

## 基于文本挖掘的民航事件风险评估

倪晓梅,王华伟<sup>†</sup>,熊明兰,王峻洲  
(南京航空航天大学民航学院,江苏南京 210016)

**摘要:**为实现“安全第一、预防为主、综合治理”的民航安全管理目标,建立了从报告中学习并评估风险等级的深度学习模型.首先采集航空安全报告系统中10年报告,根据严重度建立事件后果的量化指标,确定5个风险等级:高、中高、中、中低和低风险,并消除事件结果分布不平衡和结果多样性的影响.然后应用卷积神经网络(Convolutional Neural Network, CNN)探索非结构化的事件概要与风险等级之间的关系,通过该模型对事件进行分类,确定风险等级.风险评估模型与不同量化指标和不同方法对比,其分类准确率可达96%,优于其他指标和方法.最后应用该模型对非结构化的事件概要挖掘,对2020年事件进行快速的风险评估,预测准确率可达80%.基于CNN的民航风险评估模型可以对文本格式的事件概要充分挖掘,快速评估与主动感知风险,对支持安全预警具有重要意义.

**关键词:**民航安全;风险评估;安全预警;文本挖掘;卷积神经网络

**中图分类号:**X949

**文献标志码:**A

## Civil Aviation Incident Risk Assessment Based on Text Mining

NI Xiaomei, WANG Huawei<sup>†</sup>, XIONG Minglan, WANG Junzhou

(School of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

**Abstract:** In order to achieve the civil aviation safety management goal of ‘safety first, prevention first and comprehensive management’, a deep learning model is established to learn from reports and assess the risk level. Based on the 10-year incident reports available in the Aviation Safety Reporting System, we first establish quantitative indicators of incident consequences and classify all incidents into 5 levels according to their severity: high, moderately high, moderate, moderately low and low risk, which helps to eliminate the impact of unbalanced and intricate event consequences. Then, the relationship between the unstructured incident synopsis and the risk level is explored by convolutional neural network (CNN), and the events are classified by the model to determine the risk level. The classification model proves its superiority by comparing it with different quantitative indicators and methods, with an accuracy of 96%, which is better than the compared models. Finally, the 2020’s incident reports are predicted by this model, which enables rapid risk assessment of the synopsis of the incident, with an accuracy rate of 80%. The CNN-based civil aviation risk assessment model can fully mine the text-formatted incident synopsis, and quickly assess and actively perceive the risk level, which helps support the early warning of civil aviation safety.

**Key words:** civil aviation safety; risk assessment; safety warning; text mining; convolutional neural network

\* 收稿日期:2021-05-15

基金项目:国家自然科学基金民航联合研究基金(U1833110), Union Research of National Natural Science Foundation of China (U1833110)

作者简介:倪晓梅(1992—),女,江苏南通人,南京航空航天大学博士研究生

<sup>†</sup>通信联系人,E-mail:wang\_hw66@163.com

2021年是“十四五”的开局之年,民航局部署,到2025年,力争全国运输机场容量达到20亿人次.运输量的增加必然会给民航安全管理带来严峻的挑战.安全风险评估作为《民用航空安全管理规定》<sup>[1]</sup>的重要内容,是实现“安全第一、预防为主、综合治理”目标的保障.

民航安全发展的历史,就是从事故中学习的过程.要实现上述管理的目标,挖掘可利用的事故数据与信息,实现精准的风险评估是迫切需要解决的问题.若能深层次挖掘事故报告中描述的事故过程,拓展以往只局限于简单查询和统计的应用,克服依靠领域专家分析和研究的主观性等问题,将更好地实现通过事故进行学习的功能.

机器学习和自然语言处理的发展和应用为分析航空安全报告中的非结构化文本数据提供了工具和可能.当前在本领域,陆续有学者开展了相关研究. Alkhamisi等<sup>[2]</sup>、Zhang等<sup>[3]</sup>和Tanguy等<sup>[4]</sup>对于航空安全报告系统(Aviation Safety Reporting System, ASRS)中非结构化数据的分类和预测都采用了支持向量机(Support Vector Machine, SVM).以上采用传统机器学习的方法由于文本表示的高纬度高稀疏性,特征表达能力不强,导致预测精度不够,需要通过其他方法的集成或处理来获得更准确的预测结果.而如果风险等级评估过高,将导致实际风险管理中的“虚警”,造成资源的浪费;若风险等级评价过低,则会导致实际风险管理中的“漏警”,影响安全风险.

