 的系 行建模 或用 探索数据的模式。

单 神经网络 就是通过大量数据的 改变自身的 构 对未知数据 行 或者 别。

3 神经网络的基本结构

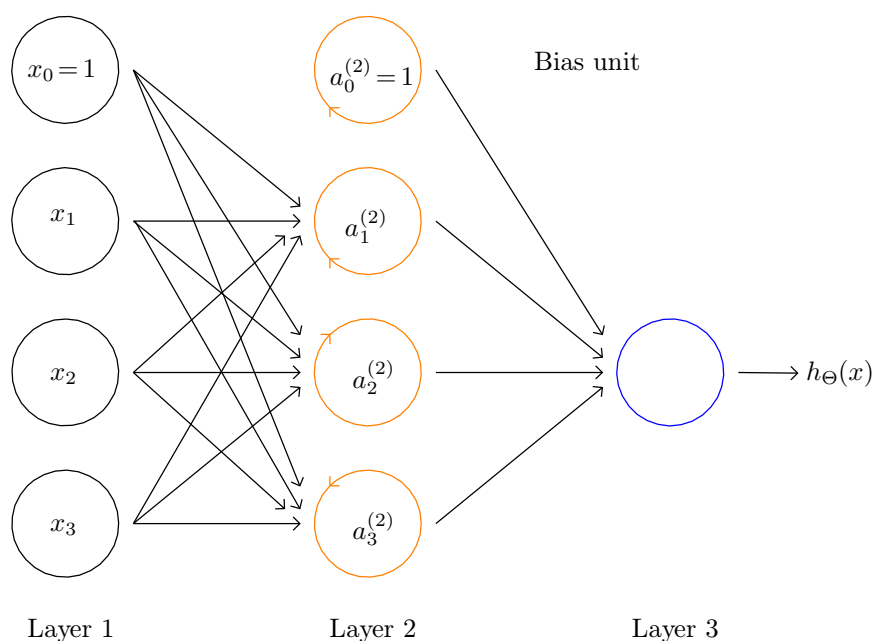
神经网络是由基本的单位神经元构成，神经元的基本结构如下所示。



- 向量  $a$  输入向量
- 向量  $\theta$  神经元的值
- $h_{\theta}(x)$  输出 一般使用符号函数

偏置单元作  $a_0=1$ ，偏置单元作 算时的常量，这样可以使神经元的输出表示 向量的乘，然后经过 递函数 算之后作 输出 便于矩 运算。

神经网络 由输入单元、输出单元以及多个 藏 成，每个 藏 由多个神经元构成，一个只有一 藏的神经网络 表示如下。



上图中Layer 1表示输入向量，Layer2 隐藏，Layer3 输出。将每个隐藏点的值均放入矩阵  $\theta$ ，以便矩阵运算，对于上述神经网络的计算过程如下。

- $a_i^{(j)}$  = 第  $j$  层的第  $i$  个单元的输出
- $\Theta^{(j)}$  = 值映射矩阵，从  $j$  到  $j+1$

$$\begin{aligned}
 a_1^{(2)} &= g\left(\Theta_{10}^{(1)}x_0 + \Theta_{11}^{(1)}x_1 + \Theta_{12}^{(1)}x_2 + \Theta_{13}^{(1)}x_3\right) \\
 a_2^{(2)} &= g\left(\Theta_{20}^{(1)}x_0 + \Theta_{21}^{(1)}x_1 + \Theta_{22}^{(1)}x_2 + \Theta_{23}^{(1)}x_3\right) \\
 a_3^{(2)} &= g\left(\Theta_{30}^{(1)}x_0 + \Theta_{31}^{(1)}x_1 + \Theta_{32}^{(1)}x_2 + \Theta_{33}^{(1)}x_3\right) \\
 h_{\Theta}(x) &= a_1^{(3)} = g\left(\Theta_{10}^{(2)}a_0^{(2)} + \Theta_{11}^{(2)}a_1^{(2)} + \Theta_{12}^{(2)}a_2^{(2)} + \Theta_{13}^{(2)}a_3^{(2)}\right)
 \end{aligned}$$

