

引用本文

吴煜,经理晗,解江.融合知识元和聚合 K 近邻算法的客舱火灾案例检索研究[J].消防科学与技术,2026,45(8):41-48.

Wu Yu, Jing Lihan, Xie Jiang. Research on cabin fire case retrieval by incorporating knowledge meta and aggregated K nearest neighbor algorithm[J]. Fire Science and Technology, 2026, 45(8): 41-48.

融合知识元和聚合 K 近邻算法的客舱火灾案例检索研究

吴煜¹, 经理晗¹, 解江²

(1. 中国民航大学 安全科学与工程学院, 天津 300300;

2. 中国民航大学 科技创新研究院, 天津 300300)

摘要:为实现对民航客舱火灾情景的准确描述和快速响应,本文构建了融合知识元和聚合 K 近邻算法的客舱火灾案例检索模型。通过挖掘客舱火灾的属性特征,作数值转换和预处理,进行相似案例检索和等级预测。结果表明:基于灾害要素解构理论可将客舱火灾案例解构为含致灾因子、承灾体、孕灾环境和应急处置的四维知识元框架,结合火灾事件特点,将客舱火灾的灾害形成、发展及应对的全过程细分为8类子知识元和12个属性特征;同时,计算随机预测目标案例与全量案例的欧式距离,采用多数投票法和聚合子模型得到最终预测值,构建1500条仿真案例数据集进行训练、验证和调优,选取24条真实案例进行测试,利用留一法交叉验证得出最佳 K 值为3,模型预测准确率达到0.9167。基于准确率、精确率、召回率和 $F1$ 分数设置多模型对比评估显示,所提出的聚合 K 近邻模型显著优于其他模型。因此,提出的客舱火灾案例检索模型,可为民航客舱火灾等级预测及应急决策提供支撑。

关键词:客舱火灾;知识元模型;聚合 K 近邻算法;案例检索;事件等级

中图分类号:X932;X915.1 **文献标志码:**A **DOI:**10.20168/j.1009-0029.2026.08.0041.08

Research on cabin fire case retrieval by incorporating knowledge meta and aggregated K nearest neighbor algorithm

Wu Yu¹, Jing Lihan¹, Xie Jiang²

(1. School of Safety Science and Engineering, Civil Aviation University of China, Tianjin 300300, China; 2. Institute of Science and Technology Innovation, Civil Aviation University of China, Tianjin 300300, China)

Abstract: To achieve accurate description and rapid response to civil aviation cabin fire scenarios, this paper constructs a cabin fire case retrieval model that integrates knowledge meta and aggregated K nearest neighbor algorithm. By mining the attribute features of cabin fires, performing numerical conversion and preprocessing, similar case retrieval and level prediction are carried out. The results indicate that based on the theory of disaster element deconstruction, cabin fire cases can be deconstructed into a four-dimensional knowledge element framework containing causative factors, disaster bearing bodies, disaster prone environments, and emergency response. Combined with the characteristics of fire events, the entire process of cabin fire disaster formation, development, and response can be subdivided into 8 sub knowledge elements and 12 attribute features; At the same time, the Euclidean distance between the randomly predicted target case and the total number of cases was calculated. The majority voting method and aggregation sub model were used to obtain the final prediction value. A dataset of 1500 simulation cases was constructed for training, validation, and optimization. 24 real cases were selected for testing, and the best K value of 3 was obtained through cross validation using the leave one method. The model's prediction accuracy reached 0.9167. Based on accuracy, precision, recall, and $F1$

基金项目:天津市技术创新引导专项基金(23YDTPJC00010);国家重点研发计划项目(2022YFB4301002)

作者简介:吴煜,中国民航大学安全科学与工程学院硕士生导师,讲师,博士,主要从事民航安全与应急管理、危险品运输等方面的研究,天津市东丽区津北公路2898号中国民航大学北23教学楼518,300300, y_wu@cauc.edu.cn。

收稿日期:2025-04-29; **修回日期:**2025-06-23

score settings, a multi model comparative evaluation shows that the proposed aggregated K nearest neighbor model is significantly superior to other models. Therefore, the proposed cabin fire case retrieval model can provide support for predicting the level of civil aviation cabin fires and emergency decision-making.

Key words: cabin fires; knowledge meta model; aggregation K nearest neighbor algorithm; case retrieval; event level

民航客舱火灾致因因素复杂,因空间密闭具有烟雾扩散迅速、应急处置困难、应急疏散受限等特点,严重威胁航空运行安全^[1]。据航空安全与危险品安全办公室(Office of Aviation Security and Hazardous Materials Safety,ASH)统计^[2],2006年至2024年7月全球发生锂电池火灾事件520起,客舱锂电池火灾事件402起,其中飞行中的客舱火灾事件238起,有68起造成航空器损失和人员伤亡。如何防控和快速应对客舱火灾事件,提升现场应急处置能力,避免事件失控升级,始终是保障民航空运安全的核心任务。挖掘客舱火灾事件案例属性并快速检索可对实时火灾事件提供等级判断和动态应急响应对策,为预案制定、快速处置、空勤人员训练等提供可行的技术路径。

