

大数据治理中数据清洗方法的思考

陈世鹏

(北京锐安科技有限公司 北京 100000)

[摘要] 在各类大数据项目中,数据清洗是数据挖掘和应用的重要保障。脏数据会导致不可靠输出,如何高效去掉脏数据、提升数据质量、对异常数据进行修复是数据治理运营的重要环节。本文通过对数据质量问题及数据清洗方法的分析与总结,提出基于NLP算法对某类大数据行业的数据进行清洗和质量提升的方法。

[关键词] 数据清洗; NLP; 数据质量; 数据挖掘

[DOI] 10.12252/j.issn.2096-627X.2020.07.955

0 引言

随着移动互联网的发展,人类正处于一个信息爆炸的时代,面对海量的数据,人们仍然经常抱怨“数据丰富,信息贫乏”,因此怎么组织和存储数据,才能挖掘提炼数据价值、高效的获取所需的信息,是目前行业迫切关注的问题^[2]。在某些大数据行业,相较于一般企业的单一或几种数据源,它汇聚了各种类型的多源异构数据,但由于数据复杂的来源,受到源头的环境或录入、传输的影响等,数据在采集和导入过程中容易引入脏数据。由于没有统一的标准规定,只能依靠数据分析师对采集的数据进行分析和清洗,从而提升数据质量,据统计,数据清洗在一个项目生命周期中占据了相当比例的时长,因此数据清洗的方法和效率显得尤为重要^[1]。

1 数据质量问题

大数据相关很多行业的数据典型特征是数据源多,由于各业务系统在功能需求和设计上的不同,经常会出现数据不一致、字段命名冲突、属性值和结构冲突等问题。例如空值问题,这类数据主要是一些应该有的信息的缺失,影响对数据的分析挖掘;不一致问题,这类问题主要是对字段取值没有约束条件或约束条件过于简单,没有进行逻辑校验,表现在字段取值超出了规定范围。作为对数据清洗结果的检验标准,综合各参考文献,数据质量主要有如下10个维度^[3]评判标准:数据规范、数据完整性、数据重复检测、数据准确性、一致性和同步、及时性和可用性、数据覆盖、可理解性和数据衰变。

2 数据清洗方法

针对不同的数据质量问题,对应有不同的数据清洗方法,常见的清洗方法有:

1) 缺失数据处理

缺失值问题在一种普遍现象,常见的处理方法是忽略对数据分析挖掘无价值的特征值;对于一些具有连续变量值属性的特征变量,可以采用均值、插值或回归分析等进行填充缺失值。基于不完备数据分析的思想,通过定义约束容差集合的差异度,从集合角度判断不完备数据对象的总体差异程度,并以聚类的方法为基础可进行缺失数据的填补^[4];针对一元回归方法的预测进度不够,可以利用多元回归^[5],配以适当惩罚函数防止过拟合,在某些场景下对缺失值进行很好的估计。

2) 相似重复数据处理

将个别字段差异或完全一致的记录称为相似重复记录,在数据源数据集成时该现象普遍存在。重复数据不仅导致数据的冗余,浪费网络带宽和存储空间,还提供用户很多相似信息,对用户有误导作用。针对相似重复记录,邻近排序算法是常用的检测方法,还有各种改进排序算法等,例如多趟排序、优先队列等,还有基于机器学习进行相似度聚类来进行重复检测。

3) 逻辑错误数据处理

数据逻辑错误指特征变量与其属性值不符,违背了业务规则或逻辑。一般都是使用业务规则对逻辑错误进行检测,在具体的应用领域,根据领域知识制定约束规则,例如规定特征变量取值的有效范围、正则化匹配等来对数据的逻辑和有效性进行检测。

4) 不一致数据处理

不一致数据的常见问题有特征值的表示方法不统一、量纲不一致等,给数据分析带来了挑战。目前,常用消除数据不一致的方法有排序、融合和基于规则等方法。

5) 异常数据处理

异常数据指不符合一般规律的数据对象^[2]。其检测一种是基于统计学的方法,

即采用数理统计方法获取数据的总体分布特征,利用箱线图等方法判断异常点;一种是基于距离的方法,例如KNN等机器学习方法,将数据按距离划分成不同的层,根据定义的距离计算各数据点到中心点的距离来判断数据是否存在异常。

