**读书报告**

王贵涛

自然语言处理的第七章主要介绍了神经网络和神经语言模型。

**神经元（Units）**

数据经过一个神经元的时候，先进行线性点乘，然后进行激活，通过一个激活函数使数据具有非线性的特征。

常见的激活函数有sigmiod、tanh、ReLU函数。

**异或问题（The XOR problem）**

在与和或问题中，均可以使用线性的方法来实现预测，但异或无论怎么分都不可能用一条直线将不同的结果分到两侧。当使用两层网络的时候，异或变的可能。

**前馈神经网络（Feed-Forward Neural Networks）**

前馈神经网络是最简单的神经网络，每一层神经元的输出直接传递给下一层，没有输出传递给上一层。前馈神经网络也被称为多层感知器（multi-layer perceptrons）。

使用单个矩阵作为整个层权值可以高效完成计算。

如果做情感分类，最后的输出是词性的概率，并且所以输出结点的总和必须为1。可以使用softmax函数创建概率分布。神经网络就像逻辑回归，不过神经网络层数更高，就像一层一层的逻辑回归分类器。不是通过模板形成特征，而是由网络前几层自己诱导特征形成。

**训练神经网络（Training Neural Nets）**

我们仍然使用交叉熵损失来作为损失函数，使用梯度下降法来优化，使用反向传播来获得偏导数。

应用函数的链式求导法则，为了计算复合函数的导数，我们需要知道每一步对上一步的导数，最后将这些导数相乘。

神经网络的优化是一个非凸问题，比逻辑回归复杂的多。在神经网络中，我们需要使用更小的随机数继续初始化，将均值和方差归一化。

正则化的方式有丢弃（dropout），随机舍弃一些神经元和他们的联系来进行训练。

**神经语言模型（Neural Language Models）**

基于神经网络的语言模型比N元语法具有更多优势，如无需平滑处理，可以处理更长的历史，在类似单词的上下文中泛化。且具有更高的预测精度。另一方面，神经网络语言模型的缺点是训练速度过慢，在许多任务中仍然时候N元语法。此外，神经语言模型是机器翻译、对话和语言生成任务的模型基础。

神经网络可以通过两个词相似的嵌入向量预测在测试集中没有出现过的搭配。