目前,客舱火灾已经形成了事件防控、应急处置和演练培训等事前、事中和事后的多阶段研究。事前阶段集中在检测预警、火灾机制等防控技术实现,如Shandhoosh V等^[3]采用机器学习故障诊断方法实现设备故障的早期识别和状态监测;Liu T Q等^[4]关注不同舱位烟雾扩散时间和温度的差异造成的火灾影响;燃料和通风同样影响舱内火灾动态变化,故可通过气体浓度获取事件预警信息^[5-6]。事中阶段则强调应急响应、处置疏散的处置方法实现,如Li X等^[7]分析了多尺度航空煤油池火灾情况下,如何合理规划应急救援路线和行动方案;杨永刚等^[8]提出基于客舱的修正社会力模型,提升了疏散效率。事后阶段则关注事故调查、演练培训等的方案实现,如Wang J等^[9]采用贝叶斯网络整合多源数据系统总结了飞机火灾调查流程;吕伟等^[10]提出模拟仿真不同状态下的客舱火灾动态发展态势,演练乘客应急疏散问题。其中,案例推理方法可以提供一种基于机器学习的相似事件推理^[11],在事中阶段采用有助于为现场应急处置提供实时对策。针对案例推理方法,现有研究集中在案例知识表征和案例推理过程两个方面。一方面,案例知识表征可以通过槽值框架表示法、情景模式、图关系结构开展案例属性特征研究。如张青松等^[12]利用框架表示法拆解客舱突发事件的结构特征;夏登友等^[13]基于“情景-应对”模式结合知识元理论构建危险化学品储罐区火灾案例的事故情景三角模型;Das R等^[14]动态提取问题相关子图,从图结构相似性与潜在推理模式进行案例表征。另一方面,案例检索作为案例推理的核心实现算法,多采用机器学习开展评估和

预测。如郭震等^[15]利用 K 近邻算法(K Nearest Neighbors,KNN)进行参数学习实现建筑火灾的评估预测;郗婕等^[16]利用梯度提升决策树(Gradient Boosting Decision Tree,GBDT)模型预测森林火灾风险;闵超等^[17]融合神经网络与决策树(Decision Tree,DT)构建油气火灾爆炸可解释预测模型;邓力等^[18]基于单模型给出集成分类模型预测低压飞机货舱环境下烟雾探测分类能力。Gigović L等^[19]构建基于支持向量机(Support Vector Machine,SVM)与随机森林(Random Forest,RF)的集成模型,预测森林火灾敏感性的空间精确度。上述研究主要集中在阶段性的方法实现,较少从系统安全出发,尚未从火灾事件推理角度形成“风险防控→应急处置→调查改进”的动态闭环研究。此外,在客舱安全领域缺少可行的火灾事件案例推理和有效的检索方法研究,空勤人员难以及时判断实时火灾等级以开展动态响应行动。

鉴于此,笔者以客舱火灾相似案例检索为研究对象,系统考虑事件发生、发展与处置的全过程,采用知识元模型开展案例知识表征,重构客舱火灾事件案例库,选择改进的聚合 K 近邻算法(Aggregation K Nearest Neighbor Algorithm,AKNN)计算案例之间的相似度,以实现实时案例的快速检索与准确预测,快速匹配当前火灾情境与历史案例,辅助事件预测和应急决策,以解决航空公司培训人员在客舱火灾应急响应时案例检索效率低、难以在短时间内实现有效决策的问题。

1 客舱火灾事件的案例表征

1.1 客舱火灾事件

客舱火灾是典型的客舱突发事件类型,本文界定的客舱火灾事件是指发生于民航客舱密闭空间内,由锂电池或其他因素引发明火、过热或烟雾现象,可能造成人员伤亡或财产损失的事故征候或事故。

参考CCAR-396《民用航空安全信息管理规定》及其配套的《事件信息填报要素及示例》等文件对事件信息格式的规定,并分析现有客舱火灾的具体事故报告,归纳得出客舱火灾突发事件案例信息具有的典型知识特征,包括事发时间、地点和民用航空器运营人;民用航空器类别、型别、国籍和登记标志;机上和地面伤亡人数,航空器损伤情况等。

事件的等级划分有助于针对性地制定专项预案、

确保安全处置。从人员伤亡和经济损失角度,遵循公安部第121号令《火灾事故调查规定》,可将民航火灾划分为特别重大火灾、重大火灾、较大火灾和一般火灾。从火灾场景角度,依据Doc 9756《航空器事故和事故征候调查手册》,火灾可划分为运行中火灾、地面火灾和坠机后火灾。从综合场景和后果角度,AC-395-AS-01《民用航空器征候等级划分办法》中将火灾事件分为运输航空严重征候、运输航空一般征候、运输航空地面征候和通用航空征候。考虑以上划分特点,为便于研究,民航客舱火灾事件可划分为Ⅰ、Ⅱ、Ⅲ3个等级。其中,Ⅰ级事件指无人伤亡,航空器未受损,不影响飞行;Ⅱ级事件指存在少量人员伤亡,航空器轻度受损,影响飞行;Ⅲ级事件指人员伤亡严重,航空器受损严重,不能持续飞行。

