

基于 TF-IDF 进化集成分类器的铁路安全故障文本分类*

高凡¹, 王富章¹, 张铭¹, 赵俊华², 李高科¹

(1. 中国铁道科学研究院, 北京 100081; 2. 北京经纬信息技术有限公司, 北京 100081)

摘要: 铁路安全问题是铁路运输保证的核心问题, 铁路安全问题非结构化文本数据量大, 文本内容无特定规律, 对于综合分析解决安全故障问题造成很高的难度。针对铁路安全问题数据智能分类, 提出进化集成分类器模型。运用 TF-IDF 算法, 通过分析接触网安全问题的数据特点提取文本特征。采用决策树作为基分类器的 Bagging 集成分类器将文本数据分类, 在 Bagging 分类过程中, 针对 Bagging 算法产生的基分类器组合解集, 提出采用遗传算法 (Genetic Algorithm) 将其优化, 产生分类结果较优的基分类器组合解集。以铁路局实际安全问题进行实验分析, 实验证明 TF-IDF+Bagging+Genetic Algorithm=Evolutionary Ensemble Classifier 进化集成分类器模型在铁路安全问题文本分类中具有较高的准确性。

关键词: 铁路安全问题; TF-IDF; 基分类器; 集成分类器; 进化集成分类器

中图分类号: TN02; U284

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.200284

中文引用格式: 高凡, 王富章, 张铭, 等. 基于 TF-IDF 进化集成分类器的铁路安全故障文本分类[J]. 电子技术应用, 2021, 47(4): 71-76, 81.

英文引用格式: Gao Fan, Wang Fuzhang, Zhang Ming, et al. Text classification of railway safety fault based on TF-IDF evolutionary integrated classifier[J]. Application of Electronic Technique, 2021, 47(4): 71-76, 81.

Text classification of railway safety fault based on TF-IDF evolutionary integrated classifier

Gao Fan¹, Wang Fuzhang¹, Zhang Ming¹, Zhao Junhua², Li Gaoke¹

(1. China Academy of Railway Science, Beijing 100081, China;

2. Beijing Jingwei Information Technologies Co., Ltd., Beijing 100081, China)

Abstract: Railway safety is the core of railway transportation guarantee. The unstructured text data of railway safety problems is large, and the content of the text has no specific rules, which makes it very difficult to comprehensively analyze and solve the safety problems. Aiming at the intelligent classification of railway safety data, an evolutionary ensemble classifier model is proposed. By analyzing the characteristics of the catenary security issues of data, TF-IDF model is adopted to realize the feature extraction. Bagging ensemble classifier which uses Decision Tree as the base classifier classifies the text data, in the process of classification of Bagging, for the combined solution set of base classifier generated by Bagging Algorithm, Genetic Algorithm is proposed to optimize it to generate the combined solution set of base classifier with better classification results. Based on the safety problem of power supply contact network of a railway bureau, the experimental analysis shows that the TF-IDF+Bagging+Genetic Algorithm=Evolutionary Ensemble Classifier model has a high classification index in the text classification of railway safety problems.

Key words: software railway safety problems; TF-IDF; base classifier; integrated classifier; evolutionary integration classifier

0 引言

安全问题历来是铁路运输的重点关注问题, 在铁路安全监控体系中, 为减少运营过程中可能存在的安全隐患, 减少设备故障, 保障系统稳定运行, 安全人员根据日常巡检、专项巡检、重大节假日排查等计划定期到现场检查存在的安全生产故障隐患, 并将问题记录为文本形

式存储。管理人员依据这些问题记录, 分析、总结与整改问题。在长期形成的海量的安全问题文本文件中, 采用文本分类技术将安全问题自动归类, 辅助管理人员更好地掌握与处理安全问题, 在铁路安全领域的重要性日渐凸显。

一般来说, 文本分类过程有以下 3 个步骤^[1]: 预处

* 基金项目: 中国铁道科学研究院集团有限公司科技研究开发计划项目 (2051DZG001, 2019YJ115, 2019YJ125); 中国国家铁路集团有限公司科研专项课题 (J2019X005); 中国铁路总公司科技研究开发计划 (P2018G050); 中国铁道科学研究院集团有限公司电子计算技术研究所产品开发课题 (2052DZ1201)

