

手写数字的数据集 MNIST

每个 image 是 28×28 的灰度图，共 400 个 pixel，dataset 中将每个 pixel 作为一个单独特征，y 值从 1-10 其中 0 被映射为 10
故每张 image 均带有 400 个 feature。

one-vs-all 的 logistic regression

如果有 num-label 个类别，那么就得训练 num-label 个 logistic regression 的分类器。当训练第 i 个 classifier 时，需要将 dataset 中的 y 做些处理：使用 $(y=i)$ 中得到 binary matrix
~~这样~~ 当用这个 classifier 做预测时，其结果表示预测该 example 为这个 label 的概率。这样重复训练 num-label 个分类器，则对于一个新的 example $x^{(1)}$ ，用 num-label 个分类器分别做预测，取结果最大那个，将 $x^{(1)}$ 预测为该类别。

简单 Neural Network

实例中的 Network 很简单只有三层，只有一层隐藏层，激活函数是 sigmoid，sigmoid 的输入 x 为所有 x 与 w 的乘积，即 linear 不带有任何高次项。但就是这么简单的模型，识别准确率能达到 97.5%，很吓人。