最新的研究表明,深度学习技术逐渐取代了传统机器学习方法,成为文本分类领域<sup>[5]</sup>的主流技术.深度学习能够更准确地表达对象,从海量数据中自动获取特征,实现端到端地解决问题,极大地提高了准确性.卷积神经网络(Convolutional Neural Networks, CNN)作为典型的深度学习方法,在文本分类领域有着较好的应用.如罗文慧等<sup>[6]</sup>、兰秋军等<sup>[7]</sup>和GAO等<sup>[8]</sup>,通过CNN构建分类模型,有效抽取句子实体关系.

基于此,本文从文本挖掘的角度出发,提出一种能够自动评估民航运营过程中安全风险等级的模型.通过CNN找到非结构化事件信息与事件结果之间的关联关系,建立对应风险等级,从而协助民航总局等相关单位进行风险量化、优先级设置、资源分配和决策制定,实现主动安全管理的目标.

## 1 民航事件结果量化

ASRS是事件报告数据库,可提供过去几十年来发生的事件/事故信息,描述了不安全事件和危险情况<sup>[9]</sup>.但由于事件结果的复杂性、数据不平衡等特点,直接建立事件概要与结果的映射在实际中需要较强的专业性,在模型算法方面的难度也较大.针对上述问题,首先建立对应的量化指标.

### 1.1 民航事件结果

搜集从2010年1月至2019年12月ASRS报告,其中与民航运营<sup>[10]</sup>相关\*的事件共有33个独立结果,如表1所示.

表1 事件结果说明表

Tab.1 Incident result description

类别	结果	
空管	发布咨询/警告 $x_1$	
	发布新起降许可 $x_2$	
	提供协助 $x_3$	
	分散交通 $x_4$	
	飞机损坏 $x_5$	
航空器	自动装置超控机组 $x_6$	
	设备问题消除 $x_7$	
	重定方向 $x_8$	
	偏离规定航线 $x_9$	
	复飞、进近 $x_{10}$	
	退出侵入的空域 $x_{11}$	
	FLC超越自动化 $x_{12}$	
	FLC遵从自动化、咨询 $x_{13}$	
	空中停车 $x_{14}$	
	备降 $x_{15}$	
	机组	紧急降落 $x_{16}$
		克服设备故障 $x_{17}$
		恢复对飞机的控制 $x_{18}$
拒绝起飞 $x_{19}$		
请求ATC协助 $x_{20}$		
返回清关 $x_{21}$		
返回出发机场 $x_{22}$		
返回登机口 $x_{23}$		
采取规避措施 $x_{24}$		
疏散 $x_{25}$		
一般事件	航班取消、延误 $x_{26}$	
	维护 $x_{27}$	
	没有报告/已采取行动 $x_{28}$	
	人员伤亡/丧失能力 $x_{29}$	
	涉及警察/保安 $x_{30}$	
	拒绝发布/不接受飞机 $x_{31}$	
	拒绝工作 $x_{32}$	
	宣布紧急状态 $x_{33}$	

注:\*民航运营相关:遵循FAR121条款,运行任务为商业乘客运输

此外还有约半数的事件有多个后果,对风险的辨识造成严重的干扰.由于结果的多样性和复杂性,直接建立评估模型难以准确判断风险水平.

1.2 数据不平衡性

民航安全风险具有一定的聚集特点,即表现为某类事件发生次数较多,某类则相对较少.