3 基于NLP算法的数据规则制定和清洗

在某些具有特殊业务应用的大数据行业中,上述常用的数据清洗方法难以起到立竿见影的效果。由于来源的多数据源特性,字段名称、属性值和结构都可能存在冲突,特征变量的名称比较混乱,无固定标准,命名不统一。例如,不同表中使用相同字段表示不同属性,不同名称的字段表示相同的属性。命名规则和标准的缺乏是该类行业数据清洗和治理的一大难点。因此这类数据清洗需要从结构层和数据层两方面进行处理。

随着人工智能的发展,自然语言识别(NLP)也有飞速的发展,充分利用前文的NLP算法,对特征变量命名进行规范,结合传统的数据清洗方案,可以有效提升类似大数据行业的数据质量。具体流程如下:

1) 整合语料库,形成初始数据元集合。目前有许多开源的语料库,由命名实体构成,命名实体一般分为3大类(实体类、时间类和数字类)和7小类(人名、地名、组织机构名、时间、日期、货币和百分比)。可对此进行整理,将命名实体进一步泛化,收集形成初始数据元集合。

2) 针对来源的多数据源,对其特征变量属性,选择合适的NLP算法来进行匹配,识别命名实体。如果该命名实体包含于1中的数据元集合中,可用该数据元进行命名;如果不存在,则对数据元集合进行扩充,并用新的命名实体进行命名。如此可对每张来源表统一命名规范,并对数据元的引用频次进行统计,根据业务需求可对数据元进行泛化扩充。

3) 在形成统一的数据规则后,然后根据数据质量问题,定义数据转换规则和有效性校验规则,采取对应的数据清洗方法对数据进行清洗,提升数据质量。

4) 根据规范后的结构和数据制定规格标准,对数据进行转换入库。

4 结语

目前,国外开发的数据清洗工具有很多,功能多样,各有优劣。数据质量问题也受到了各行各业的高度关注。本文分析总结了数据质量和质量问题的处理方法,并针对特定大数据行业的数据特点,提出了基于NLP算法的规则制定和清洗方案。当然,NLP算法也有多种,怎样选择合适的算法才能制定最佳的标准也需要进一步研究。

参考文献

- [1] 郭智慧,周傲英.数据质量和数据清洗研究综述[J].软件学报,2002.11
- [2] 宋金玉,陈炎.数据质量及数据清洗方法[J].实践与应用,2013.10第4卷第5期
- [3] McFilvray D.数据质量工程实践[M]刁兴春,曹建军,张健美,等译.北京:电子工业出版社,2010
- [4] 武森,冯小东,单志广.基于不完备数据聚类的缺失数据填补方法[J].计算机学报,2012,35(8):1726-1738
- [5] 张建新,方正,熊拥军,等.基于SNN数据清洗算法的优化[J].中南大学学报:自然科学版,2010,41(6):2240-2245
- [6] 庞雄文,姚占林,李拥军.大数据量的高效重复记录检测方法[J].华中科技大学学报:自然科学版,2010,38(2):8-11

幼儿园课堂教学有效提问的研究

陈颖

(广西玉林市玉东新区第一幼儿园 广西 玉林 537000)

[摘要] 幼儿园是学前教育的具体教育场所,它主要是针对3到6周岁的幼儿,幼儿园的主要任务就是帮助幼儿来摆脱家庭教育所带来的限制,让幼儿的身体和智力可以得到有效的开发,帮助他们提供正确的成长环境,是整个教育体系的基础,可以初步的对幼儿进行素质上的培养。在幼儿园的课堂学习过程中,提问是主要手段,教师需要从这方面进行具体的研究。本篇文章通过对幼儿园课堂教学提问中存在的问题进行阐述,分析幼儿园课堂有效提问的意义,并且就如何实现有效提问进行探讨。

[关键词] 幼儿园; 课堂教学; 有效提问

[DOI] 10.12252/j.issn.2096-627X.2020.07.956

引言

幼儿园主要以语言、科学、艺术、健康和社会五个领域构成,教学内容主要以游戏为主,在新时代下,教师需要加强教学与生活之间的联系,增强幼儿在学习中的主动性,构建和谐的课堂氛围。在幼儿园的课程结构体系上,它主要是由课程目标、课程内容、课程活动以及课程评价四个要素工程,它具有一定的基础性,是对幼儿的一种启蒙性培养,通过最浅显的一些观念进行品质和能力上的渗透。在幼儿园课堂教学中,由于幼儿在年龄以及性格上的特点,如何与他们产生有效的互动这是教师的主要研究方向,就课堂提问来说,它的设计内容、提问时间规划等方面都有着一定的欠缺,需要教师采取有效的措施看优化。

