

考虑多头注意力机制的交通事故等级分类预测模型

庄伟卿,余晗彧

(福建理工大学 互联网经贸学院,福建 福州 350014)

摘要: 针对交通事故等级预测中多因素耦合导致的分类精度不足问题,本文提出一种考虑多头注意力机制的交通事故等级分类预测模型。该模型通过优化算法自动调参,结合卷积特征提取与时序依赖处理,并引入多头注意力机制强化关键特征学习。采用 Kaggle 平台上的公开数据集对模型进行训练与测试并与其他模型进行对比。实验结果表明,该模型在4类事故的AUC值分别达到0.938(1级)、0.989(2级)、0.882(3级)和0.803(4级),均为最高值。在分类难度最大的4级事故预测中,该模型召回率表现显著领先于对比模型。在数据稀缺场景下,该模型的整体错误分类数最少且对2级事故的预测仅出现1例误判。

关键词: 道路交通;交通事故等级分类;风险预测;集成模型;机器学习

中图分类号: U491.31

文献标志码: A

文章编号: 2097-3853(2025)06-0599-08

Traffic accident classification prediction model considering multi-head attention mechanism

ZHUANG Weiqing, YU Hanyu

(School of Internet Economics and Business, Fujian University of Technology, Fuzhou 350014, China)

Abstract: To address the challenge of inadequate classification accuracy in traffic accident severity prediction caused by multi-factor coupling effects, this study proposes a severity classification model considering multi-head attention mechanism. The model features an optimization algorithm for automated hyperparameter tuning, combines convolutional operations for feature extraction with sequential dependency modeling, and leverages multi-head attention to enhance the learning of discriminative features. Experiments were conducted on a publicly available Kaggle dataset, with comparative evaluations against baseline models. Key findings demonstrate that the model attains AUC values of 0.938 (Level 1), 0.989 (Level 2), 0.882 (Level 3), and 0.803 (Level 4) in four types of accidents, outperforming all competing methods. For the most challenging Level 4 accident prediction, this model achieves significantly superior recall performance compared with other models. Under data-scarce conditions, the proposed model exhibits remarkable robustness, yielding the lowest overall misclassification rate—including only one error for Level 2 predictions—among evaluated models.

Keywords: road traffic; traffic accident classification; risk prediction; ensemble model; machine learning

世界卫生组织《2023年全球道路安全报告》显示,每年仍有约119万人死于交通事故^[1]。在此背景下,强化道路交通数据的智能化分析与科学管理,成为提升安全决策水平、预防事故发生的关键突破口。

随着科技发展和大数据的普及,越来越多的

学者开始利用机器学习技术来预测和预防道路交通事故^[2],通过分析大量交通数据来更准确地识别事故高发的区域和时间以及防范潜在事故风险^[3]。于翔海等^[4]使用合成少数过采样技术(SMOTE)对不平衡数据进行处理,采用极限梯度提升(XGBoost)建立了面向全路段和分路段的实

时事故风险预测模型,试验结果表明 XGBoost 模型具有良好的鲁棒性,但在处理含有较多噪声的数据时,模型的稳定性和准确性会受到一定影响。相比之下,卷积神经网络(CNN)的卷积操作能够自动捕获特征之间复杂的交互作用,可以较好地提取结构化数据中的潜在特征组合。因此,在交通领域的研究中,研究人员通常会将数据组织成二维或多维的形式,以使 CNN 更有效地进行特征提取^[5]。随着技术发展,集成模型应运而生,其在处理高维度、异构性强、噪声较多的数据集时具有显著的优势^[6]。Uguz 等^[7]提出了一种基于卷积神经网络和长短期记忆网络(LSTM)的混合模型 CNN-LSTM,对安塔利亚地区的交通事故数据进行建模和预测,测试结果表明混合模型在预测精度上显著优于单一模型 LSTM。尽管集成模型已经显著提高了预测性能,但仍存在一些问题,例如,集成模型在参数选择和超参数优化方面依赖于经验和反复试验,不仅耗时费力,而且可能无法找到全局最优解^[8]。2005 年,Vlahogianni 等^[9]最先将元启发优化算法应用于交通领域预测研究,

大大节约了时间成本,且实现了较为准确的预测,为后续诸如粒子群优化(PSO)^[10]、牛顿-拉夫逊优化算法(NRBO)^[11]等在交通预测领域的创新发展提供了理论基础。

