今天的线下讨论中，我们对之前已经学习过的 监督学习算法 进行了总结。其中神经网络部分的讨论由我来整理并放在论坛上。 内容基本会分为以下几个部分。

感知机简单描述

感知机的缺点

Sigmoid function 和 Logistic regression

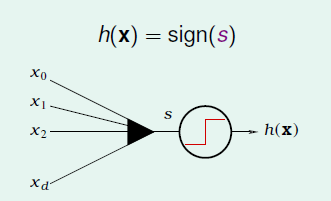
神经网络的简单描述，优缺点

神经网络的一点扩展话题

下面的陈述中有些是源于我个人的理解，犯了错误大家伙帮我指出来呀。用到的图大部分都是从 台大 林轩田老师 和 斯坦福 吴恩达 老师 在coursera上课程的ppt 里抠出来的。

## 1. 感知机

在课程中我们首先学到了感知机（perceptron）。感知机的想法来源自 真实生物系统中的神经元——感知机单元接收到很多信号，并对各路信号加权求和，与设定的threshold 比较，  


简单的图示为：  


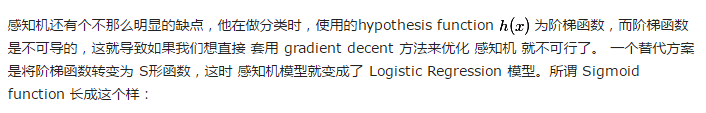
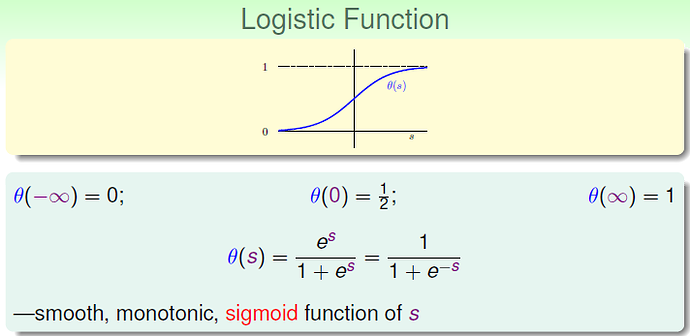


感知机的优点在于实现和训练都非常简单，并且理论上可以证明，如果问题是线性可分的（线性可分的定义已经在之前的帖子里面讨论过了），感知机 一定可以收敛。

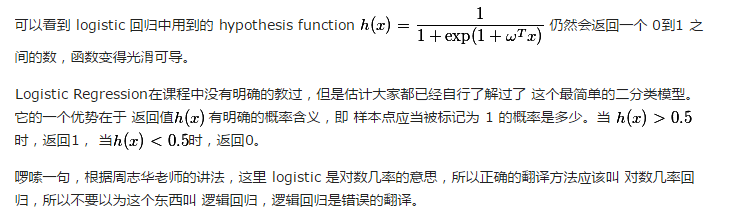
## 1.5 感知机的缺点

感知机算法有个明显的缺点，就是只能处理线性可分问题。对于不线性可分问题，一个感知机单元是永远不会收敛的。 回忆下课程中提过的 异或问题，这么简单的逻辑问题感知机都无法实现。

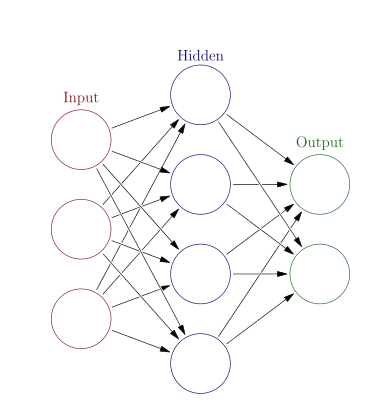
## 2 Sigmoid Function 与 Logistic Regression

[](http://cn-discussions.s3.cn-north-1.amazonaws.com.cn/original/3X/0/7/074a0a01641629142fc3feb754d12e66756ac9c4.png" \o "Sigmoid function.png)

**[Sigmoid function.png712x347 17.8 KB](http://cn-discussions.s3.cn-north-1.amazonaws.com.cn/original/3X/0/7/074a0a01641629142fc3feb754d12e66756ac9c4.png" \o "Sigmoid function.png)**



## 3. 神经网络

前面已经提到过，感知机的一大弱点就是，不能处理非线性分类问题。 而如果把一堆感知机串起来，这些‘神经元’ 就可以很好的完成 非线性的分类问题了，这就是 神经网络（Neuron Network）了。一个最简单的单隐层神经网络大概长成这个样子：  


其中隐层的每个单元代表一个 activation function，函数形式可以是sigmoid function 也可以是别的非线性函数。而每条边都对应 一个weight。

计算神经网络的输出，使用正向传播算法。而神经网络的训练使用 反向传播算法（BP）。原理很简单，就是链式法则求导，但是推导实在是恶心到吐，所以这里就不重复了。

## 3.5 神经网络的优缺点

神经网络的优势非常明显，层数可以很多，每一个weight都是拟合参数，所以神经网络可以表示的 hypothesis space大小是非常惊人的，所以能处理非常复杂的分类问题。

神经网络也有缺点，根据**[@mrbean](http://discussions.youdaxue.com/users/mrbean)** 提醒，神经网络的可解释性非常差，它能返回很好的分类结果，但是，人类并不知道它是怎么做出来的。每一个节点上究竟代表什么，人类一无所知。而且，神经网络的训练需要的运算量也是很大的，面对层数极多的神经网络（Deep Neuron Network），必须用到GPU并行计算。

## 4 神经网络的一点拓展讨论

深度学习目前被炒的火热，用到的就是很多层的神经网络模型（DNN），以及 神经网络算法的一些 变形版本，比如用在 图像识别上的卷积神经网络（CNN），以及用在语义识别 上的 循环神经网络（RNN）。感兴趣的同学们可以自行前往 Udacity 上的 deep learning 学习。

神经网络中每个中间层节点上使用的 激活函数其实不只有 sigmoid function 一种选择，只要能引入非线性，很多函数都是可以的，比如 tanh，RBF，还有做deep learning的人喜欢用的ReLu。