

中文引用格式:范博松,邵春福,王景升,等. 城市交通突发事件风险致因与后果严重程度判别模型[J]. 中国安全科学学报,2025,35(2):220-226.

英文引用格式:FAN Bosong, SHAO Chunfu, WANG Jingsheng, et al. Model on discriminating risk causes and consequence severity of urban traffic emergencies[J]. China Safety Science Journal,2025,35(2):220-226.

城市交通突发事件风险致因与后果 严重程度判别模型*

范博松¹讲师,邵春福²教授,王景升¹副教授,刘东¹副教授

(1 中国人民公安大学 交通管理学院,北京 100038;

2 新疆大学 交通运输工程学院,新疆 乌鲁木齐 830046)

中图分类号:X928.03

文献标志码:A

DOI: 10.16265/j.cnki.issn1003-3033.2025.02.0676

资助项目:中央高校基本科研业务费项目(2024JKF02ZK12);国家重点研发计划(2023YFB4302701)。

【摘要】 为提升城市交通突发事件后果严重程度判别的准确性,明确突发事件风险致因与后果严重程度的相关关系,构建改进的突发事件后果严重程度判别模型(IDM-ECS)并进行试验验证。首先,基于改进的特征选择算法(IFSA)筛选突发事件风险致因,得到列车兑现率、正点率、日路网客运量等重要风险致因;其次,采用改进的混合受限波尔兹曼机模型(HRBM)计算不同风险致因与后果严重程度的关系,通过比较概率值大小得到风险致因与后果严重程度的判别关系;最后,以轨道交通突发事件数据集作为试验样本进行验证,并从召回率、精确度、 F_1 值等方面与生成受限波尔兹曼机(GBM)、随机森林(RF)、深度森林(DF)、轻量梯度提升机(LightGBM)等4个模型进行对比。研究表明:列车兑现率、正点率、日路网客运量、5号线断面满载率、10号线断面满载率、信号故障以及车辆故障为7个最优风险致因。IDM-ECS模型平均的召回率为90.55%、精确度为91.89%、 F_1 值为91.06%,均优于对比模型。

【关键词】 城市交通; 突发事件; 风险致因; 后果严重程度; 判别模型; 改进的特征选择算法(IFSA)

Model on discriminating risk causes and consequence severity of urban traffic emergencies

FAN Bosong¹, SHAO Chunfu², WANG Jingsheng¹, LIU Dong¹

(1 School of Traffic Management, People's Public Security University of China, Beijing 100038, China;

2 School of Transportation Engineering, Xinjiang University, Urumqi Xinjiang 830046, China)

Abstract: In order to improve the accuracy of emergency consequence severity assessment, clarify the correlation between the risk causes and consequence severity in urban traffic emergencies, the improved discrimination model of emergency consequence severity (IDM-ECS) was constructed and experimentally verified. First, based on the IFSA, the risk causes of emergencies were screened to obtain the important risk causes such as train fulfillment rate, punctuality rate, and daily network passenger volume and so on.

Secondly, the improved hybrid restricted Boltzmann machine (HRBM) model was used to calculate the relationship between different risk causes and the consequence severity, and the discriminative relationship between risk causes and the consequence severity was obtained by comparing the probability values. Finally, the dataset of rail transit emergencies was used as an experimental sample for validation. The performance was compared with four models, including Generating Restricted Boltzmann Machines (GRBM), Random Forest (RF), Deep Forest (DF), and Light Gradient Boosting Machine (LightGBM), in terms of recall, precision, and F_1 value. The results show that train fulfillment rate, punctuality rate, daily network passenger volume, line 5 section full load rate, line 10 section full load rate, signal failure, and vehicle failure are the seven optimal risk causes. The IDM-ECS model has an average recall of 90.55%, precision of 91.89%, and F_1 value of 91.06%, all of which are better than those of the comparison models.

