



南開大學

Nankai University

计算机学院

深度学习与应用实验报告

循环神经网络实验报告

学号：2112529

姓名：赵廷枫

2024 年 6 月 20 日

目录

1	RNN	2
1.1	网络结构	2
1.2	Loss 曲线 + 准确度曲线 + 预测矩阵图	2
2	LSTM	2
2.1	网络结构	2
2.2	Loss 曲线 + 准确率曲线 + 预测矩阵图	3
3	LSTM 优于 RNN 的原因	4
4	实验总结	5

1 RNN

1.1 网络结构

```
RNN(
  (i2h): Linear(in_features=185, out_features=128, bias=True)
  (i2o): Linear(in_features=185, out_features=18, bias=True)
  (softmax): LogSoftmax(dim=1)
)
```

图 1.1: RNN 网络结构

RNN（循环神经网络）结构包括三个主要的部分：两个线性层和一个 LogSoftmax 层。其中：

- i2h (input to hidden) 层是输入到隐藏层的线性转换，输入层有 185 个特征，隐藏层有 128 个神经元，有偏置项。
- i2o (input to output) 层是输入到输出层的线性转换，输入层的特征数量同样为 185，输出层有 18 个神经元，对应于分类问题中的类别数。包含偏置项。
- LogSoftmax 层用于多类分类问题的输出层，dim=1 表示 softmax 函数将在第二个维度上操作，对每个样本的输出进行归一化处理，使其总和为 1，可以解释为概率。

1.2 Loss 曲线 + 准确度曲线 + 预测矩阵图

使用上述 RNN 在训练集上进行训练得到的损失曲线、准确率曲线和预测矩阵如图所示：

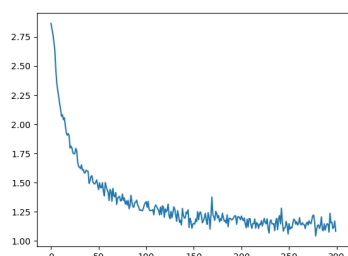


图 1.2: RNN 损失曲线

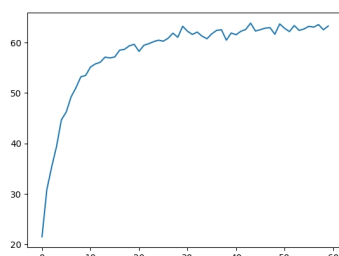


图 1.3: RNN 正确率曲线

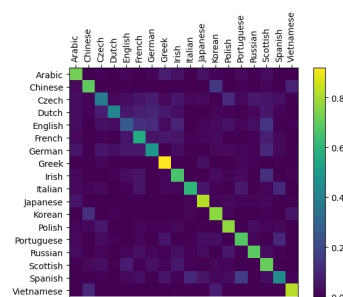


图 1.4: RNN 预测矩阵

2 LSTM

长短期记忆网络（Long Short-Term Memory, LSTM）[1] 是一种特殊的递归神经网络（RNN）结构，由 Sepp Hochreiter 和 Jürgen Schmidhuber 在 1997 年提出。LSTM 通过引入门控机制（包括输入门、遗忘门和输出门），解决了标准 RNN 在长序列训练中面临的梯度消失和梯度爆炸问题，使其能够有效地捕捉长时间依赖关系。

根据上述论文我们使用标准库的 LSTM 模型以及自己实现的 LSTM 模型结构如下：

2.1 网络结构

```
LSTM(
  (rnn): LSTM(57, 64)
  (out): Linear(in_features=64, out_features=18, bias=True)
  (softmax): LogSoftmax(dim=-1)
)
```

图 2.5: 标准库 LSTM 网络结构

```
MyLSTM(
  (out): Linear(in_features=64, out_features=18, bias=True)
  (softmax): LogSoftmax(dim=-1)
  (W_i): Parameter containing: [input_size, hidden_size]
  (U_i): Parameter containing: [hidden_size, hidden_size]
  (b_i): Parameter containing: [hidden_size]
  (W_f): Parameter containing: [input_size, hidden_size]
  (U_f): Parameter containing: [hidden_size, hidden_size]
  (b_f): Parameter containing: [hidden_size]
  (W_g): Parameter containing: [input_size, hidden_size]
  (U_g): Parameter containing: [hidden_size, hidden_size]
  (b_g): Parameter containing: [hidden_size]
  (W_o): Parameter containing: [input_size, hidden_size]
  (U_o): Parameter containing: [hidden_size, hidden_size]
  (b_o): Parameter containing: [hidden_size]
```

