

基于人工智能下的机器学习历史及展望

文/余殷博

机器学习作为最近三十年新兴起的科学学科之一,已经在现代信息生活尤其是互联网生活中实现了较为广泛的应用。机器学习是在基于严格数学理论经众多学科交叉而成,包括信息控制理论、逻辑科学、数理统计学、神经科学、计算科学等。目前,机器学习已经衍生出了众多分支,如数据挖掘、深度学习、语音识别、生物信息学、模式识别、机器人的智能控制、遥感信息安全等。综上所述,机器学习已经取得了较大发展,并且将在未来的信息社会取得更为瞩目的成就。

【关键词】机器学习 多学科交叉 数理统计 人工智能 深度学习

机器学习是人工智能领域发展最快的一个分支之一。机器人学习的本质是通过利用计算机系统强大的运算及数据处理能力,借助大量的数据训练,使计算机具有自发模拟人类学习行为,通过学习获取知识和技能,在不断改善自身性能的同时,实现人工智能的一种思想。在信息时代,特别是随着互联网技术的爆发式发展,信息爆炸现象表现的尤为明显。现代社会互联网中每天都会产生的海量数字化数据,如数字化的微博数据,数字化的聊天记录、数字化的网页浏览数据等。针对以上海量数据,大数据(BigData)处理必然成为当今社会研究的一个热点问题。在这种大的氛围下,如何对海量数据进行分析并根据严格的数学定理提取有价值的规律信息,机器学习在以后的学科发展中必将占有一席之地。

1 机器学习的发展历史

1.1 机器学习“萌芽”阶段

上世纪五、六十年代,机器学习技术研究正处于“萌芽”阶段,人们试图给机器例如大型计算机通过编程手段使其最终具备逻辑推理能力,进而使机器具有一定的智能思考和自我优化的能力。这一阶段的代表性工作主要是由A.Newell和H.Simon完成的。所做工作包含各种“逻辑”程序以及之后的“求解”程序等,这些进展在当时令人振奋。A.Newell和H.Simon也因此获得了1975年图灵奖。然而,进一步的研究证明只具有逻辑并不能使机器具

有智能“”。E.A.Feigenbaum等人认为,智能存在的前提还必须具有先验“知识”。

1.2 机器学习“发展”阶段

上世纪七、八十年代被称为机器学习的“发展”阶段。在这一时期的主流为“专家系统”。“知识工程”之父E.A.Feigenbaum凭此在1994年摘取了图灵奖章。但是,所谓“专家系统”也要面临“知识困境”,简单地说,对近乎无限的信息人类很难通过自身思维提取规则并赋予计算设备。机器自主学习的设想浮出水面。机器学习相关工作在上个世纪五十年代就已经展开,主要进行的基于神经网络的训练学习方面的研究。

在二十世纪六七十年代,多种学习技术层出不穷,例如基于决策理论的统计学习技术以及强化学习技术等,代表作品为“跳棋程序”以及“学习机器”等,统计学习理论和符号学习技术开始萌芽。

1980年,人工智能领域TopJournal《策略分析与信息系统》专门以“机器学习”为主题连续开辟三期专栏;1983年,Tioga出版社出版了R.S.Michalski、J.G.Carbonell和T.M.Mitchell等顶尖专家联合主编的图书教材《机器学习:一种人工智能途径》面世,书中汇集了超过20位学者撰写的16篇高影响因子文章,对当时多年机器学习领域的研究工作进行了系统的总结和阐述,引领了学术方向,有较大影响。

1.3 机器学习“繁荣”阶段

从二十世纪八十年代至今,机器学习成为一个独立的学科领域并开始爆发式发展、各种机器学习技术不断涌现,机器学习算法呈现多样化。机器学习研究进入“繁荣”阶段,机器学习研究在这一时期也被科学的划分成“实例学习”、“求解规划学习”、“观察发现学习”、“指令学习”等多种范畴;而E.A.Feigenbaum等专家合著的经典图书《人工智能手册》中,则把机器学习技术从另一个角度重新划分为四大范畴,包含“机械学习”、“示教学习”、“类比学习”、“归纳学习”四种。直到今天,机器学习继续蓬勃发展并演化出了众多分支,例如数据挖掘、深度学习、语音识别、生物信息学、模式识别等。大量机器学习算法被广泛应用到信息处理特别是互联网海量数据的分析处理当中。

2 机器学习的发展趋势

从当前研究的发展趋势看,机器学习今后将有如下几个热点的研究方向:

(1)从人类自身出发找出大脑本身生物学习机制,通过严格数学化应用于机器学习。

(2)在已有的人工智能方法的基础上不断优化发展和改良现有学习算法,同时展开新的研究算法的开发工作。

(3)令众多的机器学习算法走出“象牙塔”,建立实用的机器学习的算法应用系统,特别是在互联网领域开展多种学习方法集成化的研究。

(4)多种机器学习算法的同步协调使用,利用多种算法是优势规避其中的不足,改善学习系统性能。

3 结语

现有的计算机系统和人工智能系统从其硬件结构和实现原理上不具备自主学习能力,至多也只是具有非常低级的“被动”学习能力,因而不能满足科技和生产提出的新要求。对机器学习进行比较系统化的讨论和方向把握,对以后学习机器学习方面的知识和进行相应的科研工作有方向性作用。学习机器学习本质是把控研究机器学习算法的内在数学原理、建立通过数据训练使现代计算机系统具备不断学习并自动提高自身水平。近年来机器学习理论在诸多领域例如天气预报、互联网、军事等取得成功,已成为计算机科学的基础研究热点之一。

参考文献

- [1] 杨树仁,沈洪远.基于相关向量机的机器学习算法研究与应用[J].计算技术与自动化,2010,29(01):43-47.
- [2] 何清,李宁,罗文娟,等.大数据下的机器学习算法综述[C]//中国计算机学会人工智能会议.2013.
- [3] 刘鲁,刘志明.基于机器学习的中文微博情感分类实证研究[J].计算机工程与应用,2012,48(01):1-4.
- [4] 李凡长,钱旭陪,谢琳,等.机器学习理论及应用[J].计算机工程与科学,2012,34(09):F0003-F0003.
- [5] 王昊,严明,苏新宁.基于机器学习的中文书目自动分类研究[J].中国图书馆学报,2010,36(06):28-39.

作者单位

武汉市外国语中学高二(3)班 湖北省武汉市 430050