

DOI:10.3969/j.issn.1003-5060.2024.07.009

结合 Word2vec 和 BiLSTM 的 民航非计划事件分析方法

王捷¹, 周迪¹, 左洪福¹, 黄维²

(1. 南京航空航天大学 民航学院, 江苏 南京 211106; 2. 成都天奥电子股份有限公司, 四川 成都 611731)

摘要:安全是民航业的核心主题。针对目前民航非计划事件分析严重依赖专家经验及分析效率低下的问题,文章提出一种结合 Word2vec 和双向长短期记忆(bidirectional long short-term memory, BiLSTM)神经网络模型的民航非计划事件分析方法。首先采用 Word2vec 模型针对事件文本语料进行词向量训练,缩小空间向量维度;然后通过 BiLSTM 模型自动提取特征,获取事件文本的完整序列信息和上下文特征向量;最后采用 softmax 函数对民航非计划事件进行分类。实验结果表明,所提出的方法分类效果更好,能达到更优的准确率和 F_1 值,对不平衡数据样本同样具有较稳定的分类性能,证明了该方法在民航非计划事件分析上的适用性和有效性。

关键词:民航安全;文本分析;非计划事件;Word2vec;双向长短期记忆(BiLSTM)神经网络

中图分类号:TP391.1;V328 **文献标志码:**A **文章编号:**1003-5060(2024)07-0917-08

A civil aviation unplanned event analysis method combined with Word2vec and BiLSTM

WANG Jie¹, ZHOU Di¹, ZUO Hongfu¹, HUANG Wei²

(1. College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; 2. Chengdu Spaceon Electronics Co., Ltd., Chengdu 611731, China)

Abstract: Safety is the core theme of the civil aviation industry. Aiming at the problem that the analysis of civil aviation unplanned events heavily depends on expert experience and the low analysis efficiency, a civil aviation unplanned event analysis method is proposed, which combines Word2vec and bidirectional long short-term memory(BiLSTM) neural network model. Firstly, Word2vec is used to train word vectors for event text corpus, reducing the dimension of the space vector. Then, the features are automatically extracted by BiLSTM model to obtain the complete sequence information and context feature vector of the event text. Finally, the softmax function is used to classify civil aviation unplanned events. The experimental results show that the proposed method has better classification effect and can achieve better accuracy and F_1 value, and also has more stable classification performance for unbalanced data samples, which proves the applicability and effectiveness of this method in the analysis of civil aviation unplanned events.

Key words: civil aviation safety; text analysis; unplanned event; Word2vec; bidirectional long short-term memory(BiLSTM) neural network

收稿日期:2021-11-16;修回日期:2021-12-27

基金项目:国家自然科学基金联合基金资助项目(U1933202);民航大 NSF 重点基金资助项目(U1733201)

作者简介:王捷(1995—),女,回族,四川绵竹人,南京航空航天大学硕士生;

左洪福(1959—),男,湖南株洲人,博士,南京航空航天大学教授,博士生导师,通信作者,E-mail:rmstxzz@163.com.

0 引言

全球航空运输系统的复杂性和空中交通密度增加^[1],在有限的空域内,民机数量激增,系统复杂度提高,发生非计划事件的可能性增大。在民机实际运营过程中,时常伴随着多种非计划事件的发生,如由于使用环境或设计缺陷引起的非计划维修事件;鸟撞、动物入侵等随机事件等。尽管大多数与常规问题有关,但往往隐含着与安全隐患相关的信息。非计划事件数量众多、内容信息量大。目前对其处理的过程大多数是手动完成的,严重依赖于专家经验,造成效率低下、分析主观性大、结果不够客观等问题。因此采用自动化分析技术完成非计划事件文本的分析非常有必要。非计划事件的高效分析可以有效为民航工程人员执行维修、维护等任务提供准确可靠的信息,减少不必要的损失,缩小经济成本。

一般而言,提高民航安全性的策略在于“主动安全”方式,即提高运营商的准备。当非计划事件发生时,运营商能够及时采取适当的处置措施,降低非计划事件演变为危险事故的可能性。利用“主动安全”范式^[2]的重点在于加强民航工作人员应对非计划事件时,采取适当行为的能力和效率,从而降低非计划事件造成的不安全状况发展为重大安全事故的可能性。非计划事件通常由非结构化或半结构化的文本形式记录^[3],从事件文本数据中挖掘事件的描述属性、因果关系、处置方法等是一项艰巨的任务。随着人工智能的发展,将计算机与自然语言处理相结合处理文本数据成为了流行的趋势,因此本文引入自动化文本分析的概念,基于深度学习神经网络,自动提取事件特征并对非计划事件进行有效的分类和分析。

