读书报告 8.31-9.6

61518426 周之遥

·读书内容

学习第7章Neural Networks and Neural Language Models

逻辑回归其实就是一个单层神经网络，而拥有隐藏层的多层神经网络不需要像逻辑回归一样去手动获取文本特征，而是利用神经网络自动学习文本的特征。

一个简单的Feed-Forward神经网络有多层神经元，每一层是上一层的线性和并加上一个激活函数（非线性函数）。典型的非线性函数有softmax、sigmoid、tanh等。

关于Loss Function，常用的有交叉熵等。

神经网络的优化是建立在链式法则的基础上，通过逐个求导、梯度下降，获得参数的最优值。

神经网络还需要考虑一些超参的设置，比如dropout、lr、bs等。

Language Modeling是神经网络在NLP上的一个典型应用，目标是给出下一个word或整个句子的概率。

对于输入的文本，需要进行embedding然后输入神经网络，这里可以简单地做一个look-up，从已经确定的embedding矩阵中找文本对应的向量；也可以在任务中同时训练embedding。

模型在文本中移动，利用若干个word预测接下来的word，预测结果交叉熵损失反向传播优化网络。