## 循环神经网络实验报告

姓名：靳乐卿 学号：2012159

实验要求：

* 掌握RNN原理
* 学会使用PyTorch搭建循环神经网络来训练名字识别
* 学会使用PyTorch搭建LSTM网络来训练名字识别

报告内容：

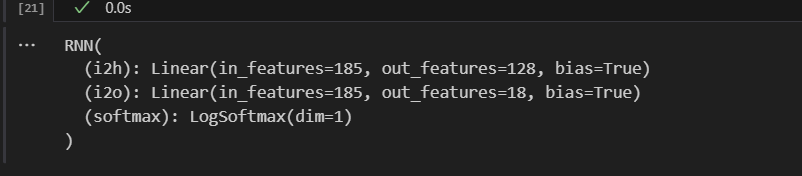
* 老师提供的原始版本RNN网络结构（可用print(net)打印，复制文字或截图皆可）、在名字识别验证集上的训练loss曲线、准确度曲线图以及预测矩阵图
* 个人实现的LSTM网络结构在上述验证集上的训练loss曲线、准确度曲线图以及预测矩阵图
* 解释为什么LSTM网络的性能优于RNN网络（重点部分）
* 格式不限

作业提交：

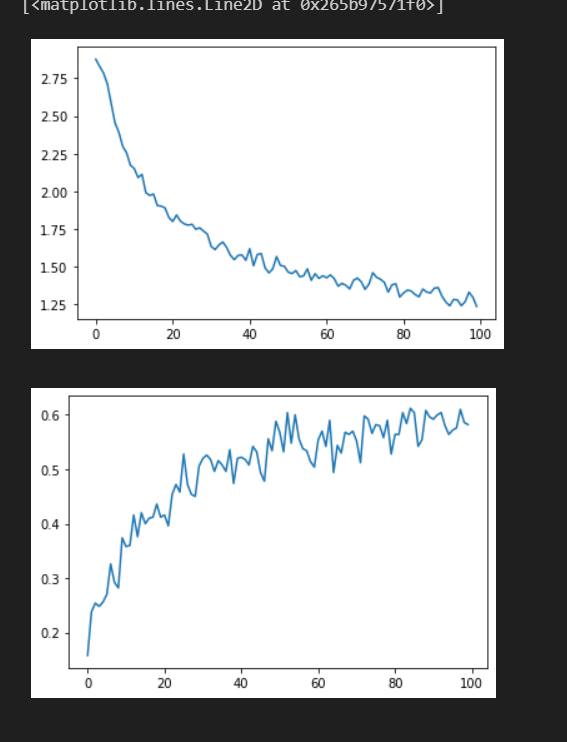
* 期末前将报告和代码（可将jupyter notebook里代码复制到一个xxx.py文件中）打包（学号+姓名.zip），提交方式另行通知
* 实验报告内容应工整
* 加分项：自己实现LSTM或者GRU