目前,国内外对民航事件信息的分析十分关注,在民航安全信息领域,文献[4]对设计和用于管理分析航空事故报告中的多种自然语言处理技术进行了调查;文献[5]使用视觉分析技术,处理航空安全报告系统数据库中的文本,针对事件叙述性文本进行处理时,文本表示和提取特征是事件文本分析的技术难点及热点;文献[6-7]指出使用工具 Word2vec 训练得到的向量低维连续,同时通过计算这些向量间的余弦距离可以判断词语之间的语义相似度;文献[8]利用 Word2vec 提取语义特征并基于支持向量机(support vector machine,SVM)进行文本分类,发现当样本数据量较大时,SVM 训练速度缓慢。

随着人工智能的发展,对于时间序列特征较突出的事件文本,循环神经网络(recurrent neural network,RNN)改进的序列建模模型长短时记忆网络(long short-term memory,LSTM)^[9]具有较强的记忆功能,有效避免了 RNN 梯度消失和爆炸问题^[10]。双向长短时记忆网络(bidirectional long short-term memory,BiLSTM)是 LSTM 的进一步改进^[11],结合前向和后向隐藏层单元,增强文本上下文之间的联系^[12],有效提取长距离特征,保留文本语义信息,自动选择特征进行分类。文献[13]集成 BiLSTM 和注意力机制,提出一种视频字幕框架,用于更好的全局特征表示,增强对视频中持久性信息的识别;文献[14]提出一种将 BiLSTM 和卷积神经网络(convolutional neural networks,CNN)与 K 均值方法相结合的神经反馈聚类算法;文献[15]采用加入注意力机制的双向 LSTM 模型抽取名词性实体间的关系,不依赖外部知识库或自然语言处理工具,捕获句子的语义信息;文献[16]对提出的 Att-BiLSTM+PI 方法进行改进,加入自注意力机制 self-Att-BiLSTM+PI,并将改进后的方法应用于航空安全事故隐式因果关系的抽取中,为航空安全事故的预测预警提供更好的语义服务。

从以上的研究发现,词嵌入模型和神经网络构造的分类器具有一定的优点。同时考虑到将 Word2vec 和 BiLSTM 相结合用于民航非计划事件分析的报道较少。因此本文提出了一种结合 Word2vec 和 BiLSTM 的非计划事件分析方法。从文本语义挖掘的角度,实现民航非计划事件的自动分类和分析,提高非计划事件分析的准确度和效率。快速辅助民航专家制定非计划事件处置策略,降低民航安全事故发生的可能性,保证民机运行安全水平。

本文的创新点概括如下:

1) 与传统方法相比,将关注点由民机运营监控数据转向至事件报告数据,从文本语义分析的角度,基于民机运营中产生的海量非计划事件文本数据展开研究。将民航非计划事件手动分析转换为自动化分析,缩短非计划事件分析时间,保证民机运行安全。

2) 结合 Word2vec+BiLSTM 方法,对民机非计划事件日志进行挖掘,得到可靠的事件分类结果。不依赖于专家经验,实现了民航非计划事件的快速定位。

3) 采用 BiLSTM 模型学习非计划事件的双

向长时依赖关系,有效提取了非计划事件的深层文本语义特征,提高民航非计划事件的分类性能。

1 民航非计划事件分析框架

结合 Word2vec 和 BiLSTM 的民航非计划事件分析总体流程框架如图 1 所示,主要包括数据预处理、文本向量化、非计划事件分类。

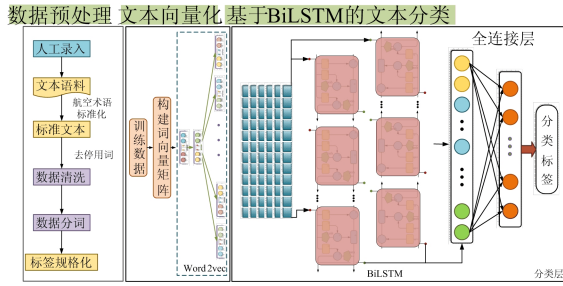


图 1 总体分析流程框架图

1.1 数据预处理

1.1.1 获取数据

在民航飞行过程中,民航工作人员通过航线移动端平台、维修控制系统将飞机重要系统的航线非计划事件传输给航空公司,由地面工作人员手动录入事件信息。在录入过程中,对语料中所有的英文文本进行小写化处理,同时对事件进行脱敏处理,保证信息安全。最终获得用于非计划事件分析的标准化事件文本语料数据。

1.1.2 文本预处理

文本预处理的目的在于去除不必要的噪声干扰,保证事件分析精度。在文本预处理步骤中,首先去除标点、停用词等无效文本,并对事件类别标签使用 LabelEncoder 函数进行映射操作,转换为数字标签;然后对事件文本进行分词操作,对于英文文本直接按照空格划分,对于中文文本使用 Jieba 分词系统的全精确模式进行分词处理。最后获得以空格划分的预处理后的文本语料。