Keywords: urban transit; emergency; risk causes; consequence severity; discrimination model; improved feature selection algorithm (IFSA)

0 引言

伴随着我国城市交通的快速发展,其运营安全问题也日益凸显。研究城市交通突发事件风险致因与后果严重程度之间的关系,对于轨道交通行业的安全生产至关重要^[1]。

目前,针对城市交通突发事件风险致因和后果严重程度的研究,主要涉及贝叶斯理论^[2]、聚类方法^[3]、参数回归模型^[4]、网络拓扑模型^[5]等,这些模型侧重于定性的相关关系分析,对于风险致因与后果状态定量关系的探究相对薄弱。近年来,伴随着深度学习技术的快速发展,使得深入挖掘数据特征与属性之间的关系成为可能,深度森林^[6]、深度神经网络^[7]等模型都成功地被应用于事故故障后果的分类与预测^[8-10]。然而,深度学习模型的研究复杂性较高,且集中在特定类型故障状态的分析^[11]。为了在小样本条件下快速实现模型的参数学习及分类结果的输出^[12],采用受限波尔兹曼机(Restricted Boltzmann Machine, RBM)进行研究,并取得了较好的识别效果^[13-15]。

因此,笔者拟通过融合改进的特征选择算法(Improved Feature Selection Algorithm, IFSA)与混合RBM(Hybrid RBM, HRBM)模型,构建改进的突发事件后果严重程度判别(Improved Discrimination Model of Emergency Consequence Severity, IDM-ECS)模型;利用实际城市交通突发事件数据训练IDM-ECS模型,并与其他判别模型对比,验证IDM-ECS模型的有效性,以期能为城市轨道交通管理人员开展风险评估提供技术支持,提升判别研究的精确度。

1 突发事件后果严重程度判别模型

1.1 模型框架

以IFSA和HRBM作为IDM-ECS的主体,其中,HRBM由生成RBM(Generating RBM, GRBM)和判别RBM(Discriminant RBM, DRBM)组成。利用循环多时间窗扫描方法构建IFSA,通过计算各个风险致因与不同后果严重程度的信息增益比,筛选出信息增益比最大的风险致因作为最优特征向量量子集。将最优特征向量量子集输入GRBM中,在GRBM中通过参数训练强化输入的特征向量量子集与后果严重程度的关系,将训练结果输入DRBM,通过逐层计算并采用改进的随机失活(Dropout)规则得到突发事件风险致因对应的后果严重程度 y_j 。模型框架如图1所示。

1.2 特征选择模型

为更有效地体现城市交通突发事件样本的首尾特征以及不同风险致因组合的关联信息,采用多时间窗扫描的方法处理突发事件样本。设突发事件样本共有 d 个风险致因,记为特征向量 $\mathbf{f} = \{f_1, f_2, \dots, f_d\}$,采用 l 种长度的时间窗 t_i 扫描切分样本的风险致因特征向量 \mathbf{f} 。为避免因首尾数据仅被扫描一次而产生偏差,将样本数据首尾相接,形成多维特征圆环,在多维特征圆环中,任意时间窗 t_i 从首个特征维度开始扫描,至 d 次的扫描任务结束,得到 l 组 $t_i \times d$ 维的特征切片向量 $\mathbf{T}, \mathbf{T} = \{T_1, T_2, \dots, T_l\}$ 。基于上述循环多时间窗扫描后得到的多风险致因组合的特征切片向量,需要利用一种特征选择算法筛选出与突发事件后果严重程度关联度更高的风险致因特征向量量子集。