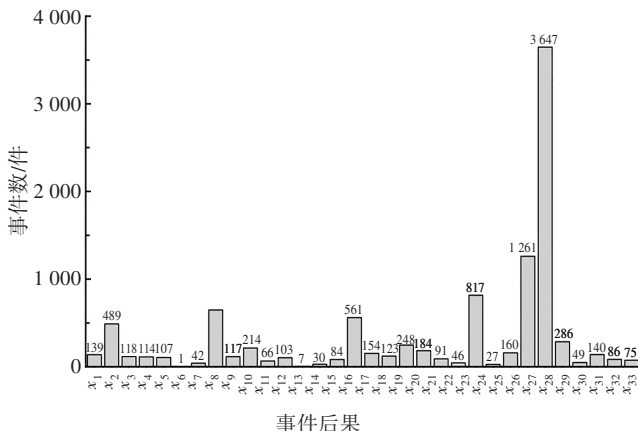


图 1 民航事件结果分布图

Fig.1 Distribution of civil aviation incident results

2010 年报告的所有事件单一后果分布如图 1 所示.

可以看出,这 33 类事件具有严重的不平衡特性,多数类与少数类的比值超过 3000,这种不平衡的类分布对假设相对平衡分布的机器学习算法提出了严峻的挑战<sup>[11]</sup>,同时难以解释结果的客观性和科学性.

1.3 风险等级量化

基于以上事件后果多样且分布不平衡两方面的问题,量化事件后果可以更直观地表现结果的严重程度,减少对专业的依赖性,降低人力成本.

国际民航组织对严重程度等级依据人员伤亡、经济损失、机场关闭时间等<sup>[12]</sup>因素分为了 5 级,故本文按照此标准,将事件后果依据风险等级划分为 5 个等级.则每个等级按照表 1 中提供的事件结果,对应风险等级如表 2 所示,表中“损失”即为国际民航组织提供的划分依据.

表 2 事件结果风险等级表

Tab.2 Incident outcome risk Level Table

严重度	风险等级	损失			对应事件结果
		人员伤亡	经济损失	运行损失	
高风险	5	超过 10 人伤亡	损失超过 1 000 万	机场关闭或停止运行	{x <sub>4</sub> , x <sub>5</sub> , x <sub>14</sub> , x <sub>29</sub> , x <sub>33</sub> }
中高风险	4	造成 1~9 人伤亡	损失 100 万~1 000 千万	跑道关闭超 24 小时/滑行道和停机坪关闭超 72 小时	{x <sub>1</sub> , x <sub>16</sub> , x <sub>18</sub> , x <sub>25</sub> }
中度风险	3	严重损伤	损失 10 万~100 万	跑道关闭超 12 小时/滑行道和停机坪关闭超 72 小时	{x <sub>2</sub> , x <sub>8</sub> , x <sub>9</sub> , x <sub>10</sub> , x <sub>17</sub> , x <sub>19</sub> , x <sub>24</sub> , x <sub>32</sub> }
中低度风险	2	轻伤	损失 1 万~10 万	延误 3 小时或航班取消	{x <sub>6</sub> , x <sub>11</sub> , x <sub>12</sub> , x <sub>15</sub> , x <sub>20</sub> , x <sub>21</sub> , x <sub>22</sub> , x <sub>26</sub> , x <sub>27</sub> , x <sub>31</sub> }
低风险	1	轻微擦伤	损失小于 1 万	不影响航班运行	{x <sub>3</sub> , x <sub>7</sub> , x <sub>13</sub> , x <sub>23</sub> , x <sub>28</sub> , x <sub>30</sub> }

对于单一后果的事件,按照风险等级划分,对于多后果的事件,根据民航局“安全隐患零容忍”的态度和“从严从实管控安全风险”的原则,遵循按风险等级高的划分,则事件分布如下图 2 所示.

风险等级的划分可以很好地解决事件不平衡的问题,并明确复杂事件的风险等级,这有助于模型准确性的提升<sup>[13-14]</sup>.