1.2 基于知识元模型的客舱火灾案例表征

1.2.1 民航客舱火灾案例的知识元模型

知识元作为知识结构的最小组成单元^[20],有助于从更微观的角度表征案例属性,具有结构化、可检索性,可支撑客舱火灾多特征分析与等级预测。由此,对客舱火灾案例进行分析和统一描述,其属性特征可表达为 $K = \{K_m, K_a\} = \{(N_m, A_m), (p_a, d_a)\}$, $\forall m \in M, \forall a \in A, \forall A \in N \in N_m$ 。其中, K 指单个客舱火灾案例的知识元集合,用状态集 K_m 和数值集 K_a 描述。 K_m 指对应知识元的概念名称集、属性的状态集。 K_a 指对应属性的取值范围、赋值的数值集。 N_m 为火灾案例对应的知识元名称集 N 和属性特征的名称集 A 。 A_m 为火灾的定量或定性属性集合,包括所有特征的属性状态集 c_a 和可统计的火灾属性特征的单位 u_a ,其中, c_a 中元素包括可描述的、可统计的; u_a 中可统计的特征都通过最小单位计算,不可进行数量统计的特征用文字描述,如火灾性状可用轻微烟、明火、烟雾弥散描述。 p_a 为火灾属性状态集,如火灾性状有3种表现:轻微烟、明火、烟雾弥散,状态归属依据案例中火灾真实情景。 d_a 为属性值集合,取值与属性状态集中特征的不同状态一一对应。 M 为所有火灾案例的集合。

从客舱火灾的防控、处置、调查的闭环改进角度,考虑客舱火灾的灾害形成、发展及应对^[21]的全过程,民航客舱火灾案例可从致灾因子 $\bar{K}$ 、承灾体 $\hat{K}$ 、孕灾环境 $\check{K}$ 、应急处置 $\tilde{K}$ 四维知识元框架进行详细表示,即 $K = (\bar{K}, \hat{K}, \check{K}, \tilde{K})$ 。

例如,致灾因子知识元可以是火灾子知识元,由火灾性状、扑灭时间和毒害性3个属性集知识元构成,故致灾因子知识元 $\bar{K}$ 可以表示为 $\bar{K} = \{[(N_{\bar{K}}, A_{\bar{K}}), (c_{a\bar{K}}, u_{a\bar{K}})], [p_{a\bar{K}}, d_{a\bar{K}}]\}$,其中, $\bar{K}$ 致灾因子知识元所包含的属性特征可以表示为 $A_{\bar{K}} = (\bar{K}_1) =$

$(\bar{K}_1^1, \bar{K}_1^2, \bar{K}_1^3)$;其他类型知识元表示以此类推。

1.2.2 民航客舱火灾案例的知识元表示

以知识元为单元,构建民航客舱火灾事故案例知识元模型,从中提取案例特征存储案例信息。每个案例包含各知识元实例化信息,为每一属性特征取值范围分配标签编码,统计分析历史事件案例,按照性状、时间、毒害性等不同角度将属性集特征描述划分为1到5的不同取值,得到属性状态数值化表示,如表1所示。

表1 客舱火灾案例知识元表示

Table 1 Representation of knowledge meta for cabin fire case

知识元类型	$K = \{K_m, K_a\}$						
	K_m				K_a		
	N_m		A_m		p_a	d_a	ω_j
	N	A	c_a	u_a			
$\bar{K}$	火灾 $\bar{K}_1$	火灾性状 $\bar{K}_1^1$	可描述的	N/A	轻微烟	1	0.109 2
					明火	2	
					烟雾弥散	3	
		扑灭时间 $\bar{K}_1^2$	可统计的	min	(0,10]	1	0.170 1
					(10,20]	2	
					(20,-)	3	
		毒害性 $\bar{K}_1^3$	可统计的	10^{-6}	[0,50)	1	0.140 7
					[50,200)	2	
					[200,-)	3	
$\hat{K}$	乘员行李 $\hat{K}_1$	事故锂电池 $\hat{K}_1^1$	可统计的	件	[0,1)	1	0.049 8
					[1,-)	2	
		违规设备 $\hat{K}_1^2$	可统计的	件	[0,1)	1	0.051 5
					[1,-)	2	
	机载设备 $\hat{K}_2$	设备位置 $\hat{K}_2^1$	可描述的	N/A	内部	1	0.056 8
					外部	2	
$\check{K}$	事故位置 $\check{K}_1$	事故位置 $\check{K}_1^1$	可描述的	N/A	行李架	1	0.069 9
					座椅	2	
					厨房	3	
					卫生间	4	
					其他区域	5	
	航班信息 $\check{K}_2$	机型 $\check{K}_2^1$	可描述的	N/A	小	1	0.081 3
					中	2	
					大	3	
$\tilde{K}$	航线 $\check{K}_2^2$	航线 $\check{K}_2^2$	可描述的	N/A	国内	1	0.056 7
					国际	2	
	人员伤亡 $\check{K}_1$	人员伤亡 $\check{K}_1^1$	可统计的	人	(0,5]	1	0.078 8
					(5,10]	2	
					(10,-)	3	
	处置时间 $\check{K}_2$	所用时间 $\check{K}_2^1$	可统计的	min	(0,5]	1	0.064 6
					(5,10]	2	
					(10,-)	3	
	经济损失 $\check{K}_3$	调查报告 $\check{K}_3^1$	可描述的	N/A	无	1	0.070 7
					有	2	

2 客舱火灾案例检索方法

2.1 方法选择和构建

民航客舱火灾事故案例采用知识元模型对火灾事故案例作统一描述,分为向量转换、数据预处理、模型验证、模型评估、模型应用,以实现客舱火灾事故案例从提取到输出的全过程,为民航客舱火灾的预测及应对提供参考。本文构建的四维框架知识元模型解构客舱火灾案例,数字化表征客舱火灾的灾害形成、发展及应对的特征,基于数字向量计算案例相似度并预测事件等级。常见的案例检索方法具有以下缺点:关键词检索仅适用于简单结构化查询,缺失语义关联;TF-IDF多用于文本相似性匹配场景^[22],难以预测事件等级;深度学习检索需海量数据,小样本下过拟合。而聚合 K 近邻算法允许为致灾因子、承灾体、孕灾环境、应急处置四个维度分别设置距离权重,可以聚合多个 K 近邻模型的预测结果,利用模型间的多样性降低单一模型的偏差与方差,有利于实现客舱火灾事故文本案例到数字向量的转化、目标案例与历史案例的相似度检索、目标案例事件级别的准确预测以及对应级别火灾事件的应急决策快速应用。

