

王琳, 杨天青, 邱鹏, 等. 基于支持向量机回归多属性在地震应急决策中的应用探讨[J]. 华南地震, 2022, 42(2): 41-45. [WANG Lin, YANG Tianqing, QIU Peng, et al. Application of Multi-attribute Regression Based on Support Vector Machine in Earthquake Emergency Decision Making [J]. South China journal of seismology, 2022, 42(2): 41-45]

基于支持向量机回归多属性在 地震应急决策中的应用探讨

王 琳¹, 杨天青², 邱 鹏¹, 郝 婧¹

(1. 贵州省工程防震研究院, 贵阳 550001; 2. 中国地震台网中心, 北京 100045)

摘要: 地震应急是为了减轻地震灾害而采取的异于正常工作程序的紧急防灾和抢险行动。传统的震后应急决策主要依据应急条例标准以及决策者的经验, 导致出队决策可能并非最优方案。随着数据挖掘和人工智能技术的发展, 目前多种推荐算法在决策推荐领域正在被广泛的应用, 对应急成员的基本情况按工龄、集合时间、应急次数3个属性进行分析, 利用遗传算法全局最优收敛的特性, 通过对支持向量机的3个属性参数进行寻优推荐, 得到最优地震应急出队成员。

关键词: 地震应急; 支持向量机回归; 寻优推荐

中图分类号: P511.2

文献标志码: A

文章编号: 1001-8662 (2022) 02-0041-05

DOI: 10.13512/j.hndz.2022.02.06

Application of Multi-attribute Regression Based on Support Vector Machine in Earthquake Emergency Decision Making

WANG Lin¹, YANG Tianqing², QIU Peng¹, HAO Jing¹

(1. Guizhou Institute of Earthquake Engineering, Guiyang 550001, China; 2. China Earthquake
Networks Center, Beijing 100045, China)

Abstract: Earthquake emergency is an emergency disaster prevention and rescue action which is different from normal working procedures to reduce earthquake disasters. The traditional post-earthquake emergency decision-making is mainly based on emergency regulations and the experience of decision makers, which may not be the optimal solution. With the development of data mining and artificial intelligence, various recommendation algorithms are widely used in the field of decision recommendation. In this paper, the basic situation of emergency response members is analyzed according to three attributes: length of service, gathering time and emergency response times. Using the characteristics of global optimal convergence of genetic algorithm, the optimal earthquake emergency team members are obtained by optimizing and recommending three attribute parameters of support vector machine.

Keywords: Earthquake emergency response; Support vector machine return; Optimization recommendation

收稿日期: 2021-03-10

基金项目: 国家重点研发计划(2018YFC1504506); 六盘水市活动断层探测项目(黔科合支撑[2020]4Y053)联合资助。

作者简介: 王琳(1994-), 女, 工程师, 硕士, 主要从事地震应急、地震地质、工程地震方面的研究。

E-mail: wanglin199@126.com.

0 引言

地震是人类共同面临的突发性较强、致灾过程复杂的自然灾害之一,我国是世界上地震活动最强烈、地震灾害最严重的国家之一。传统的震后应急决策主要是根据现行的国家和地方地震应急预案、条例、规范、标准结合决策者的个人经验等来部署现场工作队的规模、结构情况,这可能导致出队方案并非最优,且由于不同工作任务对接存在的脱节,决策部署效率低且人为因素影响较大,本文主要针对地震应急工作实施过程中存在的这些问题进行研究,拟采用支持向量机多属性回归计算模型,为决策者制定出队方案提供一定参考。