综上所述,交通事故数据的维度和复杂性不断增加,传统模型在处理此类数据时的局限性愈发突出。因此,亟需一种能够在应对复杂数据集的同时提升模型泛化能力和预测性能的方法。为了解决这些问题,本研究提出了一种考虑多头注意力机制的交通事故等级分类预测模型(NBRO-CNN-LSTM-MATT 模型),通过引入多头注意力机制(MATT)^[12],增强了模型对特征之间相关性的捕捉能力,从而在特征空间中建立更清晰的分类边界;同时还结合 NRBO 对 CNN-LSTM 网络结构进行优化,提升模型的收敛速度和计算效率。

1 模型搭建与介绍

1.1 模型搭建

NBRO-CNN-LSTM-MATT 模型结构图如图 1 所示。

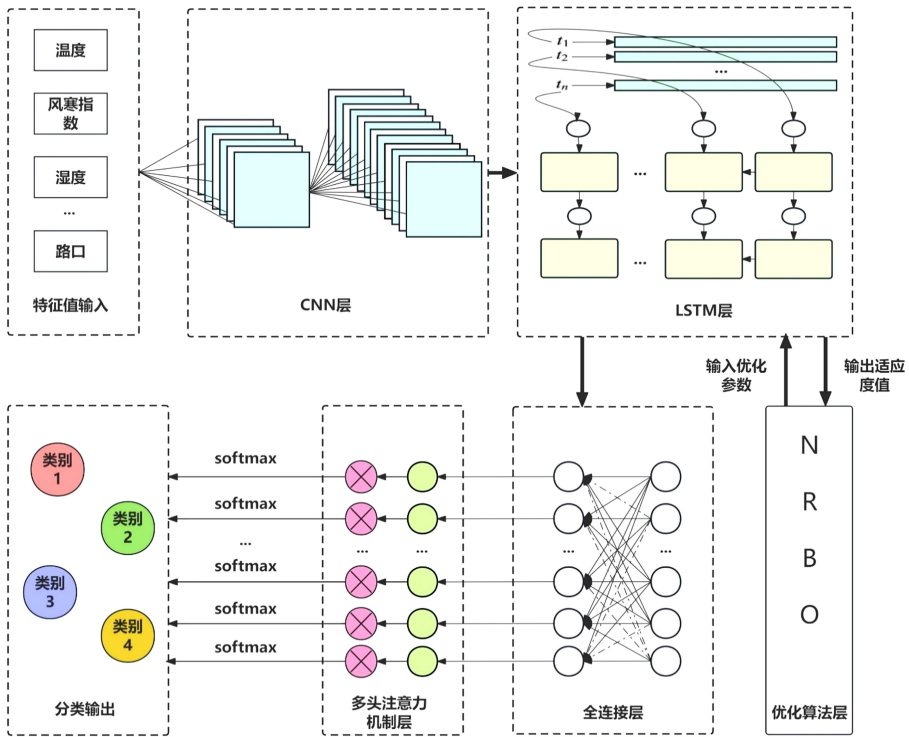


图 1 NBRO-CNN-LSTM-MATT 模型结构图

Fig.1 NBRO-CNN-LSTM-MATT model structure

首先,模型输入层接受 20 个特征的数值数据,并执行零中心化归一化处理,以确保数据分布

的标准化。为适配二维 CNN,将原始的一维特征向量重塑为尺寸为 5×4 的二维矩阵形式,并依据

特征之间的逻辑相关性或功能相似性进行排列。卷积层采用8个3×3的卷积核,在确保较低计算复杂度的同时,捕捉特征之间更细致的局部交互模式。卷积操作后的数据通过批量归一化层进行归一化处理,调整偏移量和缩放因子,确保数据在训练过程中的稳定性。

接着,对经过归一化处理的数据使用ReLU激活函数,增强模型的表达能力。随后,数据通过最大池化层,采用窗口尺寸为2×1,步长2×1的池化操作,缩减特征映射的尺寸,提取主要特征。池化后的数据被扁平化层处理成为一维向量后输入到LSTM层,输出通过全连接层映射到4个神经元,进一步提取特征。紧接着,数据进入自注意力层以捕捉特征之间的相关性,自注意力层包含4个注意力头,每个头有20个通道,能够提高模型的判别能力。

模型在训练过程中使用NRBO算法对LSTM层的隐藏层节点数、初始学习率、L2正则化系数进行优化,以提高模型的性能和收敛速度,有效地在高维搜索空间中找到全局最优解,从而提升LSTM层处理数据的效果。