1.2 文本向量化

本文采用谷歌开源工具 Word2vec 基于 Negative Sampling 算法的 Skip-gram 模型对非计划事件文本进行训练,并获得文本向量化表示。将文本语料的词语通过复杂映射到实数维向量空间上的低维词向量,使得文本特征得到最大化保留,尽量保证词向量之间包含多的语义和语法信息,输出用于 BiLSTM 模型捕获文本特征的词向量矩阵。

1.3 非计划事件分类

事件文本分类的作用是根据文本的特征表

示,训练出一个稳定的分类器,以实现不同航空运输协会 (Air Transport Association, ATA) 类别的事件达到较好的分类效果。本文采用 BiLSTM 用于事件特征抽取,构造 2 个 LSTM 网络结构从两相反的方向对事件文本进行提取,对事件序列信息开展非线性学习,捕捉文本的深层语义信息。利用 softmax 函数对非计划事件进行分类分析,得到事件分类结果。

2 相关技术

2.1 词嵌入

由于事件文本是自然语言的集合,计算机无法处理语言集合,词嵌入是将语言集合中细粒度的句子和词语映射为向量,表示文本语义信息。传统的离散表示方法存在无法体现词语间的语义联系、词向量之间维度过大且数据稀疏等问题。因此引入分布式词嵌入表示,目前较为流行的词嵌入模型之一是 Word2vec^[17],由 Mikolov 在 2013 年提出。Word2vec 具有 2 种结构,分别是 CBOW 和 Skip-gram^[18]。模型包括输入层、投影层和输出层,模型的目标在于以较小的计算量换取有效的词向量表示。Word2vec 模型结构如图 2 所示。

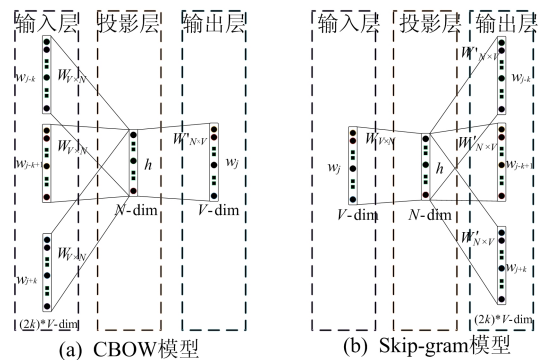


图 2 Word2vec 模型结构

CBOW 模型训练的输入是某特定词对应的上下文相关词的词向量,输出为特定词的词向量。Skip-gram 与之相反,即输入某特定词的词向量,输出其上下文相应的词向量。在训练准确度方面, Skip-gram 模型训练的词向量语义信息的表达更为准确,效果比 CBOW 模型更好。因此,本文采用 Skip-gram 模型进行非计划事件文本的词向量训练,同时由于非计划事件的描述较为简短,为避免因特殊词汇被剔除对实验结果造成影响,将所有词用于词向量训练中。

Skip-gram 模型的目标是最大化对数似然函数,采用随机梯度下降算法进行优化,训练完成后,矩阵 \mathbf{W}' 目标词对应的 $N \times V$ 的低维稠密词向量,其中 V 为词汇量大小。因此, Skip-gram 模型的本质是对输入词的输入向量和目标词的输出向量之间的余弦相似度进行 softmax 归一化处理。该模型的优化目标函数为:

$$L = \sum_{w_j \in V} \lg P(w_{j-k}, w_{j-k+1}, \dots, w_{j+k} | w_j) \quad (1)$$

其中, V 为 w_j 所在语料库。

在 Skip-gram 模型中,输入层只输入一个给定词 w_j ,以该词为中心词,设定窗口大小为 k ,得到上下文最有可能出现的 $2k$ 个单词的概率为:

$$P(w_{j-k}, w_{j-k+1}, \dots, w_{j+k} | w_j) = \frac{\exp(\mathbf{u}_j^T \mathbf{h})}{\sum_i \exp(\mathbf{u}_i^T \mathbf{h})} \quad (2)$$

其中: \mathbf{h} 为输入层矩阵的列向量,被称为中心词的输入向量; \mathbf{u}_j 为输出层矩阵的行向量,被称为 w_j 的输出向量; \mathbf{u}_j^T 为加权求和时使用的权值,因需要计算所有词的概率,因此, $j=1, 2, \dots, V$, 对应一系列的权值 $\mathbf{u}_1^T, \mathbf{u}_2^T, \dots, \mathbf{u}_V^T$ 。

