

基于

深度置信网络的 分类方法

周树森 著

DEEP

BELIEF

NE TUTORIALS



清华大学出版社

周树森 著

基于

深度置信网络的 分类方法

清华大学出版社
北京

内 容 简 介

深度置信网络作为经典的深度学习方法,可以包含较多的隐藏层,更好地学习各种复杂数据的结构和分布。本书从深度置信网络的概念、研究内容 and 应用出发,提出了一系列基于深度置信网络的分类方法,使用半监督学习和监督学习方法来提升分类性能。另外,将深度学习方法应用到手写中文识别中,提出了基于深层架构的手写识别方法,有效地提升了手写识别正确率。

全书共分为7章。第1章介绍深度置信网络的概念、发展及其研究现状;第2章介绍区分深度置信网络;第3章介绍自适应深度置信网络;第4章介绍量子深度置信网络;第5章介绍主动深度置信网络;第6章介绍主动模糊深度置信网络;第7章介绍基于深层架构的手写识别方法。

本书适合作为高等院校计算机专业高年级本科生、研究生的参考书,同时可供对机器学习比较熟悉并且对深度学习有所了解的广大科技工作者和研究人员参考。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

基于深度置信网络的分类方法/周树森著. --北京:清华大学出版社,2015

ISBN 978-7-302-41355-4

I. ①基… II. ①周… III. ①机器学习—分析方法 IV. ①TP181

中国版本图书馆CIP数据核字(2015)第209432号

责任编辑:张 玥 战晓雷

封面设计:常雪影

责任校对:时翠兰

责任印制:李红英

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦A座

邮 编:100084

社总机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:清华大学印刷厂

经 销:全国新华书店

开 本:185mm×230mm

印 张:10

字 数:131千字

版 次:2015年12月第1版

印 次:2015年12月第1次印刷

印 数:1~1000

定 价:35.00元

产品编号:066149-01

前 言

21 世纪初,机器学习领域出现了一系列基于多层神经网络架构的学习方法,称为深度学习方法。深度学习的最大特点,就是把一个复杂的认知问题分解为几层简单问题,通过对简单问题的大运算量处理,得出规律,进入下一“层”,逐层运算处理,直到触达事物本质。深度置信网络作为经典的深度学习方法,可以包含较多隐藏层,可以更好地学习各种复杂数据的结构和分布,引起了广泛关注。深度置信网络本身采用的是无监督学习机制,可以采用先以无监督学习进行抽象,后以有监督学习进行分类的方法提升分类性能。探讨如何使用深度学习方法来解决分类问题,特别是基于半监督学习的分类问题的研究工作具有重要的现实意义。

本书提出了一系列基于深度置信网络的分类方法,使用半监督学习方法和监督学习方法来提升分类性能。

第一,区分深度置信网络。深度置信网络方法在数据压缩中取得了很好的效果,但是在分类方面的性能还有待提升。本书基于深度置信网络方法,提出了一种新的半监督学习方法,称为区分深度置信网络。该方法使用一个新的深层架构来集成深度置信网络的抽象能力和指数损失函数的区分能力,可以使用少量的标注数据达到比较好的分类效果。在人工合成和现实生活中图片数据集上的实验表明,区分深度置信网络可以通过使用大量的未标注数据来大幅度提升它的分类能力。

第二,自适应深度置信网络。区分深度置信网络主要解决半监督学习问题,而没有解决如何用深度学习方法进行有效的监督学习的问题。本书基于区分深度置信网络和深度自动编码方法,提出了一种新的监督学习方法,称

为自适应深度置信网络。该方法使用无监督学习方法来构建输出层,与使用随机数初始化输出层的经典方法相比,分类结果得到了提升。在手写数字、手写字母和现实生活中图片数据集上的实验表明,自适应深度置信网络的分类结果优于经典机器学习方法和区分深度置信网络方法。

第三,量子深度置信网络。量子神经网络是一个将模糊集和传统神经网络很好地结合起来的分类方法。为了进一步提升监督学习方法的性能,本书在自适应深度置信网络和量子神经网络基础上,提出了量子深度置信网络方法。该方法使用一个新的深层架构来集成深层架构的抽象能力和量子神经网络的模糊表示能力。在多个现实应用数据集上的实验表明,量子深度置信网络的分类能力明显高于经典神经网络、模糊神经网络、区分深度置信网络和自适应深度置信网络方法。

