

基于故障树的事故分类方法^①

刘康炜^{1,2}, 万剑华¹, 靳熙芳¹

¹(中国石油大学(华东)地球科学与技术学院, 青岛 266580)

²(中国石化青岛安全工程研究院, 青岛 266071)

通讯作者: 万剑华, E-mail: wjh66310@163.com



摘要: 故障树的割集分析技术是判定事故原因的常用技术, 然而, 基于割集的技术仅能通过基本事件的组合判定事故的发生, 无法分析事故演化过程的中间事件. 本文针对事故分析报告描述的事故成因机理, 结合文本分类和故障树分析技术, 提出一种面向故障树的事故报告分类方法, 实现面向事故演化路径的事故报告的因果定位, 能够自动关联事故报告与故障树结构演化信息, 实现借鉴专家经验的事故因果演化的精确分析.

关键词: 故障树; 最近邻分类方法; 割集; 文本分类

引用格式: 刘康炜, 万剑华, 靳熙芳. 基于故障树的事故分类方法. 计算机系统应用, 2019, 28(6): 130-134. <http://www.c-s-a.org.cn/1003-3254/6849.html>

Accident Classification Method Based on Fault Tree

LIU Kang-Wei^{1,2}, WAN Jian-Hua¹, JIN Xi-Fang¹

¹(School of Geoscience, China University of Petroleum, Qingdao 266580, China)

²(SINOPEC Qingdao Research Institute of Safety Engineering, Qingdao 266071, China)

Abstract: The cut set analysis of fault tree is a common technique for judging accidents. However, the technology based on cut sets can only determine the occurrence of accidents by combination of basic events, and cannot analyze the intermediate events in the process of accident evolution. In view of the mechanism of accident analysis described in the accident analysis report, a precise classification method of accident report oriented fault tree is proposed by combining the text classification and fault tree analysis technology, and the accurate classification technology of accident oriented accident evolution path is realized, and the information of the report and the fault tree structure evolution information can be automatically correlated. We can achieve precise analysis of accident causality evolution based on expert experience.

Key words: fault tree; KNN; cut set; text categorization

由于化工生产工艺复杂、自动化控制水平高, 事故形成的机理难以捕获且多数发现事故征兆后便难以控制, 如何能够汇总历史事故原因机理, 进行事故原因的精确分析, 尽早发现事故潜在的隐患, 将有助于事故预警和预防. 然而, 虽然形成了大量事故的分析报告, 但并没有形成事故相关成因机理的提炼和汇总, 难以形成事故分析有效的知识库. 同时, 化工事故分析专家

也基于经验以故障树、FME(C)A 等安全分析模型建模了大量同类事故的成因机理^[1], 如何基于这些专家经验对事故报告进行快速准确的原因分类, 将有助于形成辅助事故分析与决策支持的知识库, 实现化工事故的早期预警和预防. 由于事故报告、故障树等事故分析模型的表达都是基于自然语言, 自然语言的分类技术在各领域已有广泛的研究, 如, 贝叶斯分类改进方法^[2]

① 基金项目: 国家重点研发计划 (2017YFC1405300); 山东省重点研发计划 (2018GGX101052)

Foundation item: National Key Research and Development Program of China (2017YFC1405300); Key Research and Development Program of Shandong Province (2018GGX101052)

收稿时间: 2018-09-08; 修改时间: 2018-10-08, 2018-11-12; 采用时间: 2018-11-14; csa 在线出版时间: 2019-05-25