2.2 双向长短时记忆神经网络模型

LSTM 神经网络模型^[19]是一种专门对时间序列型数据进行处理神经网络模型,由 Hochreiter 和 Schmidhuber 首次提出,是 RNN 的一种扩展。传统的 RNN 神经网络模型在处理序列数据时,将函数进行计算,得到输出单元,而 LSTM 将神经元转换为记忆单元再进行输出。每个单元包括输入门、遗忘门和输出门,其内部结构如图 3 所示。

LSTM 模型解决了传统 RNN 的梯度爆炸和梯度消失问题^[20],增加了对过去状态的过滤,从而选择对当前影响较大的状态^[21]。其中,输入门 i_t , 输出门 o_t , 遗忘门 f_t 和长期状态单元 c_t 、短期状态单元 h_t 各司其职,完成 LSTM 的训练任务。在 LSTM 单元里,长期状态单元用于存储长期记忆信息,保留文本序列信息。同时引入遗忘门对长期状态单元进行更新,舍弃过时信息。最后使用输出门对文本序列信息进行输出。单个 LSTM 单元在某一时刻 t 下的更新过程如下所示:

$$f_t = \sigma(\mathbf{W}_f[h_{t-1}, x_t] + \mathbf{b}_f) \quad (3)$$

$$i_t = \sigma(\mathbf{W}_i[h_{t-1}, x_t] + \mathbf{b}_i) \quad (4)$$

$$\tilde{c}_t = \tanh(\mathbf{W}_c[h_{t-1}, x_t] + \mathbf{b}_c) \quad (5)$$

$$c_t = f_t c_{t-1} + i_t \tilde{c}_t \quad (6)$$

$$o_t = \sigma(\mathbf{W}_o[h_{t-1}, x_t] + \mathbf{b}_o) \quad (7)$$

$$h_t = o_t \tanh(c_t) \quad (8)$$

其中: $\sigma(\cdot)$ 为 Sigmoid 激活函数; $\mathbf{W}_f, \mathbf{W}_i, \mathbf{W}_o, \mathbf{W}_c$ 分别表示遗忘门、输入门、输出门、记忆单元状态的权重矩阵, $\mathbf{b}_f, \mathbf{b}_i, \mathbf{b}_o, \mathbf{b}_c$ 分别表示遗忘门、输入门、输出门、记忆单元状态的偏置项。

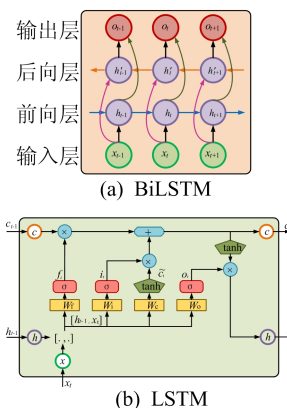


图 3 BiLSTM 模型和 LSTM 的结构

本文使用 BiLSTM 提取事件文本序列特征。与 LSTM 相比, BiLSTM 可以捕获前向和后向的上下文信息,更好地捕捉双向语义依赖,获得更全面的文本特征,增强分类性能。BiLSTM 的网络结构如图 3b 所示,其由 2 个独立的 LSTM 层组成,具有相同的输入数据,同时沿相反方向传输信息,分别计算出前向 LSTM 层和后向 LSTM 层的隐藏状态 h_t 和 h'_t 。然后将 2 个隐藏状态连接起来组成 BiLSTM 的输出。BiLSTM 的计算过程如下:

$$h_t = L_f(h_{t-1}, x_t) \quad (9)$$

$$h'_t = L_b(h_{t+1}', x_t) \quad (10)$$

$$o_t = \mathbf{W}_{ho} h_t + \mathbf{W}_{h'o} h'_t + \mathbf{b} \quad (11)$$

其中: L_f 和 L_b 表示 LSTM 前向传播和后向传播操作; \mathbf{W}_{ho} 和 $\mathbf{W}_{h'o}$ 代表前向 LSTM 和后向 LSTM 的权重矩阵; \mathbf{b} 表示输出层的偏置参数。

2.3 非计划事件分析方法

结合 Word2vec 和 BiLSTM 的非计划事件分析方法构造如图 4 所示,该分析方法整体结构由输入层、词嵌入层、BiLSTM 层、全连接层和分类层组成。每一层的输出作为下一层的输入,最终由该方法获得民航非计划事件的深度词向量特征进行非计划事件分类。

1) 输入层。将民航非计划事件已分好词的 S 条文本语料作为输入进行后续操作处理。

2) 词嵌入层。通过 Word2vec 对获取的非计划事件语料进行训练,并将词向量合并,得到

BiLSTM 的输入空间向量 $D=(x_1, x_2, \dots, x_S)$ 。

3) BiLSTM 层。使用 2 个 LSTM 层,分别捕获序列文本两个相反方向的语义特征,从整体上提取句子的长时依赖关系。利用式(3)~(11)计算得到事件文本的深层语义特征向量。