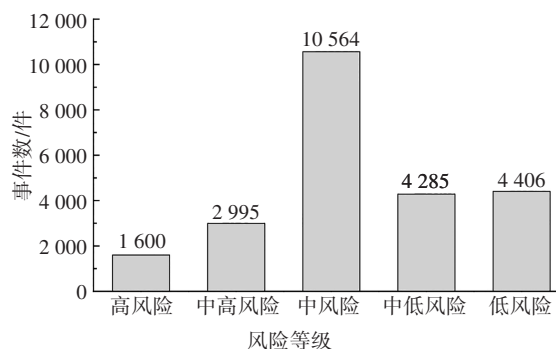


图 2 民航风险等级分布

Fig.2 Civil aviation risk level distribution

## 2 CNN 风险评估模型

在文本分类领域常用的深度学习方法主要有 CNN、循环神经网络 (Recurrent Neural Network, RNN) 及其改进和集成模型等<sup>[5]</sup>。

CNN 仿造生物的视觉机制构建, 能够很好地提取特征信息, 在分类问题上取得了较好的效果, Yoon Kim<sup>[15]</sup> 最早将 CNN 用于提取文本的特征实现文本分类。

基于 CNN 的民航事件风险评估模型主要有两大优点: (1) 文本表示自动化, 泛化性较强, 无需专业的民航背景知识即可得到较好的特征; (2) 结构灵活, 基于常用结构易于设计优化结构, 以便后续依托民航数据特点及安全风险规律改进模型。

CNN 主要由六部分组成, 输入层、卷积层、激活层、池化层、全连接层、softmax 函数等, 它与普通网络的区别就在于增加了由卷积层和池化层结合构成的特征提取器。其结构如图 3 所示。

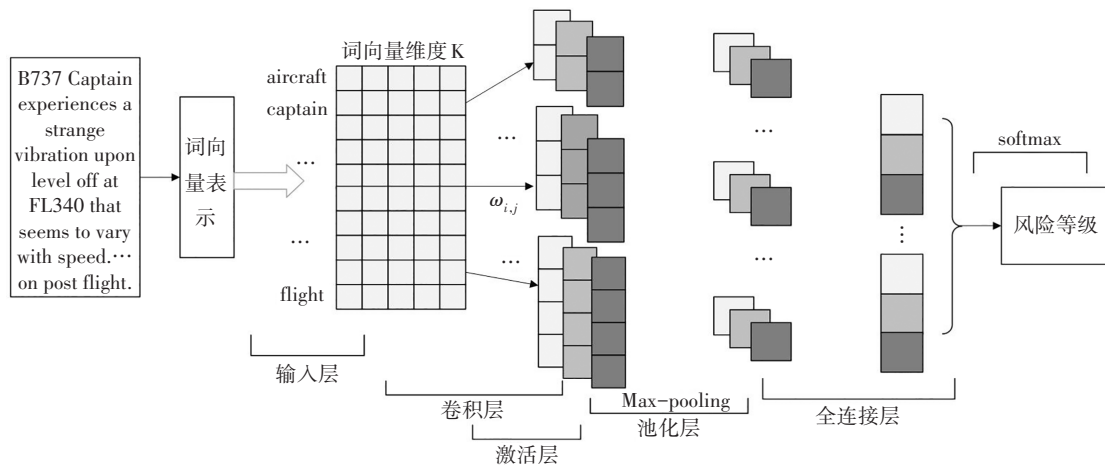


图 3 CNN 风险评估结构图

Fig.3 CNN risk assessment structure

### 2.1 文本预处理

事件概要是一种非结构化数据, 它是事件/事故的简单说明。根据 NLP 文本处理的基础步骤和民航事件的特点, 本文的预处理过程包括数据清洗, 加载风险源自定义字典、去停用词和分词 4 个环节。

1) 由于涉及到飞机型号、机场、塔台等编号, 为了便于识别, 则首先清洗数据: 将飞机型号如 (B737, A320...) 等全部转换为 aircraft, 机场编号全部转换为 airport 等。