然后,通过Softmax层对数据进行多类别分类,由Softmax函数将输出值转换为概率分布。

最后,模型通过分类输出层应用交叉熵损失函数计算分类结果。

1.2 模型相关理论

1.2.1 牛顿-拉夫逊优化算法

牛顿-拉夫逊优化算法引入牛顿-拉夫逊搜索规则(Newton-Raphson search rule, NRSR)提高了算法的探索能力和收敛速度,引入避免算子(trap avoidance operator, TAO)生成新的高质量解来增强全局搜索能力,两者结合克服了传统模型易陷入局部最优的不足,使NRBO成为解决多种优化问题的有效工具。

NRSR控制矢量,允许种群更准确地探索可行区域并获得更好的位置,从而促进勘探趋势并加快收敛。其数学公式表达如下:

$$\text{NRSR} = \text{randn} \times \frac{(\mathbf{y}_w - \mathbf{y}_b) \times \Delta x}{2 \times (\mathbf{y}_w + \mathbf{y}_b - 2 \times \mathbf{x}_n)} \quad (1)$$

$$\mathbf{y}_w = r_1 \times (\text{Mean}(\mathbf{Z}_{n+1} + \mathbf{x}_n) + r_1 \times \Delta x) \quad (2)$$

$$\mathbf{y}_b = r_1 \times (\text{Mean}(\mathbf{Z}_{n+1} + \mathbf{x}_n) - r_1 \times \Delta x) \quad (3)$$

$$\mathbf{Z}_{n+1} = \mathbf{x}_n - \text{randn} \times \frac{(\mathbf{X}_w - \mathbf{X}_b) \times \Delta x}{2 \times (\mathbf{X}_w + \mathbf{X}_b - 2 \times \mathbf{x}_n)} \quad (4)$$

式中,randn表示均值为0,方差为1的正态分布

随机数; \mathbf{y}_w 和 \mathbf{y}_b 是由向量 \mathbf{Z}_{n+1} 和 \mathbf{x}_n 生成的向量位置; \mathbf{x}_n 为当前迭代的位置向量, n 为种群中第 n 个个体的索引号; Δx 为位置增量;Mean表示计算平均值; r_1 为取值范围在(0,1)之间的随机数; \mathbf{Z}_{n+1} 表示下一次迭代的更新向量, \mathbf{X}_w 表示最差位置向量, \mathbf{X}_b 表示最佳位置向量。

TAO结合当前最佳位置和向量位置,通过随机参数生成新的解,避免陷入局部最优。其数学公式表达如下:

$$\begin{cases} \mathbf{X}'_{\text{TAO}} = \mathbf{X}_n^{t+1} + \theta_1 \times (\mu_1 \times \mathbf{x}_b - \mu_2 \times \mathbf{X}_n^t) + \theta_2 \times \delta \times (\mu_1 \times \\ \text{Mean}(\mathbf{X}_n^t) - \mu_2 \times \mathbf{X}_n^t), \text{ if } \mu_1 < 5 \\ \mathbf{X}'_{\text{TAO}} = \mathbf{x}_b + \theta_1 \times (\mu_1 \times \mathbf{x}_b - \mu_2 \times \mathbf{X}_n^t) + \theta_2 \times \delta \times (\mu_1 \times \\ \text{Mean}(\mathbf{X}_n^t) - \mu_2 \times \mathbf{X}_n^t), \text{ otherwise} \end{cases} \quad (5)$$

$$\mathbf{X}'_n = \mathbf{x}'_n - \text{randn} \times \frac{(\mathbf{y}_w - \mathbf{y}_b) \times \Delta x}{2 \times (\mathbf{y}_w + \mathbf{y}_b - 2 \times \mathbf{x}_n)} +$$

$$[a \times (\mathbf{X}_b - \mathbf{X}_n^t)] + b \times (\mathbf{X}'_1 - \mathbf{X}'_2) \quad (6)$$

式中, \mathbf{X}'_n 表示经过TAO算子处理后的当前更新位置, t 为当前迭代次数; \mathbf{X}'_{TAO} 是通过结合 \mathbf{X}_b 和 \mathbf{X}_n^t 生成的向量位置,用于在更新过程中进一步提高解的质量; \mathbf{x}_b 为当前迭代次数 t 中适应度最优的个体位置; μ_1, μ_2 分别为[0,3]和[0,1]之间的随机数, θ_1 为在(-1,1)之间的均匀随机数, θ_2 为在(-0.5,0.5)之间的均匀随机数, a 和 b 是(0,1)之间的随机数; δ 为自适应系数,用于调整步长; \mathbf{X}'_n 表示通过NRSR更新后的候选解向量; r_2 为(0,1)之间的随机数, \mathbf{X}'_1 和 \mathbf{X}'_2 分别代表从种群中随机选择的不同个体的位置向量。