4) 全连接层。将上述 BiLSTM 获取的所有非计划事件文本特征汇聚到全连接层中,得到用于最终分类的深层语义词向量特征 M 。

5) 分类层。采用 softmax 函数对全连接层获取的最终非计划事件文本特征 M 进行 ATA 类别标签的判断,并在训练过程中采用反向传播算法对网络中的参数进行梯度更新^[22]。

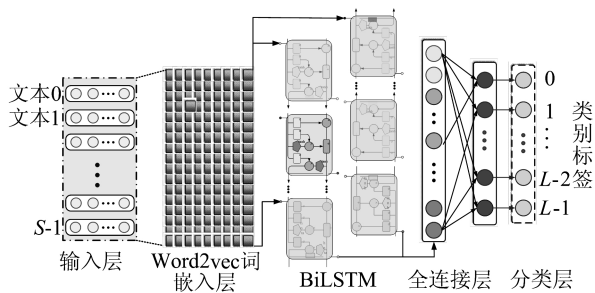


图 4 结合 Word2vec 和 BiLSTM 的非计划条件分析方法结构

3 实验与结果分析

3.1 实验数据集

本文的实验数据集是某航空公司的故障事件日志,该数据集包含 2015 年 1 月至 2019 年 12 月期间多条关于飞机典型系统及其重要部件的航线故障事件日志,共计 5 153 条原始事件日志文本,4 个 ATA 分类。由于故障日志事件数据存在不规范性,因此为保证数据质量,手动剔除故障日志中的无效文本,如故障描述为空、故障描述简写为航空公司特定代码等。

另外,对数据进行脱敏处理,保证事件信息安全,最终得到 4 901 条用于实验的数据文本,其中用于训练的事件数量共 3 920 条,用于测试的事件数量共 981 条。数据均为 UTF-8 文本格式。各 ATA 类别故障事件日志数据统计结果见表 1 所列。

表 1 故障日志统计数据

ATA	ATA 名称	原始	实验	训练	测试
A73	发动机燃油和控制	1 630	1 523	1 218	305
A74	点火装置	1 470	1 383	1 106	277
A36	气源	1 088	1 055	844	211
A75	发动机空气	965	940	752	188

3.2 评价指标

文本分类问题的性能评价指标主要有准确率、精确率、召回率及 F_1 值 4 个指标。对类别标签集合 L 中的每个类 l ,准确率 A 、精确率 P 、召回率 R 和 F_1 定义如下:

$$A_{cc} = \frac{T_P + T_N}{T_P + F_P + F_N + T_N},$$

$$P = \frac{T_P}{T_P + F_P}, R = \frac{T_P}{T_P + F_N}, F_1 = \frac{2PR}{P + R}$$
(12)

其中: T_P 为正确分类的事件数量; T_N 为正确分为负类的事件数量; F_P 表示错误分为负类的事件数量; F_N 表示错误分为正类的事件数量。

为描述多分类的整体性能,对以上指标进行平均。因考虑到文本使用的各 ATA 类别的非计划事件的数据比例相差较大,数据存在不平衡性。本文采用宏平均的优化性能指标权重 W_{F_1} 进行实验评价,具体计算如下。

$$W_{F_1} = \frac{2P_w R_w}{P_w + R_w},$$

$$P_w = \frac{\sum_{l=1}^L P_l w_l}{L}, R_w = \frac{\sum_{l=1}^L R_l w_l}{L}$$
(13)

3.3 参数设置

本文基于 Python 平台,调用 gensim、jieba 等程序包,开展了结合 Word2vec 和 BiLSTM 的民航非计划事件分析。

实验平台及环境配置如下:

操作系统为 Windows 10;CPU 为 Intel(R) Core(TM) i7-8565U;内存为 8 GiB;开发环境为 Python3. 5;深度学习框架为 Tensorflow1. 5. 0。Word2vec 词向量模型和 BiLSTM 模型参数设置见表 2 所列。

根据样本数据的平均序列长度,在实验中因采集的文本数据序列长度大部分集中于 110,设置每条事件文本的最大词数为 100,超过 100 的数据将被截断,不足 100 的用 0 补齐,固定 BiLSTM 模型的输入文本序列长度,构建文本序列矩阵。

本文使用 Sigmoid 作为激活函数,采用 Adam 优化器作为梯度更新方法,同时为防止过拟合,在训练阶段利用 Dropout 对网络参数进行约束^[23],并设置其为经典值 0. 5。由于事件数据较短和民航领域的特殊性,将词频阈值设置为 1,即训练全部的事件数据。对于词向量维度、迭代次数则通过实验确定最优值。