由此设计检索方法,其实现步骤见图1。

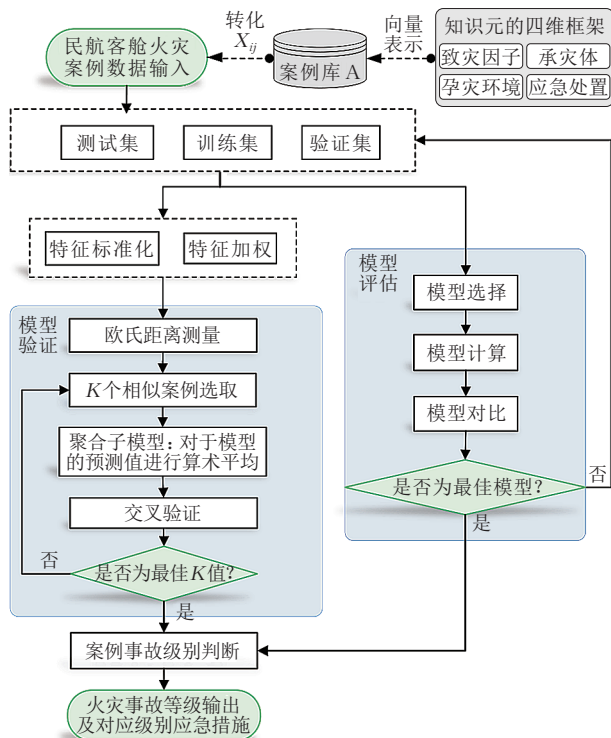


图1 基于知识元模型的民航客舱火灾案例检索步骤

Fig. 1 Procedure of civil aviation cabin fire case retrieval based on knowledge meta

2.2 方法实现

2.2.1 事件案例属性特征的转化规则

将文本案例中的属性名称提取为特征向量,如式

(1)所示。

$$A = (A_{\bar{K}}, A_{\dot{K}}, A_{\hat{K}}, A_{\check{K}}) = (\bar{K}_1, \dot{K}_1, \hat{K}_1, \check{K}_1, \bar{K}_2, \dot{K}_2, \hat{K}_2, \check{K}_2, \bar{K}_3, \dot{K}_3, \hat{K}_3, \check{K}_3) \quad (1)$$

2.2.2 特征向量提取

将第 i 个案例的属性特征向量 A 数值化,转换为 x_{ij} ,即 $x_{ij} = (x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}, x_{i6}, x_{i7}, x_{i8}, x_{i9}, x_{i10}, x_{i11}, x_{i12})$,其中, i 表示是第 i 个客舱的火灾案例; j 表示是第 i 个案例的第 j 个属性特征, $j = 1, 2, \dots, 12$ 。

2.2.3 特征标准化和权重设计

原始数据集中,不同特征取值范围相差过大,可能导致预测模型更加偏向于取值更大的特征^[23],需要对样本特征进行标准化预处理,如式(2)、式(3)所示。

$$\mu = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad (2)$$

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ij} - \mu)^2} \quad (3)$$

式中: μ 为每个属性特征的均值; σ 为每个属性特征的标准偏差; n 为案例的样本数量。

然后,运用Z-score标准化公式计算得到标准化特征,如式(4)所示。

$$Q_{ij} = \frac{x_{ij} - \mu}{\sigma} \quad (4)$$

信息增益法(Information Gain, IG)可以根据属性特征为案例所增加的信息量来衡量属性的重要性^[24],权重分配基于数据驱动,量化不同属性特征,减少预测不确定性的能力,如扑灭时间属性特征的权重最高,这是源于其与灾损能量的直接物理关联及显著统计判别性。计算特征信息熵 $H(D)$ 与特征条件信息熵 $H(D|x_j)$ 的差值,得到每个特征的信息增益 $IG(x_j)$,如式(5)所示。信息增益越大,该特征对火灾等级预测的贡献越显著。

$$IG(D, x_j) = H(D) - H(D|x_j) = -\sum_{\tau=1}^C \frac{n_{\tau}}{N} \log_2 \frac{n_{\tau}}{N} - \sum_{v=1}^V \frac{n_v}{N} \left(-\sum_{\tau=1}^C \frac{n_{\tau v}}{n_v} \log_2 \frac{n_{\tau v}}{n_v} \right) \quad (5)$$

设数据集 D 有 N 个样本, C 个级别($C=3$),统计属性 x_j 的不同取值,将 D 划分为 V 个子集 D_v ($v=1, 2, \dots, V$)。其中, $H(D)$ 表示数据集 D 的信息熵, $H(D|x_j)$ 表示 x_j 的条件信息熵; n_{τ} 表示 τ 级客舱火灾案例的样本数, n_v 表示子集 D_v 的样本总数, $n_{\tau v}$ 表示子集 D_v 中属于第 τ 级案例的样本数。为了确保所有特征权重之和为1,对每个属性特征的信息增益进行归一化,得到特征权重,如式(6)所示。

$$\omega_j = \frac{IG(D, x_j)}{\sum_{j=1}^{12} IG(D, x_j)} \quad (6)$$

计算知识元框架下的案例属性特征权重 ω_j , 结果见表1。

2.2.4 模型验证

生成仿真案例:受可被统计到的民航客舱火灾真实事件案例数量限制,在测试过程中,基于解构的客舱火灾历史事件案例属性特征,结合属性特征取值范围并考虑部分特征间的依赖关系,采用贝叶斯网络模型在严格逻辑约束下合成新案例。通过验证生成数据与假设分布的拟合度,输出 $N=1500$ 条仿真案例,支持参数调整优化真实性。