1.2.2 卷积神经网络

二维卷积通常包括输入层、卷积层、池化层、全连接层和输出层,卷积核在输入信息的两个方向上滑动,通过加权求和和非线性激活函数提取特征,从而生成特征图。其数学公式如下:

$$(\mathbf{I} * \mathbf{K})(i, j) = \sum_m \sum_n \mathbf{I}(i + m, j + n) \cdot \mathbf{K}(m, n) \quad (7)$$

式中, \mathbf{I} 是多特征数值输入矩阵, \mathbf{K} 是卷积核(滤波器), i 和 j 是输出矩阵的位置坐标, m 和 n 是卷积核的位置坐标。

1.2.3 长短期记忆网络

LSTM通过引入3个门控机制(遗忘门、输入门和输出门),能够在长时间范围内有效地记忆

和忘记信息,标准 LSTM 模型结构如图 2 所示。其中, x_t 为当前时间步的输入值, h_t 为当前时间步的隐藏状态, C_t 为当前时间步更新后的细胞状态, \bar{C}_t 为新的候选细胞状态, i_t 为输入门, f_t 为遗忘门, O_t 为输出门,tanh 为 tanh 激活函数。

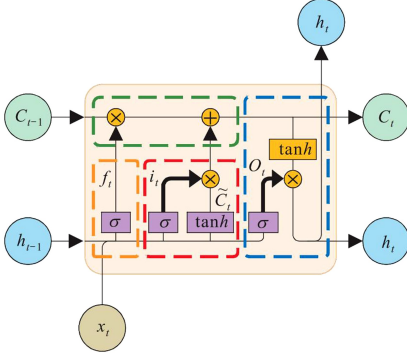


图 2 LSTM 模型结构图

Fig.2 LSTM model structure

1.2.4 多头注意力机制

多头注意力机制通过将输入数据拆分成多个部分并行计算各部分的注意力权重,再将这些结果拼接起来,从而捕捉更多的特征和模式,提高模型的稳定性和表达能力。

输入矩阵的线性变化中,对于输入矩阵 X ,通过线性变化得到查询矩阵 Q 、键矩阵 K 、值矩阵 V :

$$\begin{cases} Q = XW_Q \\ K = XW_K \\ V = XW_V \end{cases} \quad (8)$$

其中, W_Q 、 W_K 、 W_V 是可训练的权重矩阵。

单头注意力机制的计算结果如下:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (9)$$

其中,T 表示转置, $\sqrt{d_k}$ 为缩放因子,通常取键向量的维度 d_k 的平方根。

在单头注意力机制的基础上,使用多个独立的注意力头,每个头都有自己的查询、键和值矩阵。

$$h_i = \text{Attention}(QW_{Q_i}, KW_{K_i}, VW_{V_i}) \quad (10)$$

$$\text{MultiHead}(Q, K, V) = \text{Contact}(h_1, h_2, \dots, h_n)W_0 \quad (11)$$

其中,MultiHead(Q, K, V) 为所有注意力头的拼接输出, h_i 为单个头的输出, i 为索引值,Contact 表示拼接操作, n 为注意力头总个数, W_0 为输出的线性变换矩阵。

2 数据预处理与实验环境

本研究选取机器学习公开数据平台 Kaggle 中的 2020—2022 年美国全国交通事故数据。该数据集已将交通事故等级划分为 4 种严重程度,包含风速、压力、湿度等 43 个特征属性,对原数据集通过特征筛选,数据清洗等方法进行数据预处理,获得 4 个方面的数据特征,统计分析结果如表 1,表 2 所示。

表 1 事故影响因素中的连续变量统计

Tab.1 Statistics of continuous variables as accident influencing factors

连续变量	最小值	最大值	均值	标准差	众数
温度/℃	-18.89	45.56	15.42	10.33	8.00
风寒指数/℃	-22.39	45.56	14.52	11.49	22.22
湿度/%	5.00	100.00	72.96	22.20	100.00
气压/hPa	684.39	1 039.96	998.69	33.11	1 016.59
能见度/km	0	120.70	14.18	4.56	16.09
风速/ km · h ⁻¹	0	51.82	11.07	7.93	0

