

基于网络监测大数据的医学生健康信息行为研究

韩 华

昆明医科大学 云南 昆明 650500

摘要: 随着互联网技术的发展和普及,网络监测大数据逐渐成为医学生健康信息行为研究的新兴手段。本文以医学生为研究对象,采用网络监测大数据的方式,基于上网行为审计数据,通过数据挖掘和分析技术,研究了医学生健康信息行为的现状和特征,对比了医学生与非医学生在健康信息行为上的差异,分析了影响医学生健康信息行为的因素。本文的研究结果有助于深入了解医学生健康信息行为的特点和规律,对促进医学生的健康教育和健康管理以及医学教育等方面具有一定的借鉴意义。

关键词: 网络监测大数据; 上网行为审计; 医学生; 健康信息行为; 数据挖掘

1 引言

随着信息技术的迅速发展,人们可以通过网络获得大量的医学健康信息。作为医学领域的专业人士,医学生对于健康信息的需求和获取方式不同于普通人群。研究医学生健康信息行为模式,对于医学教育和健康管理等方面都具有一定的借鉴意义。

目前,健康信息行为的研究主要通过传统的调查问卷方式来实现。但是,传统调查问卷方式存在着样本量小、主观性强、数据来源单一等问题,并很难做到倚靠控制,难以全面、准确地反映医学生的健康信息行为特点。而网络监测大数据方式则可以通过对互联网上用户自发上传和分享的信息进行挖掘和分析,有效控制倚靠,从而更全面、准确地了解医学生健康信息行为的特点和规律。

2 研究方法

2.1 数据收集

本研究采用网络监测大数据的方式,以某医科大学的医学生作为研究对象,通过上网行为审计系统收集了其在校期间的上网行为数据。通过这种方式获取数据优势非常明显。

1. 数据量大: 基于网络监测大数据进行研究可以获取大量的数据,数据量远远超过传统的问卷调查和实验研究,因此可以获取更加全面和详尽的数据。

2. 无需干预: 网络监测大数据的获取过程不需要对

受访者进行干预或观察,因此可以避免传统问卷调查或实验研究中可能出现的主观干扰和信息偏差。

3. 数据质量高: 网络监测大数据的获取过程不需要对受访者进行干预或观察,因此可以避免传统问卷调查或实验研究中可能出现的主观干扰和信息偏差。

2.2 数据处理

收集到的数据包括上网时间、上网地点、浏览网页、搜索关键词、使用应用程序等在校学生的网络活动记录。数据收集时间为2021年至2022年,共涵盖了该校不同年级、不同专业的学生。数据的采集和存储都严格遵守了相关隐私保护法律法规,数据经过脱敏和最小化授权处理,实现数据可用不可见。根据研究需要,仅选取了年级、年龄、性别、专业等信息,这里的专业仅用于区分医学生和非医学生。

为保证研究结果的准确性和可靠性,对收集到的数据进行了清洗和整理。具体来说,首先筛选出与医学生健康信息行为相关的信息,如搜索关键词、应用程序名称、社交媒体讨论等。然后,对数据进行去重和去噪处理,以避免重复和无效数据对研究结果的影响。最后,对数据进行了分类和编码,以方便后续的分析 and 建模。

2.3 数据分析

在本研究中,采用了多种统计学分析方法,包括描述性统计、回归分析和方差分析等。其中,描述性统计主要用于对样本数据进行初步的统计分析,包括样本数量、均值、标准差、最大值和最小值等;回归分析则用于探究各个变量之间的线性关系,通过构建回归模型来分析变量之间的因果关系;方差分析则用于分析不同组别之间的差异性,通过比较不同组别之间的均值来判断它们是否有显著的差异。

2.4 数据挖掘

通讯作者: 姓名: 韩华, 出生年月: 1982年1月, 民族: 汉族, 性别: 男, 籍贯: 河南新野, 单位: 昆明医科大学, 职称: 工程师, 学历: 硕士, 邮编: 650500, 研究方向: 网络安全、数据分析。

基金项目: 云南省教育厅科学研究基金资助项目(编号: 2019J1212)

数据挖掘是一种通过自动或半自动的方法在大规模数据集中发现潜在模式的过程。在本研究中，使用了多种数据挖掘技术收集到的数据分析和建模。

首先，使用了聚类分析来识别不同的用户群体，并探索其行为模式。聚类分析是一种将数据样本分成不同组别的技术，目标是使同一组别内的数据样本相似，而不同组别之间的数据样本差异较大。

其次，使用了文本挖掘技术来分析用户在社交媒体上的言论和讨论，采用了自然语言处理技术来处理和分析文本，以识别出用户的情感倾向和主题。用于自动识别和提取文本中的情感信息。

3 研究结果与分析

3.1 医学生健康信息行为的现状和特征

首先，对医学生上网行为进行了分析，显示医学生在上网时最常浏览的网站类型是搜索引擎、学术网站和医疗健康网站，占有所有浏览网站的比例分别为41.1%、14.7%和12.3%。这表明医学生在上网时主要关注学习和健康相关信息。

其次，对医学生的健康信息行为进行了聚类分析，发现医学生的健康信息行为可以被分为5类，即与疫情相关的健康信息、心理健康、营养与健身、疾病相关、睡眠质量。其中，与疫情相关的健康信息是医学生在疫情期间搜索最为频繁的健康信息之一，搜索次数占比达到了16.2%。其次是心理健康相关的搜索，占比为11.5%，健康方面占比为10.3%，疾病相关占比为6.7%，睡眠质量占比为3.8%。其他类型的健康信息搜索次数占比相对较低，不予统计。

