

首先，说下多类分类和多标签分类的区别

多标签分类： 一个样本可以属于多个类别（或标签），不同类之间是有关联的，比如一个文本被划分成“人物”和“体育人物”两个标签。很显然这两个标签不是互斥的，而是有关联的

多类分类： 一个样本属于且只属于多个分类中的一个，一个样本只能属于一个类，不同类之间是互斥的，比如一个文本只能被划分成“人物”，或者被划分成“文化”，而不能同时被划分成“人物”和“文化”，“文化”和“人物”这两个分类就是互斥的

那么，如何用softmax和sigmoid来做多类分类和多标签分类呢？

## 1、如何用softmax做多分类和多标签分类

现假设，神经网络模型最后的输出是这样一个向量`logits=[1, 2, 3, 4]`，就是神经网络最终的全连接的输出。这里假设总共有4个分类

用softmax做多分类的方法：

```
tf.argmax(tf.softmax(logits))
```

首先用softmax将logits转换成一个概率分布，然后取概率值最大的作为样本的分类

这样看似乎，`tf.argmax(logits)`同样可以取得最大的值，也能得到正确的样本分类，这样的话softmax似乎作用不大

那么softmax的主要作用其实是在计算交叉熵上，首先样本集中`y`是一个one-hot向量，如果直接将模型输出logits和`y`来计算交叉熵，

因为`logits=[1, 2, 3, 4]`，计算出来的交叉熵肯定很大，这种计算方式不对，而应该将logits转换成一个概率分布后再来计算，

就是用`tf.softmax(logits)`和`y`来计算交叉熵，当然我们也可以直接用tensorflow提供的方法`softmax_cross_entropy_with_logits`来计算

这个方法传入的参数可以直接是logits，因为这个根据方法的名字可以看到，方法内部会将参数用softmax进行处理

现在我们取的概率分布中最大的作为最终的分类结果，这是多分类

我们也可以取概率的top几个，作为最终的多个标签，或者设置一个阈值，并取大于概率阈值的。这就用softmax实现了多标签分类

## 2、如何用sigmoid做多标签分类

sigmoid一般不用来做多类分类，而是用来做二分类的

它是将一个标量数字转换到 $[0, 1]$ 之间，如果大于一个概率阈值(一般是0.5)，则认为属于某个类别，否则不属于某个类别

那么如何用sigmoid来做多标签分类呢？其实就是针对logits中每个分类计算的结果分别作用一个sigmoid分类器，分别判定样本是否属于某个类别

同样假设，神经网络模型最后的输出是这样一个向量 $\text{logits}=[1, 2, 3, 4]$ ，就是神经网络最终的全连接的输出。这里假设总共有4个分类

```
tf.sigmoid(logits)
```

sigmoid应该会将logits中每个数字都变成 $[0, 1]$ 之间的概率值，假设结果为 $[0.01, 0.05, 0.4, 0.6]$ ，然后设置一个概率阈值，比如0.3，如果概率值大于0.3，则判定类别符合，那这里，样本会被判定为类别3和类别4都符合。