划分数据集:将模拟案例用于模型训练,抽取 $n_1=150$ 条进行验证与调优;其次,选择 $n_2=24$ 条具备完整事件信息的真实案例,其中 $n_{\text{test}}=23$ 条作为测试集, $n_{\text{target}}=1$ 条作为预测目标案例。

计算相似度:采用欧式距离代替历史案例 x_{ij} 与目标案例 x_{0j} 之间的相似度,如式(7)所示。

$$d(x_{0j}, x_{ij}) = \sqrt{\sum_{i=1}^n \omega_j (x_{0j} - x_{ij})^2} \quad (7)$$

聚合子模型:设置子模型数量 H 值固定为50,采用样本子集抽样,每个子模型使用原始数据90%的随机有放回样本进行训练,所有子模型共享相同的12维加权特征空间,把 H 个子模型的预测值进行算术平均,将聚合多个基于不同数据子集的KNN预测结果作为最终的预测值,如式(8)所示。

$$\hat{y} = \frac{1}{H} \sum_{s=1}^H \hat{y}_s \quad (8)$$

式中: $\hat{y}$ 为最终的预测值; $\hat{y}_s$ 为第 s 个子模型的预测值, $s=1, 2, \dots, H$ 。

留一法交叉验证:对于测试集的 $n_{\text{test}}=23$ 条样本,依次将每条样本作为验证集,其余22条样本作为训练集,训练模型并进行预测^[25],计算预测结果与真实标签之间的准确率,如式(9)、式(10)所示。

$$\text{Acc}_{\text{LOO}} = \frac{1}{n_{\text{test}}} \sum_{t=1}^{n_{\text{test}}} \text{Acc}_t \quad (9)$$

$$\text{Acc}_t = \begin{cases} 1, & y_t = \hat{y}_t \\ 0, & y_t \neq \hat{y}_t \end{cases}, t=1, 2, \dots, n_{\text{test}} \quad (10)$$

式中: Acc_{LOO} 为最终准确率,表示所有 n_{test} 次验证准确率的平均值; Acc_t 为第 t 次验证的准确率; y_t 为第 t 次留一法验证中样本的真实标签; $\hat{y}_t$ 为第 t 次留一法验证中样本的预测标签。

2.2.5 模型评估

为验证模型检索效率,分别选用KNN、SVM、DT、RF、GBT与构建的AKNN模型进行对比验证。所有模型使用相同的训练集、验证集和测试集,在相同条件下计算准确率、精确率等性能指标,对比案例检索效率与准确率,评估AKNN模型的表现。

2.2.6 模型应用

设 $\hat{y}$ 为模型预测结果,对应事件等级, E 为执行该等级下的应急决策,可构建描述性公式如式(11)所示。

$$E = f(\hat{y}) = \begin{cases} E_I, & \hat{y} = I \\ E_{II}, & \hat{y} = II \\ E_{III}, & \hat{y} = III \end{cases} \quad (11)$$

式中: $\hat{y}$ 为模型对目标案例的预测结果,可能取值为I、II、III; E 为根据预测结果执行对应级别的应急决策,借鉴 k 个相似案例中最适应当下情境中的历史案例解; E_I 为I级事件对应的应急决策; E_{II} 为II级事件对应的应急决策; E_{III} 为III级事件对应的应急决策。

3 模型评估和仿真实验

3.1 数据收集与预处理

选取具有完整事件信息报告的民航客舱火灾经典案例作为民航客舱火灾案例检索的测试集,依据前文中的规则,将案例属性信息拆解,如表2所示。

3.2 模型验证

3.2.1 相似案例检索

为便于计算,随机选取案例20为预测目标案例:其向量化表示为 $x_{0j}=(2, 2, 2, 2, 1, 1, 2, 2, 2, 1, 2, 1)$,用于最终预测。目标案例与全量案例欧式距离结果见图2。

图2中,欧氏距离最近的前三名案例均为II级火灾事件,加权距离分别为0.3599、0.4539、0.5656;4、5名案例为III级火灾事件,欧氏距离分别为0.5693、0.6101;之后的案例与目标案例之间的距离逐渐增加。选取 K 值为3,3个最近邻案例等级都为II,故目标案例为II级火灾事件。另选取 K 值为8,8个最近邻案例中II级事件和III级事件出现的次数相同,目标案例事件等级无法预测。由于不同 K 值下目标案例的等级预测可能出现不同结果,所以需要验证3是否为最佳 K 值。

3.2.2 LOO交叉验证

每次取1条真实数据作为测试样本,其余22条和1500个模拟数据作为训练集。循环23次,计算平均指标,不同 K 值下的 $F1$ 分数见表3。

从表3中可以看出,准确率在 $K=1$ 时为0.7500,随 K 值增大,准确率先增后减,在 $K=3$ 与 $K=4$ 时达到峰值0.9167,之后降到一个稳定的数值,故最佳 K 值为3,预测目标案例事件等级为II级,影响航行,应采取相应的应急措施。经过交叉验证,3是AKNN的最佳 K 值。

