交叉熵损失函数计算案例

假设有一个3分类问题，某个样例的正确答案是(1,0,0)

甲模型经过softmax 回归之后的预测答案是(0.5,0.2,0.3)

意思就是是第一类的概率是0.5，是第二类的概率是0.2，是第三类的概率为他0.3

乙模型经过softmax 回归之后的预测答案是 (0.7, 0.1, 0.2)

那么怎么判断哪个模型更好呢？

用交叉熵来判断其结果：

H(p,q) = -

P(X)表示正确答案，q(x)表示预测值

将上面的数据带入公式：

H((1,0,0), (0.5,0.2,0.3)) = -( 1 \* log0.5 + 0 \* log0.2 + 0 \* log0.3) = -log0.5 ≈ 0.301

H((1,0,0), (0.7, 0.1, 0.2)) = -(1 \* log0.7 + 0 \* log0.1 + 0 \* log0.2) = -log0.7 ≈ 0.155

交叉熵越小说明越接近正确答案，说明乙模型较好。

根据上面的说明，定义交叉熵的损失函数

Loss = - 其中 为标签值， 为预测值

用tensorflow表示就是

loss\_function = tf.reduce\_mean(-tf.reduce\_sum(y \* tf.log(pred), reduction\_indices=1))

reduce\_sum就是求和，y 是标签值，pred是预测值

reduce\_mean 是求平均值