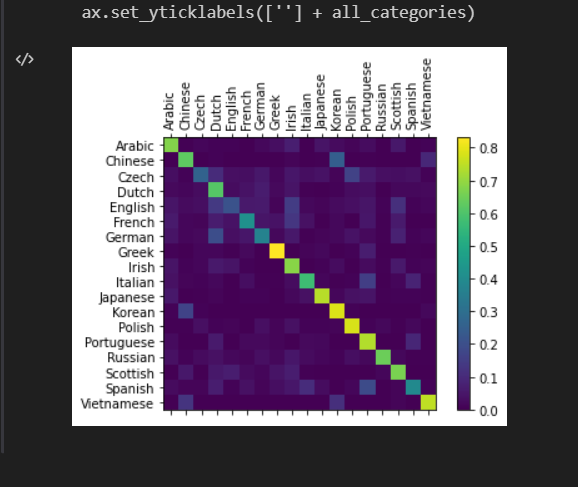
# 原始版本RNN

## 网络结构



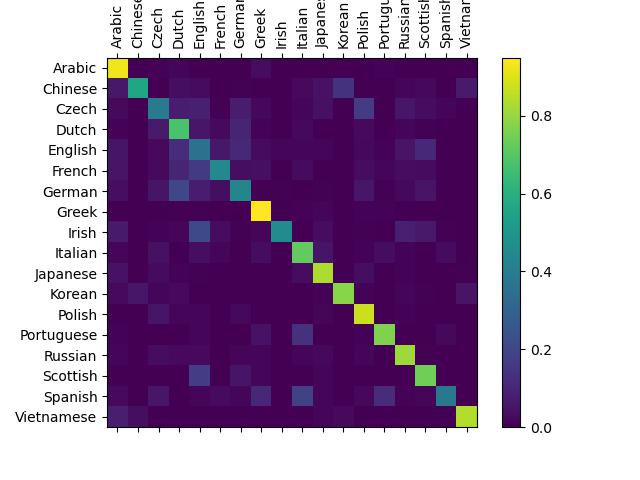
## loss曲线、准确度曲线图以及预测矩阵图

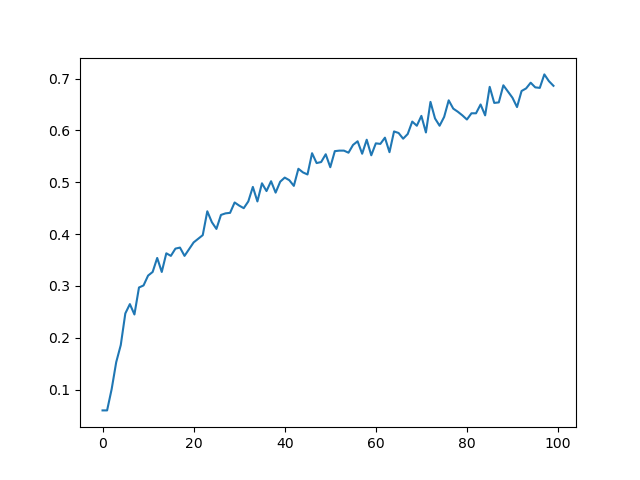


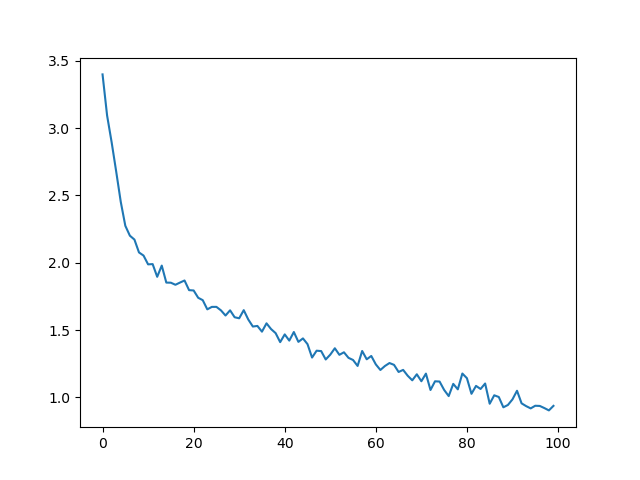


# 个人实现的LSTM

## loss曲线、准确度曲线图以及预测矩阵图

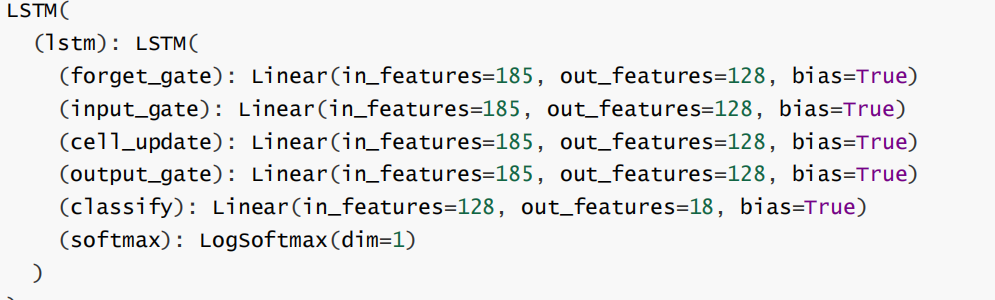


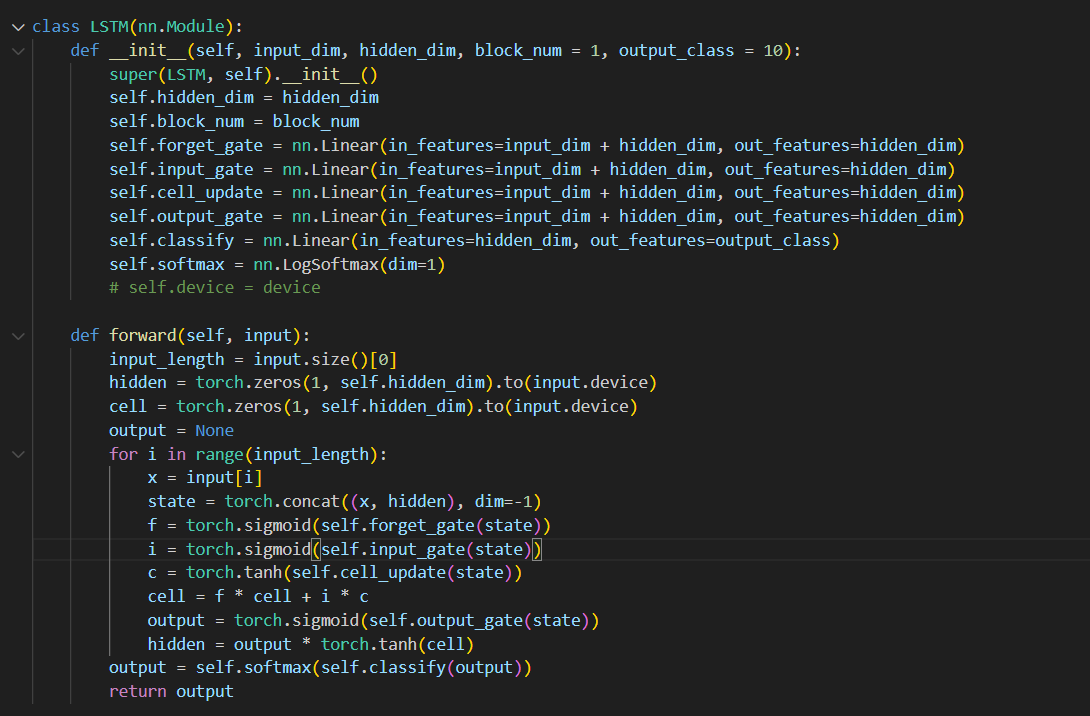




可以看到准确率是有所提升的

## 网络结构





# 为什么LSTM网络的性能优于RNN网络

LSTM优于RNN，主要有以下几个原因：

