

采用 softmax 算法进行图像分类的步骤:

1. 根据训练集样本量, 样本像素, 样本类别, 构建模型。比如训练集有  $N$  个样本, 每个样本有  $Q$  个像素, 训练集共有  $M$  个类别, 可以采用线性方法构建一个单隐含层的 softmax 分类网络。
2. 对模型的输出结果进行 softmax 运算, 得到概率分布结果
3. 采用交叉熵损失函数计算损失
4. 采用梯度下降法更新权值, 图像训练集较大时建议用小批量随机梯度下降的方法, 例如有个 5W 样本的训练数据集, 每次迭代需要计算 5W 个样本才能对参数进行一次更新, 要求得最小值可能需要多次迭代(假设这里是 10 次才能收敛到一个合适的值); 而对于小批量随机梯度下降, 每次迭代随机抽取 50 个样本, 迭代 1000 次, 相当于更新参数 1000 次, 虽然每次更新用的随机样本量少, 但是通过更多次的迭代, 也能收敛到一个合适的值。相比  $5W \times 10$  次图像运算,  $50 \times 1000$  次运算的代价少很多。