采用文本中高词频特征的对数平均计算方法,解决了文本分类中参数估计不足问题,但只是单纯使用分类技术,没有探讨文本所含领域模型问题;优化的 SVM 故障分类器^[3]采用了 RS 约简故障指标体系,去除了冗余特征,但是割裂了故障过程间的演化关系;决策树、神经网络等^[4,5]改进分类方法基于统计模型、语言模型,主要处理文本的层次化分类问题,可以实现火灾、中毒、爆炸等化工事故的粗略分类,但不能根据事故模型探讨事故成因机理,事故成因机理需要结合专家的领域经验进一步分析、分类。在化工事故分析领域,故障树是一种演绎式失效分析模型,能通过逻辑门连接事故成因关系展示事故成因的演化过程^[6]。通过割集技术可以实现事故成因的最底层原因分析、也可以实现事故的定量分析和定性分析^[7]。本文基于化工事故监管部门形成了大量的事故故障树,这些故障树已通过领域专家实现面向事故成因机理分析、并进一步实现同类故障树的成因机理的汇总,形成包含同类事故所有原因的若干标准故障树^[8]。基于标准故障树,研究事故报告的事故成因精确分类分析,不仅可以准确分类事故报告的类别,也能够对事故报告描述的事故成因进行精确分析,从而能够实现面向事故成因机理的有效分析、分类、存储和管理,逐步完善事故分析与决策支持的知识库,实现化工事故的早期预警和预防。

1 面向故障树的文本分类算法

基于故障树模型事故成因机理分析,考虑到化工事故的实际特征,故障树的逻辑门仅包含与、或两种门结构。目前本文工作中合作的化工事故监管部门已经形成了大量的事故故障树,并综合同类故障树的成因机理形成该类型的标准故障树如图 1,为自然着火型标准故障树的示意图。

事故报告是通过专家对事故进行调查分析,形成的包含事故经过、事故原因、事故处置、事故总结和事故影响五部分事故调查报告,里面含有事故形成的各种信息,但也有各种杂音信息,目前收集化工事故监管部门已经形成了 10 023 个事故报告,已经粗略分类归属于火灾、爆炸、中毒、泄漏等 4 类事故,如何根据事故成因将事故报告精确归因于某类标准故障树,是进一步研究事故成因的自动分析和识别的关键技术。

1.1 基于自然语言处理的特征提取技术

事故报告与故障树节点都是采用自然语言描述的,

需要研究文本的特征向量生成方法。由于事故报告包含大量的无关、杂音字词,首先根据百度停留词库进行去停留词处理。采用将词嵌入向量生成算法与 TF-IDF 方法结合生成分词特征向量^[9],采用 Word2Vec CBOW 模型生成 P 维词向量 $Vec(w_j)=(v_1, v_2, \dots, v_p)$ 本文取 $P=1000$ 。由于 Word2Vec 采用通用领域词汇进行向量训练,为突出领域词汇的作用,计算事故报告、故障树节点中领域词的 TF-IDF 值,作为词向量的权重,从而实现了文本向量化既考虑词的语义信息,也能考虑领域词的重要度信息。采用如下计算公式生成词的 TF-IDF 值:

$$\begin{aligned} TFIDF(d_i, w_j) &= TF(d_i, w_j) \times IDF(w_j) \\ &= TF(d_i, w_j) \times \lg\left(\frac{M}{DF(w_j)}\right) \end{aligned} \quad (1)$$

其中, $TF(d_i, w_j)$ 表示词 w_j 在文本 d_i 中出现的频率。 $IDF(w_j)$ 表示单词 w_j 的逆文档频率。 M 表示文本总数。 $DF(w_j)$ 表示出现 w_j 单词的文本个数。则包含 TF-IDF 的词向量为:

$$VecT(w_j) = Vec(w_j) \times \frac{1}{M} \sum_{i=1}^M TFIDF(d_i, w_j) \quad (2)$$

则文本 d_i 制成相应的向量为:

$$VecT(d_i) = \frac{1}{N} \sum_{i=1}^{N_i} VecT(w_{ij}) \quad (3)$$

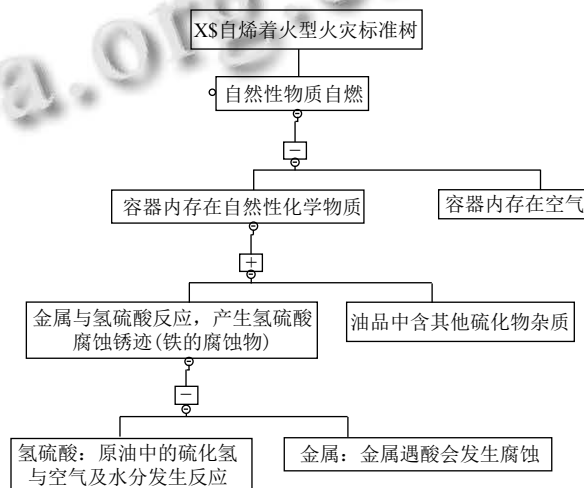


图 1 自然着火型标准故障树

1.2 扩展割集的生成算法

典型的故障树分析多通过分析故障树的最小割集和最小径集^[7,10],然而这种分析方法是分析故障树的叶

子节点,而无法分析事故成因过程,无法实现事故报告的精确归类,提出了一种面向故障树全部节点的层次

析取范式的算法,即可以计算出以基本事件为基准扩展可能的中间节点,而形成的故障树扩展割集。

27	场所	-0.068744	-0.349249	-0.020470	-0.213650	0.095787	0.040869	-0.090690	0.012222	0.356969	0.198330	0.263915	0.012506
28	管理	0.025659	-0.039479	0.112733	0.017893	-0.105488	-0.151462	-0.113766	-0.332603	0.015334	-0.319359	0.066095	-0.2077
29	组织	0.073921	-0.213884	-0.237045	0.095328	-0.140386	0.019091	-0.000573	0.157236	0.301182	-0.077438	0.053261	-0.04567
30	应急	-0.356224	0.107361	-0.129305	-0.361700	-0.182119	-0.144190	-0.007855	-0.078390	0.161459	-0.167452	0.185003	0.078
31	要求	-0.104689	-0.241905	0.139133	-0.173805	0.135221	-0.323786	-0.057312	-0.157487	0.253960	-0.120249	-0.109850	0.151
32	部门	-0.441998	-0.092363	-0.106492	0.124756	-0.218973	0.194833	-0.028210	0.201155	0.112978	-0.111792	0.138384	0.08565
33	易燃易爆	-0.133586	-0.097407	0.077149	0.253660	0.083004	-0.076713	0.042804	0.082254	0.343332	0.062448	-0.168624	-0.05
34	位置	-0.151210	-0.116913	0.129825	-0.068389	0.069050	0.025510	-0.012026	-0.119879	0.086423	0.106204	-0.112741	0.16754
35	能力	-0.029957	0.066333	-0.068434	0.043668	0.445648	-0.050357	-0.021121	-0.085356	-0.093491	-0.032293	-0.092996	0.063
36	化工	-0.093268	-0.094036	0.211252	0.260781	0.236958	-0.131462	-0.053035	0.015218	-0.048049	-0.109446	-0.078996	-0.125
37	没有	-0.091295	-0.027003	-0.092153	0.020776	-0.200669	0.250678	-0.060585	-0.076159	0.166258	-0.003610	0.081076	0.0068
38	造成	0.029571	0.067161	-0.120124	-0.032565	-0.017756	-0.328480	-0.039652	0.154438	-0.281553	-0.062260	0.012337	-0.345
39	泄漏	-0.249027	0.195972	-0.263787	-0.049364	-0.226948	-0.011053	0.150471	-0.071617	0.033440	0.096647	-0.186546	-0.083
40	问题	0.125291	-0.026819	0.370744	-0.198067	0.119883	0.064606	-0.070626	-0.209938	-0.066064	-0.166798	-0.155159	-0.135
41	发现	0.309476	-0.095456	0.075800	-0.295623	0.301594	0.292481	-0.116286	0.182237	0.132938	0.015512	0.037108	-0.168057
42	危险	-0.432166	-0.001669	-0.085942	-0.365516	-0.085660	-0.067780	-0.237426	-0.355248	0.198147	0.069731	-0.134518	-0.0
43	作业	-0.316792	0.106942	-0.190717	0.030619	-0.004766	-0.078280	0.162324	-0.038436	0.374899	0.242265	0.334999	-0.09126
44	措施	0.028955	0.079713	0.025069	-0.137996	-0.039244	-0.115696	-0.016284	0.384127	-0.293738	0.127249	0.081181	0.021257
45	车间	0.219843	-0.144600	-0.153212	0.063152	0.159887	0.255597	0.022220	0.250425	-0.008543	0.007124	0.126008	-0.042686
46	装备	-0.100807	0.046727	-0.150179	0.080432	-0.230438	-0.318536	-0.080891	0.046054	0.185562	0.511557	-0.198702	0.30576
47	操作	0.121908	0.029658	-0.039158	0.201539	-0.040352	-0.215760	-0.165871	0.026674	0.180407	0.305124	0.083826	0.111473
48	未	0.198727	0.037265	-0.372866	-0.048768	-0.205265	0.107476	-0.196909	-0.087171	0.097475	0.019941	-0.020201	0.122485
49	设备	0.236727	0.017599	-0.089038	0.128383	-0.041789	0.073495	-0.242613	-0.247655	0.229187	0.171144	-0.033836	-0.00184
50	员工	0.093456	0.035045	-0.125989	0.042659	-0.045785	0.108057	0.231747	-0.168485	-0.225910	-0.321629	0.099128	-0.08764
51	化学品	-0.401067	0.011678	-0.080393	-0.335297	-0.128359	0.024650	-0.337351	-0.213374	0.119927	-0.072906	-0.446365	0.1
52	石油化工	-0.256647	0.027436	0.040816	0.291040	0.413000	0.106531	0.012117	-0.124954	0.168766	-0.148155	-0.159916	0.156
53	排查	0.265148	-0.236462	-0.107052	0.064009	-0.052351	-0.048087	0.026588	0.329278	0.590293	0.169180	0.001045	-0.085537
54	气体	-0.344681	0.157181	-0.016510	0.441039	0.102618	-0.092434	0.055834	0.027324	0.146966	0.331377	-0.006089	-0.024191
55	死亡	-0.188127	-0.367717	0.025349	-0.269538	-0.074430	0.397166	-0.425497	-0.006520	0.116229	-0.060416	0.173223	-0.012
56	装置	0.439578	-0.161540	-0.036773	-0.071396	-0.207118	-0.115194	0.128631	0.084625	0.335124	-0.192335	0.135514	0.27687

