

10.4-10.12 周报

> Author: 彭日骏
> Edit Time: 2025/10/4-2025/10/12

Plan

- 完成CNN, RNN, LSTM, GRU, Transformer的基础概念学习以及算法推导
- 读懂Attention Is All You Need

Extra

- 尝试自主编写了CNN, RNN, LSTM代码

其中我了解到，似乎我只需要定义好

```
criterion = nn.CrossEntropyLoss() # MSE交叉熵损失
optimizer = optim.Adam(model.parameters(), lr=learning_rate) # 优化器追踪梯度
```

训练过程中我就可以通过前向计算后得到的结果outputs，直接计算交叉熵损失，并通过optimizer直接自动计算梯度，更新模型参数

```
loss = criterion(outputs, labels) # 计算MSE交叉熵损失
optimizer.zero_grad() # 清空之前梯度
loss.backward() # 自动计算当前梯度
optimizer.step() # 根据梯度更新模型参数
```

直接就可以完成反向传播（~~感觉我啃天书一样的反向传播数学公式被喂了狗了~~）。

学习笔记见[人工智能算法学习笔记](#)，里面更新了我这周的学习笔记，体现了我这周对CNN、RNN、LSTM、GRU、Transformer算法的了解掌握程度。

同时学习的过程中我有手痒痒做了一点实践，从最基础的之前涉猎过的CNN开始，尝试参考其他代码，完全自己手写了一个MNIST数据集的训练程序，在这个过程中我又增进了更多的概念理解，同时也发现了很多实践和理论上的区别，并对自己的知识点做了补充。过程详见[Learning_CNN](#)文件夹。

Summary

本周Plan基本达到预期，通过观看其他人写的中文的文章，再回头看论文，很多地方虽然仍有困惑，但是至少对Transformer架构初步有了一个完整的理解，接下来的概念补充时日可待。

我计划继续对不了解的概念——[人工智能算法学习笔记](#)与[Learning_CNN](#)中均还有疑惑的地方，我用**蓝色荧光笔**和**why?**都标记了出来，以方便我后续的补充——在下一周深化了解，补充相关的算法理解。例如，相关的乱七八糟的优化是具有共性的，比方说**残差网络**这个概念我尚有不明白，而他的推广性是很强的，我需要去细化学习这些东西。

而学习方法大概是通过检索相关的文章进行阅读，相关的视频观看，然后通过在实践中进行探索。所以我想要向老师您申请尝试下一周的目标是自己自主训练一个RNN, LSTM, GRU模型，在训练中加深模型了解，在实践中对有用的算法进行学习并记录，然后如果允许可以尝试也自己训练一个Transformer，不过可能一周的时间会不太够，我可能期望两周时间能训练完Transformer。

Next Week Plan

根据我2025/10/4-2025/10/12的学习情况，我下周的目标订立如下：

- 自主训练一个RNN, LSTM, GRU, Transformer（时间尚许）模型
- 训练模型过程中，结合各种文章和视频，做笔记记录下遇到的优化算法

希望老师能给点训练模型实践中的建议，比方说MNIST数据集是个CNN训练很经典的数据集，但是关于RNN等其他的模型，我可能不太清楚用什么数据集去作为新手实践是比较友好的，请老师批评和指正！