

《防灾减灾中的数据挖掘》教学案例库建设实践

李忠¹, 李姗姗², 单维锋¹

(1. 防灾科技学院应急技术与管理学院, 河北廊坊 065201;

2. 防灾科技学院计算机科学与工程学院, 河北廊坊 065201)

[摘要] 随着信息技术的飞速发展, 数据挖掘技术在各个领域的应用日益广泛。在防灾减灾领域, 数据挖掘技术能够从海量数据中提取有价值的信息, 为决策提供科学依据。本论文以数据挖掘技术在防灾减灾领域中的若干实际应用作为案例, 详细介绍了研究背景、采用的数据挖掘技术、教学案例库构建、案例库在教学中的作用及效果, 并得出结论。旨在通过系统化的案例库建设, 提高这门课程的教学质量, 并为相关领域的研究和实践提供参考。

[关键词] 教学案例库; 防灾减灾; 数据挖掘技术; 教学效果

[作者简介] 李忠(1966—), 男, 山东诸城人, 防灾科技学院应急技术与管理学院教授, 工学博士, 研究方向: 人工智能、数据挖掘等。李姗姗(1981—), 女, 河南洛阳人, 防灾科技学院计算机科学与工程学院副教授, 工学博士, 研究方向: 数据挖掘技术。单维锋(1977—), 男, 山东聊城人, 防灾科技学院应急技术与管理学院教授, 工学博士, 研究方向: 大数据技术、人工智能等。

[基金项目] 本文系河北省研究生专业学位精品教学案例(库)建设项目(项目编号: KCJPZ2023060); 河北省高等教育教学建设与改革项目(项目编号: 2022GJJG484)。

[DOI] <https://doi.org/10.62662/kxwxz0208027>

[中图分类号] G642

[本刊网址] www.oacj.net

[投稿邮箱] jkw1966@163.com

人类社会是在同自然灾害和人为事故的长期斗争中而不断地发展的, 从远古一直走到现在。随着全球气候变化和城市化进程的加快, 自然灾害和人为事故的发生频率和强度不断升级, 给人类社会带来了巨大的损失和安全隐患, 防灾减灾已成为人类社会的共识。数据挖掘技术作为一种从海量数据中提取有价值信息的技术, 在防灾减灾领域具有广阔的应用前景。例如, 利用数据挖掘技术, 可以从气象、地质、水文、土壤等多源数据中提取出规律性的知识, 预测灾害的发生, 从而为防灾减灾提供科学依据。

作为大学里多个专业的一门课程, 数据挖掘技术的教学中, 在有关理论、算法教学基础上, 多以信息欺骗、网络攻击、商业或者金融等数据分析为例加以说明, 而很少介绍防灾减灾领域的数据挖掘技术应用案例。这严重限制了大数据、云计算、物联网、人工智能、移动互联等现代信息技术在灾害监测、防灾减灾、应急救援等领域的推广应用。因此, 面向防灾减灾的教学案例库建设将数据挖掘技术和实际相结合, 注重培养学生解决复杂工程问题的实践能力。从实际教学过程来看, 传统教学主要是

教师授课、学生被动接受的形式, 难于达成“以生为本、因材施教”的目标, 缺少了沉浸感, 不能充分调动学生的主观能动性, 教学效果难达预期。现有教材大多侧重理论介绍和简单例子说明, 非常生涩、枯燥, 缺少对实际案例的深入解析, 使得学生难以对数据挖掘技术有比较透彻的认识。因此, 在数据挖掘技术的实际教学中, 很有必要增加相应的实际案例, 以帮助学生理解原理、方法和应用, 促使学生通过对实际案例的模仿和优化, 提高学习效率和对相关知识掌握的广度与深度。为此, 本文将依托多年来利用数据挖掘技术解决地震监测预警、防灾救灾等的工程案例, 建立《数据挖掘技术》教学案例库, 为广大师生和防灾救灾从业者提供数据挖掘应用经验。