1. 幼儿园课堂教学提问中存在的问题

1.1 提问内容的问题

对于现阶段幼儿园课堂教学来说,很多教师都没有认识到提问的重要性,在问题内容的设计上没有很强的目的性,没有对幼儿进行一定的引导,往往是教师随口的提问,也没有对幼儿的回答进行分析和评价。其次,有些内容在设计的过程中过于直白,没有给幼儿思考的空间,无法形成思维模式上的培养,就是单纯的回答问题,无法真正的实现课堂提问的作用。实际上对于幼儿来说,这个阶段的思维塑

造和习惯养成是非常重要的,如果内容过于简单,不利于幼儿个性化的培养。

1.2 提问时机的问题

对于幼儿来说,他们在课堂上的注意力并不集中,没有集体的意识,做任何事情都是随着自己的心情,有些教师在提问上没有注意时机的重要性,没有注意对幼儿情绪上的调动,也没有注意提问过程中的层次性,这就导致有些幼儿根本无法对教师的问题进行理解,不会实现课堂提问的效果。

1.3 提问方法的问题

有些教师在对幼儿进行提问的方法设计上过于单一,往往就是采用一问一答的形式,没有设计一些有趣的提问环境,也没有一定的创新性,很难激发幼儿的兴趣,而且有些不同的幼儿有着不同的性格特点,如果采用统一的提问方式,许多幼儿会不敢表达自己真正的想法。

2. 幼儿园课堂有效提问的意义

2.1 集中幼儿的注意力

对于幼儿园阶段的学习来说,教师需要在课堂教学中加强与幼儿的互动,吸引他们的注意力,提高课堂提问的有效性就是在侧面上激发他们的学习兴趣,使他们的思维方式处于一个高度集中的状态,让幼儿可以带着求知欲进入到学习中,感受

到问题所带来的乐趣,让幼儿可以主动的参与到其中,这对于他们的品质培养来说是非常重要的。

2.2 开拓幼儿的智力水平

在幼儿园的教学中,它主要是帮助幼儿通过真实的活动以及相关经验提升自身的智力水平,有效的课堂提问可以帮助幼儿获得创造性的思维,在幼儿的常规思想中,一个问题只有一个正确答案,但是在课堂提问中,教师可以引导幼儿对同一问题进行多角度的研究,用不同的方法来思考问题,实现知识在横向、纵向之间的联系,这是对幼儿大脑智力开发的一种提升,也是创造性思维的初步培养。

2.3 明确幼儿的水平

除了对幼儿本身能力的提高之外,有效的课堂提问还可以帮助教师明确幼儿的整体水平,获得相应的信息反馈,为教师后续的课堂教学计划提供一定的依据,从而改善幼儿的学习方式,提高课堂教学效率,切实的改进教学质量。

3、幼儿园课堂教学有效提问的实施方案

3.1 加强问题内容设计的悬念性

设置悬念提问,有助于幼儿思维的发展和良好学习心理的培养。在故事教学中,教师常在幼儿了解故事的名称后或讲到某些关键时刻时,故意不说出故事如何发展、人物要说的话或要做的动作,从而造成悬念,让幼儿根据情节的发展进行合理的想象,使他们从不同方位、不同途径来探索问题的多种可能性。例如,在“小花猫过河”的教学中,教师出示故事的前半部分,让幼儿想想小花猫会如何过河?故事中出现的老牛、小狗、小鸭会给小花猫什么帮助?这些问题能激发幼儿的想象力,假设出各种可能性,有利于幼儿摆脱单一呆板的思维方式,培养幼儿丰富的想象能力。加强内容设计上的悬念性还可以让幼儿具有一定的好奇心,激发他们在这方面的兴趣,从而构建和谐的课堂气氛。

3.2 扩大课堂提问的范围

对于幼儿来说,许多教师都有着刻板的印象,在问题的设计上过于局限,认为幼儿的思维模式比较简单,在提问的范围上大多都是非常容易的,但是从课堂提问的效果以及幼儿的接受程度来说,教师应该合理的扩大课堂提问的范围,可以适当的尽心内容上的拓展,既可以激发学生的求知欲望,又可以减少对现有的提问内容进行创新。首先,教师应该对教材的相关内容进行深入的研究,对提问的难易程度进行选择,给予幼儿一定的思考的时间,一定要提高他们在课堂上的主动性。另外,如果有幼儿觉得题目难度过高,不能立即回答出来,教师可以灵活的设计一些铺垫性问题,增强幼儿的举一反三能力,增强学习动力,也要在此过程中加强对他们的鼓励,不要打压幼儿在课堂中的积极性。