这就是前向传播算法，如果用矩阵运算表示前向传播算法，那么对于一个有两层隐藏层的神经网络，他的矩阵算法如下。

#### 算法 前向传播算法

输入向量  $x$ :

$$\begin{aligned}
 a^{(1)} &= x \\
 z^{(2)} &= \Theta^{(1)}a^{(1)} \\
 a^{(2)} &= g(z^{(2)}) \quad (\text{add } a_0^{(2)}) \\
 z^{(3)} &= \Theta^{(2)}a^{(2)}
 \end{aligned}$$

$$\begin{aligned}a^{(3)} &= g(z^{(3)}) \quad (\text{add } a_0^{(3)}) \\z^{(4)} &= \Theta^{(3)} a^{(3)} \\a^{(4)} &= h_{\Theta}(x) = g(z^{(4)})\end{aligned}$$

## 4 神经网络的训练

神经网络 采用后向 传播算法 在介 后向 传播算法之前 先 看一下神经网络 的花 函数。

Cost function:

$$J(\Theta) = -\frac{1}{m} \left[ \sum_{i=1}^m \sum_{k=1}^K y_k^{(i)} \log(h_{\Theta}(x^{(i)}))_k + (1 - y_k^{(i)}) \log(1 - h_{\Theta}(x^{(i)}))_k \right] + \frac{\lambda}{2m} \sum_{l=1}^{L-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (\Theta_{ji}^{(l)})^2$$

函数的右边 Regularization，目的是避免神经网络 陷入过拟合的情况，这里相当于对参数 行惩罚，注意不对参数 中的偏置 行惩罚。

函数的左边  $m$  个输入  $k$  向量的情况下 根据符号函数 算的。

符号函数的定义如下

$$g(z) = \frac{1}{1 + e^{-z}} (\text{Logistic / Sigmoid Function})$$

那么对于一个符号函数 他的花 函数  $J(\theta)$  然后根据符号函数的性 将  $J(\theta)$  化

$$\begin{aligned}J(\theta) &= \frac{1}{m} \sum_{i=1}^m \text{Cost}(h_{\theta}(x), y) \\ \text{Cost}(h_{\theta}(x), y) &= \begin{cases} -\log(h_{\theta}(x)) & \text{if } y=1 \\ -\log(1-h_{\theta}(x)) & \text{if } y=0 \end{cases} \\ J(\theta) &= \frac{1}{m} \sum [-y \log(h_{\theta}(x)) - (1-y) \log(1-h_{\theta}(x))]\end{aligned}$$

有了花 函数，

### 算法 后向传播算法

Training set  $\{(x^{(1)}, y^{(1)}), \dots, (x^{(m)}, y^{(m)})\}$

Set  $\Delta_{ij}^{(l)} = 0$  (for all  $l, i, j$ ). (used to compute  $J(\Theta)$ )

For  $i = 1$  to  $m$

Set  $a^{(1)} = x^{(i)}$

Perform forward propagation to compute  $a^{(l)}$  for  $l = 2, 3, \dots, L$

Using  $y^{(i)}$ , compute  $\delta^{(L)} = a^{(L)} - y^{(i)}$

Compute  $\delta^{(L-1)}, \delta^{(L-2)}, \dots, \delta^{(2)}$

$\Delta_{ij}^{(l)} = \Delta_{ij}^{(l)} + a_j^{(l)} \delta_i^{(l+1)}$

$$\begin{aligned}D_{ij}^{(l)} &= \frac{1}{m} \Delta_{ij}^{(l)} + \lambda \Theta_{ij}^{(l)} \text{ if } j \neq 0 \\ D_{ij}^{(l)} &= \frac{1}{m} \Delta_{ij}^{(l)} \text{ if } j = 0 \\ \frac{\partial}{\partial \Theta_{ij}^{(l)}} J(\Theta) &= D_{ij}^{(l)}\end{aligned}$$

上面算法中的  $\delta$  的 算方法如下。

算法 后向传播算法中 $\delta$ 的计算

对第四

$$\begin{aligned}\delta^{(4)} &= a^{(4)} - y \\ \delta^{(3)} &= (\Theta^{(3)})^T \delta^{(4)} .* g'(z^{(3)}) \\ \delta^{(2)} &= (\Theta^{(2)})^T \delta^{(3)} .* g'(z^{(2)})\end{aligned}$$