第四,主动深度置信网络。为解决半监督学习中的有效样本选择问题,本书将自适应深度置信网络和主动学习方法相结合,提出了一种新的半监督学习方法,称为主动深度置信网络。该方法使用同一个深层架构来进行半监督学习和主动学习,使深层架构在主动学习过程中进行迭代训练,逐步提高抽象和分类能力。本书将主动深度置信网络成功地应用到情感分类任务中,实验结果表明,主动深度置信网络的性能优于经典半监督学习方法、主动学习方法,对本书前面提出的深度学习方法也有性能上的改进。

第五,主动模糊深度置信网络。为了进一步提升主动深度置信网络的性能,本书将主动深度置信网络与模糊学习方法相结合,提出了主动模糊深度置信网络。该方法继承了深层架构优异的数据抽象能力和模糊集优异的分类能力。本书将该方法成功地应用到情感分类任务中,实验结果表明,主动模糊深度置信网络获得了所有对比方法中最好的性能。

最后,本书将深度学习方法应用到手写中文识别中,提出了一个有效的手写中文识别方法,称为基于深层架构的手写识别方法。该方法将深层架构的抽象能力和指数损失函数的分类能力相结合,使用深层架构进行粗分类,

然后使用改进的二次分类函数进行细分类。在三个手写中文识别数据库上的实验表明,基于深层架构的手写识别方法可以通过深层架构来提升分类性能,识别正确率优于经典的手写中文识别方法,从而进一步验证了基于深度置信网络的分类方法解决大规模实际应用中的分类问题的能力。

本书的相关研究和出版得到了国家自然科学基金(No. 61300155, 61170161)、鲁东大学校基金(No. LY2013004)资助。另外,本书的编写还得到鲁东大学信息与电气工程学院院长邹海林教授、哈尔滨工业大学深圳研究生院王晓龙教授和陈清财教授的大力支持,在此对这些同志的鼓励和帮助表示衷心的感谢。

限于作者的学识水平,书中难免有不妥和疏漏之处,恳请各位专家、同仁和读者不吝赐教和批评指正。

作者邮箱 zhoushusen@gmail.com。

周树森

2015年6月于烟台

摘 要

最近,机器学习领域出现了一系列基于多层神经网络架构的学习方法,称为深度学习方法。深度置信网络作为经典的深度学习方法,可以包含较多隐藏层,可以更好地学习各种复杂数据的结构和分布,引起了广泛关注。但由于深度置信网络本身采用的是无监督学习机制,因此,探讨如何使用深度学习方法来解决分类问题,特别是基于半监督学习的分类问题的研究工作还比较少。

本书提出了一系列基于深度置信网络的分类方法,使用半监督学习和监督学习方法来提升分类性能。研究内容主要包括以下几方面:

第一,区分深度置信网络。深度置信网络方法在数据压缩中取得了很好的效果,但是在分类方面的性能还有待提升。本书基于深度置信网络方法,提出了一种新的半监督学习方法,称为区分深度置信网络。该方法使用一个新的深层架构来集成深度置信网络的抽象能力和指数损失函数的区分能力,可以使用少量的标注数据达到比较好的分类效果。

第二,自适应深度置信网络。区分深度置信网络主要解决半监督学习问题,而没有解决如何用深度学习方法进行有效的监督学习的问题。本书基于区分深度置信网络和深度自动编码方法,提出了一种新的监督学习方法,称为自适应深度置信网络。该方法使用无监督学习方法来构建输出层,与使用随机数初始化输出层的经典方法相比,分类结果得到了提升。

第三,量子深度置信网络。量子神经网络是一个将模糊集和传统神经网络很好地结合起来的分类方法。为了进一步提升监督学习方法的性能,本书在自适应深度置信网络和量子神经网络基础上,提出了量子深度置信网络方

法。该方法使用一个新的深层架构来集成深层架构的抽象能力和量子神经网络的模糊表示能力。

第四,主动深度置信网络。为解决半监督学习中的有效样本选择问题,本书将自适应深度置信网络和主动学习方法相结合,提出了一种新的半监督学习方法,称为主动深度置信网络。该方法使用同一个深层架构来进行半监督学习和主动学习,使深层架构在主动学习过程中进行迭代训练,逐步提高抽象和分类能力。

第五,主动模糊深度置信网络。为了进一步提升主动深度置信网络的性能,本书将主动深度置信网络与模糊学习方法相结合,提出了主动模糊深度置信网络。该方法继承了深层架构优异的数据抽象能力和模糊集优异的分类能力。

最后,本书将深度学习方法应用到手写中文识别任务中,提出了一个有效的手写中文识别方法,称为基于深层架构的手写识别方法。该方法将深层架构的抽象能力和指数损失函数的分类能力相结合,使用深层架构进行粗分类,然后使用改进的二次分类函数进行细分类。

关键词: 深度学习,监督学习,半监督学习,主动学习,文本分类,图像分类,手写识别

Abstract

Recently, deep learning methods, one kind of learning method which based on multilayer neural networks, were proposed. Deep belief networks (DBN) is a typical deep learning method, which is usually composed of many hidden layers, has better capability of learning complex structures and distributions of real data sets than traditional neural networks. However, because of the unsupervised learning mechanism used in DBN, there is still few research on using deep learning method to solve classification problems, especially semi-supervised classification problems.

