

本科毕业论文（设计）

学 院： 数学与统计学院

专 业： 信息与计算科学

学生姓名： 王圳川

学 号： 20271022

指导教师： 周声龙

**北京交通大学**

2024年4月

学士论文版权使用授权书

本学士论文作者完全了解北京交通大学有关保留、使用学士论文的规定。特授权北京交通大学可以将学士论文的全部或部分内容编入有关数据库进行检索，提供阅览服务，并采用影印、缩印或扫描等复制手段保存、汇编以供查阅和借阅。

（保密的学位论文在解密后适用本授权说明）

学位论文作者签名： 指导教师签名：

签字日期： 年 月 日 签字日期： 年 月 日

学士论文诚信声明

本人声明所呈交的毕业论文（设计），题目 深度神经网络的稀疏优化方法 是本人在指导教师的指导下，独立进行研究工作所取得的成果。尽我所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得北京交通大学或其他教育机构的学位或证书而使用过的材料。

申请学位论文与资料若有不实之处，本人承担一切相关责任。

本人签名： 日期：

中文摘要

**摘要：**中文摘要应将论文的内容要点简短明了地表达出来，约400字左右，字体为宋体小四号。内容应包括工作目的、研究方法、成果和结论。要突出本论文的创新点，语言力求精炼。为了便于文献检索，应在本页下方另起一行注明论文的关键词（3-5个），如有可能，尽量采用《汉语主题词表》等词表提供的规范词。图X幅，表X个，参考文献X篇。

**关键词：**

ABSTRACT

**ABSTRACT:**与中文摘要内容要相对应。

**KEYWORDS：**

目 录

[中文摘要 i](#_Toc414268225)

[ABSTRACT ii](#_Toc414268226)

[目 录 iii](#_Toc414268227)

[1 引言 1](#_Toc414268228)

[2 稀疏优化方法（小四、黑体、左对齐、空1字符） 2](#_Toc414268229)

[2.1 量度优化（小四、宋体、首行缩进1字符） 3](#_Toc414268230)

[2.1.1 绝对值优化（小四、宋体、首行缩进2字符） 3](#_Toc414268231)

[3 4](#_Toc414268232)

[3.1 4](#_Toc414268233)

[3.1.1 4](#_Toc414268234)

[4 5](#_Toc414268235)

[4.1 5](#_Toc414268236)

[4.1.1 5](#_Toc414268237)

[5 结论 6](#_Toc414268238)

[参考文献 7](#_Toc414268239)

[致 谢 8](#_Toc414268240)

[附 录 8](#_Toc414268240)

目录说明：

（小四、黑体、左对齐、空1字符）

（小四、宋体、首行缩进1字符）

（小四、宋体、首行缩进2字符）

1. 引言

深度神经网络在许许多多的领域内（例如：图像分类、机器翻译、语音合成）展现出了最先进的水平。虽然一方面模型的质量在随着模型和训练数据的增大而提升，但是随之而来的训练和部署方面的成本也水涨船高。像图像分类和机器翻译类的神经网络的参数量是以千万计的，并且通常预测一个结果就需要百亿量级的浮点数运算。这就导致这些领域内的先进模型的部署、预测成本到达一个不可接受的程度，尤其是移动和嵌入式设备上的模型部署、预测。一些近年来大热的巨型模型像目标检测领域的先进模型Inception-V3[4]，一次预测需要进行57亿次运算，拥有2,700,000个参数；NLP领域的GPT-3模型需要175,000,000个参数（假设其参数为16 bits精度的话大约需要350GiB的存储空间）。除了部署与预测，训练这些大模型成本更高，每一次训练大概需要数百万美元。

为了解决上述问题，稀疏优化成为了一个可行手段。通过稀疏优化，我们可以使得参数的一个子集置为0，而任何与权重0相乘的运算将被跳过，并且由于参数是稀疏的，可以采取稀疏存储的方式节省内存空间，极大程度上降低了模型的部署和预测成本[2]。比如Minerva模型[5]通过对激活值较小的参数的舍弃，大约节约了训练过程中50％的能源消耗；Eyeriss模型[6]通过将激活值置为0节约计算成本；Cnvlutin[7]也是通过类似的方法达到了与优化之前的模型13倍的速度差距和降低10倍的能源消耗。由此可见，对模型参数的稀疏优化的确能够大幅度降低模型训练、预测、部署的成本。

虽然对模型的稀疏优化能够极大程度上降低成本，但是有许多人认为这种大幅度的减少模型参数的方法势必也会伴随模型性能的大幅度下降。但是事实证明并非如此。有研究已经证明深度神经网络对参数的高稀疏程度具有相当强的容忍度[1]。甚至有研究发现深度神经网络中的95％的参数能够通过剩余的5％的参数预测出来[8]。这些都说明了一定程度的稀疏优化不会对模型的性能造成明显的影响。事实上模型性能与稀疏程度的关系通常如下图所示，一定的稀疏程度非但不会使模型的性能下降，反而会让模型的性能得到提升。一些研究和奥卡姆剃刀原则都能够解释这种现象：提高模型的稀疏程度能够降低模型中存在的噪声。



在过去的几年里，学术界提出了很多稀疏优化的方案，比如变分丢弃法[3]、正则法、结构剪枝。然而这些方法并非对模型参数的直接稀疏优化，而是通过训练过程中产生的梯度或激活值为指标，丢弃掉神经网络中的一些结构，从而达到稀疏优化的目的。或者是采取在损失函数中增加正则项的方式，通过训练隐式地让模型规模减小，间接地达到稀疏优化的目的。这些方案往往以直觉为导向，缺乏足够的数学理论支持。而本设计计划将深度神经网络抽象为最优化问题，藉由硬阈值方法在训练的过程中迭代式地将模型中一部分参数置为0，直至收敛。并计划将此方法运用于Transformer模型，应用于情感预测领域，同时提供理论和经验结果证明该方法的有效性。

It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience.