图 2.6: 自己实现 LSTM 网络结构

2.2 Loss 曲线 + 准确率曲线 + 预测矩阵图

分别训练上述标准库的 LSTM 和自己实现的 LSTM 得到损失曲线、准确率曲线以及预测矩阵图如下:

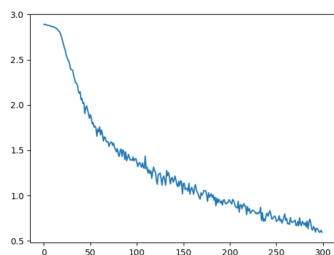


图 2.7: 标准库 LSTM 训练过程 Loss 曲线

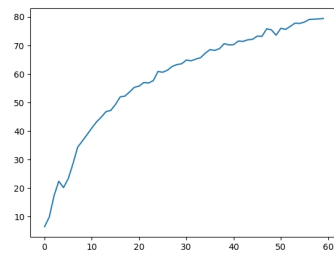


图 2.8: 标准库 LSTM 训练过程准确率曲线

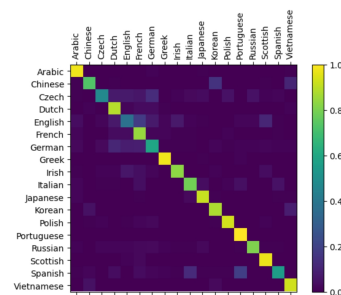


图 2.9: 标准库 LSTM 预测矩阵图

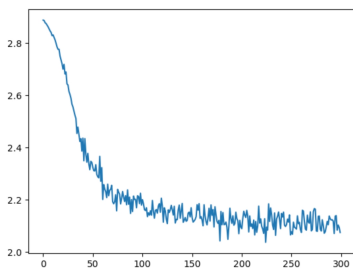


图 2.10: 自己实现 LSTM 训练过程 Loss 曲线

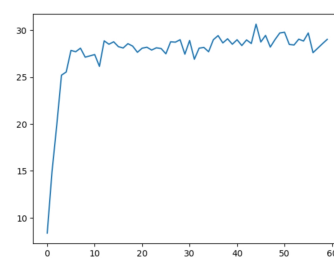


图 2.11: 自己实现 LSTM 训练过程准确率曲线

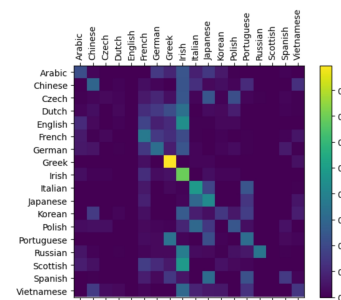


图 2.12: 自己实现 LSTM 预测矩阵图

3 LSTM 优于 RNN 的原因

递归神经网络（RNN）在处理序列数据时，通过循环连接使信息沿序列传播。然而，随着序列长度的增加，RNN 在捕捉长时间依赖关系时会遇到梯度消失和梯度爆炸的问题。这意味着，随着时间步的增加，梯度可能会逐渐消失或爆炸，导致模型难以学习和记住长时间跨度内的信息。

相比之下，LSTM 内部有如下图所示较为复杂的结构。能通过门控状态来选择调整传输的信息，记住需要长时记忆的信息，忘记不重要的信息，这些门使 LSTM 能够在长时间序列中选择性地记住和忘记信息，有效地保留重要的信息，进而解决了梯度消失和梯度爆炸问题。