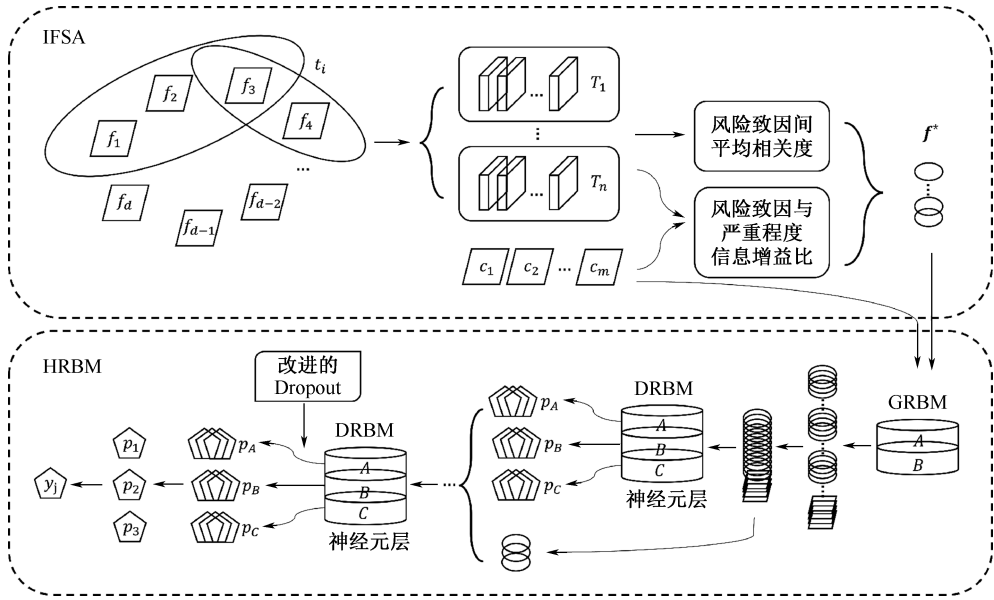


图1 IDM-ECS 模型框架

Fig.1 IDM-ECS modeling framework

在概率论和信息论中,常见的特征选择算法主要通过计算最大信息增益作为标准来筛选,信息增益表征了一个特征能够为整个系统带来多少信息,带来的信息越多,该特征越重要。系统有无某个特征,信息量将发生变化,而前后信息量的差值就是这个特征给系统带来的信息量。所谓信息量,就是熵;熵和条件熵的计算如下。

设 A 为突发事件样本 D 的一个风险致因,且取有限个值 $x_i(i=1, 2, \dots, n)$,其概率分布为:

$$P_r(A = x_i) = p_i, i = 1, 2, \dots, n \quad (1)$$

式中 p 为风险致因取不同值的概率。

A 的熵为:

$$E(A) = - \sum_{i=1}^n p_i \log_2 p_i \quad (2)$$

设 Y 为突发事件样本 D 的后果严重程度,且取有限个值 $y_j(j=1, 2, \dots, m)$,即后果严重程度分为 m 类,其概率分布为:

$$P_r(Y = y_j) = q_j, j = 1, 2, \dots, m \quad (3)$$

式中 q 为后果严重程度取不同值的概率。

Y 的熵为:

$$E(Y) = - \sum_{j=1}^m q_j \log_2 q_j \quad (4)$$

为分析不同风险致因与后果严重程度的关系,需要计算在 A 的条件下 Y 的条件熵,见下式:

$$E(Y|A) = \sum_{i=1}^n p_i E(Y|A = x_i) = - \sum_{i=1}^n P(x_i) \sum_{j=1}^m p(y_j | x_i) \log_2(p(y_j | x_i))$$

$$= - \sum_{i=1}^n \sum_{j=1}^m p(y_j, x_i) \log_2(p(y_j | x_i)) \quad (5)$$

Y 与 A 的信息增益 $G(Y, A)$ 为 Y 的熵 $E(Y)$ 与条件熵 $E(Y | A)$ 之差,表示在选择风险致因 A 的条件下使得 Y 的分类判别不确定性减少的程度,见下式:

$$G(Y, A) = E(Y) - E(Y|A) \quad (6)$$

信息增益越大,不确定性减少的越多,突发事件样本分类的不确定性就越低,表示风险致因 A 对后果严重程度 Y 的分类判别效果越好。不过信息增益越大,越可能导致算法一直选择某一个对信息增益贡献度最大的风险致因,因此,为校正信息增益非常大的极端情况,引入信息增益比(Gain Ratio, G_R),用信息增益与风险致因 A 的熵的比值表示,见下式:

$$G_R(Y, A) = \frac{G_R(Y, A)}{E(A)} \quad (7)$$

然而,信息增益比虽然能够避免去选择对信息增益贡献最大的风险致因,但也只考虑了风险致因与后果严重程度的相关性,并没有考虑不同风险致因之间的关联关系。事实上,在城市轨道交通突发事件中,不同风险致因相互作用,风险致因间的关联度也一定程度上影响了特征向量子集的选择。因此,提出一种改进传统信息增益比的计算方法,来平均风险致因间的关联度,以降低风险致因间关联度对特征选择的影响偏差。

设 B 为突发事件样本 D 中不同于 A 的一个风险致因,且取有限个值 $z_k(k=1, 2, \dots, s)$ 。 A 与 B

的相互作用程度用 2 个风险致因相互重叠的信息量 $I(A; B)$ 衡量, 见下式:

$$I(A; B) = \sum_{i=1}^n \sum_{k=1}^s p(x_i, z_k) \log_2 \frac{p(x_i, z_k)}{p(x_i)p(z_k)} \quad (8)$$