表2 民航客舱火灾案例属性信息

Table 2 Attribute information of civil aviation cabin fire cases

案例序号	案例基本信息	等级判断	案例属性特征 x_{ij} 数值化												
	时间-航班-致因-后果		x_{i1}	x_{i2}	x_{i3}	x_{i4}	x_{i5}	x_{i6}	x_{i7}	x_{i8}	x_{i9}	x_{i10}	x_{i11}	x_{i12}	y
1	2012-南航 CZ6199-空气循环机水雾误判-无人员伤亡	I	1	1	2	1	1	1	5	3	1	2	2	2	1
2	2013-南航 CZ6917-充电宝自燃-无人员伤亡,继续航行	I	1	1	1	2	1	1	1	2	1	1	1	1	1
.....															
20	2024-西南航 3316-手机自燃-1人轻微不适,更换飞机和机组	II	2	2	2	2	1	1	2	2	2	1	2	1	2
21	2024-海航 HU438-鸟击起火-无人员伤亡,航班返航	II	2	3	2	1	1	2	5	2	2	1	1	2	2
22	2025-国泰航 CX811-设备故障-无人员伤亡,航班返航	III	3	3	3	1	1	1	5	2	2	1	2	2	3
23	2025-港航 HX115-充电宝自燃-无人员伤亡,航班迫降	II	2	2	2	2	1	1	1	2	1	1	2	1	2
24	2025-美联航 UA1513-烤箱起火-无人员伤亡,更换飞机和机组	I	1	2	2	1	1	1	3	2	2	1	3	2	1

表4 各算法模型下的测试结果表

Table 4 Test results table of various algorithm models

模型	准确率	精确率	召回率	F1 分数
KNN	0.875 0	0.866 7	0.891 7	0.876 8
SVM	0.708 3	0.688 9	0.718 1	0.671 6
RF	0.750 0	0.818 2	0.784 7	0.746 6
GBT	0.875 0	0.875 0	0.847 2	0.834 2
DT	0.833 3	0.755 6	0.870 8	0.791 3
AKNN	0.916 7	0.904 8	0.958 3	0.922 2

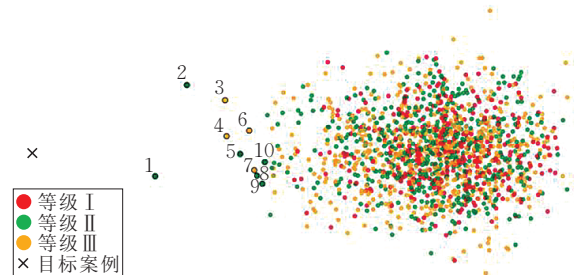


图2 目标案例与全量案例欧式距离结果

Fig.2 Euclidean distance between the target case and all cases

表3 LOO交叉验证划分数数据集准确率

Table 3 Accuracy of the data set divided by LOO validation

K 值	1	2	3	4	5
准确率	0.750 0	0.791 7	0.916 7	0.916 7	0.875 0
K 值	6	7	8	9	10
准确率	0.875 0	0.875 0	0.875 0	0.875 0	0.875 0

3.3 模型评估

3.3.1 模型测试结果

依据2.2节中构建的模型,将划分的训练集、验证集和测试集带入不同模型进行验算,最终得到各预测模型的准确率、精确率、召回率、F1分数计算结果,见表4。

AKNN模型在分类任务中表现最优,F1值为0.922 2,显著高于其他模型。与SVM和GBT相比,AKNN在保持较高准确率0.916 7的同时,平衡了精确率0.904 8和召回率0.958 3,尤其在处理不平衡数据时优势明显。其优越性源于聚合对噪声的鲁棒性和参数优化后的模型泛化能力。通过Bagging集成策略,AKNN有效降低了KNN的方差,提升了模型稳定性,适合用来预测客舱火灾事件等级,检索相似案例。

AKNN模型的各项得分远高于其他模型,准确率、召回率等均为最高,对火灾事件等级的捕捉能力最强。KNN模型的准确率、精确率低于AKNN,整体识别能力不足;DT、GBT模型在多指标上也不及AKNN;RF、SVM的各项指标普遍偏低,稳定性差。在火灾事件预测中,AKNN模型的高准确率确保了整体的预测可靠性,高召回率使其更擅长挖掘真实火灾事件案例,减少漏判,更适合对事件等级需精准识别的场景,在综合性能上比其他模型更具优势。

3.3.2 混淆矩阵分析

在决策边界处的识别局限性、过度关注高权重特征、样本不足导致边界学习不充分等原因,导致每个模型都存在真实标签预测有误的情况。从测试集混淆矩阵看,AKNN在事件等级预测中表现更优:等级I的预测正确14例,误判2例,处于最优水平;等级II的预测全部正确;等级III的预测全部正确,优于其他模型,如图3所示。

整体而言,AKNN模型在预测目标民航客舱火灾事件等级时存在如下优点:对不同事件等级的预测更均衡,尤其在等级III上优势明显,而其他模型在部分等级上的预测存在较多误判;等级预测准确性、全