图2 事故报告的词向量特征生成部分结果

本文扩展下行法、采用广度优先逐层替代计算扩展割集。核心思想:根据故障树逻辑门符号一层层向下判定,即遇到“与”门,将“与”门的上层节点替换为“与”门的所有下层节点,并将替换后的结果加入范式集合;遇到“或”门,将“或”门的上层节点依次替换为“或”门的下层节点,并将替换后的结果加入范式集合。求解故障树的扩展割集CS,算法步骤如下:

Step 1. 对故障树每个节点进行预处理。读入树节点 T_i , 将每个树节点 T_i 表示为该子树的根节点的文本 W_i 、孩子节点个数 N_i 、逻辑门符号 S_i 、孩子节点的文本的集合 C_j 。其中 $0 \leq j \leq N_i$, 即 $T_i = (W_i, N_i, S_i, C_j)$, $0 \leq j \leq N_i$ 。

Step 2. 读入故障树的根节点 T_1 , 并根据其逻辑门 S_1 进行处理。若是“与”门,则根节点加入集合、将根节点替换为“与”门所有的下层节点 C_j , 并加入集合, $CS = (W_1, C_1, C_2, \dots, C_{N_1})$; 若是“或”门,则根节点加入集合、将根节点依次替换为“或”门的下层节点 C_j , 并依次加入集合, $CS = (W_1, C_1, C_2, \dots, C_{N_1})$ 。

