Logistics回归相关的内容可以参考之前的两篇笔记《分类问题的两种模型》和《Logistics回归与线性回归对比》，Logistics回归相对于线性回归，有比较好的表现，可以弥补线性回归的一些不足，但是Logistics回归也有很强的局限性，我们下面分析它的局限性并提出一种改进的方法，从而引出神经网络的概念。

**1** **回归的局限性**

考虑下面一种情况，如表1所示，是训练数据的特征向量（2维），，代表训练数据的类别。用坐标图表示出来如右图所示：



表1 训练数据及Label

|  |  |  |
| --- | --- | --- |
|  |  | Label |
| 0 | 0 | Class2 |
| 0 | 1 | Class1 |
| 1 | 0 | Class1 |
| 1 | 1 | Class2 |

图1 训练数据分布

如果用Logistics回归对上面的数据做分类的话（如图2），能否做到？



图2 Logistics回归训练

显然不能，因为Logistics的分类面是一条直线，无论你怎么改变参数都无法将红色和蓝色的点分开，如图3。



图3 无法用线性分类面分开不同的类

如果一定要用Logistics方法做分类，可以通过特征变换的方法让其线性可分，其实这个和前面讲的中的非线性特征变换类似，都是为了将不可分的特征变为可分的特征。

**2 特征变换(Feature Transform)**

假设变换后的特征为，代表到的距离，代表到的距离，则特征变换后的特征如图4所示：



图4 特征变换后的数据分布

可以看出，经过特征变换后可以用Logistics方法分类出不同的点，但是问题的关键在于：**要找到一个好的变换式(transformation)是非常困难的。**所以我们希望“找一个好的变换式”这件事情由机器自动完成。

**3 神经网络(Neural Networks)**

我们提到“找一个好的变换式”这件事情希望由机器完成，那机器如何能完成呢？其实**只需要将Logistics regression的模型级联起来**，就可以完成这件事情。



图5 Logistics regression model级联

图5中的2个Logistics regression model可以完成变换特征向量到可分空间这件事情。原因如下：

原数据的特征向量总是可以通过调整下式参数得到使得原本线性不可分的数据线性可分，如图4所示，所以图5中前面的框图可以完成变换特征向量。



后面框图中的Logistics regression model则是完成Logistics分类任务，我们把这样多个Logistics regression model级联起来的模型网络叫做神经网络(Neural Networks，如图6所示)，它可以完成逻辑运算中的与、或、非和异或四种运算，即对线性或非线性的运算都是可以做的，或者说理论上可以完成各种情况的分类，只要你的Neural Networks够多。



图6 Neural Network



2017.08.16.