

深度学习技术在党校工作中的应用初探

史转转¹, 周靖博², 寇冰²

(1. 中共武汉市硚口区党校, 湖北 武汉 430000; 2. 战略支援部队兴城特勤疗养中心, 辽宁 兴城 125105)

摘要:为解决党校系统缺乏对党务工作数据进行深层分析的功能从而制约了党校教学事业发展的的问题, 文章对深度学习技术在党校教学和管理领域的应用进行了深入探索。文章对党务工作现状进行了分析, 对深度学习的概念与相关技术进行了介绍, 从需求、数据、技术三个维度对深度学习技术在党校工作中的应用基础进行了说明, 文章列举了深度学习技术在党校工作领域的三个应用场景, 最后还对应用过程进行了详细阐述。深度学习技术的应用对于提升党校数据利用率、拓展相关功能、合理配置师资资源等方面具有积极意义。

关键词:深度学习; 数据挖掘; 党校; 模型构建; 模型训练

中图分类号: G647; TP399 文献标识码: A

文章编号: 1009-3044(2020)27-0034-03

DOI: 10.14004/j.cnki.ckt.2020.2860

开放科学(资源服务)标识码(OSID):



A Preliminary Study on the Application of Deep Learning Technology in Party School Work

SHI Zhuan-zhuan¹, ZHOU Jing-bo², KOU Bing²

(1. Party School of Qiaokou District Committee of Wuhan, Wuhan 430000, China; 2. Xingcheng Special Service Sanatorium of Strategic Support Force, Xingcheng 125105, China)

Abstract: In order to solve the problem that the party school system lacks the function of in-depth analysis of party affairs work data, which restricts the development of party school teaching, the article explores the application of deep learning technology in the field of party school teaching and management. The article analyzes the current situation of party affairs, introduces the concept of deep learning and related technologies, and explains the application basis of deep learning technology in party school work from the three dimensions of demand, data, and technology. The article lists deep learning Three application scenarios of technology in the field of party school work, and finally the application process is elaborated. The application of deep learning technology has positive significance in improving the utilization of party school data, expanding related functions, and rationally arranging teacher resources.

Key words: deep learning; data mining; party school; model construction; model training

1 背景

党校是我国教育培训党员干部的主渠道和研究宣传马克思主义的主阵地。党校的职责是, 通过一系列有计划地培训, 来提高广大党员干部用马克思主义的立场、观点和方法来处理实际问题的能力^[1]。党校系统的信息化建设, 是我国干部教育事业发展的必然要求, 也是新时代党员教育事业应对科技快速发展形势的必然选择^[2]。近些年来, 我国党校系统的信息化建设取得了跨越式发展, 工作效率和教学能力都有了显著提高, 但在基于党校需求方面的数据分析研究领域基本处于空白。目前, 各党政机关使用的应用系统均缺乏对党务工作数据进行深层分析的功能, 各党校的日常教学工作也是如此, 制约了党校教学事业发展。

当前, 党校工作模式正由传统经验模式向以教学与科研有机融合为核心的模式转变, 因此引入深度学习技术对于提升党校数据利用率、创新党校办学机制、优化党校师资资源、加强党

校内涵建设、提高党校教学实效等方面具有积极的意义。

2 深度学习介绍

深度学习(Deep Learning)这一概念是Hinton等人于2006年提出的, 属于人工智能研究领域中的一个新的分支。深度学习可以通过各种深度神经网络模型来对现实世界中的不同问题进行建模, 依托大数据对模型进行训练, 从而进行数据间的相关性发现, 以此来实现现实问题的预测分析和推理分析^[3]。经过多年的发展, 以深度学习为代表的人工智能技术, 已经在很大程度上颠覆了传统学科的研究方法。依托强大的感知能力、拟合能力和推理能力, 深度学习能适用于很多的应用场景, 例如: 自然语言处理、计算机视觉、生物医疗、推荐系统等。鉴于深度学习在学术界和工业界的巨大影响力, 2013年, 深度学习技术被研究人员评为世界十大技术突破之首。