2) 为保障文本处理的质量, 按照文献[16]中专业术语的缩写构建自定义词典。

3) 停用词和分词按照常用英语停用词, 则转换后的语言如图 4 所示。

### 2.2 CNN 结构

#### 1) 卷积层

卷积层的输入为  $n \times n$  的矩阵, 卷积核大小为  $m \times m$ ,  $n \geq m$ , 权重为  $\omega$ , 卷积后的输出为  $c$ , 则对应位置  $(i, j)$  的输出为:

$$c_{i,j} = \sum_{k=1}^m \sum_{l=1}^m \omega_{k,l} x_{i+k-1, j+l-1}, j \leq n-m+1 \quad (1)$$

式中:  $x_{i,j}, \omega_{i,j}$  分别为  $x, \omega$  在位置  $(i, j)$  的值。

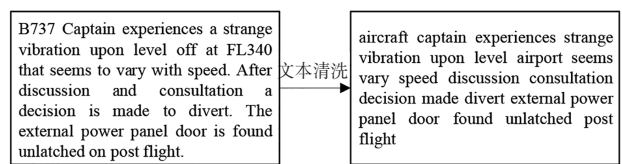


图 4 文本清洗

Fig.4 Text cleaning

#### 2) 池化层

本文在池化层的构造中采用 Max-pooling 方法, 则对应位置  $(i, j)$  的输出为:

$$p_{i,j} = \max\{c_{i,j}, c_{i,j+1}, \dots, c_{i,j+a-1}, \dots, c_{i+a-1, j+q-1}\} \quad (2)$$

#### 3) 输出层

全连接层后生成向量  $f = (d_1, d_2, \dots, d_g)$ , 其中  $d_g$  为特征值,  $g$  为  $f$  的维度, 将  $f$  输入到 softmax 分类器中进行风险等级分类, 由于风险等级有 5 级, 则选用多分类交叉熵损失函数作为目标函数  $L$ :

$$L = -\frac{1}{N} \sum_{s=1}^N \sum_{c=1}^5 y_{s,c} \ln p_{s,c} \quad (3)$$

式中:  $y_{s,c}$  为样本  $s$  的标签, 与等级  $c$  为同一风险等级时则为 1, 不同为 0,  $p_{s,c}$  为样本  $s$  属于等级  $c$  的概率.

### 3 实验

实验平台为 3.00 GHz, Intel (R) Core (TM) i7-9700 CPU, 8GB 内存, 在 Google Colab 中实施, 实验程序基于开源软件库 Tensorflow2.0 和 keras.

#### 3.1 实验结果

10 年报告共 28351 条信息, 选取其中 25% 为测试数据, 其余为训练数据. CNN 风险评估模型结构和参数如图 5 所示.

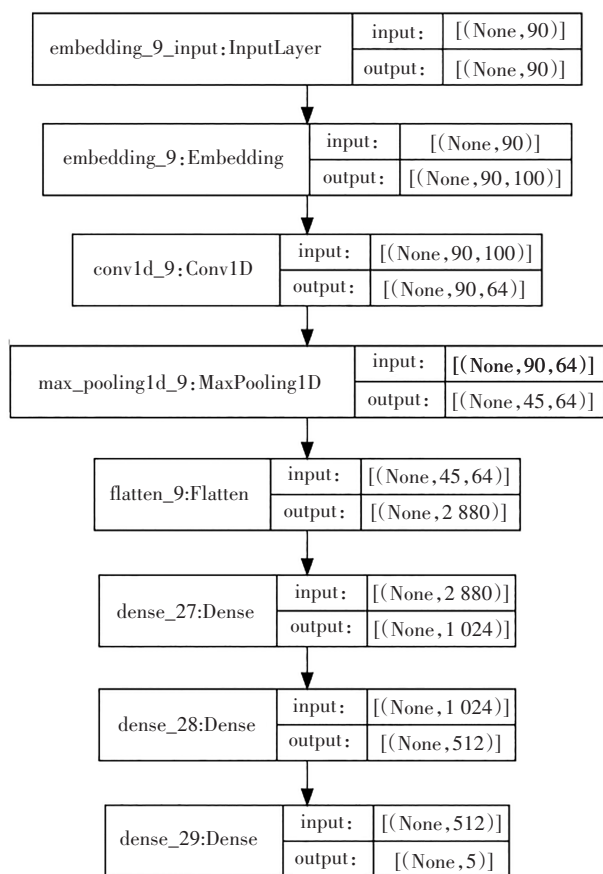


图 5 CNN 结构图  
Fig.5 CNN structure

结构图说明最大词向量维度为 90, 卷积层采用 Conv1D.