表 2 Word2vec 参数和 BiLSTM 模型设置

Word2vec 参数	设定值	BiLSTM 模型	设定值
词向量维度	200	词向量维度	200
窗口大小	5	LSTM 隐藏层大小	128
词频阈值	1	学习率	0.001
模型类别	1	Dropout	0.5
优化算法	0	训练批次	64

3.4 实验结果及分析

3.4.1 参数选择实验

为选定最优的词向量维度、确定最大迭代次数,本文采用单因子变量实验,确定非计划事件分析的最佳参数。参数选择的实验结果分别如图 5 和图 6 所示。

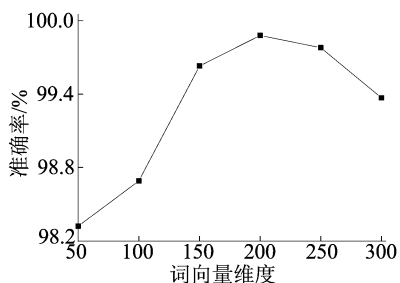


图 5 基于词向量维度的准确率变化

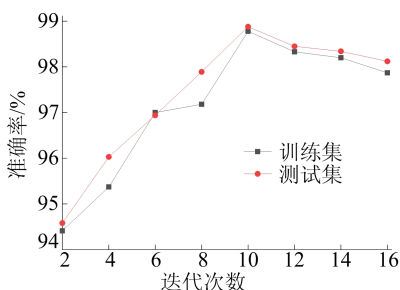


图 6 基于迭代次数的准确率变化

由图 5 可知,起初随着词向量维度增大,准确率不断提高,直到词向量维度达到 200 时,准确率最高。进而随着词向量维度增加,准确率逐渐降低。这是由于当词向量维度在一定范围内增加时,Word2vec 通过调整模型参数获得更好的拟合文本数据分布,从而精确地把握文本语义信息。当词向量维度过大将会导致参数过多,调整的参数初始值与真实值相差过大,模型在训练时间内很难将参数调整完全,从而导致训练的词向量不能获取文本信息。因此当词向量维度超过某个阈值后,模型的准确率呈下降趋势。因此本文设置词向量维度为 200。

迭代次数对于文本分类效果具有较大影响。

理论上迭代次数越多,模型对数据的学习越充分,拟合情况越好。但当迭代次数过多时,可能会导致过拟合,文本分类的准确率将会降低。由图 6 可知,当迭代次数为 10 时,事件分类的准确率最高。因此,本文设置迭代次数为 10。

3.4.2 不同文本表示方法的对比实验

本文通过一组实验研究不同文本表示对 BiLSTM 分类性能的影响。比较用 Tfidf 代替 W-BiLSTM 中的 Word2vec 词向量方法(用 Tfidf-BiLSTM 表示),其结果见表 3 所列。

表 3 不同文本表示方法的分类结果 %

方法	A	P	R	W_{F_1}
Tfidf-BiLSTM	98.45	98.68	98.53	98.72
W-BiLSTM	98.88	98.98	98.78	98.88

从表 3 可以看出,在民航非计划事件分类分析中,结合 Word2vec+BiLSTM 方法在各项指标上均高于 Tfidf+BiLSTM 方法。实验证明利用 Word2vec 训练能有效表示词语间的语义关系,提高非计划事件的分类精度。

3.4.3 不同分类方法的对比实验

为充分验证本文所提出的结合 Word2vec 和 BiLSTM 的民航非计划事件分析方法的有效性,增强对比性。将本文方法与结合 Word2vec 的机器学习模型 SVM(W-SVM)、神经网络模型 CNN(W-CNN)以及 LSTM(W-LSTM)进行建模对比,同时在民航非计划事件四分类的数据集上进行验证,每个实验重复 10 次以减少随机性,10 次实验的准确率如图 7 所示。

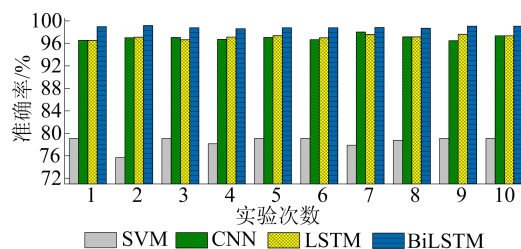


图 7 10 次实验准确率

计算出平均准确率和标准差,具体实验结果见表 4 所列。

由表 4 可知,深度学习模型方法比传统模型方法在各项指标上都有很大的提升,其主要原因是深度学习模型可以有效地从输入数据中自动学习基本特征,而 SVM 等传统模型在很大程度上

需要手动提取特征,复杂的特征选择和降维操作需要耗费大量的时间成本。

表 4 不同分类方法的分类结果 %

方法	准确率		P	R	W_{F_1}
	平均值	标准值			
W-SVM	79.10	1.08	68.27	79.10	72.11
W-CNN	97.04	0.45	97.44	96.94	97.05
W-LSTM	97.15	0.28	97.17	96.94	97.16
W-BiLSTM	98.88	0.18	98.98	98.78	98.88

