

基于层次化类别嵌入的电力 ICT 系统故障分类

李建桂¹, 梁 越¹, 高鹏飞¹, 刘绍华², 马应龙¹

(1. 华北电力大学 控制与计算机工程学院, 北京 102206; 2. 北京邮电大学 电子工程学院, 北京 100876)

摘要: 为解决电力信息通信客服系统在故障研判时存在故障分类准确率低甚至误分的问题, 提出基于层次化类别嵌入的文本分类方法, 进行电力信息通信系统故障识别. 首先, 基于电力信息通信系统故障的用户保修工单文本数据构建电力信息通信系统层次化电力故障标签; 其次, 提出了基于层次化深层金字塔卷积神经网络和基于层次化中断循环神经网络 2 种层次化文本分类方法, 采用层次化类别嵌入方法逐层进行故障类型分类. 实验结果表明, 基于层次化深层金字塔卷积神经网络的方法效果最优, 可以提供高效、准确的故障识别服务.

关键词: 电力信息通信客服系统; 电力文本分类; 层次化文本分类; 类别嵌入

中图分类号: TP39

文献标志码: A

A Hierarchical Category Embedding Based Approach for Fault Classification of Power ICT System

LI Jian-gui¹, LIANG Yue¹, GAO Peng-fei¹, LIU Shao-hua², MA Ying-long¹

(1. School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China;

2. School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: To solve the low classification accuracy or even misclassification issue in fault diagnosis, a text classification method based on hierarchical category embedding is proposed in information and communication technology (ICT) customer service systems. First, a hierarchical label system is constructed for the failure data in power ICT systems based on the textual data of the work orders. Then, hierarchical deep pyramid convolutional neural networks (HDPCNN) and hierarchical disconnected recurrent neural networks are proposed, which adopt hierarchical category embedding technique for level-by-level fault type classification. The experimental results show that the hierarchical text classification algorithm HDPCNN has the best classification accuracy, which can provide efficient and accurate services for fault type recognition.

Key words: power information and communication technology customer service system; power text classification; hierarchical text classification; category embedding

电力信息通信 (ICT, information and communication technology) 客服系统是电力企业生产、经营、管理的重要支撑手段, 涉及企业人力资源、财务、物

资及专业应用等诸多领域, 可为电力企业内、外部用户提供各类信息通信系统故障、使用策略等在线咨询, 并可针对具体的咨询内容及时有效地提供

收稿日期: 2020-12-21

基金项目: 国家重点研发计划项目 (2018YFC0831404; 2018YFC0830605)

作者简介: 李建桂 (1996—), 女, 硕士生.

通信作者: 马应龙 (1976—), 男, 教授, E-mail: yinglongma@ncepu.edu.cn.

解决方案,以保证国家电网信息系统业务的正常运行. 确保 ICT 客服系统安全稳定运行,提供优质服务,根据用户的故障描述,高效准确地确定故障类型,并针对性地提出解决方案是 ICT 客服系统要实现的最终目标,随着公司信息化水平的提升,ICT 系统故障的类型也越来越多,人工判断故障类型并提出解决方案的服务方式已不能满足需求,因此,实现故障类型的自动识别和定位迫在眉睫. 文本分类技术是自然语言处理中一个重要的研究内容,目前已得到很多研究者的青睐,并广泛应用于电力系统的数据处理^[1]. 例如,Xi 等^[2]提出了一种基于集成学习的多标签分类方法以解决 ICT 客服系统积累的工单故障. 支持向量机(SVM, support vector machines)^[3]、双向长短期神经网络(BiLSTM, bi-directional long short-term memory)^[4]等算法也在电力系统中已被广泛应用.

使用多标签分类方法时一般忽略标签之间的层次关系,把所有的标签都平面化,而 ICT 系统的故障类别之间存在层次关系,属于层次多标签分类场景^[5]. 层次多标签文本分类是一个比较复杂的问题. 基于层次化类别嵌入的电力 ICT 系统故障识别方法面向电力领域,是采用逐层的基于类别嵌入的一种层次化文本分类局部算法,该算法为每一层训练一个分类器,以实现准确高效的 ICT 系统故障类型识别. 笔者的贡献包括:①提出了基于层次化故障类别体系的多标签故障类型识别方法,构建了一个用于层次化多标签故障类型识别的深度学习训练集;②提出了 HDPCNN(hierarchical deep pyramid convolutional neural networks)和 HDRNN(hierarchical disconnected recurrent neural networks)2种基于层次化类别嵌入的层次化多标签文本分类算法,用于面向层次化故障类别体系多标签故障类别分类;③与基线算法进行比较的实验结果显示,HDPCNN 算法具有更好的分类准确率,可以有效提升 ICT 客服系统的故障类型识别性能,有助于自动识别电力 ICT 系统故障,及时进行故障定位,提高 ICT 客服系统的客户满意度.

