## 代码部署

分别将<https://github.com/Winnie-Qi/Surname_Generation_with_RNNs> 无条件姓氏生成和有条件姓氏生成两个文件移植至本地环境执行。数据集surnames\_with\_splits.csv（与作业文档09 RNN姓氏分类任务为同一个数据集）在学习通-资料-数据。

同样两种部署方式可选：

1、保留\*\*\*.py文件，将运行结果截图放至README.md中。

2、新建\*\*\*.ipynb文件（需合理划分代码单元格），移植代码仓库中的内容，提交时结果应显示在单元格下方。无需README截图。

## 2、无条件姓氏生成

运行Model1\_Unconditioned\_Surname\_Generation.py，生成**10个**通用姓氏（即不考虑国家信息）。

## 3、有条件姓氏生成

运行Model2\_Conditioned\_Surname\_Generation.py，根据学号尾数生成**5个**特定国家风格的姓氏。

学号尾数为0：阿拉伯语（Arabic）

学号尾数为1：捷克语（Czech）

学号尾数为2：荷兰语（Dutch）

学号尾数为3：英语语（English）

学号尾数为4：法语（French）

学号尾数为5：德语（German）

学号尾数为6：葡萄牙语（Portuguese）

学号尾数为7：意大利语（Italian）

学号尾数为8：俄语（Russian）

学号尾数为9：西班牙语（Spanish）

## 回答问题

① 两个模型的核心差异体现在什么机制上？

A. 字符编码方式不同

B. 是否考虑国家信息作为生成条件

C. RNN单元类型不同（GRU/LSTM）

D. 损失函数计算方式不同

② 在条件生成模型（Model2\_Conditioned\_Surname\_Generation）中，国家信息通过什么方式影响生成过程？

A. 作为额外的输入特征拼接

B. 作为GRU的初始隐藏状态

C. 作为注意力机制的key

D. 作为输出层的偏置项

③ 文件2中新增的nation\_emb层的主要作用是：

self.nation\_emb = nn.Embedding(num\_nationalities, rnn\_hidden\_size)

A. 将字符索引映射为稠密向量

B. 将国家标签转换为隐藏状态初始化向量

C. 生成姓氏的长度控制参数

D. 计算交叉熵损失的辅助参数

④ 对比两个文件的sample\_from\_model函数，文件2新增了哪个关键参数？

A. temperature

B. nationalities

C. device

D. max\_length

## 4、成果提交

将代码&README推送至GitHub。要求新建仓库或子文件夹，提交至腾讯共享文档的目录下只能包含两个代码文件和README文档。

提交内容需包含09 RNN姓氏分类和10 RNN 姓氏生成两个任务文档的内容。