进一步采用对称不确定性^[16] 指标 $S_U(A, B)$ 来衡量风险致因之间的关联度, 见下式:

$$S_U(A, B) = 2 \times \frac{I(A; B)}{E(A) + E(B)} \quad (9)$$

其中, $S_U(A, B)$ 的取值范围为 $[0, 1]$ 。当 $S_U(A, B) = 0$ 时, 表示 A 和 B 为 2 个相互独立的风险致因, 当 $S_U(A, B) = 1$ 时, 表示 A 和 B 为 2 个完全相关的风险致因。则可得风险致因 A 与其他 $d-1$ 个风险致因的平均相关度, 见下式:

$$\overline{S_U(A)} = \frac{\sum_{f_u \in \mathcal{F}_{u \neq A}} S_U(A, f_u)}{d}, \quad u = 1, 2, \dots, d \quad (10)$$

综上, 在计算任意风险致因 A 与突发事件后果严重程度关联度时, 既要考虑风险致因 A 与后果严重程度 Y 有最大相关性, 又要考虑其与其他风险致因 f_u 有最小关联度, 最后得到调整后的信息增益比 G_R' , 见下式:

$$G_R'(Y, A) = \frac{G_R(Y, A)}{E(A) + \overline{S_U(A)}} \quad (11)$$

基于处理后的突发事件样本的风险致因切片向量 T , 计算出每个风险致因切片中所有风险致因的信息增益比大小, 挑选出每个切片中信息增益比最大的风险致因, 去重后得到 g 个最优风险致因组成最优特征向量集 $f^* = \{f_1^*, f_2^*, \dots, f_g^*\}$ 。

1.3 HRBM 模型

RBM 是一种无向生成模型, 拥有一层输入层和一层隐藏层, 层与层之间互相连接, 层内无连接。通过训练, 隐藏层可以学习到输入特征的概率分布。将 IFSA 得到的最优风险致因特征向量集与样本的后果严重程度一同输入 RBM, 就能使 RBM 通过参数学习得到风险致因和后果严重程度的联合分布, 即得到 GRBM。不过, 为精准分类判别输入样本, 将 GRBM 的联合概率分布替换为条件概率分布, 进而得到样本的后果严重程度, 即得到 DRBM。因此, 将 GRBM 和 DRBM 二者结合起来, 实现对样本数据进行参数学习和分类判别的 2 个重要目标, 进而构建起判别后果严重程度的 HRBM, 能够有效提高模型的代表学习能力和泛化能力。

1) GRBM。假设 GRBM 有 b 个隐节点, 将 IFSA 筛选出的最优风险致因特征向量集 f^* 和对应后果严重程度的突发事件样本作为模型输入, 得到三者的概率分布, 见下式:

$$P(x = x_\alpha, y = y_\beta, h = h_\gamma) = \frac{\exp(-V(x_\alpha, y_\beta, h_\gamma))}{\sum_{x, y, h} \exp(-V(x_\alpha, y_\beta, h_\gamma))}, \quad \alpha = 1, 2, \dots, g; \beta = 1, 2, \dots, m; \gamma = 1, 2, \dots, b \quad (12)$$

式中: $x = \{x_1, x_2, \dots, x_g\}$ 为输入的 g 个风险致因向量; $y = \{y_1, y_2, \dots, y_m\}$ 为输入的 m 个后果严重程度向量; $h = \{h_1, h_2, \dots, h_b\}$ 为 b 个隐藏层的节点; $V_\varphi(x_\alpha, y_\beta, h_\gamma)$ 为 GRBM 的能量公式, 如下:

$$V(x_\alpha, y_\beta, h_\gamma) = -x_\alpha \mathbf{W} h_\gamma^T - x_\alpha \boldsymbol{\eta}^T - \boldsymbol{\mu}^T h_\gamma - \boldsymbol{\theta}^T y_\beta - h_\gamma^T \mathbf{U} y_\beta \quad (13)$$

式中: \mathbf{W} 为输入层和隐藏层之间的权重系数矩阵; \mathbf{U} 为隐藏层和后果严重程度数据之间的权重系数矩阵; $\boldsymbol{\eta}, \boldsymbol{\mu}, \boldsymbol{\theta}$ 分别为输入层、隐藏层和后果严重程度数据的偏置系数向量。