——Albert Einstein

1. 稀疏优化的最优性条件

我们先假设数据集记数据集的标签，模型的预测值为。

那么我们要优化的问题是：

这里的，是神经网络所有参数的范数，即神经网络所有层中的非零参数的个数。

如果令，为神经网络的层数，则原优化问题可写作含有个约束条件的问题：

我们假设神经网络每一层之间相互独立，那么我们可以推得原问题在前层上每一层是一个有稀疏约束条件的优化问题，不失其一般性，我们假定为第层：

在第层上为无约束优化问题：

* 1. 有稀疏约束条件的优化问题

设，是指中绝对值最大的第个元素，即

* + 1. 基本可行解

**定义3.1.** 是的一个基本可行解，如果有：

**定理3.1.** 如果令是的最优解，那么是一个基本可行解

**证明.** 如果，那么有，，否则存在使得与的最优性相悖。因此我们有。如果该等式在显然也成立。

* + 1. -

对于问题

这里是闭凸可行域，是一个连续可导的函数，可能是非凸的。那么该问题的一个解可以被称为是稳定点，当：

如果是该问题的最优解，那么显然也是稳定点，因此，稳定点是一个必要最优性条件。许多解决上述非凸问题的优化方法都只能保证收敛到稳定点。

实际上述性质可以转化为，解是一个稳定点，当且仅当

这里假设，记作上的正交投影，即

我们将上式推广至有稀疏约束的问题*P*中，记为-稳定性

**定义3.2.** 一个解是问题的一个*L*-稳定点，如果满足关系

这里由于并不是一个凸集，所以投影运算并不止得到一个值，特别地，对于就是包含中绝对值最大的*s*个位置的参数矩阵。

接下来将证明在一个合适的李普希茨条件下，-稳定性是一个必要最优性条件，首先，我们需要用一种更简洁的方式描述条件。

**引理3.1.** ，满足条件当且仅当和

**证明：（）**。假设满足条件。如果，那么由条件可知，所以。如果，那么，又因为，所以  
（）。假如满足条件。如果，那么，则；因此，在这种情况下，。如果，那么并且，由该条件可知

因此，包含中绝对值最大的*s*个元素，其他的元素要么小于或者等于他们，故满足条件。

**假设3.1.** 目标函数的梯度是李普希茨连续，设其李普希茨常数，即

该假设对目标函数成立，取，

**引理3.2. (Decent Lemma)** 假设连续可导且满足假设3.1，

**引理 3.3.** 如果假设3.1成立，。对于任意，y是一个参数矩阵，满足

我们有

**证明：**可以写成，

上述最小化问题可以看作

则有

由引理3.2可知

又因为

则可以得到

**定理3.2.** 假设3.1成立的前提下，，假设是问题的最优解，那么

1. 是一个-稳定点
2. 只包含一个元素

**证明：**假设存在，并且与不同，由引理3.3可知

与最优解的性质相悖，因此只包含一个元素

1. 结论

论文的结论是最终的、总体的结论，不是正文中各段的小结的简单重复。结论应该准确、完整、明确、精练。如果不可能导出应有的结论，也可以没有结论而进行必要的讨论。可以在结论或讨论中提出建议、研究设想、仪器设备改进意见以及尚待解决的问题等。

参考文献

参考文献是文中引用的有具体文字来源的文献集合。按照GB 7714《文后参考文献著录规则》的规定执行。

参考文献以文献在整个论文中出现的次序用[1]、[2]、[3]……形式统一排序、依次列出。

参考文献的表示格式为:

**著作:**

[序号] 作者.译者.书名.版本.出版地.出版社.出版时间.引用部分起止页

**期刊:**

[序号] 作者.译者.文章题目.期刊名.年份.卷号(期数).引用部分起止页

**会议论文集：**

[序号]作者.译者.文章名.文集名 .会址.开会年.出版地.出版者.出版时间.引用部分起止页

**学位论文：**

[序号]作者.题名[学位论文]（英文用[Dissertation]）.保存地点.保存单位.年份.引用部分起止页

**专利:**

[序号] 专利申请者.题名.国别.专利文献种类.专利号.发布日期.引用部分起止页

**技术标准:**

[序号] 起草责任者.标准代号.标准顺序号—发布年.标准名称.出版地.出版者.出版年份.引用部分起止页

致 谢

放置在参考文献页后，对象包括：1）国家科学基金，资助研究工作的奖学金基金，合同单位，资助或支持的企业、组织或个人。2）协助完成研究工作和提供便利条件的组织或个人。3）在研究工作中提出建议和提供帮助的人。4）给予转载和引用权的资料、图片、文献、研究思想和设想的所有者。5）其他应感谢的组织和个人。

附 录

附录A 程序代码

附录是作为论文主体的补充项目，并不是必须的。

论文的附录依序用大写正体英文字母A、B、C……编序号，如：附录A。

附录B 工程图纸