理、文本表示和文本分类。其中,文本预处理包括分词、去除停顿词、去除不规则数据等。文本表示即将提取的文本特征转换为计算机可处理的数据类型(通常用向量表示)。文本表示方法包括词袋模型(BOW)、词频-逆向文件频率(Term Frequency-Inverse Document Frequency, TF-IDF),以及基于深度学习的 Word2vec 表示方法等^[2]。在构造分类器时使用的技术大体包括单个的基分类器、集成分类器以及深度学习模型^[3]。集成分类器主要包含 Bagging 和 Boosting 两种, Bagging 的主要算法是随机森林, Boosting 的主要算法是梯度提升树。在深度学习领域,有基于深度信念网络、卷积神经网络^[4-5]、循环神经网络等完成文本分类任务^[6]。

目前,国内学者在铁路文本分析领域研究较多。王广采用贝叶斯网络的预测模型和自适应增强算法(Adaptive Boosting, AdaBoost)的预测模型实现天气相关的铁路道岔故障预测^[7];赵阳等以故障文本信息为依据,针对高铁信号车载设备,提出贝叶斯结构学习算法(HDBN_SL)^[8];李佳奇等将面向 Agent 的分布式人工智能技术引入到信号设备故障诊断系统中^[9];杨连报针对信号故障不平衡数据,采用 SVM-SMOTE 算法对 TF-IDF 转换后的小类别文本向量数据随机生成,分别采用基分类器和集成分类对数据进行分类^[10]。本文处理安全问题为事故故障发生前人员检查时发现的危险、隐患内容,通常包括安全问题发生的时间、地点、问题描述等关键

要素^[11],所以针对每一类安全问题,存在特定的关键词,本文在文本特征抽取上采用 TF-IDF 表示,针对 Bagging 集成分类器的基础上,采用遗传算法优化,提升分类准确性^[12]。

1 进化集成分类器模型

安全问题文本分类属于文本挖掘中的分类问题,应用文本分析的处理流程,进化集成分类器模型分为 3 层:数据预处理层、进化优化层以及智能决策层^[13],如图 1 所示。基于 TF-IDF,数据预处理层对安全问题描述文本进行特征提取和特征向量化表示,通过基分类器决策树(Decision Tree)对上述文本进行分类并建立相应的基模型,再基于 Bagging 集成分类器对预处理结果进行采样训练,产生集成分类器分类模型^[14]。在数据预处理层 Bagging 训练过程中产生的每个基模型的解的基础上,进化优化层使用遗传算法优化集成分类器并最终得到具有更高分类准确度的基分类器组合。最上层智能决策层为用事先定义好的指标评价模型,经调整和改进行后的模型直接用于文本自动分类^[15]。

2 TF-IDF 法数据处理

安全问题中存在的大量铁路专用术语,如低压过流、闭锁、红光带等,本文采用基于字典的结巴分词工具,构建了包括通用语料及专业语料的铁路安全问题语料库,以便对安全问题准确分词。