本文所提出的 W-BiLSTM 分析方法在民航非计划事件数据集上均比其他方法获得了更好的结果。W-BiLSTM 在非计划事件测试集上的平均准确率为 98.88%。与 W-SVM、W-CNN、W-LSTM 相比,W-BiLSTM 的准确率提高了 19.78%、1.84%、1.73%。由上可以得出神经网络模型之间的准确率相差较小,数值变化不大,其原因是本文收集的事件数据大部分为人工简化处理的短文本序列,各方法提取的特征较为相似,因此得到的准确率结果较接近,符合工程实际情况。与 W-CNN 相比,W-BiLSTM 更擅长处理文本序列问题,对上下文序列特征更加敏感,特征提取能力强。与 W-LSTM 相比,W-BiLSTM 可以提取事件文本前后的双向信息,比单向传播提取的文本特征更加有效,因此在民航非计划事件的分类结果上表现更优。

考虑到样本的不平衡性,引入权重 F_1 值综合评估所提出方法的分类性能,W-BiLSTM 的 W_{F_1} 为 98.88%,相比其他 3 种方法,分别提升了 26.77%、1.83%、1.72%。以上验证结果均证明了本文方法在民航非计划事件分析上的有效性和优越性。

不同模型的 F_1 值如图 8 所示。从图 8 可以看出,在各个类别下,W-BiLSTM 模型的 F_1 值高于其他模型,说明所提出的 W-BiLSTM 方法在总体上具有相对稳定的分类性能,几乎不受不平衡数据的影响,在民航非计划事件分析上优势突出,具有良好的实用性。

不同分类模型在航空故障事件日志分类任务中的混淆矩阵如图 9 所示,混淆矩阵用于反映每类标签具体的分类信息。混淆矩阵中每一行代表真实标签,每一列代表预测标签,对角线的数字代表每个标签中被正确分类的样本数。从图 9 可以看出,BiLSTM 中正确分类的样本数量要比其他模型中的样本数量多,证明该模型在民航非计划

事件分析中的有效性和鲁棒性,有利于进一步提升民航非计划事件分类的精度和准确率,适用于民航非计划事件分析任务。

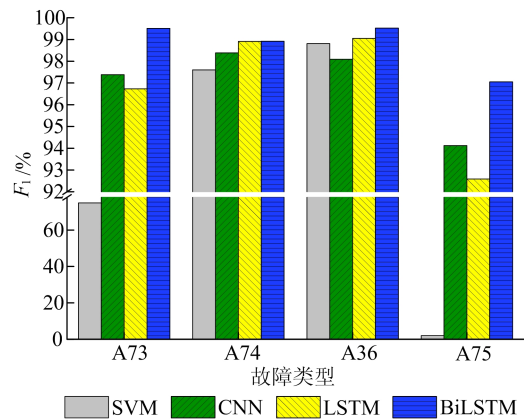


图 8 不同模型的 F_1 值

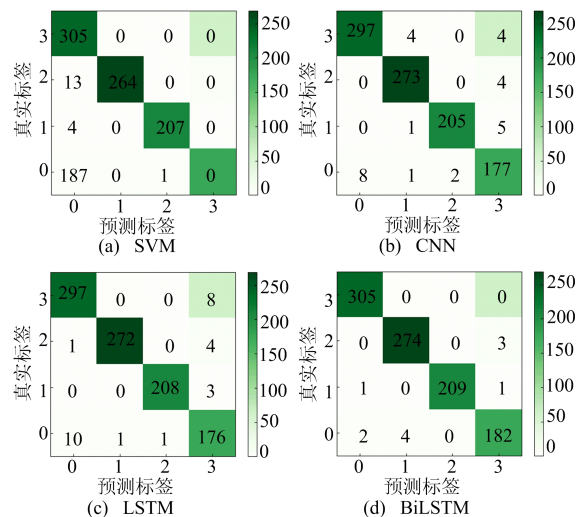


图 9 不同模型的混淆矩阵

4 结 论

针对民航非计划事件分析费时且效率低下的问题,本文提出了一种结合 Word2vec 和 BiLSTM 的民航非计划事件分析方法。首先对事件文本进行预处理,规范化表示事件;然后针对事件文本特征稀疏的特点,利用 Word2vec 降低词向量维度,获得文本向量表示;最后为充分挖掘事件局语义特征、缓解梯度爆炸和消失问题,采用对序列化数据具有优势的 BiLSTM 模型获取整个事件文本特征并进行分类。