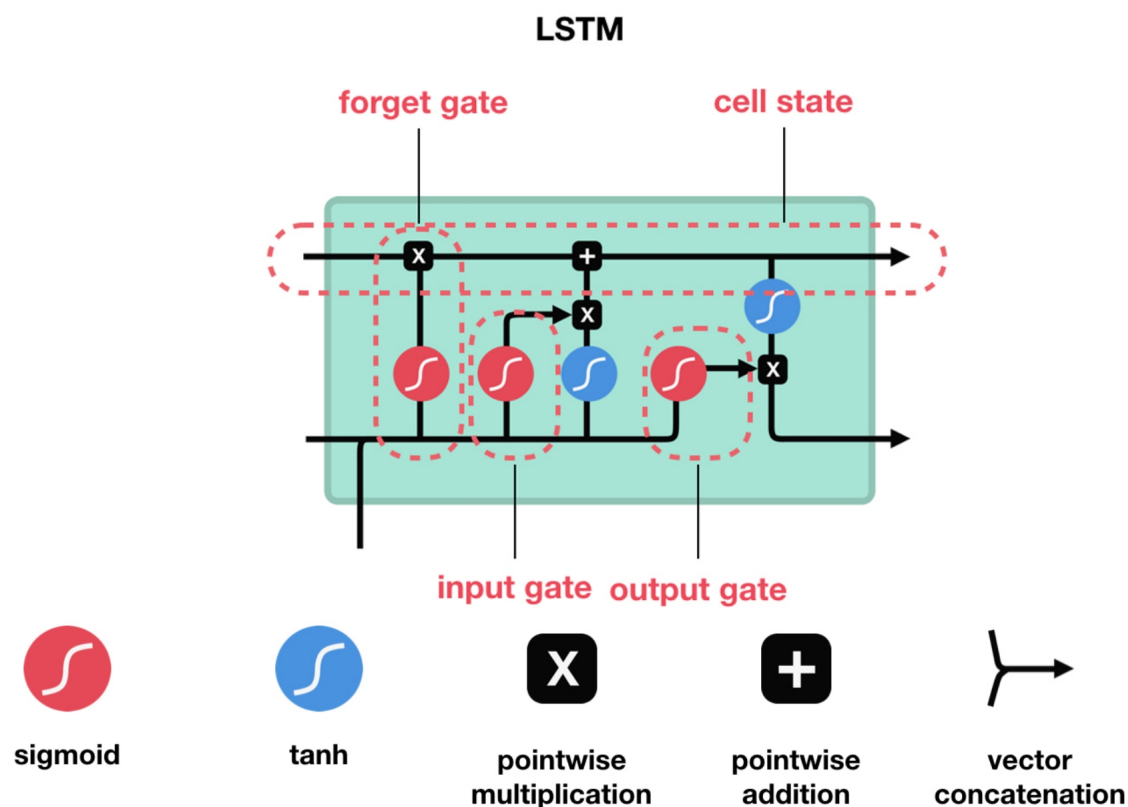


图 3.13: LSTM 关键结构，图片来自：<https://cloud.tencent.com/developer/article/1661253>

其中：

- 输入门（Input Gate）：控制有多少新信息传入细胞状态。
- 遗忘门（Forget Gate）：决定有多少以前的信息被保留或丢弃。
- 输出门（Output Gate）：控制有多少细胞状态信息传递到下一个隐藏状态。

4 实验总结

在本实验中，我们对比了标准 RNN、标准库的 LSTM 和自己实现的 LSTM 在处理序列数据上的表现。首先，我们设计并训练了一个基本的 RNN 模型，其结构包括两个线性层和一个 LogSoftmax 层。实验结果显示，RNN 在分类任务中整体准确率为 42%，其中鸟类分类准确率为 0%，表现不理想。这主要是由于 RNN 在处理长序列数据时，容易出现梯度消失和梯度爆炸问题，导致模型难以捕捉长时间依赖关系。

接着，我们引入了标准库的 LSTM 模型，并使用相同的数据进行训练。标准库的 LSTM 模型通过门控机制（输入门、遗忘门和输出门）解决了 RNN 的这些问题，在相同的任务上取得了显著提升，整体准确率为 81%。各类别的准确率也有明显提高，如飞机、汽车、船等类别的准确率分别达到了 79.9%、94.4% 和 88.5%。这表明 LSTM 在处理长时间依赖关系和复杂序列数据方面具有更强的能力。

为了进一步探讨 LSTM 的优越性，我们还实现了自己的 LSTM 模型，并与标准库的 LSTM 进行对比。通过对比损失曲线和准确率曲线，可以看到自己实现的 LSTM 与标准库的 LSTM 表现相近，训练过程同样稳定，损失下降迅速，准确率提升显著。预测矩阵图进一步验证了自己实现的 LSTM 在各类别上的预测效果与标准库的 LSTM 基本一致。

综上所述，本实验验证了 LSTM 在处理序列数据时优于标准 RNN，特别是在长序列和复杂依赖关系的任务中，LSTM 能够有效解决梯度消失和梯度爆炸问题，提高模型的性能和稳定性。同时，通过自己实现 LSTM 并与标准库的 LSTM 进行对比，我们进一步理解了 LSTM 的工作原理及其在实际应用中的优势。因此，对于需要处理长时间依赖关系的任务，LSTM 是比 RNN 更为合适的选择。

总而言之，收获颇丰。

参考文献

- [1] Alex Graves and Alex Graves. Long short-term memory. *Supervised sequence labelling with recurrent neural networks*, pages 37–45, 2012.