To improve the classification performance of semi-supervised and supervised learning, this book proposes several classification methods based on deep belief networks. The main contents of this book are as follows.

First, discriminate deep belief networks are proposed. DBN has reached good performance in data compression, however, the data classification performance is need to be improved. We propose a new semi-supervised learning method, discriminate deep belief networks, which uses a new deep architecture to integrate the abstraction capability of DBN and the discriminative capability of the exponential loss function. It helps the method reaching good classification performance by using fewer labeled data.

Second, adaptive deep belief networks are proposed. Discriminate deep belief networks are used to solve semi-supervised learning problems, and cannot solve supervised learning problems with deep learning method

efficiently. We propose a supervised learning method based on deep learning. This supervised learning method is referred to as adaptive deep belief networks, and is based on discriminate deep belief networks and deep autoencoder. This method uses an unsupervised learning method to construct the output layer. Compared with classical random initialization methods, this method can obtain a higher classification accuracy.

Third, quantum deep belief networks are proposed. Quantum neural networks is a classification method that combines the merits of fuzzy sets and traditional neural networks methods. To further improve the performance of supervised learning, we propose the quantum deep belief networks, based on adaptive deep belief networks and quantum neural networks. This method uses a new deep architecture to integrate the abstraction ability of deep architecture and fuzzy representation ability of quantum neural networks.

Fourth, active deep belief networks are proposed. To solve the effective sample selection problem, we propose a new semi-supervised learning method, active deep belief networks, based on adaptive deep belief networks and active learning method. This method uses the same deep architecture for semi-supervised learning and active learning, and iteratively trains it in active learning procedure, improves the abstraction and classification ability step by step.

Fifth, active fuzzy deep belief networks are proposed. To improve the performance of active deep belief networks method continuously, we propose a new semi-supervised learning method, active fuzzy deep belief networks, based on active deep belief networks and fuzzy learning method. This method inherits the powerful abstraction ability of deep architecture and powerful fuzzy classification ability of fuzzy sets.

At last, we apply deep learning methods in handwritten Chinese character recognition mission, and propose an effective handwritten Chinese character recognition method, handwriting recognition based on deep architecture. This method integrates the abstraction ability of deep learning method and discriminative ability of exponential loss function, uses deep architecture for coarse classification and modified quadratic discriminant function for fine classification.

Key words: Deep learning, Supervised learning, Semi-supervised learning, Active learning, Text classification, Image classification, Handwriting recognition

目 录

第 1 章 绪论	1
1.1 引言	1
1.2 机器学习方法	2
1.2.1 监督学习方法	3
1.2.2 半监督学习方法	4
1.3 深度学习方法	6
1.4 本书的研究内容	9
1.5 本书的结构安排	12
第 2 章 区分深度置信网络方法	15
2.1 引言	15
2.2 图像分类	16
2.3 区分深度置信网络	17
2.3.1 半监督学习问题描述	17
2.3.2 区分深度置信网络结构	18
2.3.3 区分深度置信网络的无监督学习方法	19
2.3.4 区分深度置信网络的监督学习方法	21
2.3.5 区分深度置信网络算法流程	25
2.4 区分深度置信网络实验	26
2.4.1 区分深度置信网络实验设置	26

2.4.2	在小规模人工数据集上的实验	27
2.4.3	在中规模图片数据集上的实验	28
2.4.4	在大规模手写数据集上的实验	30
2.4.5	在不同规模和深度的深层架构上的实验	31
2.5	本章小结	35
第3章	自适应深度置信网络方法	36
3.1	引言	36
3.2	自适应深度置信网络	37
3.2.1	监督学习问题描述	37
3.2.2	自适应深度置信网络结构	37
3.2.3	自适应深度置信网络的无监督学习方法	38
3.2.4	自适应深度置信网络的监督学习方法	40
3.2.5	自适应深度置信网络算法流程	40
3.3	自适应深度置信网络实验	42
3.3.1	自适应深度置信网络实验设置	42
3.3.2	在中规模图片数据集上的实验	43
3.3.3	在中规模手写字母数据集上的实验	43
3.3.4	在大规模手写数字数据集上的实验	46
3.4	本章小结	50
第4章	量子深度置信网络方法	51
4.1	引言	51
4.2	量子深度置信网络	52
4.2.1	量子深度置信网络结构	52
4.2.2	量子深度置信网络的监督学习方法	55