1 计算方法简介

目前数据挖掘和人工智能技术的快速发展,多种推荐算法在辅助决策、个性化智能推荐等方面得到了广泛应用,如我们所熟知的购物平台商品智能推送,新闻媒体、公众号的信息推荐、音乐播放平台的每日歌单推荐等。遗传算法是建立在遗传学与自然选择基础上的自适应搜索过程,作为解决复杂问题的一种有效手段,从一个初始化的群体出发,选取性能优良的个体,在这些优良个体之间进行生物化的繁殖,使它们一代代地进化到解空间中越来越好的区域,从而最终得到最优解。K最近邻分类算法(k-nearest neighbor, 简称 KNN 推荐算法)的原理是一个样本在某一特征空间中,与其最相似的 k 个样本中的大部分属于某一类,则该样本也属于这一类。

地震的发生随机性强,破坏性地震致灾范围广、强度大,给人民的生命财产安全、社会秩序的稳定造成重大威胁,因此地震现场应急工作时间紧任务重,对应急现场工作人员的应急应变能力,专业知识储备情况等要求较高。本文依据应急预案、条例等,对应急现场队成员的情况进行分析后,发现应急现场队成员参与地震应急工龄、实际现场应急次数、到单位的最快集合时间为三个较为重要的,代表现场队员地震应急处置能力的属性,本文利用遗传算法全局最优收敛的特性对支持向量机的上述三个属性参数进行寻优,采用支持向量机回归的方法拟合不同属性,有效解决三个不同属性之间的关联性问题,再采用 KNN 算法推荐得到最优应急团

队^[1-4],以得到最优解的支持向量机推荐现场队员。

2 支持向量机推荐模型的构建

(1)由于应急团队不同属性的性质不同,通常具有不同的数量级,若各属性间的水平相差很大时直接用原始属性值进行分析,就会突出数值较高的属性在综合分析中的作用,相对削弱数值水平较低属性的作用。因此,为了保证结果的可靠性,需要对原始属性数据进行标准化处理,再利用标准化的数据进行数据分析。本文研究中的备选应急队员在工作年龄、应急次数、应急集合时间三种属性之间关联性较小,因此采用 min-max 标准化处理方法,该方法称之为离差标准化,对这三种属性进行数据标准化线性变换,目的是将结果值映射到0到1之间^[2],函数公式如下:

$$x = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

式(1)中: x_{\min} 为样本数据最小值; x_{\max} 为样本数据最大值; x 为样本数据。

(2)支持向量机的基本概念:首先通过非线性变换将输入空间变换到一个高维空间,然后在这个新空间中求取最优线性分类面,而这种非线性变换是通过定义适当的内积函数来实现的。支持向量机求得的分类函数形式上类似于一个神经网络,其输出是若干中间层节点的线性组合,而每一个中间层节点对应于输入样本与一个支持向量的内积,因此也被叫做支持向量网络^[5-8]。结构如图1所示:

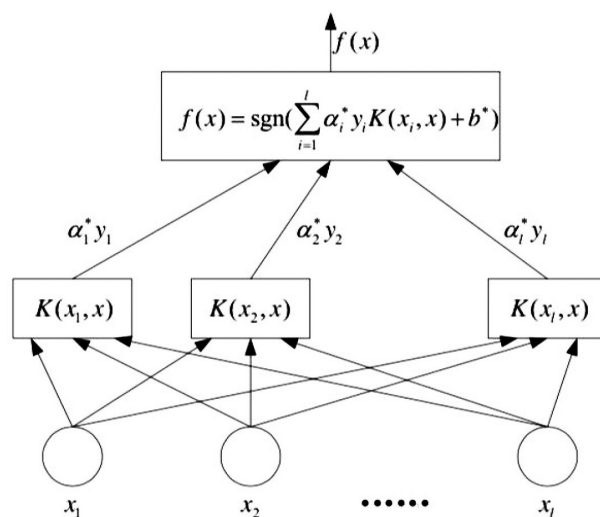


图1 支持向量机概念图

Fig.1 Concept diagram of support vector machines

SVM 应用于回归拟合分析时,不再是寻找一个最优分类面使得两类样本分开,而是寻找一个最优分类面使得所有训练样本离该最优分类面的误差最小^[9],概念图如图2。

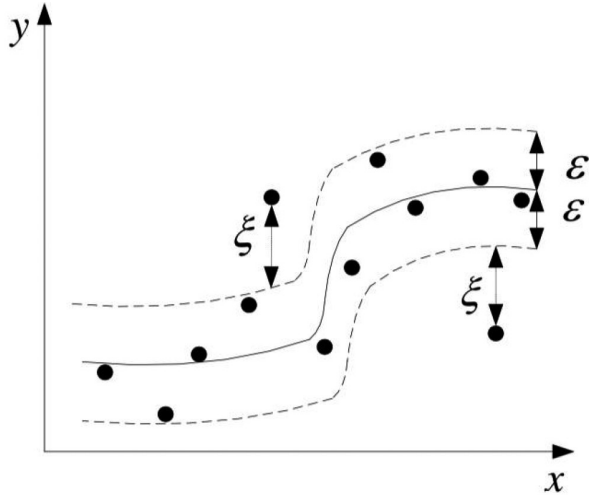


图2 回归支持向量机概念图

Fig.2 Concept diagram of regression support vector machines

应急团队的学习样本集表示为 $\{(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i)\}$ ($x_i \in R^n$ 为输入变量, y_i 为与 x_i 对应输出变量), 回归问题是指从样本中学习一个函数, 然后给定一个 x_i 值, 能够预测 y_i 的值, SVR 函数公式如下^[10-11]:

$$f(x) = \omega \Phi(x) + b \quad (2)$$

式(2)中 $\Phi(x)$ 为 R^n 到高维空间的一个非线性变换。确定适当的参数值 ω 和 b 的值, 使回归风险系数最小化, 回归风险系数公式如下^[10]:

$$R_{\text{reg}}(f) = C \sum_{i=1}^l r(f(x_i) - y_i) + \frac{1}{2} \|\omega\|^2 \quad (3)$$

式(3)中: $r(x)$ 是损失函数, 为提高模型的鲁棒性, 引入惩罚因子 $c > 0$ 。

利用 Lagrange 乘子向量、对偶原理及核函数方法, 将 SVM 的约束问题最终转化得到如下回归函数:

$$f(x) = \sum_{i=1}^l (a_i - a_i^*) K(x, x_i) + b \quad (4)$$

对于任何 $i=1, 2, \dots, l$, 式中 $0 \leq a_i, a_i^* \leq c$, 且 $a_i \times a_i^* = 0$ 为拉格朗日算子; $K(x, x_i)$ 为核函数, 本文选用常用的高斯核函数来构造支持向量机。高斯核函数公式为:

$$K(x, x_i) = \exp\left(-\frac{\|x - y\|^2}{\sigma^2}\right) \quad (5)$$

式(5)中 σ 为径向基核带宽的调节参数, 结合公式

(4)得到支持向量机推荐模型。

$$f(x) = \sum_{i=1}^l \left(a_i - a_i^* \right) \exp\left(-\frac{\|x - y\|^2}{\sigma^2}\right) + b \quad (6)$$

式(6)可以看出, 支持向量机模型的性能和泛化能力主要取决于参数 c 和 σ , 常采用试验法来确定参数, 但容易陷入局部解, 因此考虑引入采用遗传算法来确定参数 c 和 σ , 遗传算法具有全局最优解的特点, 这样能够保证模型的有效性和准确性。