表 2 事故影响因素中的分类变量统计

Tab.2 Statistics of categorical variables as accident influencing factors

类型	分类变量	频数	占比/%
道路设施	便利设施	54	0.88
	减速带	2	0.03
	人行横道	583	9.50
	让行标志	25	0.41
	路口	300	4.89
	无出口	10	0.16
	铁路	58	0.95
	环岛	1	0.02
	车站	112	1.83
	停车标志	178	2.90
交通控制	交通降速	5	0.08
	交通信号	857	13.97
时间信息	调头标志	1	0.02
	白天	4 032	65.72

不同严重等级的事故具有不同的特征分布,其复杂程度为严重等级 4>严重等级 3>严重等级 1>严重等级 2。具体而言,严重等级 1 的事故主要发生在高温、高湿度的单一气候环境中且多发生在设施较为齐全的区域,特征之间的复杂交互作用使得该类别的特征较为多样性。严重等级 2 的事故通常发生在稳定气候的条件下,且道路设施和交通信号较少,这使得该类别在特征分布上较为集中,便于区分。严重等级 3 的道路设施比例较低,但事故天气条件波动略大,特别是在湿度和能见度上的波动较为显著,导致其特征表现出一定的分散性。严重等级 4 的事故发生在波动性较大的环境中,有高比例事故发生在有道路设施和交通控制的路段,且温度、风寒指数和气压等多个连续变量的标准差均高于平均水平,这使得该类别的特征分布最为分散,极其难以预测。

因严重等级为 4 级的数据样本量较少,存在严重的类别不平衡问题,为了防止严重等级 4 级的样本量占比过小而出现训练不足和预测失败的情况,本研究通过数据随机采样方法平衡数据集的各个类别,具体操作如下:

(1)特征筛选,对原数据中与交通事故影响因素无关的因素进行剔除,如城市邮政编码、机场代码等。

(2)数据清洗,检查数据中特征值为空的情况,并清除重要特征缺失的数据。将道路设施、交通控制中的分类变量特征值 TRUE 和 FALSE 值用 1 和 0 进行替换,将时间信息中的 Day 和 Night 用 1 和 0 进行替换。

(3)数据缩减,从 3 个数据量较大的类别中,分别随机抽取与类别 4 样本数量近似的数据。通过随机抽样,保证抽取样本的代表性,同时避免引入选择偏差。

本研究严格按年度划分数据集,以 2020-2021 年的数据作为训练集,2022 年的数据作为测试集,两者在时间上无任何重叠。该划分策略不仅确保了时间序列数据的独立性,也充分考虑到了不同年份环境差异可能对模型预测性能产生的影响,从而更真实地反映了模型的泛化能力。训

练集包含 4 321 条数据,其中严重等级为 1 的样本量有 1 130 条,严重等级为 2 的样本量有 1 092 条,严重等级为 3 的样本量有 1 090,严重等级为 4 的样本量有 1 009;测试集包含 1 813 条数据,其中严重等级为 1 的样本量有 477 条,严重等级为 2 的样本量 469 条,严重等级为 3 的样本量 437 条,严重等级为 4 的样本量 430 条,预处理过后的数据特征如表 3 所示。

本研究所提 NRBO-CNN-LSTM-MATT 交通事故等级分类预测模型是在 MATLAB R2024a 软件以及 AMD Ryzen 5 4600H with Radeon Graphics 3.00 GHz,内存 16.0 GB 硬件环境下开展实验。

表 3 预处理后的数据特征

Tab.3 Characteristics of data after preprocessing

编号	温度/°C	风寒指数/°C	湿度/%	...	风速/ km · h ⁻¹
1	22.78	22.78	94	...	0
2	8.00	7.33	93	...	5.63
3	26.67	26.67	54	...	0
⋮	⋮	⋮	⋮	⋮	⋮
6133	7.00	5.22	81	...	9.33
6134	25.00	25.00	88	...	4.83
6135	5.00	1.28	79	...	18.51

3 实验结果和讨论

3.1 分类结果对比

3.1.1 混淆矩阵分类结果对比

LSTM、NRBO-LSTM、NRBO-LSTM-MATT 和 NRBO-CNN-LSTM-MATT 等 4 种模型的混淆矩阵如图 3 所示。图 3 中,中间矩阵中的黑色部分表示被正确分类的样本数,其余浅灰色部分表示被错误分类的样本数;右侧矩阵的黑色部分为召回率、浅灰色部分为漏报率;底部矩阵的黑色部分为精确率、浅灰色部分为错误发现率。4 个模型中 LSTM 模型在第 2 类上的分类表现良好,召回率和精准率分别为 96.1% 和 82.5%,但在第 4 类上表现较差,召回率和精准率分别为 43.2% 和 50.6%,如图 3(a) 所示。图 3(b) 中 NRBO-LSTM