1. 处理长期依赖关系：RNN 网络在处理长序列时，会面临梯度消失或梯度爆炸的问题，导致无法有效捕捉到长期依赖关系。而 LSTM 网络通过引入门控机制，特别是遗忘门和输入门，能够更好地管理和控制信息的流动，从而更好地处理长期依赖关系。通过门控机制，LSTM 可以选择性地保留或遗忘过去的信息，有效地解决了梯度消失和梯度爆炸的问题。

2. 学习多个时间尺度的特征：LSTM 网络能够学习和利用多个时间尺度的特征，这对于处理时间序列数据非常重要。通过 LSTM 的记忆单元和各种门控机制，网络可以灵活地选择性地保留和利用不同时间尺度上的信息。这种能力使得 LSTM 网络能够更好地捕捉序列数据中的长期和短期模式。

3. 避免梯度消失和梯度爆炸：RNN 网络在反向传播过程中容易出现梯度消失或梯度爆炸的问题，导致模型难以收敛。LSTM 网络通过引入门控机制，有效地控制梯度的流动，避免了梯度的不稳定性问题。这使得 LSTM 网络更容易训练和优化。

LSTM 网络通过引入门控机制和记忆单元，能够更好地处理长期依赖关系，学习多个时间尺度的特征，并且避免了梯度消失和梯度爆炸的问题。这些特性使得 LSTM 网络在许多序列建模任务中表现出更好的性能，相对于基本的 RNN 网络而言更具优势。