Step 3. 读入故障树的下一个节点 $T_i = (W_i, N_i, S_i, C_j)$, 并根据逻辑门 S_i 进行处理。若是“与”门,则寻找集合CS中存在的与节点文本 W_i 相同的树节点 T_e , 并用“与”门的所有下层节点 C_j 替换集合中的树节点文本 W_e , 替换后的结果加入结果集CS, $CS = (W_1, W_2, \dots, W_e) = CS = (W_1, W_2, \dots, C_1, C_2, \dots, C_{N_i})$ 。若是“或”门,则

寻找集合CS中存在的与节点文本相同的树节点 T_e , 并用“或”门的下层节点 C_j 依次替换集合中的树节点文本 W_e , 替换后的结果加入结果集CS, $CS = (W_1, W_2, \dots, W_e) = CS = (W_1, W_2, \dots, C_1, C_2, \dots, C_{N_i})$ 。

Step 4. 重复 Step 3, 直到所有的节点 T_i 都读完, 即能求出故障树故障发生的所有扩展割集 $CS = (W_1, W_2, \dots, W_e)$ 。

X\$自然着火型火灾标准树
 自然性物质自燃
 容器内存在自然性化学物质 容器内存在空气
 金属与氢硫酸反应, 产生氢硫酸腐蚀锈迹(铁的腐蚀物) 容器内存在空气
 油品中含其他硫化物杂质 容器内存在空气
 氢硫酸: 原油中的硫化氢与空气及水分发生反应 金属: 金属遇酸会发生腐蚀 容器内存在空气

图3 算法生成的图1故障树的扩展割集

每一个故障树都通过扩展割集算法的计算, 形成了大量扩展割集形成的词库集合, 通过文本向量生成算法, 计算出每个扩展割集的向量空间。这些向量空间也是事故报告分类的中心, 需要采用合适的分类算法经每个事故报告分类到相应的扩展割集中, 从而能够精确分析到事故报告描述的事故成因机理。本文在计算事故报告提取文本特征与扩展故障树割集的相似度计算方法采用的是, 较为成熟的KNN方法^[2,3,9]。

1.3 基于权重调整的KNN-W的事故报告分类方法

本文将事故分析报告和故障树扩展割集的余弦相似度作为距离衡量标志, 余弦相似度越大说明两类向

量距离越近. 标准的 KNN 分类方法 (记为 KNN-S), 当样本数量分布不平衡时, 实验结果明显偏重于数量多的类^[11]. 由于故障树的扩展割集不唯一, 而且每个故障树的割集个数都不相同, 会导致样本分布不均匀, 从而会导致 KNN 分类方法的准确度下降. 改进的 KNN-W 分类方法, 充分解决了 KNN-S 分类方法中非平衡化样本干扰的难题, 并在 KNN-S 分类方法的基础上, 根据采集样本所属的大类别数目和标准树所包含的扩展割集数目设计权重约束, 实现了对 KNN-S 分类方法中样本分布不平衡的调节, 还在一定程度上提高了事故报告分类的精确度. 采用基于权重计算的 KNN-W 的算法步骤如下:

Step 1. 利用式 (3) 计算事故分析报告文本向量 $VecT(d)=(v_1, v_2, \dots, v_K)$ 与每类故障树的扩展割集向量 $VecT(CS_j)=(c_{j1}, c_{j2}, \dots, c_{jK})$ 的余弦相似度 $\cos\theta_j$.

$$\cos\theta_j = \frac{\sum_{i=1}^P (v_i \times c_{ji})}{\sqrt{\sum_{i=1}^P (v_i^2) \times \sum_{i=1}^P (c_{ji}^2)}} \quad (4)$$

Step 2. 将所求得的余弦相似度 $\cos\theta_j$ 设置权重 $weight$, 加权结果为 $\cos\theta_j = \cos\theta_j \times weight$, 其中 $weight = \frac{Count_j(D)}{Count_j(CS)}$, $Count_j(D)$ 为属于扩展割集 CS_j 所在大类的事故报告数, $Count_j(CS)$ 为属于扩展割集 CS_j 所在故障树的扩展割集数, 即权重与属于类别的事故报告数成正比, 与所属类别的故障树扩展割集成反比.