安全问题以文本形式记录在信息系统中。表 1 列举

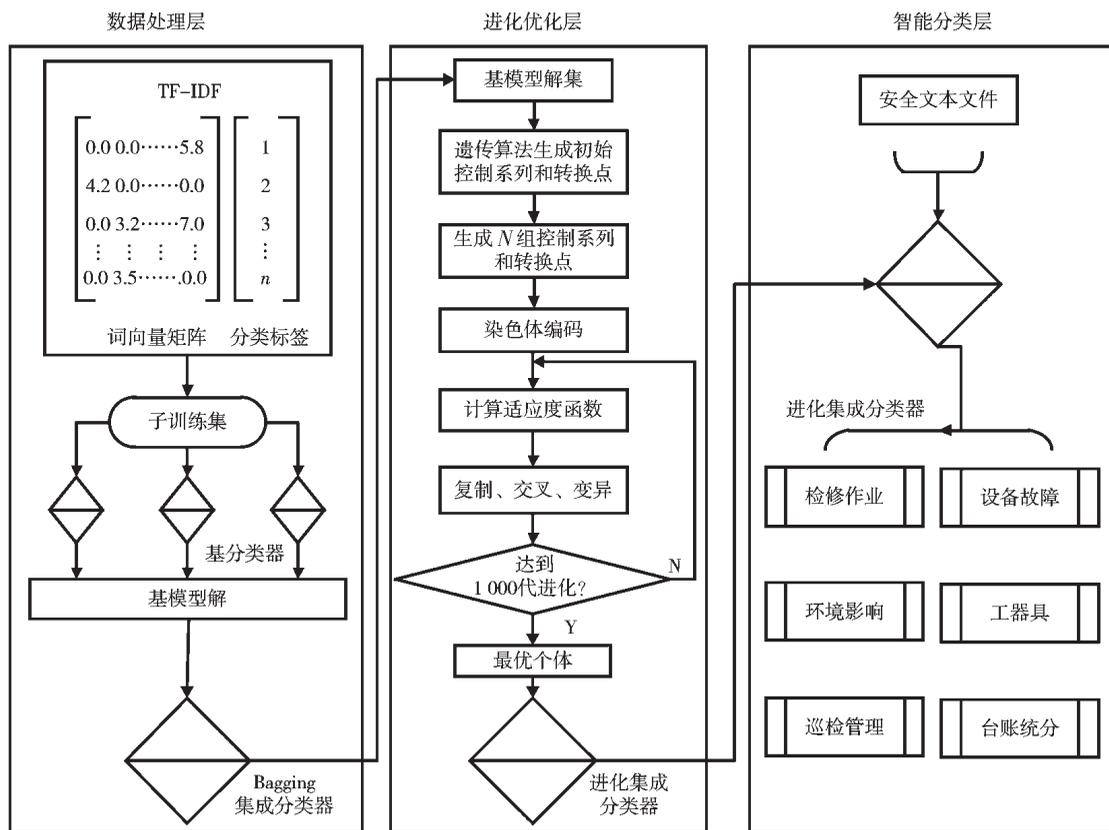


图 1 进化集成分类器模型整体框架

表1 铁路信号安全问题表述

序号	安全问题描述	类别
1	2018年5月28日2时28分,衡北出发场-大浦街上行天窗作业发现梯车1组高空作业人员郑海洋在进行分相检修时未对分相相关节进行绝缘距离测量,未执行接触网检修工艺	检修作业
2	2017年4月27日5:40参加韶关集中修跟班作业,发现1211084轨道车驾驶室外堆放有工具材料,其中球头挂环、绝缘子、大绳、水桶未采取固定措施,存在掉落线路影响行车安全风险	工器具
3	2016年12月28日13:10在军田站检查军田站分相供电改造施工情况,发现所改造供电电线的绝缘子绝缘爬距不足,影响设备供电安全。工区配合施工安全监管不力,未及时发现并制止施工单位违规施工	设备故障
4	2017年5月24日11时在河冬间检查龙怀高速公路跨京广线施工现场时发现,施工现场贯通线架空改电缆后,电缆两侧的电缆终端杆上均没有悬挂“高压危险、禁止攀登”的安全警示标识牌,存在路人触电伤害安全隐患	环境影响
5	6月28日在坪石接触网工区检查防洪重点地段巡视记录本,5月份工区没有按供电段广供电段调发[2017]29号文件要求,工区每周巡视检查记录不少于一次	巡检管理
6	5月31日9:20至10点巡视检查原高桥站施工点,发现配合人员未及时填写施工安全监督员写实本及在工作票中签名	台账统计

了铁路领域各类别安全问题的典型实例,根据对安全问

题的表述,可见虽然铁路安全问题类别较多,但在各类别中都有各自特征的高频词,同时在其他安全问题类别中呈低频出现,如图2所示。

在本文的模型中,应用TF-IDF将经过分词的词语过滤低频词后再计算词频。

词频(Term Frequency, TF)和逆向文件频率(Inverse Document Frequency, IDF)均表示搜索关键词在所有文档中出现的频率,前者频率越高,证明该关键词与文档相关性越高;后者频率越高,表示该关键词与文档相关性越低。所以,为了对文档中某一词语的出现次数进行归一化处理,可以通过分析该词语的词频。对于词语 t_i ,在某个文档中的重要程度可表示为:

$$TF_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (1)$$

式中, $TF_{i,j}$ 为词频, $n_{i,j}$ 是词语 t_i 在 d_j 文件中检索到的次数, $\sum_k n_{k,j}$ 是 d_j 中所有字、词出现的次数之和。

