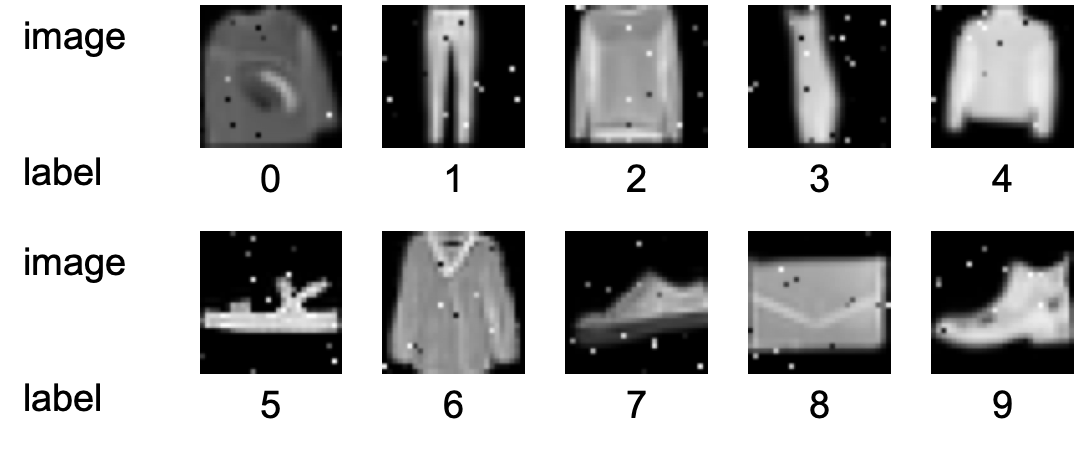
第二次大作业（2019年秋）

本次大作业要求通过深度学习解决图像分类问题，以在线比赛的形式进行，可在下述网页查看比赛的具体描述，进行数据下载和结果提交：

<https://www.kaggle.com/t/3cdecaa3c50a42f89b2d81c7e203d6a7>

**任务描述：**

本次大作业要求完成10类图片的分类问题，图像示例及类别如下：



本次作业的数据包含30000张图片组成的训练集（train.npy和train.csv）以及5000张图片组成的测试集（test.npy），均可从kaggle网站Data栏目下下载。npy文件可通过numpy.load()函数读取，每个npy文件包含一个N\*784的矩阵，N为图片数量。矩阵每行对应一张28\*28的图片，同学可在预处理环节自行将每行784维的向量转换成28\*28的图像。train.csv文件包含训练集的标签，含image\_id和label两列，共30000行，image\_id对应矩阵中的行下标，label为该图片的类别标签。

在预测环节，需要利用训练好的模型对测试集中的5000张图片进行分类，预测结果应生成submit.csv文件，同样包含image\_id和label两列，共5000行，每行对应一张图片的结果。在kaggle网站上提交该文件后会看到自己的分数（指标为catergorization accuracy）及排名,提交文件格式可参考kaggle网站的Data栏目下的samplesummission.csv。

注意：若两名同学提交完全相同的csv文件，会被网站自动查重，导致两人账号均被封，两位同学本次作业均按抄袭处理，请同学们不要抄袭！

**编程语言要求：**

编程语言原则上应使用Python，深度学习框架可从pytorch/tensorflow(keras)/caffe中任选其一。

若同学们还有其他疑问，请邮箱联系助教，或在讨论区进行讨论。

大作业需提交以下三份材料：

1. 源代码，通过网络学堂提交；
2. 实验报告，通过网络学堂提交；
3. 预测结果（csv文件），通过kaggle网页提交，在线实时排名。

大作业的最终成绩将由源代码成绩、实验报告成绩、排名成绩三部分按一定比例组合而成。