模型在第 2 类上有较好表现,其召回率和精准率分别为 95.1% 和 84.6%,但在第 4 类上的分类效果不佳,召回率仅为 22.3%。图 3(c) 中 NRBO-LSTM-MATT 模型在第 2 类上的分类性能接近 NRBO-CNN-LSTM-MATT,其召回率为 95.8%,精准率为 85.5%,但在第 3 类和第 4 类上表现欠佳。相比之下,图 3(d) 中的 NRBO-CNN-LSTM-MATT

模型在第 2 类和第 4 类分类预测上的表现最好,召回率分别达到了 98.4% 和 58.1%,精准率分别为 87.3% 和 55.1,整体上优于其他模型;在第 1 类上,该模型的召回率和精准率分别为 71.7% 和 77.1%,同样显示出较强的分类能力;而在第 3 类上精准率也高达 69.9%,为最优值。综上,NRBO-CNN-LSTM-MATT 模型的整体分类性能最佳。

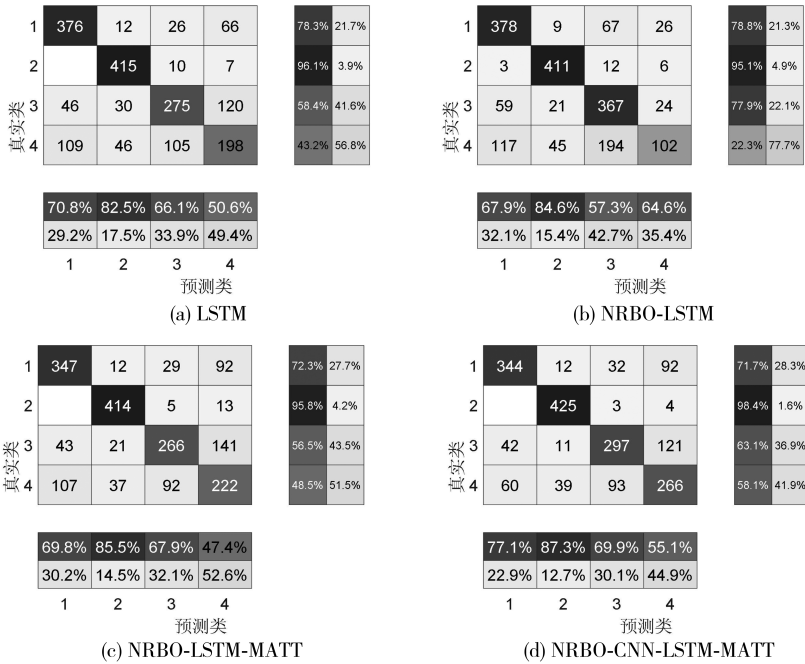


图 3 不同模型混淆矩阵对比图

Fig.3 Comparison of confusion matrices of different models

3.1.2 ROC 曲线分类性能对比

7 种模型的受试者工作特征 (receiver operating characteristic, ROC) 曲线如图 4 所示,横轴表示假阳性率,纵轴表示真阳性率。图中的数值表示各个模型的曲线下面积 (area under the curve, AUC),AUC 越高,模型的分类判别能力越强。由图 4 可知,NRBO-CNN-LSTM-MATT 模型在类别 1 和类别 2 上的 AUC 值分别为 0.937 66 和 0.989 39,均优于其他模型;在类别 3 上的 AUC 值为 0.882 21,相较于次优的 NRBO-LSTM-MATT 模型提高了约 1%,相较于表现最差的 KELM 模型提高约 5.3%;在类别 4 上的 AUC 值为 0.802 64,比次优的 KELM 模型相对提升了约 2.1%,相较于表现最差的 MLP 模型的 0.746 40 提高了约 7.5%。

综上所述,NRBO-CNN-LSTM-MATT 在 4 种分类上的 AUC 值均为最高值,表明该模型在不同类别的决策阈值下,整体预测的准确性和可靠性更高,能够有效减少错误分类的发生。

3.2 稳健性检验

为了评估各模型在小样本量条件下的预测性能是否与大样本量情况下保持一致,本研究从训练集中随机抽取 350 个样本进行训练,并从测试集中随机抽取 150 个样本进行分类预测,重复 30 次实验取其平均值来进行小样本稳健性检验。图 5 为不同模型的小样本量分类预测结果对比图,灰色实线表示实际分类结果,圆圈表示模型的预测分类结果。