一、数据挖掘技术课程主要内容概述

(一) 数据特征分析与预处理技术

作为数据挖掘的前导, 需要对数据集进行统计分析, 如最大值、最小值、中值、1/4 值和 3/4 值等“五数概要”, 了解数据的分布特点, 为后续数据挖掘算法的选择提供基础。

另外, 数据挖掘技术课程需要介绍一些基本的

预处理方法,如除噪方法、归一化方法、数据集成方法等,为获得合适的数据集做准备。

(二) 回归分析

作为一种传统的数据分析方法,回归分析在数据分类、趋势预测等领域应用广泛。数据挖掘技术课程主要介绍线性回归、非线性回归和逻辑回归,在防灾减灾领域具有广泛的应用。

(三) 分类与预测

之所以将分类和预测放在一起,是因为二者具有相通性。按照定义,对离散数据的操作是分类,而对连续数据的操作是预测。这也是数据挖掘中最基本的技术之一,其目的是将数据集中的数据分成不同的类别,在模式识别、趋势预测、异常点查找等领域具有广泛的应用。常用的分类算法包括决策树、贝叶斯推理、kNN、时间序列分析、人工神经网络等。

(四) 聚类

聚类是将数据集中的数据分成若干个簇的过程,使得同一簇内的数据相似度较高,不同簇间的数据相似度较低,在防灾减灾中可以用于灾害区域的划分和灾害影响的评估等方面。常见的聚类算法有 K-means、K-means++、BIRCH 算法、CURE 算法、ROCK 算法、DBSCAN 算法、自组织神经网络、一趟聚类算法等。

(五) 关联分析

关联分析是发现数据集中不同数据项之间的关联关系的过程,如购物篮中不同商品的关联、致灾因素之间的关联性等。常见的关联分析算法有 Apriori、FP-Growth、CARMA 等。

(六) 异常点挖掘

异常点挖掘是发现数据集中与其他数据显著不同的数据点的过程,是分类挖掘的具体应用。常见的异常点挖掘算法有基于统计的方法、基于距离的方法、基于密度的方法和基于聚类的方法等。异常点挖掘在防灾减灾中的应用包括查找地震前兆数据、挖掘各种地质灾害发生前的异常现象,以便对潜在的危害进行预警。

(七) 文本挖掘与 Web 挖掘

文本挖掘与 Web 挖掘分别是文本数据和 Web 数据中提取有价值信息的过程,常见的文本挖掘技术有文本分类、文本聚类和主题模型等,常见的 Web 挖掘技术有内容挖掘、结构挖掘和日志挖掘等。通过文本挖掘或者 Web 挖掘,可以发现灾害事件的发生和影响、舆情传播等。

二、教学案例库构建

对于教学案例库的结构,不同的课程、不同的

教师对此有不同的解读。如有人认为教学案例内容要包括案例名称、首页注释、摘要和关键词、背景信息、案例正文、结语、案例思考题、使用说明、参考文献、附件,以及与案例相关的其他资料;还有人从案例入库原则、案例结构、案例库内容体系设计、案例库建设方法、案例库管理与维护等几个方面建设;也有人给出了教学案例库的建设流程:案例素材的来源、选用原则、构成要素。

本文在综合不同教师的教学案例库建设经验基础上,删繁就简,提出教学案例库的内容建设,包括目的、方法描述、数据来源及预处理、过程说明与结果分析、案例使用说明等几个方面。

下面用数据分类挖掘的例子加以说明。

目的:利用 BP 神经网络对唐山市岩溶塌陷区进行分类评估。

理论分析及模型:岩溶塌陷预测工作,首先要确定由哪些参数构成预测评价指标体系,从而也就确定了输入参数。根据唐山市地质构造特征和岩溶塌陷实际情况,建立了如图 1 的预测评价指标体系,其中包括 8 个二级指标和 16 个三级指标。