3.3 注意提问的方式和时机

在幼儿园的课堂提问中,教师需要注意提问的时机,最大限度的调动幼儿的兴趣,这样才能起到事半功倍的效果。首先,教师可以选择在幼儿注意力最集中的时候,让他们主动的融入其中,引发在思维上的共鸣,增强师生之间的有效互动,弥补所讲知识上的不足,起到查漏补缺的作用。另外,在提问过程中,教师要避免同一提问方式的多次使用,要注意方式上的灵活性,可以让幼儿自由讨论和发言,教师进行适当的引导,活跃课堂气氛,也可以通过情境设置的方式,给幼儿创造一定的空间环境,但是教师需要注意提问过程中的目的性,最好具有一定的教学意义。

4、结语

综上所述,幼儿园需要为学生提供一定的良好环境,教师在课堂提问的设计上应该认识到现状的不足,然后加强内容上的悬念设计,注意提问的方式和时机,扩大课堂提问的范围,从根本上调动幼儿的积极性,从而提高课堂学习效果。

参考文献

[1]曹鑫莉.我国幼儿园集体教学活动教师提问研究回顾与反思[J].河南教育(幼教),2019(10):30-35.

中小学生学习安全教育存在问题与对策

罗海华

(广西北流市清水口镇清水口中心小学 广西 北流 537400)

[摘要]近年来中小学生的安全问题时有发生,影响中小学生的安全因素越来越复杂,对于学校的安全保护措施带来了一些冲击。中小学生的安全教育问题是实施教学的前提,保护中小学生的安全是学校、教师、家长都必须要注意的问题,要充分发挥家校作用,全方位的保护学生的安全,保护每一个家庭的完整与健康。

[关键词]小学生;安全教育;应对策略

[DOI] 10.12252/j.issn.2096-627X.2020.07.957

据相关的数据显示,我国每年都有近1.5万的中小学生学习遭受非正常死亡的伤害,其中包括交通事故、意外、校园霸凌等等。其中发现出现事最多的安全事件的隐患。而分析了安全事故隐患发生的原因则是因为学校缺乏对学生进行安全意识的培养以及不注意学生的安全问题,这才导致了安全问题成为事故的主要原因所在。中小学生的学习能力强,可塑性高,是培养安全意识和防火、防灾意识的关键时期。

一、中小学生学习形成安全问题的原因

1. 社会没有给予足够重视

随着社会经济发展和变革的产生,中小学生的安全问题也随之发生变化。据全国中小学生学习安全情况报告显示,在2018年公布的《2019年全国中小学生学习安全形势分析》指出报告不跟进社会的发展也不像社会公开进行说明,这说明此种类型的报告仅流于表面注重形式而不是真正的注重中小学生的安全问题。也说明了现在社会大众对我国中小学生的安全问题没有一个清晰的认识。

目前,各地中小学生的安全问题时有发生,但是我们在报纸或是新闻媒体上却不常见,这是因为当地人逃避承担责任,隐瞒事实真相、封锁消息,使得违法之人逍遥法外,受害人得不公正的处理。现在,我国法律明确要求公安机关要对学校周边区域增设巡逻点,及时的发现并消除安全隐患等问题。但是从具体的情况看,这一条似乎并没有落实。如在某市发生的学校门口杀人惨案,究其根本就是安全问题没有的到保障,致使悲剧的发生。

2. 学校内部的管理问题

政府部门积极落实国家的安全发展战略,学校也积极相应政府的号召,对学校的安全问题加以重视。但是仍有一部分人,对于学校的安全管理问题不注意。比如,某些教师的安全意识不到位,对学生的安全教育力度不够,安全演习较少,常规的检查不落实等等。或是在校内出现骑车载人,在校外家长用车子超载载人等情况时有发生。教师不光要提升学生的安全意识,还要对家长的安全意识加以培养,造成全员安全意识的提高。