逆向文件频率由总文件数除以包含某词语的文件数取对数得到,用于度量一个词语普遍重要程度,其公式如下:

$$IDF_i = \log \frac{|D|}{|j:t_i \in d_j|} \quad (2)$$

式中, IDF_i 为逆向文件频率, $|D|$ 为总文件数, $|j:t_i \in d_j|$ 为包含某一词语的文件数。

高权重 $w_{i,j} = TF_{i,j} \times IDF_i$,由文档内的高词语频率乘以该词语在整个文档集合的低文件频率。

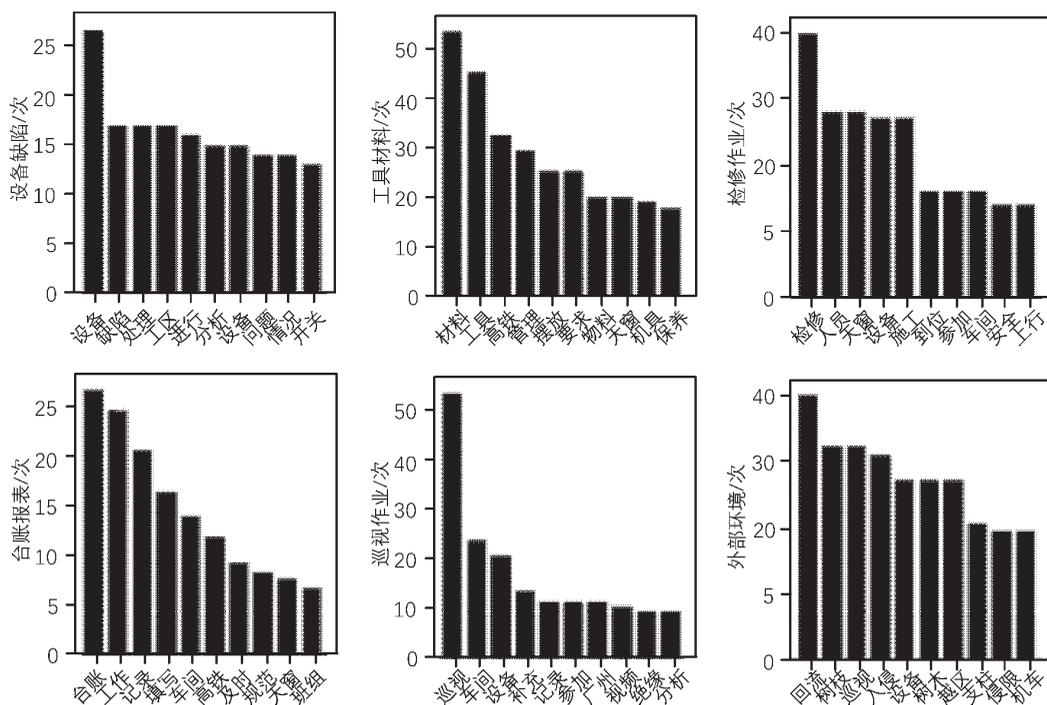


图2 安全问题内容特征分析

数据预处理流程如图3所示。首先在既有语料库的基础上,根据不同的安全问题进行分词,同时统计特征词语所在的文档及其出现次数,然后通过TF-IDF方法计算该特征词语的词频并过滤低频词,最终形成矩阵X和矩阵Y,其中矩阵X代表文档特征的权重,矩阵Y代表分类类别。

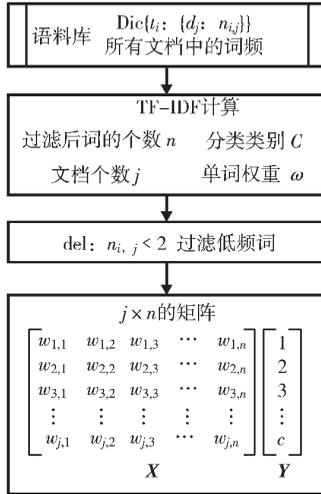


图3 数据预处理流程

3 基于遗传算法的进化集成分类器

3.1 决策树基分类器

决策树分类器是训练一棵大规模的树结构,对该树进行减枝,直到达到合适的规模和分类效率。本文采用ID3算法,ID3依据信息增益选择属性。其中属性TF-IDF过滤后的单词,若属性w的值将样本集T划分成T₁,T₂,...,T_m,共m个子集,其信息增益表示为:

$$gain(w) = info(T) - \sum_{i=1}^m \frac{|T_i|}{|T|} \times info(T_i) \quad (3)$$

式中,|T|为T的样本个数,|T_i|为子集T_i的样本个数。其中info(T)的计算公式为:

$$info(T) = - \sum_{j=1}^s freq(C_j, T) \times \log_2(freq(C_j, T)) \quad (4)$$

式中,freq(C_j, T)为T中的样本属于C_j类别的频率,s是T中的样本的类别数量。

3.2 Bagging 集成分类器

本文采用决策树作为基分类器。假设原始训练数据量是n。在原始数据集随机抽取训练数据,抽取的数据量为n'(n' ≤ n),用抽取的数据训练第一个决策树,用同样的方法做m次有放回抽取,训练m个决策树,将测试集放到每一个决策树中预测,最后通过最优分类器与其他分类器的加权投票组合,选出表现性能最优的组合集成分类器,其学习过程如下:

(1)输入训练集

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\};$$

(2)过程:

for t=1, 2, ..., T do

$$h_t = \xi(D, D_{ts})$$

end for

其中,ξ为基分类器算法;T为训练轮数。

(3)输出:

$$H(x) = \arg \max_{y \in Y} \sum_{t=1}^T \prod (h_t(x) = y)$$

3.3 进化集成分类器

遗传算法是一种自适应全局优化的搜索算法,通过对初始解进行遗传迭代,每一次迭代通过适应度函数对当前解进行优劣进行评价,在评价的基础上产生优秀的下一代。进化集成分类器是建立在给定的某个Bagging集成分类器上,通过遗传算法优化该集成分类器并最终得到的具有更高分类准确度的基分类器组合。进化集成分类器的核心主要为基因编码及种群繁殖过程。

3.3.1 基因编码机制

进化集成分类器搜索最优解的目标为得到基分类器的最优组合。根据这一原则,本文设定50个基分类器,种群(个体)规模为25,为了模拟基分类器是否被选择,设定个体大小(基因长度)与基分类器数量一致,采用0、1编码基因,0表示基分类器未选择,1表示被选择。所以初始个体基因由0、1随机编码长度为50的二进制符号串组成。

3.3.2 种群繁殖过程

种群繁殖过程是将种群内个体不断地经过选择、交叉、变异的操作,把适应度较高的个体遗传到下一代,通过这一迭代过程最终得到一个优良个体的过程。可以认为,这一最终得到个体的表现型是最接近该问题的最优解。本文中迭代次数G=100。

(1)初始适应度计算:初始适应度计算以每个基分类的准确度为准,准确度越高的其适应度越高。p_i(0 < i < 50)代表每个基分类器(决策树)的准确度,̄p代表全部基分类器的均值,σ_p代表全部基分类器的方差,个体适应度公式为:

$$\frac{p_i}{(p_i + (p_i - \bar{p}) \times 10^3) \times ((p_i - \bar{p}) / \sigma_p)^2} \quad (5)$$

(2)选择:选择过程即根据个体适应度,选择优秀的个体遗传到下一代。本文以个体灵敏度随机选择要交叉的两个个体,个体初始灵敏度p_s=0.95,个体灵敏度计算过程如下:

$$\textcircled{1} s = \min((p_s + c \times 0.010), p_s)$$

$$\textcircled{2} \text{for } i=1, 2, \dots, n \text{ do}$$

$$\textcircled{3} p_s = ((s-1)/(s \times \text{len}(a))^2 - (1)) \times ((s \times \text{len}(a))^2 - i)$$

其中,c反映经过迭代后其准确度是否有提升,如果有提升加1。经过计算,准确度越高的个体其灵敏度就越低,越不容易被选择变异。a为个体总数。

通过上述计算,本文将适应度最大的精英个体保留

下来。

(3)交叉:对于种群中的所有个体,以随机原则将其两两搭配成对,对于每一对个体组合,设定一定的概率来交换两个个体间的部分染色体,本文设定交叉概率 $P_c=0.75$ 。

(4)变异:针对个别满足变异的个体,进行随机变化。变异过程可产生新的个体,不仅增加了种群的多样性,也为寻找最优解提供了潜在的可能,本为的变异概率 $P_m=0.05$ 。

(5)评价函数:经过变异后的种群用评价函数来评价个体的好坏,每个个体为一个基分类器组合(0为选择,1为不选择),通过评价函数式(6)计算该基分类器组合的准确度,从而得到最优解,按照准确度从大到小的原则筛选出准确度最大的个体,并将其作为精英个体保留下来。

$$H(x)=\arg \max \sum_{t=1}^T \prod_{y \in Y} (h_t(x)=y) \quad (6)$$

本文设置的终止条件为总的进化代数 ng 超过设定代数 G 。

4 实验结果与分析

为分析本文提出的进化集成分类器进化过程,以某铁路局供电接触网安全问题为实例,采用精准率(Precision)、召回率(Recall)和精确值和召回率的调和均值(F-score)作为模型评价和对比的指标验证分类结果。实现数据包括6种问题类别2400条数据。

