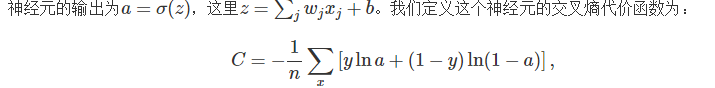
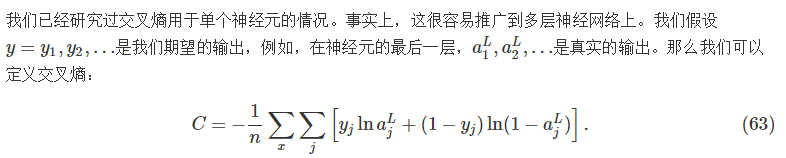
# 一、Loss损失函数

## binary\_crossentropy（亦称作对数损失，logloss）



事实上，如果输出神经元是sigmoid神经元的话，交叉熵都是更好的选择。而且，在自编码器中，即使原数据不是0,1的二值数据，但是是[0,1]之间的标准化数据，也可以用sigmoid最为最后一层的激活函数。

## categorical\_crossentropy



## mean\_squared\_error或mse

当我们的输出层是线性神经元（linear neurons）的时候使用均方误差，不会产生速度下降的问题。

# optimizers最优化函数

optimizer = ['SGD', 'RMSprop', 'Adagrad', 'Adadelta', 'Adam', 'Adamax', 'Nadam']

一般来说,sgd算法迭代到最优解附近时收敛很慢，而ADAM、ADADELTA、RMSPROP和，ADAM和ADADELTA收敛速度确实快，但是最终效果可能不如SGD和RMSPROP。

具体选择哪种优化方法应该是与模型和任务相关的。Sjl的Minist\_mlp实验，RMSPROP和ADADELTA取得了相近的效果，而sgd用更多的迭代,效果依然不如。

## Sgd

## keras.optimizers.SGD(lr=0.01, momentum=0.0, decay=0.0, nesterov=False)

* lr：大于0的浮点数，学习率,控制每批（batch）结束时更新的权重
* momentum：大于0的浮点数，动量参数,控制上次权重的更新对本次权重更新的影响程度
* decay：大于0的浮点数，每次更新后的学习率衰减值
* nesterov：布尔值，确定是否使用Nesterov动量

Examples:

lr = [0.001, 0.01, 0.1, 0.2, 0.3]

momentum = [0.0, 0.2, 0.4, 0.6, 0.8, 0.9]

# Metrics性能评估

性能评估函数类似与目标函数, 只不过该性能的评估结果讲不会用于训练

**表. tensorFlow实现的Metrics。**

注：除fbeta\_score额外拥有默认参数beta=1外,其他各个性能指标的参数均为y\_true和y\_pred

|  |  |
| --- | --- |
| binary\_accuracy | 对二分类问题,计算在所有预测值上的平均正确率 |
| categorical\_accuracy | 对多分类问题,计算再所有预测值上的平均正确率 |
| sparse\_categorical\_accuracy | 与categorical\_accuracy相同,在对稀疏的目标值预测时有用 |
| top\_k\_categorical\_accracy | 计算top-k正确率,当预测值的前k个值中存在目标类别即认为预测正确 |
| mean\_squared\_error | 计算预测值与真值的均方差 |
| mean\_absolute\_error | 计算预测值与真值的平均绝对误差 |
| mean\_absolute\_percentage\_error | 计算预测值与真值的平均绝对误差率 |
| mean\_squared\_logarithmic\_error | 计算预测值与真值的平均指数误差 |
| hinge | 计算预测值与真值的hinge loss |
| squared\_hinge | 计算预测值与真值的平方hinge loss |
| categorical\_crossentropy | 计算预测值与真值的多类交叉熵(输入值为二值矩阵,而不是向量) |
| sparse\_categorical\_crossentropy | 与多类交叉熵相同,适用于稀疏情况 |
| binary\_crossentropy | 计算预测值与真值的交叉熵 |
| poisson | 计算预测值与真值的泊松函数值 |
| cosine\_proximity | 计算预测值与真值的余弦相似性 |
| matthews\_correlation | 计算预测值与真值的马氏距离 |
| precision | 计算精确度，注意percision跟accuracy是不同的。percision用于评价多标签分类中有多少个选中的项是正确的 |
| recall | 召回率，计算多标签分类中有多少正确的项被选中 |
| fbeta\_score | 计算F值,即召回率与准确率的加权调和平均,该函数在多标签分类(一个样本有多个标签)时有用,如果只使用准确率作为度量,模型只要把所有输入分类为"所有类别"就可以获得完美的准确率,为了避免这种情况,度量指标应该对错误的选择进行惩罚. F-beta分值(0到1之间)通过准确率和召回率的加权调和平均来更好的度量.当beta为1时,该指标等价于F-measure,beta<1时,模型选对正确的标签更加重要,而beta>1时,模型对选错标签有更大的惩罚. |
| fmeasure | 计算f-measure，即percision和recall的调和平均 |