用 Φ 指代 3 个偏置系数, $\Phi = \varphi(\boldsymbol{\eta}, \boldsymbol{\mu}, \boldsymbol{\theta})$ 。以最小化负对数似然函数作为目标函数, 见下式:

$$\min L_\varphi(x, y, h) = -\lg P_r(x, y, h) \quad (14)$$

针对特定的突发事件样本 $x_\alpha, y_\beta, h_\gamma$, 使用 mini-batch 梯度下降法最小化目标函数。

2) DRBM。因为 GRBM 与 DRBM 的区别在于风险致因数据与后果严重程度数据是并列关系还是条件关系, 所以基于 GRBM 参数学习的结果, 重新输入突发事件样本数据, 将风险致因作为输入, 基于条件概率判别样本数据的后果严重程度。依然以最小化负对数似然函数作为目标函数, 见下式:

$$\min L_\varphi'(x, y, h) = -\lg p(y, h | x) \quad (15)$$

同样, 使用 mini-batch 梯度下降法最小化目标函数, 见下式:

$$\frac{\partial \lg p(y_\beta | x_\alpha)}{\partial \varphi} = -E'_{h | x_\alpha, y_\beta} \left[\frac{\partial}{\partial \varphi} V(x_\alpha, y_\beta, h_\gamma) \right] + E'_{y, h | x_\alpha} \left[\frac{\partial}{\partial \varphi} V(x_\alpha, y, h) \right] \quad (16)$$

依然采用对比散度的方法计算梯度, 不过由于第 2 项计算的是后验概率, 因此, 不必考虑输入风险致因数据的重构, 通过调整输入层与隐藏层间的参数最小化求解误差 e , 见下式:

$$e = \frac{1}{2} (h^* - \sum_{\alpha=1}^d \delta_\alpha (\sum_{\gamma=1}^b h_\gamma \mathbf{W}^T + \boldsymbol{\eta}^T) x_\alpha)^2 \quad (17)$$

式中: h^* 为隐藏层目标值; δ_α 为Dropout规则,服从概率为 p 的伯努利分布,即 $\delta_\alpha \sim \text{Bernoulli}(p)$ 。

3) HRBM。HRBM的目标函数由式(14)和式(16)组成,见下式:

$$L_H(x, y, h) = \rho L_\varphi(x, y, h) + \tau L_\varphi'(x, y, h) \quad (18)$$

式中: ρ 和 τ 分别为GRBM与DRBM对整个模型的影响比例。当 $\rho > \tau$ 时,表示模型更偏重于GRBM,即模型在小样本数据的情况下能够更好地学习到风险致因与后果严重程度的关系;反之,当 $\rho < \tau$ 时,表示模型更偏重于DRBM,即模型在大样本数据的情况下能够更好地分类,避免风险致因与后果严重程度关系的过拟合。

为改进模型本身,而不是通过调节模型的使用率来避免过拟合,引入改进的Dropout规则。不同于Dropout规则中每个输入节点 x_α 以相同概率 p 被遮盖,改进的Dropout规则提出被遮盖的概率 p 应该是自适应的,其值取决于输入层和隐藏层之间的权重 w_α 的值,一般权重越大,被遮盖的概率越高。因此,每个节点被遮盖的概率服从一个新的伯努利分布 $\delta_\alpha' \sim \text{Bernoulli}(\zeta(w_\alpha x_\alpha))$,其中, ζ 为激活函数。那么模型误差见下式:

$$e' = \frac{1}{2} (h^* - \sum_{\alpha=1}^d \delta_\alpha' (\sum_{\gamma=1}^b h_\gamma \mathbf{W}^\top + \boldsymbol{\eta}^\top) x_\alpha)^2 \quad (19)$$

通过最小化模型误差,求解使混合模型目标函数 $L_H(x, y, h)$ 达到最优的参数,代入式(11)和式(12),得到不同风险致因所导致的各类后果严重程度的概率 $P_r(y | x_\alpha)$,取概率值最大的后果严重程度作为最终的判别结果 c ,见下式:

$$c = \text{argmax} P_r(y | x_\alpha) \quad (20)$$

2 判别模型实证分析

2.1 突发事件数据描述

城市交通突发事件的样本数据来源于2014—2018年北京市轨道交通突发事件日志,样本数据包括线路故障、屏蔽门故障等8个直接致因,以及影响列车运行的4类27个间接致因,包括列车计划完成情况、列车晚点情况、路网客流情况、环境因素等。间接致因中,列车计划完成情况包括实际开行列车数、列车兑现率;列车晚点情况包含列车正点率、2分晚点列车数;路网客流情况包括日路网客流量、各条线路的断面满载率;环境因素包含工作日与否、天气状况、突发事件所在线路。从直接致因与间接致因共35个致因入手分析轨道交通风险致因对突发事件后果严重程度的影响。参照中华人民共和国交通运