深度学习从概念提出至今, 15年的时间里经历了发展期、爆发期, 中间新提出或衍生出了很多新的算法原理, 这些算法

收稿日期: 2020-06-25

作者简介: 史转转(1986—), 女, 山西汾阳人, 讲师, 硕士, 主要从事习近平新时代中国特色社会主义思想、政治经济学、党规党纪、基层党建等理论研究和教学工作。

在多项任务中都表现出色,典型的深度学习模型有:卷积神经网络(CNN)、循环神经网络(RNN)、深度神经网络(DNN)、自编码器(AE)等^[4],其中CNN、RNN、DNN属于监督学习,AE属于无监督学习。

目前,深度学习技术应用正逐渐渗透到我们生活的不同领域,我们日常使用的很多技术都融入了深度学习的相关技术,比如:图像识别、医疗数据分析、文字翻译、网络购物、广告推送等,且都取得了非常好的应用效果。特别是2016年3月,DeepMind公司开发的AlphaGo与围棋世界冠军、职业九段棋手李世石进行围棋人机大战,以4比1的总比分获胜。

3 应用基础

3.1 需求基础

党校系统经过多年的发展,虽然形成了相对完善的教学制度和管理理念,但工作模式相对固化、传统思维掣肘,不利于党校事业的进一步发展。为贯彻落实习总书记关于党校工作的重要指示精神,提高党校的智能化管理水平,党校系统也需积极地探索大数据,迎接人工智能时代。虽然目前以深度学习为代表的的人工智能技术在党校工作中的应用基本属于空白,但深度学习技术在党校领域应用潜力巨大,具有很强的可行性和实用性。在党校的教学培训、理论研究、日常管理等环节中,均可以利用深度学习技术进行党校数据的二次开发,提高管理服务水平。

3.2 数据基础

数据是数字时代的新“石油”,是深度学习三大核心要素之一,由于党校的正规化、信息化建设比较早,因此各党校在日常的教学、管理等过程中产生了大量的信息数据,包括结构化的数据和非结构化的数据。结构化的数据指的是党校信息系统在使用过程中产生的系统数据,例如党校电子文案系统、党校人员管理系统、党校财务管理系统等;非结构化数据指的是课程教案、公文资料、任职通知、工作记录、年终总结、人事档案等纸质资料。这些数据非常珍贵,为深度学习在党校系统内的应用提供了非常好的数据基础。

3.3 技术基础

这些年,深度学习发展很快,新模型层出不穷,深度神经网络、卷积神经网络、深度信念网络、生成对抗网络、循环神经网络等模型在不同领域都取得了非常好的应用效果。其中,在图像分类方面,基于CNN的技术已经超过了人类的准确率^[6];在语音识别方面,基于DNN的技术达到了95%的准确率^[7];在机器翻译方面,基于DNN的技术已经和人类的平均翻译水平大致相当^[8]。此外,面向深度学习的可用开源框架近些年也迎来了爆发式增长,例如:Tensorflow^[9]、Caffe、MXNet、Theano、Torch和PyTorch等,这些为深度学习技术在党校数据挖掘中的应用提供了非常好的技术基础。

4 应用分析

4.1 在个性化教学中的应用

个性化教学是近些年为了提高党校教学质量而提出的新方案,具体是针对不同类型的学员特点,在授课方式(讲授式教学、互动式教学、研讨式教学、体验式教学等)、授课内容(理论、案例、实操等)、授课时长等方面制定个性化的教学活动,具有

针对性强、教学效果好的优点。在实施过程中,利用深度学习技术可以分析得出不同特征学员的特点和需求,以此来制定针对性的教学方案。例如,黄宏涛^[10]基于BP神经网络提出了一种认知诊断方法,探讨该方法在个性化教学中的应用,结果表明:在小样本诊断中该方法能够为个性化补救教学提供依据;牛亚男^[11]基于神经网络技术设计并实现了一个教学资源个性化推荐系统,应用效果良好。