面性表现最优,准确率、精确率、召回率的综合分数也最高。因此,相比其他模型,AKNN模型既能减少误

判,又能提升真实事件的识别率,更适配火灾事件精准预测的需求。

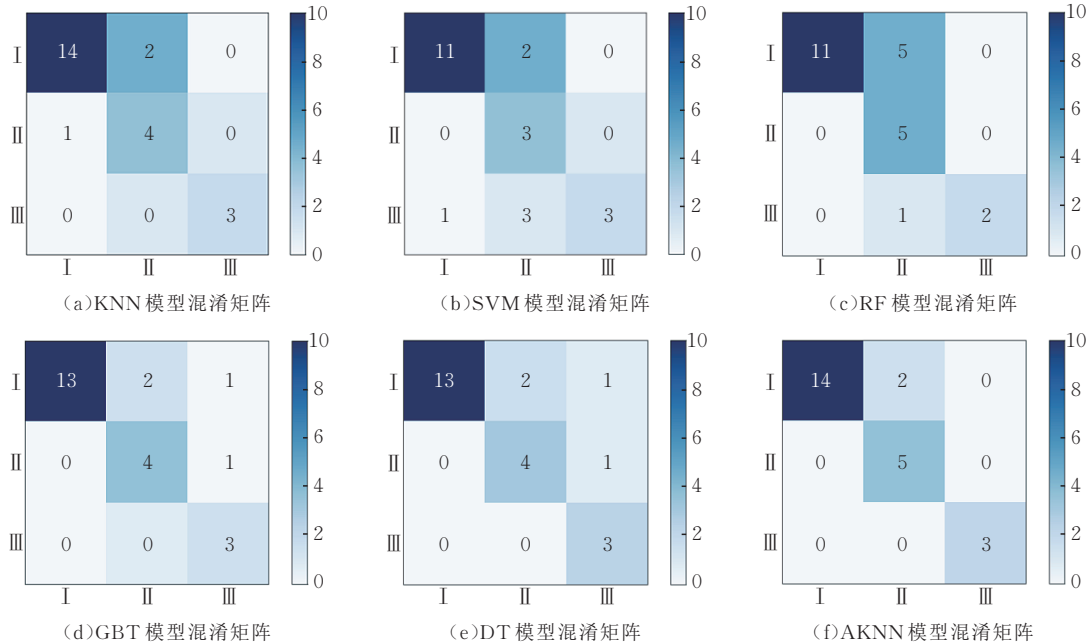


图3 各机器学习模型混淆矩阵图

Fig. 3 Confusion matrix diagrams of various machine learning models

4 结论

采用知识元模型开展案例知识表征,重构客舱火灾事件案例库,选择改进的聚合 K 近邻算法提出了客舱火灾案例检索模型,可实现实时案例的快速检索与准确预测,为民航客舱火灾等级预测及应急决策提供支撑。

1)综合考虑了客舱火灾突发事件的属性特征,构建了基于四维框架的致灾因子、承灾体、孕灾环境、应急处置的知识元表征,形成了对火灾事件的全过程案例描述,便于形成案例反推的闭环研究。

2)提出了一种面向客舱火灾事件的聚合 K 近邻算法检索方法。测试表明,通过该模型检索目标案例与相似历史案例,预测目标案例事件等级为II,并经过LOO交叉验证得出最佳 K 值为3,预测准确率为0.8750提高到0.9167,效果显著。

3)重点研究了通过文本形式获取事件案例,未来可开展对视频、语音等多模态事件的数据采集,构建多模态数据融合框架;并可根据事件等级和案例情境开展智能实时决策,提供动态应急方案。

参考文献:

- [1] 吴煜, 杨萌, 张青松. 基于云模型的客舱火灾应急处置协同能力测评方法[J]. 中国安全科学学报, 2023, 33(12): 206-213.
Wu Y, Yang M, Zhang Q S. Evaluation method of collaborative capacity of cabin fire emergency response based on cloud model

- [J]. China Safety Science Journal, 2023, 33(12): 206-213.
- [2] Lithium battery incidents[DB/OL]. [2025-3-16]https://www.faa.gov/hazmat/resources/lithium_batteries/incidents.
- [3] Shandhoosh V, Venkatesh S N, Chakrapani G, et al. Intelligent fault diagnosis for tribo-mechanical systems by machine learning: Multi-feature extraction and ensemble voting methods[J]. Knowledge-Based Systems, 2024, 305: 112694.
- [4] Liu T Q, Liu K N, Jia R H. Smoke and flame temperature propagation characteristics during cabin fire in Airbus A380[J]. Thermal Science, 2025, 29(1 Part A): 215-226.
- [5] Li C, Xu W B, Jin Y K, et al. Numerical investigations on combustion characteristics of full-size cabin fire with typical material pyrolysis and burnout under different fire source positions[J]. Case Studies in Thermal Engineering, 2024, 54: 104075.
- [6] Shi S, Lyu N W, Jiang X, et al. Hydrogen gas diffusion behavior and detector installation optimization of lithium ion battery energy-storage cabin[J]. Journal of Energy Storage, 2023, 67: 107510.
- [7] Li X, Zhang M X, Zhong Y, et al. Upwind thermal radiation prediction of multi-scale aviation kerosene pool fire on the flight deck for emergency rescue: Experiment, numerical simulation and machine learning[J]. International Journal of Thermal Sciences, 2023, 190: 108308.
- [8] 杨永刚, 王璟璟, 宋炜. 飞机客舱应急疏散的修正社会力模型[J]. 中国安全科学学报, 2022, 32(10): 201-206.
Yang Y G, Wang J J, Song W. Modified social force model for aircraft cabin emergency evacuation[J]. China Safety Science Journal, 2022, 32(10): 201-206.