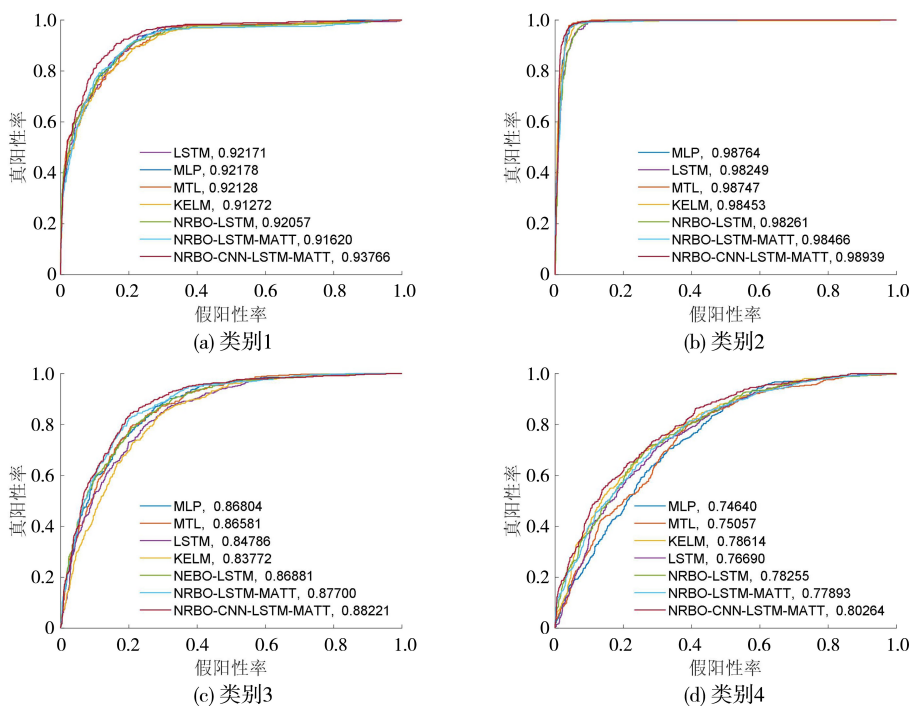


图 4 不同模型的 ROC 曲线对比图

Fig.4 Comparison of ROC curves of different models

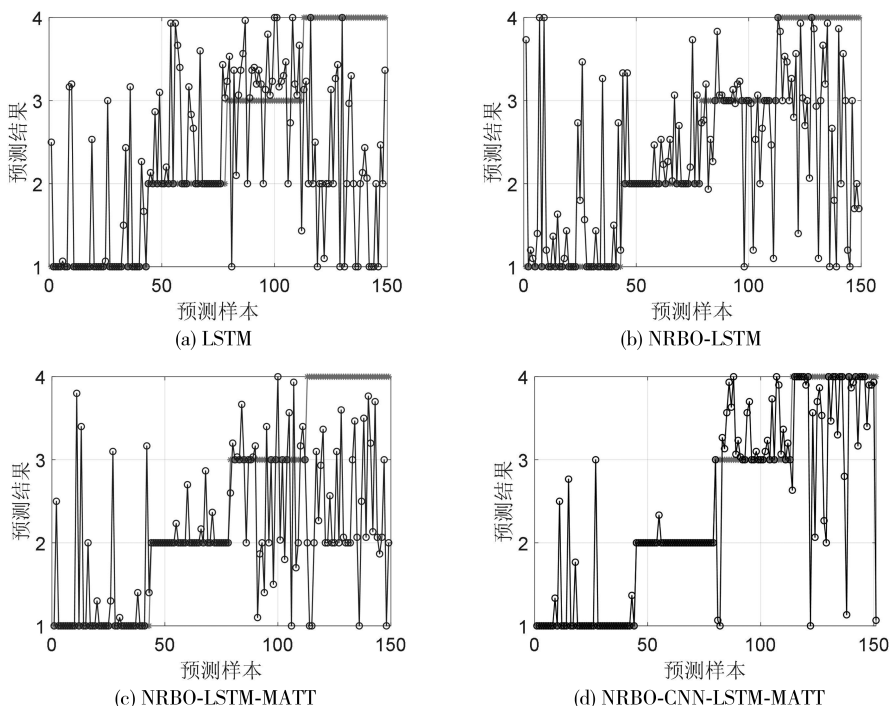


图 5 不同模型小样本量分类预测结果对比图

Fig.5 Comparison of classification prediction results of different models on small sample sizes