表1 健康信息行为聚类分析

序号	搜索关键词	搜索次数占比
1	新冠疫情（病毒、防控措施、口罩）	16.2%
2	心理健康（压力、焦虑、抑郁）	11.5%
3	健康（营养、饮食、减肥、健身）	10.3%
4	疾病（体检、疾病预防、疾病筛查）	6.7%
5	睡眠（失眠、睡眠质量、入睡困难）	3.8%

最后，研究了医学生上网行为和获取健康信息行为之间的相关性。结果表明，医学生浏览学术网站和医疗健康网站的频率与其健康信息获取行为有显著相关性，而浏览社交媒体和娱乐网站的频率则与健康信息获取行为无显著相关性。

3.2 医学生与非医学生健康信息行为的差异

首先，对医学生和非医学生样本数据进行了方差分析。结果显示，在健康知识获取、健康信息评估、健康信息利用等维度上，医学生的得分均显著高于非医学

生。并对数据在整体健康信息行为得分上进行了t检验。结果表明，医学生的整体得分显著高于非医学生，相比于非医学生，医学生更关注自身的健康信息行为。医学生和非医学生在健康信息行为上存在着一些差异。具体来说，医学生更加关注疾病的预防和治疗，而非医学生更加注重保健和健康饮食。

其次，通过对医学生与非医学生获取健康信息的文本数据进行了情感分析可以看到明显的差异，在健康信息行为中，医学生消极情感占比12.5%，而非医学生的消极情感占到21.8%，可以看到非医学生的消极情感明显高于医学生。这样的结果可能是由于医学生的专业背景和职业需求，更加积极地寻求和接受健康信息，相对而言，非医学生可能没有这样的意识和需求。医学生在学习和实践中接触到更多的健康问题，对健康信息的态度更加积极，相比之下，非医学生可能没有这样的背景和经历，因此可能更容易产生消极情感。医学生可能更容易从专业渠道中获得健康信息，例如医学书籍、学术期刊等，这些渠道通常提供比较科学和准确的健康信息，相比之下，非医学生可能更多地从互联网和社交媒体等非专业渠道中获取健康信息，这些渠道存在一定的误导性和不准确性，因此可能导致更多的消极情感。

表2 医学生与非医学生健康信息行为情感分析对比

情感 \ 分类	医学生	非医学生
积极	35.6%	23.7%
中性	51.9%	65.5%
消极	12.5%	21.8%

3.3 医学生健康信息行为的影响因素

在本研究中，探究了可能影响医学生健康信息行为的因素，并进行了统计学分析。结果表明，以下几个因素可能会影响医学生的健康信息行为。

3.3.1 年级：研究结果显示，医学生在大一和大二的时候，其健康信息行为显著高于大三和大四。这可能是因为医学生在前两年学习基础医学知识时对健康信息的认识更多，而在后两年接触临床实践后可能会转向更注重临床操作技能。大一和大二的学生可能还处于适应大学生活和学习的阶段，对自身的健康问题更为敏感，随着时间的推移，可能逐渐适应并降低对健康信息的关注。因此，这些原因可能导致大一和大二学生的健康信息行为显著高于大三和大四。

3.3.2 性别：研究结果表明，女性医学生的健康信息行为得分显著高于男性医学生。这可能是因为女性更注重自身健康，并更愿意花费时间和精力去获取和学习健

康信息,女性对健康的重视程度可能比男性更高,因为女性身体的生理结构和功能特点决定了她们在保持健康方面需要更加关注。因此,女性医学生可能更容易受到健康信息的影响,更积极地寻求相关信息

3.3.3 疫情:新冠疫情期间,由于新冠疫情的突发性和快速传播,一些传统的信息渠道可能无法及时更新,医学生健康信息获取渠道受到了很大的影响。因此,医学生在疫情期间更倾向于使用实时更新的互联网信息来了解健康知识和相关政策,健康信息行为也受到了影响。由于医学生与疫情直接相关,他们需要承受更大的心理压力和焦虑感。同时,由于疫情的不确定性,医学生可能会出现恐慌、恐惧等情绪,这些情绪也可能影响他们的健康行为和获取信息行为。

4 结论与展望

本研究通过使用上网行为审计系统对医学生的健康信息行为进行了分析,揭示了医学生的健康信息行为现状、特征和影响因素,并与非医学生进行了比较,得出了不同类型学生的健康信息行为的差异。同时,研究还探究了疫情对医学生健康信息行为的影响。研究结果为促进医学生和社会公众的健康教育和预防医疗提供了重要的参考。

虽然本研究取得了一些有益的结果,但也存在着一些局限性和不足之处。未来的研究可以进一步深入探究这些问题,以实现更加精准、有效的健康信息管理和预防医疗。可以进一步探究这些方法和技术在健康信息行为领域的应用,并开发更加智能化和高效的健康信息管理系统。

参考文献

- [1]杨霞,李泉政,耿娜娜,等.国内健康信息行为研究进展[J].电脑知识与技术,2022,18(6):73-75.
- [2]黄凤仪,宋慧军.河北某高校医学生健康信息素养调查分析[J].医学研究与教育,2020,37(2):63-69. DOI:10.3969/j.issn.1674-490X.2020.02.010.
- [3]杨阳,王鹭,王荣.医学院校大学生健康信息搜寻行为研究[J].医学信息学杂志,2020,41(9):47-50. DOI:10.3969/j.issn.1673-6036.2020.09.009.
- [4]尹博.健康行为改变的跨理论模型[J].中国心理卫生杂志,2007,21(3):194-199. DOI:10.3321/j.issn:1000-6729.2007.03.015.
- [5]金浏河.新冠肺炎疫情影响下大学生网络媒介素养提升方案[J].科教文汇,2020(26):49-50. DOI:10.16871/j.cnki.kjwhb.2020.09.024.