- [9] Wang J, Tao Z X, Yang R, et al. A review of aircraft fire accident investigation techniques: Research, process, and cases[J]. *Engineering Failure Analysis*, 2023, 153: 107558.
- [10] 吕伟, 汪京辉, 李加伟, 等. 火灾场景下宽体客机人员疏散风险分析[J]. *中国安全科学学报*, 2023, 33(3): 141-146.
Lyu W, Wang J H, Li J W, et al. Analysis of evacuation risk of wide-body aircraft in fire scene[J]. *China Safety Science Journal*, 2023, 33(3): 141-146.
- [11] Chehade M H, Sylla A. Coupling case-based reasoning (CBR) and machine learning for manufacturing time estimation[J]. *IFAC-PapersOnLine*, 2024, 58(19): 941-945.
- [12] 张青松, 魏祥宇, 吴煜, 等. 基于CBR的客舱突发事件应急决策模型研究[J]. *消防科学与技术*, 2023, 42(6): 855-859.
Zhang Q S, Wei X Y, Wu Y, et al. Research on emergency decision method of cabin emergency based on CBR[J]. *Fire Science and Technology*, 2023, 42(6): 855-859.
- [13] 夏登友, 胡人元, 朱毅, 等. 基于知识三元组的危化品储罐区火灾情景构建[J]. *安全与环境学报*, 2020, 20(3): 938-945.
Xia D Y, Hu R Y, Zhu Y, et al. On the fire scenario construction of the dangerous chemical tank area based on the knowledge triad[J]. *Journal of Safety and Environment*, 2020, 20(3): 938-945.
- [14] Das R, Godbole A, Naik A, et al. Knowledge base question answering by case-based reasoning over subgraphs[C]//International conference on machine learning. PMLR, 2022: 4777-4793.
- [15] 郭震, 贾笑岩, 李富民, 等. 基于机器学习的建筑火灾蔓延快速预测[J]. *中国安全科学学报*, 2023, 33(11): 117-125.
Guo Z, Jia X Y, Li F M, et al. Fast prediction for building fire spread based on machine learning[J]. *China Safety Science Journal*, 2023, 33(11): 117-125.
- [16] 郝健, 傅微. 基于机器学习的流域尺度森林火灾灾害风险预测[J]. *自然灾害学报*, 2024, 33(1): 89-98.
Xi J, Fu W. Watershed-scale forest fire risk prediction based on machine learning[J]. *Journal of Natural Disasters*, 2024, 33(1): 89-98.
- [17] 闵超, 张乾, 黄鑫, 等. 油气场站火灾爆炸风险的神经支持决策树识别与预测[J]. *安全与环境学报*, 2024, 24(7): 2574-2584.
Min C, Zhang Q, Huang X, et al. Risk recognition and prediction of fire and explosion in oil and gas stations based on neural-backed decision trees[J]. *Journal of Safety and Environment*, 2024, 24(7): 2574-2584.
- [18] 邓力, 吴丹丹, 周进, 等. 低压货舱多参数火灾探测集成模型的优化选择[J]. *安全与环境学报*, 2024, 24(5): 1840-1849.
Deng L, Wu D D, Zhou J, et al. Low pressure environments multi-sensor fire detection algorithm ensemble model of quadratic optimization choice[J]. *Journal of Safety and Environment*, 2024, 24(5): 1840-1849.
- [19] Gigović L, Pourghasemi H R, Drobnjak S, et al. Testing a new ensemble model based on SVM and random forest in forest fire susceptibility assessment and its mapping in Serbia's Tara National Park[J]. *Forests*, 2019, 10(5): 408.
- [20] 赵瑞, 王宁, 张磊. 基于多案例融合的突发事件推演方法[J]. *系统工程*, 2025, 43(5): 18-29.
Zhao R, Wang N, Zhang L. The method of emergency deduction based on multi-case fusion[J]. *Systems Engineering*, 2025, 43(5): 18-29.
- [21] 吴煜, 杨萌, 张青松. 基于CBTA-CRM的客舱锂电池火灾VR演练平台设计及应用[J]. *消防科学与技术*, 2025, 44(1): 88-96, 102.
Wu Y, Yang M, Zhang Q S. Establishment and application of VR emergency drill platform of cabin lithium battery fire based on CBTA-CRM[J]. *Fire Science and Technology*, 2025, 44(1): 88-96, 102.
- [22] Park S, Oh S, Park W. Automated classification model for elementary mathematics diagnostic assessment data based on TF-IDF and XGBoost[J]. *Applied Sciences*, 2025, 15(7): 3764.
- [23] Tan X H, Cai Z P, Quan Z B, et al. Standardized extraction of air traffic control hazard features based on expert knowledge[J]. *Aerospace*, 2025, 12(2): 94.
- [24] 张涛, 翁康年, 张倩帆, 等. 基于情境案例推理的播前收视率预测方法[J]. *管理工程学报*, 2020, 34(6): 156-164.
Zhang T, Weng K N, Zhang Q F, et al. Audience rating prediction before broadcasting based on context case-based reasoning[J]. *Journal of Industrial Engineering and Engineering Management*, 2020, 34(6): 156-164.
- [25] Ali Üncü Y, Danışman T, Özdoğan H. Predicting $(n, 3n)$ nuclear reaction cross-sections using XGBoost and Leave-One-Out Cross-Validation[J]. *Applied Radiation and Isotopes*, 2025, 219: 111714.

(责任编辑:毛 星)