参数设置如下: epoch 次数为 60, 即共训练 60 轮, 迭代次数为 1000, 学习率为 0.02, 训练后准确率可达 96.3%, 结果如图 6 所示.

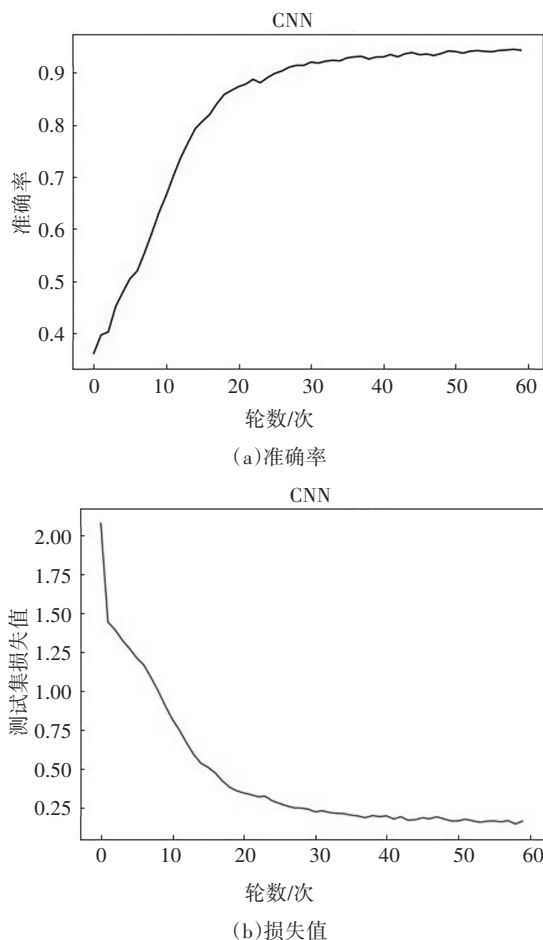


图 6 CNN 风险分类结果图

Fig.6 CNN risk classification result

#### 3.2 对比实验

为验证该方法的优越性, 分别通过不同量化指标和不同模型两方面的对比说明, 如下图 7 所示.

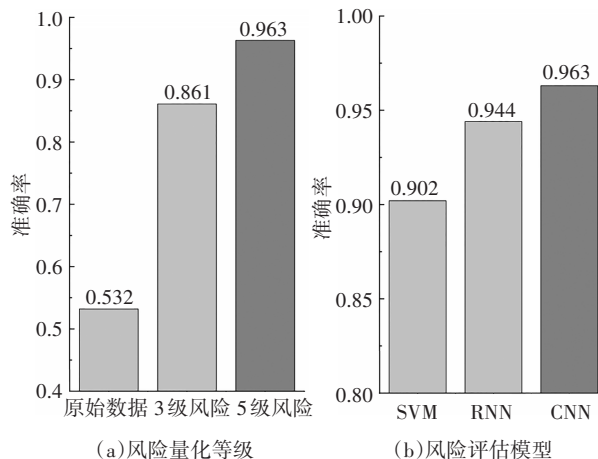


图 7 对比模型准确率

Fig.7 Accuracies of comparison models

其中(a)表示应用 CNN 模型分别对事件后果无



分类(即保留原始数据中的 33 种后果)、参考文献 [2] 将事件后果风险等级分为高,中,低三等(本文即将中高、中低等风险全部归结为中等风险)和本文提出的 5 级对比. 结果表明同样使用 CNN, 5 等级的量化指标不仅符合国内和国际的标准, 准确度也更高.

在 5 等级的量化标准下, 图 7(b) 表示分别应用 SVM、CNN 和 RNN 进行文本分类. 结果说明: CNN 模型的分类效果最好; 同时也表明, 对于民航事件信息, 以 CNN 和 RNN 为代表的深度学习模型准确度高于以 SVM 为代表的传统机器学习模型.