Step 3. 将加权余弦值排序, 统计 K 个余弦值 $\cos\theta_j$ 当中每类故障树所占的比重. 比重最大的故障树即为所求的该事故报告的类别.

上述基于权重计算的 KNN-W 相比于 KNN-S 分类方法, 充分考虑了样本所属领域的大类别数目, 结合样本数据中故障树模型的扩展割集结构分布, 利用权重调剂不平衡分类的比例, 降低已有类别的故障树割集比重, 增加待分类的样本比例权重, 使分类方法能够适应不同事故数据集.

2 实验结果与分析

随机选取事故监管部门提供的 1000 份事故报告作为事故原因分析待分类的对象, 由于报告中的事故处置、事故总结和事故影响不能支撑事故原因分析, 首先对事故报告进行预处理, 只获取事故经过和事故原因分析两部分文本内容进行文本特征提取, 通过贝

叶斯方法配合人工确认对这 1000 篇事故已经进行了粗略分类, 分类结果如表 1 所示.

基于事故监管部门提供的 11 类标准故障树, 利用扩展割集算法计算 11 个标准故障树的扩展割集, 表 2 为标准的结构信息和通过算法计算的扩展割集信息. 这些扩展割集包含事故形成的充分条件, 如果事故报告中隐含了相应的割集信息, 则表示该事故包含该事故成因信息, 这样可以通过分析事故成因来确定事故报告的精确分类.

表 1 事故报告粗略分类

序号	事故分类	文档数
C1	中毒类	142
C2	爆炸类	405
C3	火灾类	314
C4	泄漏类	139

表 2 标准故障树的结构信息

序号	标准树名	结点数	扩展割集数	所属大类
T1	中毒事故标准树	124	117	C1
T2	着火破坏型爆炸标准树	69	47	C2,C3
T3	反应失控型爆炸标准树	48	43	C2,C3
T4	反应失控型火灾标准树	36	25	C3
T5	传热型蒸汽爆炸标准树	53	39	C2
T6	平衡破坏型蒸汽爆炸标准树	27	19	C2
T7	自燃着火型爆炸标准树	14	7	C2, C3
T8	自燃着火型火灾标准树	8	5	C3,
T9	超压引起的物理爆炸标准树	29	20	C2,C3
T10	泄露着火型爆炸标准树	68	43	C4, C2,C1
T11	泄露着火型火灾标准树	59	42	C4, C3,C1

在算法设计中, 首先采用本文提出的结合树形复杂性和文章分类数的权值调整 KNN-W 算法, 取 $K=20$ 进行实验, 表 3 给出识别的每类标准树所包含的通过算法识别数目, 实际应该包含的数目, 正确识别的数目, 并给出每类故障树的识别的准确率、召回率和 F1 值^[3,4,10], 通过分析发现在样本集较大, 故障树结构较复杂的标准树中, 整体性能要好于标准树结构简单的分类, 算法总体的准确性为 89.6%.

为了验证本文提出的的权值调整算法的性能, 在对比实验中, 采用标准 KNN-S 方法和 2.3 算法进行对比实验, 两类的 K 值分别为 10、20、50 进行分析, 每一类的 F1 值的对比结果如图 3 所示. 表 4 给出每次实验的各故障树类的平均 F1 值, 从算法对比分析, 可以看出 KNN-W 算法整体性能较 KNN-S 有不同程度的提高, 从 K 值选择分析可以看出, $K=20$ 较

$K=10$ 有较大提升, 但 $K=50$ 较 $K=20$ 提升较少.