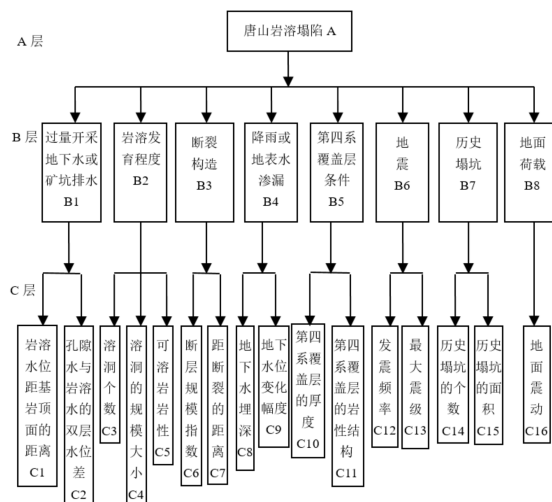


图 1 岩溶塌陷分类评估主控指标体系

人工神经网络(Artificial Neural Network, ANN)是模拟人脑思维方式的数学模型,用来模拟人类大脑神经网络的结构和行为,反映了人脑功能的基本特征,如学习、联想、并行处理、模式分类、记忆等。其中,BP神经网络是一种非线性数据处理的监督训练模型,采用多层前馈网络的反向传播(Back Propagation, BP)学习算法,具有算法简单、容错性好、自适应能力强等优点,能够将高维数据经过处理后变为低维数据,常用于数据预测领域,在科学研究、工程设计中应用广泛。

数据来源及预处理:从上述16个参数可以看出,有些是定量的数值型,有些是定性的标称量,需要从多种渠道获得这些数据,包括实地测量、相关地质资料、水文地质资料、图纸等等,这些数据的获取往往耗费很大精力和财力。将获得的原始数据保存在数据库中备用。

数据预处理:原始数据往往存在噪声、缺失、重复、量纲不统一、不同参数数值差别巨大等问题,这就需要进行数据预处理操作。

处理过程与结果:传统的BP神经网络方法易陷入局部极值,且计算效率不高,因此实际中常采用改进的BP神经网络方法,如增加学习因子 η 、动量因子 α 和陡度因子 λ ,以提高网络的收敛速度。经过多次实验,选取学习因子 $\eta=0.8$,动量因子 $\alpha=0.6$,陡度因子 $\lambda=0.8$,构建完成岩溶塌陷预测的BP网络挖掘模型。

利用5条测试样本数据对该模型进行精确度测试,获得100%的正确率,因此该模型可行。

利用上述构造的三层BP网络岩溶塌陷挖掘模型,将各样本点数据代入预测挖掘模型,得到各个测点的稳定性评价结果。然后将此结果与GIS地图进行集成,形成工作区的岩溶塌陷预测成果图,如图2所示。

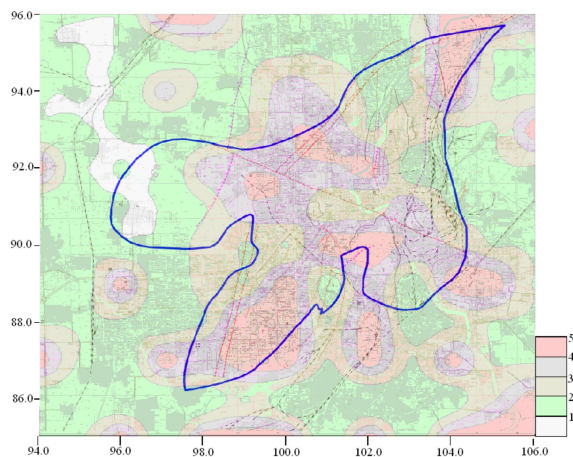


图2 神经网络预测成果图

在图2中,图示1、2、3、4、5分别对应岩溶塌陷的稳定性评价等级:稳定、基本稳定、难塌、易塌、极易塌。可以看出岩溶塌陷易塌区集中在工作区的中间位置,颜色越红表明越易塌陷。

案例使用说明:

