

《人工智能导论》实验三设计说明

主讲人：马少平教授

助教：李祥圣

清华大学计算机系人工智能研究所

1. 任务简介

日常生活中，我们经常需要在填写信件、银行开卡填写表单的时候手写大量的数字，如身份证号、手机号、邮编等等。事实上，专业人员在过去录入这些信息非常费时费力，而且还容易出现错误。现在请你设计一个手写体数字照片智能分类程序，省去人工识别 0-9 数字的麻烦。

在 Kaggle 竞赛平台上有 Digit Recognizer 的经典任务，并且提供了上万张手写体数字照片的灰度数据（可以借助 Matlab 还原照片）。目前，该任务已经被许多科研人员设计的人工智能程序完美解决，达到 100%测试精度。请大家在熟悉 Kaggle 平台的使用之后，直接通过该平台完成 Digit Recognizer 的竞赛任务，与一线科学家一较高下；并将你在从事竞赛过程中的心得体会、实验流程，以及最佳测评结果截图写入实验报告，同时在附录中递交你的代码。

链接：<https://www.kaggle.com/c/digit-recognizer>

2. 实验数据

实验数据可通过 Kaggle 平台进行下载，数据提供了训练集与测试集。没有验证集，所以需要自己再训练集上进行手工划分。注意这个数据与经典的 mnist 数据不太一样，所以切忌从外引入实验数据。

3. 实验要求

- 1) **账号注册**：请同学们以自己的姓名学号：如 lxs2017310123，注册 Kaggle，并以自己的账号在 Kaggle 上提交结果，在排名榜单上看到自己的排名，完成竞赛任务。
- 2) **模型实验**：使用 CNN（卷积神经网络）模型完成该任务是必须要实现的；也鼓励大家比较 CNN 与本课程其他模型（LSTM, MLP 等，作为 baseline）的效果，并写入实验报告。参考课堂上的 PPT。
- 3) **实现方法**：代码的语言不限，可借助深度学习的框架实现（theano, TensorFlow, keras 等），严禁抄袭。

4. 实验报告内容

- 1) 模型的结构图，以及流程分析。
- 2) 实验的结果，准确率是本实验的评价指标，报告中请给出竞赛的**最佳精度**（Kaggle 平台给出）以及**提交总次数**。
- 3) 试比较实验中使用的不同参数效果，并分析原因。
- 4) 比较 CNN 模型与不同 baseline 的模型效果，并进行分析。（如果有实现）
- 5) 问题思考
- 6) 心得体会

5. 问题思考

- 1) 实验训练什么时候停止是最合适的？简要陈述你的实现方式，并试分析固定迭代次数与通过验证集调整等方法的优缺点。
- 2) 实验参数的初始化是怎么做的？不同的方法适合哪些地方？（现有的初始化方法为零均值初始化，高斯分布初始化，正交初始化等）
- 3) 过拟合是深度学习常见的问题，有什么方法可以方式训练过程陷入过拟合。
- 4) 试分析 CNN（卷积神经网络）相对于全连接神经网络的优点。

6. 评价方式

程序结果与代码：80%

实验报告：20%（baseline 不一定要实现，但实现可根据难度加分）

7. 提交方式

在网络学堂上提交，需要提交的必要材料如下：

- 1) 实验报告，以学号_姓名.pdf 命名；
- 2) 实验代码以及程序运行导引（README）。

8. 联络方式

助教：李祥圣

手机：13763361656（微信同）

电邮：lixsh6@gmail.com