3. 学生与家长的安全意识匮乏

学生赖以信任的不光有教师还有他们的家长,所以有时家长的安全意识也直接或间接影响着学生的安全意识。有的家长对安全意识不重视或是缺乏安全教育意识,没有辅导学生安全意识的能力。例如,安全用电和防火意识是大家都知道的常识,在学生玩火或是用电危险的时候,家长只会责骂而不进行正确的引导正确的安全意识也会导致学生对安全意识没有具体的概念。有些家长只关心孩子的成绩,把剩下所有的教育都推给学校或是社会,这也会造成安全意识的缺乏。还有的家长带着孩子闯红灯,破坏孩子安全意识和遵守规则的意识。对于中小学生的安全意识的培养,家长有着不可推卸的责任。

二、提高中小学生学习安全教育的必要性

提高中小学生的安全意识是以以人为本教育理念为基础,也是教育回归基础的重要使命,加强学生的生存能力的有效方法。

1. 提高中小学生的安全意识

中小学生的年纪较小,易冲动行事,对自身的安全意识不强对社会的了解也不全面。一般表现为:学校内活泼好动,不计较后果与得失,对社会了解不够,缺乏辨别是非善恶的能力;不注重用火、用电和出行的安全,侥幸心理较高;在打架、盗窃等方面意识认识不够给自身和他人带来伤害。所以全方位,立体的角度进行多角

度教育的培养,是提高学生安全意识与正确行为能力的方法。

2. 加强中小学生学习面对复杂社会形势的需要

由于现在信息发展速度较快,学生的社会化现象比较明显,以及校园的周边出现一些社会性的闲散人员或是一些辍学的失学人员会做出一些社会性的行为,例如抽烟、喝酒、骂人、打架斗殴等等问题的出现。出现一些以大欺小,索要保护费等的情况。还会出现一些反社会的人出现,校园暴力的一再发生。加强安全教育对中小学生学习复杂的社会情况是现在迫切需要的。

3. 落实中小学生学习全面发展教育的需要

对学生进行全面安全教育,培养合格人才的产生。从学校角度开展安全教育可以全面提高学生的整体素质,为学生的健康成长和营造良好的文化氛围做出贡献。中小学生的自我约束以及遵守纪律和保障校内的安全都是小学生实现全面发展的关键所在。中小学生的全面发展在这两方面得到了体现。

三、如何加强安全教育

中小学阶段学生的可塑性高,学习能力强,在学校、家庭、社会三个方面都是进行安全教育的有效场所,能够养成良好学习习惯与安全意识。在面对突发事件时,在应对方面会有积极正面的作用。

1. 学校教育是中小学生学习安全教育的主要力量。

一个专业素质能力强的安全教育的团队,对于中小学生学习安全意识以及能力的培养有着积极的作用。学校在构建专业的安全教育团队时,需要专业能力强、素质高的团队以及一个管理能力强的团队。

具体如何开展中小学生的安全教育,可以做好以下几个方面:

(1) 将学校进行的相关活动所要注意的事项提前告诉学生。例如,在进行学校中的实践活动,或是要进行的体育测试的注意事项等等。

(2) 校园周边存在的安全问题以及注意事项告知学生。例如,学校哪里进行施工,路过请注意或是周边道路在进行修路等等。

(3) 研究当地的每年所发生的一些情况,如暴雨、或是暴雪等等,挖掘安全素材,根据安全素材进行演练,如一些多发生地震的地区,对于地震的演习必须进行演练保证学生安全有效的撤离。

2. 家庭教育是中小学生学习安全教育的主要力量

父母对于孩子的教育是终身的,会影响孩子一生的发展。家庭教育也是中小学生学习的一个重要方面。是对学校教育的补充。虽然家庭教育不如学校教育那样的系统和专业,但是对一个中小学习品格的形成会非常重要的,都说孩子是父母的影子,家长对于孩子的示范作用一直影响着孩子的一生。家长要善于利用一切机会教育和引导孩子进行正确的安全教育。

四、总结

重视中小学生的安全教育是对一个家庭负责任,更是我们每一个人的义务。中小学生学习是国家的未来,民族的希望,他们的安全关乎着社会的稳定和家庭的幸福。为了每一位中小学生的健康成长,必须将安全教育贯彻于教育的每一个方面,为他们创建一个安全的学习环境和一个良好的学习氛围。

参考文献

[1]柯艺灿.浅谈新形势下小学生安全教育[J].福建论坛:社科教育版,2008(04).