3 支持向量机回归多属性模型

地震现场应急队员的工龄、应急次数、应急集合时间, 是决定应急团队最优解的关键, 但这些属性是非线性的, 因此需要构建一个 SVR 多属性回归模型^[5-6], 以解决三个不同属性之间的关联性问题, 回归模型可构建为:

$$r_i = \omega \cdot (m(A_k|i), m(B_s|i), m(C_u|i)) + b \quad (7)$$

式(7)中: r_i 应急队员的综合评分, $m(A_k|i)$ 是应急队员的工龄断评分, $K \in \{0, 1, 2, \dots, 17\}$, 共 18 个工龄。 $m(B_s|i)$ 是应急队员参与过地震应急的次数段评分, $u \in \{0, 1, 2, \dots, 8\}$, 共 9 个参与地震应急救援次数。 $m(C_u|i)$ 是应急集合时间评分, $v \in \{0, 1, 2, \dots, 9\}$, 共 10 个应急集合时间。

支持向量机回归通过核函数 $K(x, x_i)$ 将数据从空间非线性变换到高维空间, 从而使得数据在该空间中被转换成线性组合^[11], 用支持向量机回归的方法构建非线性回归模型公式如下:

$$r_i = \sum_{i=1}^l (a_i - a_i^*) K((m(A_k|i), m(B_s|i), m(C_u|i), m(D_v|i)), x) \quad (8)$$

使用序列最小优化方法求解该支持向量机回归问题。

4 实验结果分析

本文以贵州省地震局目前的地震应急出队分组作为参考, 按地震应急预案的要求将应急队伍分为 ABCD 四类, 其中 A 为震情监视跟踪组, B 为通信与后勤保障组, C 为烈度评定与灾害评估组, D 为综合协调组, 按顺序对组员进行编号 A01、02... 09; B01、01... 10; C01、02... 18; D01、02... 04, 分别统计了每位成员的工龄、现场工作次数和集合时间(到单位的最快时间), 将应急团队的学习样本集表示为 $\{(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i)\}$ (x_i 为输入变量, y_i 为与 x_i 对应输出变量), 使用

SVM 工具训练基于支持向量机回归的用户多属性推荐模型，尝试寻求最优的 c 值，得到最优分类面，计算出的各个应急队员在上述三个属性条件下离最优分类面最小值 ω 结果详见表 1， ω 值越小，则表示该成员为最优出队选择，反之则相反，根据计算得到的结果可以筛选出所需的应急队员，完成队员的推荐。

表 1 应急队员多属性推荐表
Table 1 Multi-attribute recommendation table for emergency team members

编号	工龄	现场工作次数	集合时间	ω	分组
A01	11	3	30	0.572 272 666	A---震情监视跟踪组
A02	5	9	40	0.666 584 849	
A03	4	2	18	0.747 310 683	
A04	2	0	15	0.86 919 656	
A05	2	0	25	0.83 762 522	
A06	3	2	45	0.740 794 073	
A07	5	4	30	0.681 692 129	
A08	4	0	20	0.740 562 473	
A09	0.5	0	30	0.92 925 664	
B01	11	2	50	0.519 530 746	B--通信与后勤保障组
B02	20	10	40	0.33 195 936	
B03	1	0	30	0.902 235 666	
B08	0.5	0	30	0.92 925 664	
B09	0.5	0	30	0.92 925 664	
B10	0.5	0	30	0.92 925 664	
C01	10	4	30	0.581 870 423	C--烈度评定与灾害评估组
C02	5	7	30	0.681 692 129	
C03	8	6	30	0.607 508 188	
C04	10	8	30	0.581 870 423	
C05	15	4	40	0.509 725 941	
C06	1.5	0	10	0.934 429 148	
C07	8	1	40	0.592 400 908	
C08	3	1	40	0.755 120 303	
C09	0.5	2	30	0.92 925 664	
C10	1.8	0	30	0.843 545 964	
C11	0.5	1	15	0.96 830 971	
C12	0.5	0	15	0.96 830 971	
C13	1.5	0	50	0.811 843 868	
C14	12	10	60	0.423 371 539	
C15	5	3	10	0.751 535 489	
C16	0.5	0	30	0.92 925 664	
D01	20	7	60	0.20 736	D--综合协调组
D02	1	0	15	0.941 288 736	
D03	1.5	1	50	0.811 843 868	
D04	2	2	30	0.83 014 349	

