深度学习小结

为期七周的深度学习小组学习，在假期来临前结束了。从第一周的研三师兄师姐的代码分享，第二周的博士赵哲焕师兄的keras框架分享，第三周开始由我们研一开始分享了CNN、RNN及其变种LSTM的基础知识和Attention机制等等，预示着正式拉开学习序幕。

下面仅就在学习小组中分享的内容做一些汇报。

通过何志勇同学分享了的CNN，从理论上了解了卷积和池化操作，并演示了应用于情感分类的任务，也是有tensorflow实现。

彭钰莹同学主讲RNN、双向RNN以及多层RNN的相关概念和代码实现。

我主讲了LSTM的原理和代码分析部分。从RNN的应用开始，谈到了它的问题——长词依赖，从而引出LSTM来解决这一问题。 LSTM作为RNN变种，从结构上增加了门和记忆单元两个概念，模拟了人大脑记忆的更新，其适用于NLP。随后，我分享了一篇关于RNN正则化的论文，主要提出了在层间加Dropout的方法，来预防过拟合问题。同时，通过分析在tensorflow框架的相关代码，对代码中的每一步做了解释。从代码中学习了Clip方法，该方法用于防止梯度爆炸。

杨培同学主要分享了机器翻译中的Encoder-Decoder模型，然后推广到应用Attention机制的E-D模型，并阐述了该机制可以应用于我们实验室任务中的可能性。