输部颁布的《城市轨道交通运营安全风险分级管控和隐患排查治理管理办法》(简称《管理办法》)中给出的后果严重程度等级取值表及总体判断标准评估故障状态。后果严重程度等级取值见表1。

表1 后果严重程度取值

后果严重程度	取值
特别严重 (Extremely Serious, ES)	10
严重 (Serious, S)	5
较严重 (Relatively Serious, RS)	2
不严重 (Not Serious, NS)	1

根据《管理办法》,后果严重程度的判断标准涉及人员伤亡、社会影响2个方面。文中所选用的突发事件样本数据包含当日人员伤亡情况以及当日列车调整、列车延误2个社会影响类指标。

列车调整是列车遇到突发状况时管理人员采取的停运、通过、清人、掉线和中途折返等调整措施,列车延误则是列车遇到突发状况时到站晚点5 min及以上情况,在城市轨道交通系统日常的运营过程中,以上2种情况的产生都会对乘客的出行产生干扰,导致乘客产生不必要的出行延迟以及更换出行线路等社会影响。因此,用人员伤亡、列车调整、列车延误3个指标表征突发事件后果严重程度。

不同的突发事件会对路网列车的运行产生不同程度的影响,因而也对应着不同的后果严重程度。根据表1所划分的4种后果严重程度,将模型判别的结果进行分类,各风险等级判别结果的分类矩阵见表2,后果严重程度等级取值区间为[1, 10],将突发事件样本数据中的人员伤亡、列车调整、列车延误指标值分别归一化处理。由于人员伤亡、列车调整、列车延误3个指标相互独立,因此,对于任一条突发事件,3个指标对应的后果严重程度等级取值的最高值,可以看作该突发事件的后果严重程度值,对应得到后果严重程度等级。

表2 各风险等级判别结果的分类矩阵

Table 2 Classification matrix of prediction results of each risk level

真实分类	判别结果			
	NS	RS	S	ES
NS	a_{11}	a_{12}	a_{13}	a_{14}
RS	a_{21}	a_{22}	a_{23}	a_{24}
S	a_{31}	a_{32}	a_{33}	a_{34}
ES	a_{41}	a_{42}	a_{43}	a_{44}

注: a_{ij} 为针对实际样本的后果严重程度 i ,模型得到判别结果为 j 的样本数。

基于以上数据处理,先将风险致因输入IFSA

中,得到最优风险致因特征向量量子集。将样本数据划分为训练集和验证集,取训练集中的最优特征向量量子集与后果严重程度值输入 HRBM 模型中,训练得到最优的判别模型。利用验证集对得到的判别模型进行验证。

2.2 风险致因特征提取

将 35 个风险致因中的分类变量进行独热编码 (One-Hot) 处理,数值变量进行归一化处理,以此作为突发事件的初始风险致因特征向量 f 代入 IFSA, 基于不同时间窗 t_i 计算得到不同特征向量切片 T 中各个风险致因的信息增益比 $G_{R'}$ 。选择所有特征向量切片中信息增益比最大的风险致因,去重后得到突发事件风险致因最优特征向量量子集,共包含 7 个最优风险致因,分别为列车兑现率、正点率、日路网客运量、5 号线和 10 号线断面满载率、信号故障、车辆故障等。由结果可知:不论是日路网客运量越大,还是实际开行列车的数量与计划开行列车的数量相差越大,或是列车正点到达的比率较低,都表明有较为严重的突发事件产生。5 号线作为连接天通苑大型居住区与中心城区的南北通勤动脉,10 号线作为路网长度最长、站点最多、客流量最大的环形线路,长时间承受着较高的乘客压力,因此,当断面满载率较高时,容易出现突发状况。作为近 80% 突发事件的致因,信号故障和车辆故障这 2 个直接致因与突发事件的产生紧密相关。总体来说,城市轨道交通系统每日的列车运行情况越不理想、线路满载率越高,越容易发生后果严重程度高的突发事件。