精准率(Precision):

$$\text{Precision}=\text{TP}/(\text{TP}+\text{FP}) \quad (7)$$

召回率(Recall):

$$\text{Recall}=\text{TP}/(\text{TP}+\text{FN}) \quad (8)$$

F-score:

$$\text{F-score}=\frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision}+\text{Recall}} \quad (9)$$

式中,TP代表被正确检索到的样本,FP代表被误检索到的样本,FN代表属于此类而被误检索为其他类的样本。

图4中显示了模型进化过程中种群的多样性(Diversity)、种群分类的平均值(AverageScore)、方差(StandardScore)以及最优值(BestScore)。由实验结果可见,当种群迭代次数达到45次时,种群繁殖过程中的平均值、方差和分类最大值均达到最大值,种群多样性保持在80%~100%之间。

为了对比不同类型分类器对同一安全问题的分类结果,本文在以TF-IDF方法表示安全问题文本的基础上,分别采用基分类器、Bagging集成分类器及进化集成分类器对同一安全问题进行分类,最终结果如表2所示。从表2中可以看出,进化集成分类器分类指标最高,Bagging集成分类器分类指标次之,基分类器分类指标最差。

表2 分类模型结果对比

分类器	精准率	召回率	F1
基分类器	0.736	0.683	0.694
集成分类器	0.815	0.882	0.854
进化集成分类器	0.874	0.878	0.871

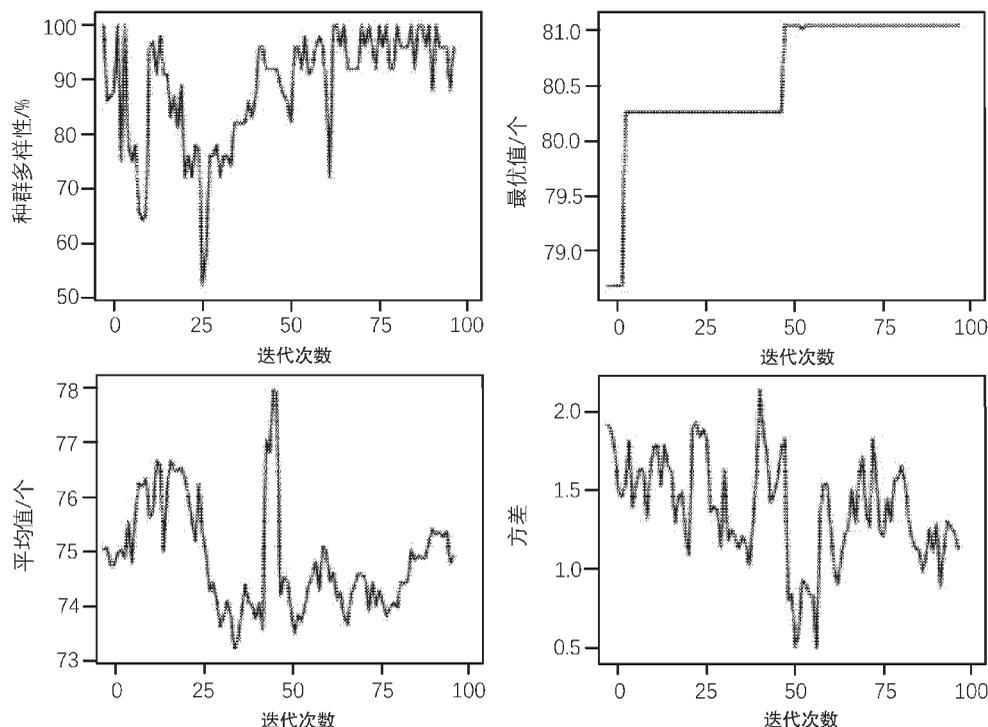


图4 种群繁殖过程

5 结论

本文采用 TF-IDF 方法表示安全问题文本,通过遗传算法优化集成分类器并最终得到具有更高分类准确度的基分类器组合,提出了 Evolutionary Ensemble Classifier 文本分类模型,应用某路局安全接触网问题进行验证,实验证明 Evolutionary Ensemble Classifier 模型在安全接触网问题具有较高的准确性,也可以为处理铁路其他文本分类问题提供参考模型。