该案例属于数据挖掘技术中的数据分类挖掘内容,因此在讲授数据分类挖掘内容时可以采用该案例进行讲解。案例中采用BP人工神经网络分类

模型,让学生了解BP神经网络的拓扑结构、输入端数据、输出端类别及可视化的图形图像,掌握模式分类识别的原理、过程、结果和结论的全过程。

三、案例库在教学中的作用及效果

(一)案例库在教学中的作用

通过实际应用案例的教学,学生可以更好地理解数据挖掘技术在防灾减灾中的应用,提升学习兴趣,增强实践能力,掌握数据挖掘技术的实际应用,提升综合素质,培养解决实际问题的能力。

(二)案例库在教学中的效果

从三个方面说明。

首先学生反馈情况。通过对学生的问卷调查,发现学生对案例库的教学效果给予了高度评价。90%的学生认为案例库提升了他们的学习兴趣,85%的学生认为案例库增强了他们的实践能力,80%的学生认为案例库提升了他们的综合素质。

其次教学效果评估情况。通过对学生的考试成绩和实验报告进行评估,发现案例库的教学效果显著,学生的考试成绩平均提高了10%,在上机实验时学生兴趣大增,实验报告的质量有了显著提升。

再次教学案例库的推广情况。目前来自实际工程的案例已达到20多个,按照“目的、方法描述、数据来源及预处理、过程说明与结果分析、案例使用说明”等几个方面进行整理,拟出版发行,更好地推广传播,惠及更多的学生、教师和工程技术人员等。

四、结论

本文以面向防灾减灾的《数据挖掘技术》教学案例库建设实践为研究对象,详细介绍了研究背景、数据挖掘技术课程内容、教学案例库构建、案例库在教学中的作用及效果,得出结论如下:

(1)教学案例库建设,提升了学生的学习兴趣、实验能力和报告编写水平,有效提高了教学质量和效果。通过系统化的案例库建设,提升了数据挖掘技术在防灾减灾领域的教学效果。

(2)通过不断地累积实际科研和工程案例,按照一定的规则整理成册并出版发行,惠及更多的读者,为防灾减灾研究和实践提供参考。

参考文献:

- [1] 吕辉.“多传感器信息融合技术”研究生课程案例库建设与教学实践[J].黑龙江教育,2024(10):51-53.
- [2] 韩耀振,王常顺,侯明冬,等.专业硕士计算机控制教学案例库建设[J].高教学刊,2024(34):84-88.
- [3] 张红军.农业管理专业学位研究生教学案例库建设研究——以《现代农业经营与管理案例分析》课程为例[J].现代畜牧科技,2025(1):178-180.

[4] 鱼明.课程思政背景下经管类专业《数据挖掘》课程
教学改革与实践[J].湖北开放职业学院学报,2025,38(9):
90-92,95.

[5] 陆永来.《Python 数据分析》课程教学案例库建设与
应用研究[J].办公自动化,2024(19):41-44.

Construction of Teaching Case Library for “Data Mining Technology in Disaster Prevention and Reduction”

LI Zhong¹, LI Shan-shan², SHAN Wei-feng¹

- (1. School of Emergency Technology and Management, Institute of Disaster Prevention, Langfang Hebei 065201;
2. School of Computer Science and Engineering, Institute of Disaster Prevention, Langfang Hebei 065201, China)

Abstract: With the rapid development of information technology, data mining technology is increasingly being applied in various fields. In the field of disaster prevention and reduction, data mining technology can extract valuable information from massive data and provide scientific basis for decision-making. This paper takes several practical applications of data mining technology in the field of disaster prevention and reduction as case studies, and provides a detailed introduction to the research background, adopted data mining technology, construction of teaching case library, the role and effectiveness of case library in teaching, and draws conclusions. The aim is to improve the teaching quality of the course through the construction of a systematic case library, and provide reference for research and practice in related fields.

Key words: teaching case library; disaster prevention and reduction; data mining techniques; teaching effectiveness