2.3 模型评价结果分析

以 2014—2018 年共 9 817 条突发事件数据作为输入样本,将样本数据以 8 : 2 比例随机组合,分别作为训练集和测试集输入 IDM-ECS 模型中进行训练和测试,计算得到模型评价结果。为比较分析,同时选择 DRBM、RF、DF 及轻量梯度提升机 (Light Gradient Boosting Machine, LightGBM) 分类模型进行对比试验,得到各个模型的判别结果。

针对各类突发事件的后果严重程度,选用召回率 R 、精确度 P 、 F_1 值等 3 个数值指标评价模型的判别能力,评价指标计算方法见下式:

$$R = \frac{a_{ii}}{\sum_j a_{ij}} \quad (21)$$

$$P = \frac{a_{ii}}{\sum_i a_{ij}} \quad (22)$$

$$F_1 = \frac{2 \times P \times R}{P + R} \quad (23)$$

式中: R 表示针对某一类突发事件,模型的判别结果符合真实分类的样本占这一类突发事件样本的比例; P 表示对于某一类突发事件判别结果,模型的判别结果符合真实分类的样本占判别结果为该类突发事件样本的比例; F_1 表示 P 和 R 的加权调和平均数。

对比 IDM-ECS 模型及 DRBM、RF、DF、LightGBM 模型的 R 、 P 、 F_1 值评价指标,结果见表 3—表 5。

表 3 模型的 R 对比

Table 3 Comparison of recall of the models

模型	不严重	较严重	严重	特别严重	均值
DRBM	87.26	82.87	89.00	75.00	83.53
RF	85.02	82.32	87.12	62.50	79.24
DF	90.22	86.70	89.29	87.50	88.43
LightGBM	90.32	89.15	93.05	87.50	90.00
IDM-ECS	91.63	88.70	94.36	87.50	90.55

表 4 模型的 P 对比

Table 4 Comparison of precision of the models

模型	不严重	较严重	严重	特别严重	均值
DRBM	93.31	73.82	83.33	75.00	81.36
RF	92.85	70.43	83.15	55.56	75.50
DF	95.13	78.70	86.29	70.00	82.53
LightGBM	95.85	79.94	89.80	87.50	88.27
IDM-ECS	95.73	82.16	89.88	100.00	91.89

表 5 模型的 F_1 值对比

Table 5 Comparison of F_1 value of the models

模型	不严重	较严重	严重	特别严重	均值
DRBM	90.18	78.08	86.07	75.00	82.34
RF	88.76	75.91	85.09	58.82	77.15
DF	92.61	82.51	87.77	77.78	85.16
LightGBM	93.00	84.29	91.40	87.50	89.05
IDM-ECS	93.64	85.30	91.96	93.33	91.06

由表 3—表 5 可知:IDM-ECS 模型的整体判别情况最优。相较于 LightGBM, IDM-ECS 模型在“不严重”分类的 P 和“较严重”分类的 R 判别上略显劣势,但整体判别结果具有优势;相较于 DRBM, IDM-ECS 模型“特别严重”的评价值都有明显提升,考虑到“特别严重”数据量较小,表明模型对小样本的判别具有较大优势。

3 结论

1) 将 IFSA 应用于城市轨道交通突发事件风险

致因的筛选中,计算出不同风险致因的信息增益比,得到列车兑现率、正点率、日路网客运量、5号线断面满载率、10号线断面满载率、信号故障以及车辆故障等7个与路网突发事件后果严重程度关系显著的最优风险致因。运营管理人员可以通过重点关注这几项致因的变化,有针对性的指导风险防控工作。

2) 将 HRBM 应用到对风险致因与突发事件后果严重程度映射关系的剖析中,实现对城市轨道交通突发事件后果严重程度的判别。与传统的 DRBM、RF、DF、LightGBM 模型对比, IDM-ECS 模型平均的召回率为 90.55%、精确度为 91.89%、 F_1 值为 91.06%,均优于对比模型,显示出模型在判别效果方面的优势。