表3 KNN-W 算法性能分析

标准树类	实际数	识别数	准确数	准确率 (%)	召回率 (%)	F1-Score(%)
T1	97	90	87	89.7	96.7	93.1
T2	55	60	48	87.3	80.0	83.5
T3	72	74	61	84.7	82.4	83.5
T4	131	125	119	90.8	95.2	92.9
T5	46	46	41	89.1	89.1	89.1
T6	127	123	119	93.7	96.7	95.2
T7	98	115	92	93.9	80.0	86.4
T8	269	244	241	89.6	98.8	94.0
T9	33	31	25	75.8	80.6	78.1
T10	43	57	39	90.7	68.4	78.0
T11	29	35	24	82.8	68.6	75.0

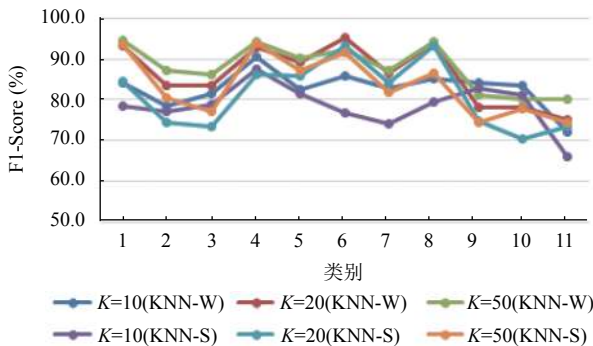


图4 KNN-W 与 KNN-S 的 F1-Score 对比结果

表4 平均指标数据对比

算法	K 值	结果 (%)
KNN_W	10	2.8
	20	86.3
	50	86.8
KNN_S	10	78.5
	20	81.3
	50	81.6

3 总结

直接应用机器学习分类方法难以有效分析领域问题, 针对化工事故分析领域, 提出一种结合领域专家经验的分析模型进行事故精确分析的思路, 基于故障树分析的事故模型, 设计故障树的扩展割集, 使故障成因的充分必要条件不仅反馈在叶子节点的集合, 也反馈

到包含中间节点的所有节点组合, 这样可以检测出事故报告中描述的事故演化过程, 在分类算法上, 选择 KNN 方法, 并根据分类特点改进 KNN 算法使其更能有效进行分类, 实验结果表明所设计算法能够实现较精确的分类.

参考文献

- 1 吴济民. 国外化工企业工艺安全技术管理概述. 中国安全生产科学技术, 2011, 7(7): 192-198.
- 2 张杰, 陈怀新. 基于归一化词频贝叶斯模型的文本分类方法. 计算机工程与设计, 2016, 37(3): 799-802.
- 3 张鑫, 胡瑾秋, 张来斌, 等. 基于 RS 和 SVM 的化工过程高精度故障诊断方法. 石油学报 (石油加工), 2017, 33(4): 777-784. [doi: 10.3969/j.issn.1001-8719.2017.04.024]
- 4 Anava O, Levy KY. k^* -nearest neighbors: From global to local. Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona, Spain. 2016. 4923-4931.
- 5 Goudjil M, Koudil M, Bedda M, et al. A novel active learning method using SVM for text classification. International Journal of Automation and Computing, 2018, 15(3): 290-298. [doi: 10.1007/s11633-015-0912-z]
- 6 Ruijters E, Stoelinga M. Fault tree analysis: A survey of the state-of-the-art in modeling, analysis and tools. Computer Science Review, 2015, 15-16: 29-62.
- 7 Boutros T, Liang M. Detection and diagnosis of bearing and cutting tool faults using hidden Markov models. Mechanical Systems and Signal Processing, 2011, 25(6): 2102-2124. [doi: 10.1016/j.ymssp.2011.01.013]
- 8 蒋骏. 基于哨兵事件数据库的炼化重大事故预警系统的构建[硕士学位论文]. 青岛: 青岛科技大学, 2016.
- 9 Trstenjak B, Mikac S, Donko D. KNN with TF-IDF based framework for text categorization. Procedia Engineering, 2014, 69: 1356-1364. [doi: 10.1016/j.proeng.2014.03.129]
- 10 王超学, 张涛, 马春森. 改进 SVM-KNN 的不平衡数据分类. 计算机工程与应用, 2016, 52(4): 51-55, 103. [doi: 10.3778/j.issn.1002-8331.1403-0109]
- 11 Ruijters E, Stoelinga M. Fault tree analysis: A survey of the state-of-the-art in modeling, analysis and tools. Computer Science Review, 2015, 15-16: 29-62.