在本实验中, CNN 模型在准确度上略优于 RNN, 这是因为 CNN 模型的词向量最大为 90, 不属于长序列, CNN 模型足够处理. 而民航词汇的专业性让 CNN 能够专注于捕捉民航事件文本的语义信息.

### 3.3 预测

为进一步验证模型的通用性, 将 2020 年共 2111 份报告输入 CNN 分类模型, 预测 2020 年的风险等级, 对事件进行风险评估, 结果表明准确率可达到 79.6%, 混淆矩阵如图 8 所示:

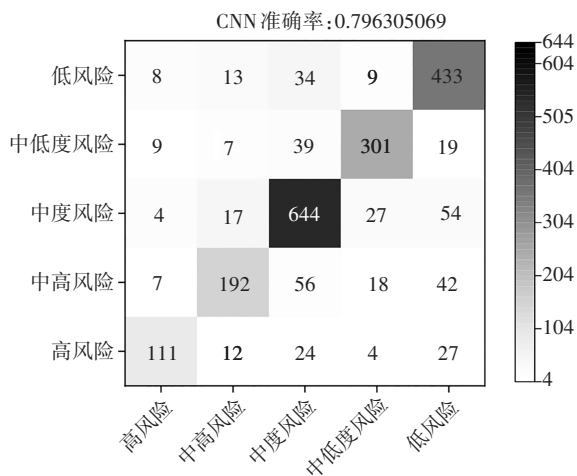


图 8 CNN 预测混淆矩阵

Fig.8 CNN prediction confusion matrix

其中单个事件的风险评估过程如图 9 所示:

图 9 说明对于单个事件, 当采集到一线人员的事件描述时, 即可通过本模型进行风险等级的预测, 对于上述事件, 风险等级的预测结果为 3 级, 即中度风险, 对比原事件信息, 对应的事件后果为发布新的起降许可, 符合中等风险的损失标准, 说明预测正确.

同时预测结果的有效性也说明: 采用该模型, 后续在工程中, 采集到一线人员对事件的描述后, 可以无需专家分析而直接对事件进行评估, 确定风险水平, 以便民航各个部门快速响应, 针对风险等级采取对应措施, 从而减少财产损失和人员伤亡.

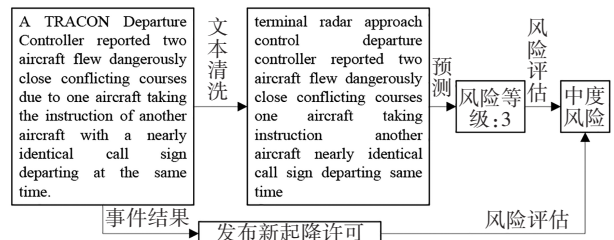


图 9 单个事件风险评估图

Fig.9 Single event risk assessment chart

实验结果表明 CNN 风险评估模型可以有效对非结构化的事件报告分类, 而对事件结果按照风险等级的量化可以有效提高分类的准确率, 最新的实验数据进一步验证了模型的通用性.

## 4 结论

本文通过提取 ASRS 事件信息, 针对其非结构化、不平衡等特点, 通过对事件风险等级的量化分级与映射, 设计了风险评估模型. 提高了事件概要的分类与预测准确度, 为实现主动风险管理提供支撑. 主要结论如下:

- 1) 结合实际风险管理需要, 将所有事件结果分成 5 级, 在工程上可直观评估事件风险, 同时在算法上可解决数据不平衡问题;
- 2) 应用 CNN 网络建立民航事件风险评估模型, 专注于捕捉民航事件文本的语义信息, 准确率可达 96.3%, 通过与不同量化指标和模型的对比说明了优越性;
- 3) 在事件分类的基础上, 进一步拓展其功能, 开展风险预测, 通过 2020 年的报告进行评估预测, 新事件的风险等级基本可被准确评估.