4.2 探索教学质量因素方面的应用

采用深度学习技术,可以从历史的教学数据和教学结果之间的因素,分析教学管理中授课信息之间的隐藏关系,探索不同授课方式、不同授课对象对教学质量的影响,为科学评估教学质量和教学管理提供决策支持。例如,温新^[12]提出了一种基于改进遗传算法和BP神经网络的教学质量评价模型;于权^[13]基于神经网络技术,结合军校教学工作的实际特点和部队教育发展对教学质量的要求开发了一种有效的军校教学质量评价系统。

4.3 在合理配置师资资源中的应用

党校教学资源有限,通过构建基于深度学习的教学资源优化模型,来合理地配置教师、教具、场地、财力、信息等资源,从而达到党校教学资源的最大化利用、最优化配置。例如,徐晓明^[14]在分析任务需求的基础上,提出了一种基于SVM的仿真资源调度模型,取得了整体最优的调度结果;夏锋等^[15]提出了一种基于神经网络的教学资源推荐方法,取得了较好的应用效果。

5 应用过程

5.1 数据准备和数据预处理

党校存储的丰富数据是深度学习在该领域应用的基础。数据从党校各个部门采集,包括结构化的电子数据和非结构化的纸质数据。不过为了满足深度学习数据分析处理的基本要求,需要提前对源数据进行数据预处理。对于以纸质文档形式留存的数据,比如:教学档案、满意度调查表、任职通知、工作记录、年终总结、人事档案等,需要进行数字化处理;对于系统内的数据库数据,主要包括以下操作:基于数据集成的数据抽取,基于数据审查的数据清洗,基于数据筛选的数据规约,基于数据序列化、归一化的数据转换,基于文本内容的关键词提取等等。

5.2 模型构建及数据训练

针对不同的应用场景,构建不同的深度学习模型,例如:在探索个性化教学中的应用时,考虑到授课时间长的因素,可以选用长短时间记忆网络。模型构建涉及:深度学习模型的选择和搭建以及参数的选择和调优等过程,相关参数包括输入序列、层数、层节点数、激活函数和损失函数等等。训练模型的过程中可以采用10折交叉验证(10-fold Cross Validation)与留一法(Leave-One-Out)、分层采样(Stratification)来测试模型的准确性。

5.3 功能实现及系统开发

模型在达到一定的精度之后即实现了相关功能,此时需要考虑基于该模型的软件系统的开发,系统在开发过程中需要与党校工作人员广泛交流,结合他们的工作经验,以期达到实用、科学的目的,最终实现拓展党校服务内涵,提高教学管理水平

的目标。

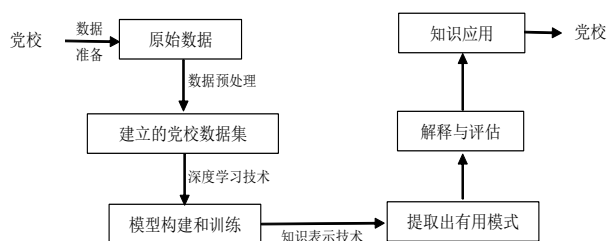


图1 深度学习技术在党校工作中的应用流程

6 结束语

现阶段,党校在数据分析研究方面基础薄弱,目前党校部门使用的系统均缺乏对数据进行深层分析以及对党校知识进行自动获取的功能,制约了党校工作的进一步发展。而深度学习技术的使用,为我们开展党校教学和管理工作大数据分析提供了技术基础,因此,引入深度学习技术对于党校系统具有非常重要的意义。虽然目前深度学习技术在党校教学领域的研究尚处于研究阶段,但随着相关工作的开展,深度学习技术在党校领域的发展趋势必将越来越深入、越来越广泛。

参考文献:

- [1] 邱志奎. 新时代加强党员干部党性教育的若干思考[J]. 才智, 2019(32): 44.
- [2] 李志清. 广州干部教育培训信息化建设研究——以广州行政学院为例[J]. 科技风, 2015(14): 225-226.
- [3] 李盼盼. 基于深度学习的个性化特勤疗养相关技术研究[D]. 长沙: 国防科技大学, 2017.
- [4] 段龙云, 辛婷婷. 递归神经网络多目标跟踪技术[J]. 电脑知识与技术, 2017, 13(26): 176-177, 188.

