

人工智能深度学习概念研究与综述

杜艳明 王秋红

浙江工业职业技术学院, 浙江 绍兴 312000

摘要: 首先阐述深度学习的基础概念, 讨论了深度学习的发展历程, 并简要介绍了深度学习的基础模型, 进而讨论了深度学习的应用场景及最新研究成果, 在结束语中指明了深度学习仍是今后一段时间研究的热点和重点, 展望了今后深度学习的研究和发展方向。

关键词: 人工智能; 深度学习; 机器学习

中图分类号: TP391.4; TP181

文献标识码: A

文章编号: 1671-5780 (2016) 22-0072-01

1 引言

深度学习是机器学习研究中的一个新的领域, 其研究动机在于建立、模拟人脑进行分析学习的神经网络, 它期待模仿人脑的机制来解释数据, 例如图像, 声音和文本。类似机器学习方法一样, 深度机器学习方法也分为监督学习与无监督学习之分。不同的学习框架下建立的学习模型会有很大的不同。例如, 卷积神经网络 (Convolutional neural networks, 简称 CNNs) 就是一种深度的监督学习下的机器学习模型, 而深度置信网 (Deep Belief Nets, 简称 DBNs) 就是一种无监督学习下的机器学习模型。含多隐层的多层感知器就是一种深度学习结构。深度学习通过组合低层特征形成更加抽象的高层表示属性类别或特征, 以发现数据的分布式特征表示。

2 深度学习的概念及发展

机器学习的发展历程可以大致分为 2 个阶段: 浅层学习和深度学习。直到近些年, 大多数机器学习的方法都是利用浅层结构来处理数据, 这些结构模型最多只有 1 层或者 2 层非线性特征转换层。典型的浅层结构有: 高斯混合模型 (GMMs)、支持向量机 (SVM)、逻辑回归等等。在这些浅层模型中, 最为成功的就是 SVM 模型。但是当处理更多复杂的真实世界的问题时, 比如涉及到自然信号的人类语音、自然声音、自然语言和自然图像以及视觉场景时他们的模型效果和表达能力就会受到限制, 无法满足要求。早在 1974 年 Paul Werbos 提出了反向传播 (back propagation, BP) 算法, 解决了由简单的神经网络模型推广到复杂的神经网络模型中线性不可分的问题, 但反向传播算法在神经网络的层数增加的时候参数优化的效果无法传递到前层, 容易使得模型最后陷入局部最优解, 也比较容易过拟合。在很长一段时间里, 研究者们不知道在有着多层全连接的神经网络上怎样高效学习特征的深度层次结构。2006 年, Hinton 提出了深度置信网络 (deep belief network, DBN)^[1], 这个网络可以看作是由多个受限玻尔兹曼机 (restricted boltzmann machines, RBM) 叠加而成。从结构上来说, 深度置信网络与传统的多层感知机区别不大, 但是在有监督学习训练前需要先无监督学习训练, 然后将学到的参数作为有监督学习的初始值。正是这种学习方法的变革使得现在的深度结构能够解决以往的 BP 不能解决的问题。

3 深度学习的应用场景

深度学习目前在很多领域都优于过去的处理方法, 下面根据所处理数据类型的不同, 对深度学习的应用进行分类简要介绍。

3.1 深度学习在语音识别、合成及机器翻译中的应用

微软研究人员使用深度信念网络对数以千计的 senones (一种比音素小很多的建模单元) 直接建模, 提出了第 1 个成功应用于大词汇量语音识别系统的上下文相关的深层神经网络—隐马尔可夫混合模型 (CD-DNN-HMM), 比之前最领先的基于常规 CDGMM-HMM 的大词汇量语音识别系统相对误

差率减少 16% 以上。

3.2 深度学习在图像分类及识别中的应用

A. Krizhevsky 等首次将卷积神经网络应用于 ImageNet 大规模视觉识别挑战赛中, 所训练的深度卷积神经网络在 ILSVRC—2012 挑战赛中, 取得了图像分类和目标定位任务的第一。其中, 图像分类任务中, 前 5 选项错误率为 15.3%, 远低于第 2 名的 26.2% 的错误率。从深度学习首次应用于 ILSVRC 挑战赛并取得突出的成绩, 到 2014 年挑战赛中几乎所有参赛队伍都采用深度学习方法, 并将分类识错率降低到 6.7%, 可看出深度学习方法相比于传统的手工提取特征的方法在图像识别领域具有巨大优势。

3.3 深度学习在人脸识别中的应用

基于卷积神经网络的学习方法, 香港中文大学的 DeepID 项目以及 Facebook 的 DeepFace 项目在户外人脸识别 (labeled faces in the wild, LFW) 数据库 [6] 上的人脸识别正确率分别达 97.45% 和 97.35%, 只比人类识别 97.5% 的正确率略低一点点。DeepID 项目采用 4 层卷积神经网络 (不含输入层和输出层) 结构。之后, 采用基于卷积神经网络的学习方法, 香港中文大学的 DeepID2 项目将识别率提高到 99.15%, 超过目前所有领先的深度学习和非深度学习算法在 LFW 数据库上的识别率以及人类在该数据库的识别率。

4 结束语

深度学习展现的强大的学习能力和泛化能力, 表明它将是机器学习领域中研究的重点和热点。深度学习算法在大规模数据集下的应用取得突破性的进展, 但仍有一些问题值得进一步研究如无标记数据的特征学习, 模型规模与训练速度、训练精度之间的权衡及与其他方法的融合等。

参考文献

- [1] Sarikaya R, Hinton G E, Deoras A. Application of deep belief networks for natural language understanding[J]. IEEE Transactions on Audio, Speech and Language Processing, 2014, 22 (4): 778-784.
- [2] Wan L, Zeiler M, Zhang S X, et al. Regularization of neural networks using dropconnect[C]//Proceedings of the 30th International Conference on Machine Learning. Atlanta: IMLS, 2013: 2095-2103.
- [3] Bengio Y, Lamblin P, Popovici D, et al. Greedy layerwise training of deep networks[C]//Proceedings of 20th Annual Conference on Neural Information Processing Systems. Vancouver: Neural information processing system foundation, 2007: 153-160.
- [4] Palm R B. Prediction as a candidate for learning deep hierarchical models of data[D]. Technical University of Denmark, Denmark, 2012.