将本文所提出的方法与其他方法进行对比,实验结果表明,本文所提出的 Word2vec+BiLSTM 方法增强了语义理解并提高了民航非计划事件的分类准确率和 F_1 值,同时对不平衡数据

也展示出良好的分类性能。Word2vec 比 Tf-idf 词表示方法更优。BiLSTM 处理前后向信息的效果比 CNN、LSTM 更好,解决了 CNN 和 LSTM 提取特征不全的问题。另外还与传统 SVM 模型相比,本文提出的方法不用人为设计和提取特征,具备自动化学习和提取特征的能力。

[参 考 文 献]

- [1] NG K, LEE C, CHAN F. A robust optimisation approach to the aircraft sequencing and scheduling problem with runway configuration planning[C]//IEEE International Conference on Industrial Engineering & Engineering Management. [S. l.]: IEEE, 2017: 40-44.
- [2] ZHANG X, MAHADEVAN S. Ensemble machine learning models for aviation incident risk prediction [J]. Decision Support Systems, 2019, 116(1): 48-63.
- [3] BILAL M, OYEDELE L O, QADIR J, et al. Big data in the construction industry: a review of present status, opportunities, and future trends[J]. Advanced Engineering Informatics, 2016, 30(3): 500-521.
- [4] LUDOVIC, TANGUY, NIKOLA, et al. Natural language processing for aviation safety reports: from classification to interactive analysis-science direct[J]. Computers in Industry, 2016, 78: 80-95.
- [5] ROBINSON D S. Visual representation of safety narratives [J]. Safety Science, 2016, 88: 123-128.
- [6] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space[EB/OL]. (2013-09-07). <https://arxiv.org/abs/1301.3781v1>.
- [7] 赵明, 杜亚茹, 杜会芳, 等. 植物领域知识图谱构建中本体非分类关系提取方法 [J]. 农业机械学报, 2016, 47(9): 278-284.
- [8] LILLEBERG J, YUN Z, ZHANG Y. Support vector machines and word2vec for text classification with semantic features[C]//2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing. [S. l.]: IEEE, 2015: 136-140.
- [9] WANG J, PENG B, ZHANG X. Using a stacked residual LSTM model for sentiment intensity prediction[J]. Neurocomputing, 2018, 322: 93-101.
- [10] GREFF K, SRIVASTAVA R K, KOUTNIK J, et al. LSTM: a search space odyssey[J]. IEEE Transactions on Neural Networks and Learning Systems, 2016, 28(10): 2222-2232.
- [11] CHEN T, XU R, HE Y, et al. Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN [J]. Expert Systems with Applications, 2017, 72: 221-230.
- [12] LIU G, GUO J. Bidirectional LSTM with attention mechanism and convolutional layer for text classification [J]. Neurocomputing, 2019, 337: 325-338.
- [13] BIN Y, YANG Y, SHEN F, et al. Describing video with attention-based bidirectional LSTM[J]. IEEE Transactions on Cybernetics, 2018, 49(7): 2631-2641.
- [14] YANG F, LIU G S, MENG K, et al. Neural feedback text clustering with BiLSTM-CNN-Kmeans[J]. IEEE Access, 2018, 6: 57460-57469.
- [15] ZHOU P, SHI W, TIAN J, et al. Attention-based bidirectional long short-term memory networks for relation classification[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. [S. l. : s. n.], 2016: 207-212.
- [16] 王红, 祝寒, 林海舟. 航空安全事故因果关系抽取方法的研究[J]. 计算机工程与应用, 2020, 56(11): 265-270.
- [17] MIKOLOV T, SUTSKEVER I, CHEN K, et al. Distributed representations of words and phrases and their compositionality[C]//Advances in Neural Information Processing Systems. [S. l. : s. n.], 2013: 3111-3119.
- [18] JATNIKA D, BIJAKSANA M A, SURYANI A A. Word2Vec model analysis for semantic similarities in English words [J]. Procedia Computer Science, 2019, 157: 160-167.
- [19] 孙运森, 林锋周. 激流长短时记忆网络在移动场景中的应用研究进展[J]. 现代计算机(专业版), 2017(35): 10-15.
- [20] 梁军, 柴玉梅, 原慧斌, 等. 基于极性转移和 LSTM 递归网络的情感分析[J]. 中文信息学报, 2015, 29(5): 152-159.
- [21] 王玲. 基于 Word2Vec 词嵌入和双向长短时记忆网络的文本分类实现[J]. 电子技术与软件工程, 2020(15): 70-71.
- [22] 黄贤英, 刘广峰, 刘小洋, 等. 基于 word2vec 和双向 LSTM 的情感分类深度模型[J]. 计算机应用研究, 2019, 36(12): 3583-3587, 3596.
- [23] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors[J]. Computer Science, 2012, 3(4): 212-223.

(责任编辑 吴 亮)