未来随着事件分析维度的增加和事件数量的积累, 为更好地发挥模型通过事件学习的功能, 或可在原 CNN 模型的基础上通过预训练、正则化和集成等方法进一步提高模型性能.

## 参考文献

- [1] 中华人民共和国交通运输部令2018年第3号.民用航空安全管理规定[S].北京:中华人民共和国国务院公报,2018.  
Decree No. 3, 2018 of the ministry of transport, PRC Civil aviation safety management regulations [S]. Beijing: Bulletin of The State Council of the People's Republic of China, 2018 (In Chinese).
- [2] ALKHAMISI A O, MEHMOOD R. An ensemble machine and deep learning model for risk prediction in aviation systems [C]// 2020 6th Conference on Data Science and Machine Learning Applications (CDMA). Riyadh, Saudi Arabia: IEEE, 2020: 54-59.
- [3] ZHANG X G, MAHADEVAN S. Ensemble machine learning models for aviation incident risk prediction[J]. Decision Support Systems, 2019, 116: 48-63.
- [4] TANGUY L, TULECHKI N, URIELI A, *et al.* Natural language processing for aviation safety reports: from classification to interactive analysis[J]. Computers in Industry, 2016, 78: 80-95.
- [5] MINAE S, KALCHBRENNER N, CAMBRIA E, *et al.* Deep learning: based text classification [J]. ACM Computing Surveys, 2021, 54(3): 1-40.
- [6] 罗文慧, 蔡凤田, 吴初娜, 等. 基于文本挖掘的道路运输安全风险源辨识模型[J]. 西南交通大学学报, 2021, 56(1): 147-152.  
LUO W H, CAI F T, WU C N, *et al.* Text-mining based risk source identification model for transportation safety [J]. Journal of Southwest Jiaotong University, 2021, 56(1): 147-152. (In Chinese)
- [7] 兰秋军, 李卫康, 刘文星. 不同情境下中文文本分类模型的表现及选择[J]. 湖南大学学报(自然科学版), 2016, 43(4): 141-146.  
LAN Q J, LI W K, LIU W X. Performance and choice of Chinese text classification models in different situations [J]. Journal of Hunan University (Natural Sciences), 2016, 43(4): 141-146. (In Chinese).
- [8] GAO X, XU X B, LI D C. Accuracy analysis of triage recommendation based on CNN, RNN and RCNN models [C]// 2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers. Dalian, China: IEEE, 2021: 1323-1327.
- [9] NASA. ASRS program briefing [EB/OL]. US: Aviation Safety Reporting System, 2020 [2021-04-10]. <https://asrs.arc.nasa.gov/overview/summary.html>.
- [10] Air carrier certification: 14 CFR-121 [S]. USA: Federal Aviation Administration, 2020.
- [11] BUDA M, MAKI A, MAZUROWSKI M A. A systematic study of the class imbalance problem in convolutional neural networks [J]. Neural Networks, 2018, 106: 249-259.
- [12] MOHAMED C. Safety management system [M]. Cairo: International Civil Aviation Organization, 2018: 34-39.
- [13] 郎宪明, 李平, 曹江涛, 等. 基于非平衡数据处理的管道泄漏检测与定位研究[J]. 湖南大学学报(自然科学版), 2018, 45(2): 110-118.  
LANG X M, LI P, CAO J T, *et al.* Study on pipeline leak detection and location based on imbalance data processing [J]. Journal of Hunan University (Natural Sciences), 2018, 45(2): 110-118. (In Chinese).
- [14] ZHU X H, WANG J Z, HONG Z H, *et al.* Empirical studies of institutional federated learning for natural language processing [C]// Findings of the Association for Computational Linguistics: EMNLP 2020. Stroudsburg, USA: Association for Computational Linguistics, 2020: 625-634.
- [15] KIM Y. Convolutional neural networks for sentence classification [C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Stroudsburg, USA: Association for Computational Linguistics, 2014: 1746-1751.
- [16] FAA. Approved NOTAM contractions [EB/OL]. USA: Federal Aviation Administration, 2021 [2021-04-10]. <https://www.notams.faa.gov/downloads/contractions.pdf>