4.2.3	量子深度置信网络算法流程	56
4.3	量子深度置信网络实验	57
4.3.1	量子深度置信网络实验设置	57
4.3.2	在小规模花数据集上的实验	58
4.3.3	在小规模诊断数据集上的实验	60
4.3.4	在大规模手写数据集上的实验	61
4.4	本章小结	65
第5章	主动深度置信网络方法	66
5.1	引言	66
5.2	情感分类	67
5.3	主动深度置信网络	70
5.3.1	主动学习问题描述	70
5.3.2	主动深度置信网络的主动学习方法	71
5.3.3	主动深度置信网络算法流程	73
5.4	主动深度置信网络实验	75
5.4.1	主动深度置信网络实验设置	75
5.4.2	主动深度置信网络性能	76
5.4.3	主动学习效果	78
5.4.4	损失函数效果	79
5.4.5	使用不同数量标注数据实验	80
5.5	本章小结	82
第6章	主动模糊深度置信网络方法	83
6.1	引言	83
6.2	模糊深度置信网络	84
6.2.1	模糊深度置信网络结构	85

6.2.2	模糊参数提取	86
6.2.3	模糊深度置信网络算法	88
6.2.4	使用模糊深度置信网络分类	89
6.3	主动模糊深度置信网络	91
6.3.1	主动模糊深度置信网络算法描述	91
6.3.2	使用主动模糊深度置信网络分类	92
6.4	主动模糊深度置信网络实验	93
6.4.1	主动模糊深度置信网络实验设置	93
6.4.2	模糊深度置信网络性能	94
6.4.3	主动模糊深度置信网络性能	95
6.4.4	使用不同数量的标注数据实验	97
6.4.5	本书所提出的各种方法的训练时间	99
6.5	本章小结	101
第 7 章	基于深度学习的手写中文识别	102
7.1	引言	102
7.2	手写识别	103
7.3	使用深层架构的手写识别系统	104
7.4	手写识别系统中用到的深层架构	105
7.5	手写中文识别实验	106
7.5.1	在 HIT-OR3C 数据库上的实验	107
7.5.2	在 CASIA-OLHWDB1 数据库上的实验	111
7.5.3	在 SCUT-COUCH2009 数据库上的实验	111
7.6	本章小结	115
结论		116
参考文献		120

Contents

Chapter 1	Introduction	1
1.1	Background and Motivation	1
1.2	Machine Learning	2
1.2.1	Supervised Learning	3
1.2.2	Semi-supervised Learning	4
1.3	Deep Learning	6
1.4	Research Contents and Main Contributions of This Book	9
1.5	Organization of This Book	12
Chapter 2	Research on Discriminate Deep Belief Networks Methods	15
2.1	Introduction	15
2.2	Current Status of Image Classification	16
2.3	Discriminative Deep Belief Networks	17
2.3.1	Problem Formulation of Semi-supervised Learning	17
2.3.2	Architecture of Discriminative Deep Belief Networks	18
2.3.3	Unsupervised Learning of Discriminative Deep Belief Networks	19
2.3.4	Supervised Learning of Discriminative Deep Belief Networks	21
2.3.5	Algorithm Procedure of Discriminative Deep Belief Networks	25

2.4	Experiments for Discriminative Deep Belief Networks	26
2.4.1	Experimental Setting of Discriminative Deep Belief Networks	26
2.4.2	Experiments on Small-scale Artificial Dataset	27
2.4.3	Experiments on Middle-scale Image Dataset	28
2.4.4	Experiments on Large-scale Handwritten Dataset	30
2.4.5	Experiments on Different Scales and Depth of Deep Architecture	31
2.5	Summary of This Chapter	35
Chapter 3	Research on Adaptive Deep Belief Networks Methods	36
3.1	Introduction	36
3.2	Adaptive Deep Belief Networks	37
3.2.1	Problem Formulation of Supervised Learning	37
3.2.2	Architecture of Adaptive Deep Belief Networks	37
3.2.3	Unsupervised Learning of Adaptive Deep Belief Networks	38
3.2.4	Supervised Learning of Adaptive Deep Belief Networks	40
3.2.5	Algorithm Procedure of Adaptive Deep Belief Networks	40
3.3	Experiments for Adaptive Deep Belief Networks	42
3.3.1	Experimental Setting of Adaptive Deep Belief Networks	42
3.3.2	Experiments on Middle-scale Image Dataset	43
3.3.3	Experiments on Middle-scale Handwritten Alphabets Dataset	43