由图 5(a)与图 5 (b)可知,LSTM 与 NRBO-LSTM 在小样本分类预测任务中表现出较大的波动性,类别混淆严重。相比于前两者,图 5(c)

中的 NRBO-LSTM-MATT 模型在类别 2 上的误分类现象明显减少,但在类别 3、4 上的预测准确性有所下降。图 5(d)中的 NRBO-CNN-LSTM-

MATT 在类别 2 中分类效果最佳,被错分的数量仅为 1 个;在类别 3 上的预测值跳跃较小,预测结果集中于 3、4 之间;在类别 1 和类别 4 的预测表现相较于其他模型均体现出明显优势。综上所述,即使是小样本量数据训练, NRBO-CNN-LSTM-MATT 模型在分类预测任务中的表现依旧是最佳。

4 结论

1) NRBO-CNN-LSTM-MATT 模型在多维数据处理及类别区分上表现卓越,结合了 NRBO 的优化策略、CNN 的高效特征提取与 LSTM 的数据处

理能力以及 MATT 的精细化信息处理,不仅有效降低了模型误分类率,而且显著增强了分类结果的精确性和可靠性。

2) 模型的高 AUC 值进一步确认了其在各种决策阈值下维持预测准确性的能力,强化了其在实际操作中的可靠性和价值,也突显了模型在面对复杂实际应用场景时的适用性和效率。

3) 虽然模型在性能评估上表现优异,但在最高严重等级 4 级的事事故分类预测中的召回率和精准率依然有待提高。未来研究将会探索更加复杂的数据增强技术或者更深层次的特征工程,以期进一步优化模型在极端类别数据上的表现。

参考文献:

- [1] 严利鑫, 胡鑫辉, 刘清梅, 等. 道路交通事故严重程度预测及致因分析 [J]. 华东交通大学学报, 2024, 41(5): 65-73.
- [2] ZHANG K, GUO Y M. Attention-based residual dilated network for traffic accident prediction [J]. Mathematics, 2023, 11(9): 2011.
- [3] JIN Z X, NOH B. From prediction to prevention: leveraging deep learning in traffic accident prediction systems [J]. Electronics, 2023, 12(20): 4335.
- [4] 于翔海, 白佃国, 于光, 等. 基于 XGBoost 模型的城市道路实时交通事故风险预测研究 [J]. 公路交通科技, 2023, 40(4): 237-247.
- [5] 陈喜群, 沈楼涛, 李俊懿, 等. 基于异构数据特征的城市轨道交通 OD 客流短时预测方法 [J]. 交通信息与安全, 2024, 42(2): 158-165.
- [6] ÇEVEN S, ALBAYRAK A. Traffic accident severity prediction with ensemble learning methods [J]. Computers and Electrical Engineering, 2024, 114: 109101.
- [7] UGUZ S, BUYUKGOKOGLAN E. A hybrid CNN-LSTM model for traffic accident frequency forecasting during the tourist season [J]. Tehnicki Vjesnik-Technical Gazette, 2022, 29(6): 2083-2089.
- [8] XU X C, JIN X F, XIAO D Q, et al. A hybrid autoregressive fractionally integrated moving average and nonlinear autoregressive neural network model for short-term traffic flow prediction [J]. Journal of Intelligent Transportation Systems, 2023, 27(1): 1-18.
- [9] VLAHOGIANNI E I, KARLAFTIS M G, GOLIAS J C. Optimized and meta-optimized neural networks for short-term traffic flow prediction: a genetic approach [J]. Transportation Research Part C: Emerging Technologies, 2005, 13(3): 211-234.
- [10] QIU C Y, WANG C L, FANG B X, et al. Amultiobjective particle swarm optimization-based partial classification for accident severity analysis [J]. Applied Artificial Intelligence, 2014, 28(6): 555-576.
- [11] SOWMYA R, PREMKUMAR M, JANGIR P. Newton-raphson-based optimizer: a new population-based metaheuristic algorithm for continuous optimization problems [J]. Engineering Applications of Artificial Intelligence, 2024, 128: 107532.
- [12] CHENG S Y, LIU Y A. Research on transportation mode recognition based on multi-head attention temporal convolutional network [J]. Sensors, 2023, 23(7): 3585.

(责任编辑: 方素华)