参考文献

- [1] 康东.中文文本挖掘基本理论与应用[D].苏州:苏州大学,2014.
- [2] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space[J]. Computer Science, arXiv: 1301.3781[CS.CL], 2013.
- [3] BREIMAN L. Random forests[J]. Machine Learning, 2001, 45(1): 5-32.
- [4] 张笑铭, 王志君, 梁利平. 一种适用于卷积神经网络的 Stacking 算法[J]. 计算机工程, 2018, 44(4): 243-247.
- [5] 陶宏曜, 梁栋屹. 基于卷积神经网络的职位描述文本分类方法[J]. 电子技术应用, 2017, 38(6): 30-34.
- [6] 王伟, 孙玉霞, 齐庆杰, 等. 基于 BiGRU-Attention 神经网络的文本情感分类模型[J]. 计算机应用研究, 2019, 36(12): 3558-3564.
- [7] 王广. 基于文本数据的铁路道岔故障预测[D]. 北京: 北京交通大学, 2017.
- [8] 赵阳, 徐田华. 基于文本挖掘的高铁信号系统车载设备故障诊断[J]. 铁道学报, 2015(8): 53-59.
- [9] 李佳奇, 党建武. 基于 MAS 电务故障诊断模型的研究[J]. 铁道学报, 2013, 35(2): 72-80.
- [10] 杨连报, 李平, 薛蕊, 等. 基于不平衡文本数据挖掘的铁路信号设备故障智能分类[J]. 铁道学报, 2018(2): 59-66.
- [11] 张磊, 王喆. 基于铁路安全管理信息报告的文本挖掘技术研究[J]. 铁路计算机应用, 2018, 27(8): 9-12.
- [12] 任勉, 甘刚. 基于双向 LSTM 模型的文本情感分类[J]. 计算机工程与设计, 2018, 39(7): 272-276.
- [13] 朱菲, 金炜东. 一种 SVM-DS 决策融合方法及在高铁故障中的应用[J]. 华中科技大学, 2018, 44(7): 127-130, 134.

(下转第 81 页)

(上接第 70 页)

- [5] FOROUZANFAR M, FEGHHI R, BASERI J, et al. High Efficiency 8.8-9.6 GHz class J power amplifier[C]. Proceedings of the 16th Mediterranean Microwave Symposium, 2016.
- [6] LEE Y H, KIM M W, KIM C W. A CMOS band-pass delta sigma modulator and power amplifier for class-S amplifier applications[J]. The Journal of Korean Institute of Communications and Information Sciences, 2015, 40(1): 9-15.
- [7] NIKANDISH G, BABAKRPUR E, MEDI A. A harmonic termination technique for single-band and multi-band high-efficiency class-F MMIC power amplifiers[J]. IEEE Transactions on Microwave Theory & Techniques, 2014, 62(5): 1212-1220.
- [8] MOHADESKASAEI S A, LIN F, ZHOU X, et al. Novel design theory for high-efficiency and high-linearity microwave power amplifier based on 2nd harmonic: enhanced class-J[J]. Progress In Electromagnetics Research M, 2017, 57: 103-111.
- [9] Li Guohua, Li Sudong, Cheng Zhiqun, et al. High-efficiency broadband GaN HEMT power amplifier based on harmonic-tuned match[J]. International Journal of RF and Microwave Computer-aided Engineering, 2020, 30(2): e22097.1-e22097.8.
- [10] SHARMA T, SRINIDHI E R, DARRAJI R, et al. High-efficiency input and output harmonically engineered power amplifiers[J]. IEEE Transactions on Microwave Theory Techniques, 2017, 66: 1002-1014.
- [11] KLOCKENHOFF H, BEHTASHR, WURFL J, et al. A compact 16W X-band GaN-MMIC power amplifier[J]. IEEE MTT-S International Microwave Symposium Digest, 2006: 1846-1849.
- [12] SAGAR K D, TUSHAR S, Zhu Ning, et al. Input-harmonic-controlled broadband continuous class-F power amplifiers for Sub-6-GHz 5G applications[J]. IEEE Transactions on Microwave Theory & Techniques, 2020, 68(7): 3120-3133.
- [13] KARTHIKEYAN K, RAMAKRISHNA V, KELLER S, et al. Broadband GaAs MESFET and GaN HEMT resistive feedback power amplifiers[J]. IEEE Journal of Solid-State Circuits, 2000, 35(9): 1285-1292.
- [14] RAAB F. H. Maximum efficiency and output of class-F power amplifiers[J]. IEEE Transactions on Microwave Theory & Techniques, 2001, 47(6): 1162-1166.
- [15] 徐兴福. ADS 2011 射频电路设计与仿真实例[M]. 北京: 电子工业出版社, 2014.