从表1可以得出,若每个组只选择一名成员。则A-震情监视跟踪组安排出队的最佳人员是A01, B-通信与后勤保障组的最佳出队人员是B02, C-烈度评定与灾害评估组的最佳出队人员是C14, D-综合协调组的最佳出队人员是D01,在实际的出队安排中,各队选出的最佳人员可作为每队的组长(副组长)候选人,供指挥决策者参考。

通过与传统的贵州省地震应急现场工作队出队情况相比,本文采用的基于支持向量机回归的用户多属性推荐得出的最佳成员在发生在贵州省境内的2019年10月沿河 $M_s4.9$ 级地震与2020年7月赫章 $M_s4.5$ 级地震中担任队长或副队长的角色,是各组中地震应急现场工作经验和技术知识最为丰富,熟知应急工作的流程、内容,综合能力相对突出的成员,这也从侧面映射出地震现场工作对应急技术人员的专业知识储备、应急应变能力要求较高。

5 结论

经过多年的经验积累与探索研究,我国已形成了一套适宜的应急工作体系,但受人员专业知识背景、工作基本能力等的影响,在近年来的应急工作中,具体实施的环节之间出现了不同程度的任务对接脱节,导致决策部署效率低的现象。

本文采用支持向量机多属性回归计算模型,完善决策者安排出队人员,实验结果与队员实际出队能力相符,证明该方法用于筛选地震应急出队人员是可行的,实现应急团队的智能推荐,不仅能提高地震应急工作与人工智能的结合,推进地震事业走向智能化、信息化,也能完善和弥补传统决策模式导致的决策部署效率低、任务工作脱节等问题。因此通过不断完善队伍人员信息,增加人员属性参数,不断优化,可进一步实现队伍成员的筛选,从决策制定的角度出发,可以为震后应急现场工作快速高效部署、有序实施的提供重要基础支撑。

参考文献

- [1] ALPAYDINE. 机器学习导论[M]. 范明, 咎红英, 牛常勇, 译. 北京:机械工业出版社, 2009.
- [2] wuxD, KuMARV. 数据挖掘十大算法[M]. 李文波, 吴素研, 译. 北京:清华大学出版社, 2014.
- [3] 王君贤. 基于随机森林与支持向量机的心血管疾病预测研究[D]. 天津:天津大学, 2018.
- [4] 阙晓辉. 基于面部特征的人脸识别技术研究[D]. 武汉:武汉理工大学, 2007.
- [5] LIN Chen, XIE Runquan, Guan Xinjun, et al. Personalized news recommendation via implicit social experts[J]. Information Sciences, 2014 (254): 1-18. DOI: 10.1016/j.ins.2013.08.034..
- [6] 施凤仙, 陈恩红. 结合项目区分用户兴趣度的协同过滤算法[J]. 小型微型计算机系统, 2012, 33(7): 1533-1536.
- [7] 王宏宇, 糜仲春, 梁晓艳, 等. 一种基于支持向量机回归的推荐算法[J]. 中国科学院研究生院学报, 2007, 24(6): 742-748.
- [8] LIU J, DOLAN P, PEDERSEN E R. Personalized news recommendation based on click behavior[C]//Proc. 15th international conference on intelligent user interfaces. New York City: ACM Press, 2010.
- [9] 王红军, 周宇, 王伦文. 基于SVR-Kriging插值的矿井工人二维指纹定位数据库构建算法[J]. 电子与信息学报, 2017, 39(11): 2571-2578.
- [10] 纪振平, 高志强. 基于SVM的连铸二冷目标控制模型研究[J]. 沈阳理工大学学报, 2018, 37(01): 47-50+86.
- [11] 赵广杰, 尹四清. 基于支持向量机回归多属性智能电视电影推荐[J]. 电子技术及信息科学, 2015, 39(06): 32-35.