- [5] 李盼盼, 谭庆平, 曾平, 等. 机器学习技术在疗养数据挖掘中的应用初探[J]. 中国医疗设备, 2018, 33(4): 104-107.
- [6] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [7] 文铭. 基于深度神经网络的语音识别前端处理[D]. 合肥: 中国科学技术大学, 2019.
- [8] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate[C]. Computer Science, 2014.
- [9] 李河伟. 一种移动式TensorFlow平台的卷积神经网络设计方法[J]. 电脑知识与技术, 2017, 13(22): 179-182.
- [10] 黄宏涛, 李世珍, 李世玉, 等. 基于BP神经网络的认知诊断方法在个性化教学中的应用[J]. 中国远程教育, 2019(1): 86-91.
- [11] 牛亚男. 教学资源个性化推荐系统的设计和实现[D]. 大连: 大连理工大学, 2013.
- [12] 温新. 基于改进遗传算法和BP神经网络的教学质量评价模型研究[D]. 哈尔滨: 东北林业大学, 2019.
- [13] 于权. 基于BP神经网络军校教学质量评价系统[D]. 大连: 大连理工大学, 2016.
- [14] 徐晓明. 基于SVM机器学习的仿真网格资源调度模型[J]. 武汉理工大学学报(信息与管理工程版), 2013, 35(4): 555-559.
- [15] 夏锋, 牛亚男, 孔祥杰. 一种基于神经网络的教学资源个性化推荐方法[EB/OL].[2019-12-20]. <https://patents.google.com/patent/CN103353872B/zh>.

【通联编辑: 谢媛媛】

(上接第27页)

达到要求,性能上也接近了原生的HDFS方案,可以应用于实际的生产部署。

但方案仍有待完善之处,后续将从以下几个方面开展工作:

第一,对于网络I/O的问题,可以在此计算与存储分离架构下进行缓存模块设计,进一步提升系统性能。

第二,Ceph对象存储由于采用CRUSH算法,没有元数据服务器来存储相关的目录结构,故在做涉及rename命令操作的时候会影响到系统的性能,之后的工作可以重设rename机制,设计一个元数据管理模块来对rename的源文件与目标文件进行管理。

第三,我们也对业界的商业存储方案进行了调研,发现华为的大数据分离存储方案能够较好地解决上述问题,后续将会对华为存储开展相关功能与性能测试。

参考文献:

- [1] 马一力,傅湘林,韩晓明,等. 存储与计算的分离[J]. 计算机研究与发展, 2005, 42(3): 520-530.

- [2] Bernstein P A, Reid C W, Das S, et al. Hydr - A Transactional Record Manager for Shared Flash[C]// Conference on Cidr. 2011.
- [3] Aguilera M K, Merchant A, Shah M A, et al. Sinfonia: a new paradigm for building scalable distributed systems[J]. ACM Transactions on Computer Systems, 2009, 27(3): 5.
- [4] Ghemawat S, Gobioff H, Leung S T. The Google file system[J]. ACM SIGOPS Operating Systems Review, 2003, 37(5): 29-43.
- [5] 李大江. HDFS纠删码机制的优化研究[D]. 哈尔滨: 哈尔滨工业大学, 2018.
- [6] O'Driscoll A, Dugelaite J, Sleator R D. 'Big data', Hadoop and cloud computing in genomics[J]. Journal of Biomedical Informatics, 2013, 46(5): 774-781.
- [7] Jain P, Goel A, Gupta S C. Monitoring checklist for ceph object storage infrastructure[J]. Computer Science and Its Applications, 2015: 456: 611-623.
- [8] 翟永东. Hadoop分布式文件系统(HDFS)可靠性的研究与优化[D]. 武汉: 华中科技大学, 2011.

【通联编辑: 梁书】