(收稿日期: 2020-12-24)

作者简介:

邵煜伟(1993-), 男, 硕士研究生, 主要研究方向: GaN 功率 MMIC 研究。

因电子车牌未能普及,所以主要验证视频图像处理结果,利用设计的基于 MATLAB GUI 的车流量检测系统进行数据处理,实验数据如表 1 所示,从表中可以看出对于车流量较小的场合,本系统算法准确率较高。在车流量密度较大的场合,多个车辆互相遮挡,导致车辆前景分割时发生粘连,则多个车流量被认为是一个运动目标,分割不准确,进而导致计数错误。

表 1 交通流参数计算(处理时间为 1 min)

数据	图 6(a)视频段	图 6(b)视频段
实际车流量/辆	10	86
计算车流量/辆	10	82
准确率/%	100	95.3

4 结论

本系统融合 RFID、数字图像、无线通信网络等技术,设计集成射频识别技术和视频技术的交通流参数测量方法及系统,既解决了传统视频监控的破坏性、不稳定性等诸多缺陷,又解决了新兴 RFID 技术的隐蔽性和不可视化,真正有效地提高了车辆识别的效率和精度,为道路交通流参数测量提供了一种新的测量方法。

参考文献

- [1] 李钦铭.面向智能交通的交叉路口车流量预测技术研究[D].北京:北京邮电大学,2017.
- [2] 柳长源,曹园园,罗一鸣.基于视频图像的车辆实时检测系统[J].计算机工程,2019(8):265-269,277.
- [3] 邱梦.基于视频图像处理的车流量检测理论研究[D].沈阳:东北大学,2016.
- [4] 陆键,项乔君.关于我国智能运输系统 ITS 产业化发展方向的思考[J].东南大学学报(自然科学版),2002(3):488-494.
- [5] 韩学昌.基于视频图像处理的车流量传感器的研究与设计[D].沈阳:东北大学,2015.
- [6] 周吉天白.基于超高速下超高频射频识别信道仿真模型的研究[J].科学技术与工程,2019,19(34):234-238.
- [7] 夏莹杰,欧阳聪宇.面向高速公路抛洒物检测的动态背景建模方法[J].浙江大学学报(工学版),2020,54(7):1-7.
- [8] 王海涛,侯静静.基于混合高斯模型融合背景减差的猪只运动跟踪[J].浙江农业科学,2020,61(2):329-332.
- [9] 刘云霄.基于图像信息的道路车流量检测算法研究与实现[D].南京:南京航空航天大学,2019.
- [10] 王腾.基于视频检测技术的交通车流量研究[D].银川:宁夏大学,2014.
- [11] 李婷俐.基于多特征级联分类器的道路车辆检测方法研究[D].长沙:湖南大学,2012.
- [12] 胡宏宇.基于视频处理的交通事件识别方法研究[D].长春:吉林大学,2010.
- [13] 吴志伟.基于视频的车流量检测技术研究[D].广州:华南理工大学,2012.
- [14] 谢衍祺.基于 MATLAB 的车辆识别系统[J].电子技术与软件工程,2017(9):65.
- [15] 汪惠兰,林航飞,李贻龙.基于 Matlab 实时运动目标跟踪检测系统[J].计算机安全,2012(3):17-20.
- [16] 彭春华,刘建业,刘岳峰,等.车辆检测传感器综述[J].传感器与微系统,2007,26(6):4-7.
- [17] 陈曦,殷华博.基于视觉运动目标跟踪技术分析[J].无线电工程,2011,41(6):22-24.
- [18] 普伟.基于视频的交通流量参数检测[D].西安:西安电子科技大学,2007.

(收稿日期:2020-06-16)

作者简介:

刘海艳(1982-),女,硕士,讲师,主要研究方向:传感器技术、虚拟仪器。

李俊敏(1970-),女,硕士,副教授,主要研究方向:PLC 应用、机电控制。

(上接第 76 页)

- [14] 李保利.基于类别层次结构的多层文本分类样本扩展策略[J].北京大学学报(自然科学版),2015,51(2):357-366.
- [15] Zhai Chengxiang, LAFFERTY J. A study of smoothing methods for language models applied to information retrieval[J]. ACM Transactions on Information Systems, 2004, 22(2): 179-214.

(收稿日期:2020-04-08)

作者简介:

高凡(1987-),通信作者,女,博士研究生,副研究员,主要研究方向:交通信息工程、计算机控制,E-mail:right870202@163.com。

王富章(1962-),男,博士,研究员,主要研究方向:交通信息工程及控制、铁路应急、大数据、铁路信息化。

张铭(1979-),女,博士,研究员,主要研究方向:交通信息工程及控制、城轨大数据。

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所