参考文献

- [1] YAN Hongyan, GAO Ce, ELZARKA H, et al. Risk assessment for construction of urban rail transit projects[J]. Safety Science, 2019, 118: 583-594.
- [2] WANG Jinghong, YAN Wenyu, XU Han, et al. Investigation of the probability of a safe evacuation to succeed in subway fire emergencies based on Bayesian theory[J]. KSCE Journal of Civil Engineering, 2018, 22(3): 877-886.
- [3] ZHENG Xuanchuan, WEI Yun, QIN Yong, et al. Classification method of urban rail transit emergencies based on improved K-means algorithm[J]. Journal of Transportation Systems Engineering and Information Technology, 2019, 19(3): 134-140.
- [4] SINGH R, HORCHER D, GRAHAM D J, et al. Decomposing journey times on urban metro systems via semiparametric mixed methods[J]. Transportation Research Part C: Emerging Technologies, 2020, 114: 140-163.
- [5] 马壮林, 邵逸恒, 舒兰, 等. 多层网络视角下地铁网络脆弱性分析[J]. 中国安全科学学报, 2023, 33(8): 164-172.
MA Zhuanglin, SHAO Yiheng, SHU Lan, et al. Vulnerability analysis of metro network from perspective of multi-layer network[J]. China Safety Science Journal, 2023, 33(8): 164-172.
- [6] ZHOU Zhihua, JI Feng. Deep forest[J]. National Science Review, 2019, 6(1): 74-86.
- [7] NICOLA P, LOUISE C, GENSERIK R. Learning about risk: machine learning for risk assessment[J]. Safety Science, 2019, 118: 475-486.
- [8] 李志慧, 孙雅倩, 陶鹏飞, 等. 交通事故后的交通运行风险状态等级预测方法[J]. 吉林大学学报:工学版, 2022, 52(1): 127-135.
LI Zhihui, SUN Yaqian, TAO Pengfei, et al. Prediction method of traffic operation risk level after traffic accident[J]. Journal of Jilin University: Engineering and Technology Edition, 2022, 52(1): 127-135.
- [9] 姜万录, 李满, 张培尧, 等. 基于全矢增强深度森林的旋转设备智能故障诊断方法[J]. 中国机械工程, 2022, 33(11): 1324-1335.
JIANG Wanlu, LI Man, ZHANG Peiyao, et al. Intelligent fault diagnosis method for rotating equipment derived from full vector enhanced deep forest[J]. China Mechanical Engineering, 2022, 33(11): 1324-1335.
- [10] 邵怡韦, 陈嘉宇, 林翠颖, 等. 小训练样本下齿轮箱故障诊断:一种基于改进深度森林的方法[J]. 航空学报, 2022, 43(8): 118-132.
SHAO Yiwei, CHEN Jiayu, LIN Cuiying, et al. Gearbox fault diagnosis under small training samples: an improved deep forest based method[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(8): 118-132.
- [11] 范冰倩, 董秉聿, 王彪, 等. 基于深度学习的地铁施工作业人员不安全行为识别与应用[J]. 中国安全科学学报, 2023, 33(1): 41-47.
FAN Bingqian, DONG Bingyu, WANG Biao, et al. Identification and application of unsafe behaviors of subway construction workers based on deep learning[J]. China Safety Science Journal, 2023, 33(1): 41-47.
- [12] 吴海涛, 刘月, 杜慧敏. 小样本条件下地铁运营事故致因推理模型[J]. 中国安全科学学报, 2023, 33(3): 134-140.
WU Haitao, LIU Yue, DU Huimin. Research on model of subway operation accident's cause under small sample condition[J]. China Safety Science Journal, 2023, 33(3): 134-140.
- [13] YANG Jun, DENG Jiandong, LI Shujuan, et al. Improved traffic detection with support vector machine based on restricted Boltzmann machine[J]. Soft Computing, 2017, 21(11): 3101-3112.
- [14] LI Chaopeng, WANG Jinlin, YE Xiaozhou. Using a recurrent neural network and restricted Boltzmann machines for malicious traffic detection[J]. NeuroQuantology, 2018, 16(5): 823-831.
- [15] ALDWAIRI T, PERERA D, NOVOTNY M A. An evaluation of the performance of restricted Boltzmann machines as a model for anomaly network intrusion detection[J]. Computer Networks, 2018, 144: 111-119.
- [16] 孙兴斌, 孙彦赞, 郑小盈, 等. 面向多类不平衡网络流量的特征选择方法[J]. 计算机应用研究, 2017, 34(2): 568-571, 594.
SUN Xingbin, SUN Yanzan, ZHENG Xiaoying, et al. Feature selection for multi-class imbalanced internet traffic[J]. Application Research of Computers, 2017, 34(2): 568-571, 594.

作者简介: 范博松 (1996—),男,山西运城人,工学博士,讲师,主要从事城市交通安全方面的研究。
E-mail:fanbosong@ppsuc.edu.cn。