**北京师范大学珠海分校**

**本科生毕业论文**

论文题目 **深度学习模型研究及应用**

学 院　　 应用数学学院

专 业 数学与应用数学

学 号 1717010022

学 生 姓 名 黄继安

指导教师姓名 李艳

指导教师单位 北师大珠海分校应用数学学院

2020年11月28日

深度学习模型研究及应用

文献综述

* 1. 深度学习模型算法发展历史

深度学习(Deep Learning)是一种对输入数据进行表示学习或者进行非线性变换的技术，它是不少于2个隐含层的神经网络。因其特殊的层级连接方式可以实现渐进抽象的非线性处理，所以在处理从原始输入数据到期望输出结果的复杂非线性变换上特别出色，并由此实现对原始数据进行表示学习或者非线性建模学习。深度学习特别强调“端到端(end-to-end)”的学习方式，也就是直接从原始输入数据到输出结果的学习过程，这一点相比于传统的机器学习算法需要花费大量的人力资源在特征处理上节省了大量的时间，让研究人员可以把更多时间专注于模型上，因此深度学习(Deep Learning)通常也称之为表示学习。

深度学习(Deep Learning)本质上是包含了多个隐含层的人工神经网络，而人工神经网络的研究历史并非一帆风顺，具体可追溯到上世纪40年代。Pitts和McCulloch在1943年提出了首个神经元数学模型[1]，Hebb在1949年提出神经元的学习准则[2]，Rosenblatt于1957年提出了感知机(perceptron)模型[3]，自此开启了人工神经网络的第一波热潮。但是因为人工智能领域的知名学者Minsky等指出感知机模型是线性模型，它无法解决异或问题，导致人工神经网络的研究第一次进入了低谷时期。带动人工神经网络第二次研究热潮的是1986年Williams、Rumelhart和Hinton在自然杂志《Nature》发表的著名误差逆传播算法(back propagation,BP)算法[4]，用于训练含有多隐含层的人工神经网络，BP算法使得求解具有非线性学习能力的多层感知机模型变为可能的事情。事实上在我们的认知里，BP算法作为人工神经网络训练的标准算法，一直是沿用到今天。Hornik、Stinchcombe和White在1989年在理论上证明了，神经网络可以逼近任意复杂的连续函数[5]，这一证明进一步的刺激了非线性感知机的研究发展。那么模型有了，训练问题看似也被BP算法解决了，可是不要忽视了其他方面的局限，由于多层神经网络大量的参数需要优化，这决定了训练数据的规模越大越好，但遗憾的是当时的数据规模都还很小，难以支撑多层神经网络的学习，这就导致了多层神经网络在许多领域的表现还没有浅层模型表现的好，研究的热潮仅仅持续了几年便又褪去。

参考文献

[1] McCulloch W S, Pitts W. A logical calculus of the ideas immanent in nervous activity[J]. Bulletin of Mathematical Biophysics,

[2] Hebb D O. The organization of behavior[M]. New York: Wiley,

[3] Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain[J]. Psychological Review

[4] Rumelhart D E, Hinton G E, Williams R J. Learning internal representations by error propagation[J]. Nature

[5] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators[J]. Neural Networks