1 ICT 客服系统数据集处理

1.1 ICT 客服系统工单故障记录文本特点

实验数据集是国家电网 ICT 客服系统工单故障记录,工单内容主要是记录、解决和监控用户的服务

请求和故障报告. 原始工单故障记录主要有3个特点:①由于受到 ICT 客服系统客服人员知识面以及相关经验的影响,故障分类准确率低;②不同客服人员的记录习惯导致记录的故障会有语义上的误差;③工单记录信息不完整,ICT 客服系统中,每个类别故障发生的次数不均衡,所以对应故障类别的记录数量就会不同,在实验中会遇到样本不均衡问题.

1.2 构建 ICT 系统层次化故障标签

ICT 客服系统工单故障记录文本分类属于典型的层次多标签分类场景,例如,工单故障记录“计算机浏览器故障,因电脑 IE 浏览器由 IE6 升级为 IE8,导致打开国家电网首页浏览器字体大小显示有误”,这条故障记录中涉及桌面运维故障、计算机故障、软件故障,而这3个故障类别之间存在层次关系,其中计算机故障是桌面运维故障的子类别,而软件故障是计算机故障的子类别. 结合 ICT 系统故障类别特点和大量的背景调研以及结合电力领域专家的建议,构建了完整层次结构的 ICT 系统故障类别,在构建 ICT 系统层次化故障标签时,采用了手工构建为主,相关处理脚本为辅的方式,表1所示为数据集完整的标签体系和完整标签分布.

1.3 层次化多标签训练集半自动化构建

1.3.1 数据预处理

在处理原始 ICT 客服系统电力故障数据集时首先要去掉原工单数据中联系人、申请人、来电号码等部分内容,完成对原工单数据中无效数据的初步筛选,留下要进行文本分类的初样本;其次对初样本进行数据清洗,去除噪声数据,如无效的数据记录,或者故障类别定位有明显错误的记录等;最后采用 Jieba 分词对数据集进行分词、去停用词等操作,生成最后的目标数据集,用于后续的实验.

1.3.2 ICT 系统故障类别标签挖掘

ICT 系统故障类别没有标准的故障类别体系,需要从 ICT 客服系统工单故障文本中挖掘出候选故障类别标签集. 将数据预处理后已进行分词和去停用词的词作为候选关键词,采用 TextRank 算法^[6]根据重要度选取最重要的 N 个关键词作为 ICT 系统故障类别的参考关键词,最后再根据样本数据的词集与关键词关联,构建故障类别. 将 TextRank 算法用于 ICT 系统故障文本关键词提取,需要构建关键

表1 WOFR数据集标签分布 %

第 1 层	第 2 层	第 3 层
桌面运维故障(47)	计算机故障(12.6)	操作系统故障(3.2)
		软件故障(3.2)
		无线网卡故障(3.2)
		硬件故障(3.0)
	安全故障(8.3)	杀毒软件故障(4.3)
		安全客户端故障(4)
	协同办公故障(9.4)	系统问题故障(2.1)
		打印复印故障(1.8)
		扫描仪故障(2.2)
		电话传真故障(1.8)
		企业门户故障(1.5)
	网络故障(12.7)	内网故障(2.5)
		无线外网故障(2.4)
		有线外网故障(2.5)
		内网邮件故障(2.6)
外网邮件故障(2.7)		
桌面类其他故障(4.0)	移动组件故障(2.8)	
	其他应用故障(1.2)	
其他系统故障(53)	主数据故障(6.3)	财务故障(2.4)
		人资故障(1.9)
		主数据物资故障(2.0)
	物资故障(5.6)	ERP-MM 故障(1.8)
		电子商务故障(2.0)
		物资辅助故障(1.8)
	生产故障(6.0)	电能监测故障(6.0)
	安监一体化管理故障(8.0)	系统管理故障(8.0)
	纪检监察管理故障(8.0)	账号管理故障(8.0)
	其他故障(19.1)	工会管理故障(2.0)
		合作管理故障(1.4)
		经济法律故障(1.7)
		科技管理故障(4.6)
		综合管理故障(2.7)
		车辆管理故障(2.0)
政工管理故障(2.3)		
智能审计故障(2.4)		

词图 $G=(V,E)$, 其中 V 是由 ICT 系统故障文本经过分词去停用词处理后构成的词集合, E 表示词与词之间的关系. 关键词 V_i 的 TextRank 值 (TR 值) 为

$$S(V_i) = (1-c) + c \sum_{V_j \in I(V_i)} \frac{w_{ji}}{\sum_{V_k \in O(V_j)} w_{jk}} S(V_j) \quad (1)$$

其中: $I(V_i)$ 为关键词 V_i 的所有入链集合, 在关键词图中对应入度的点; $O(V_j)$ 为与关键词 V_j 存在边的其他关键词集, $|O(V_j)|$ 即与 V_j 存在边的关键词个数; c 为阻尼系数; w_{ji} 为权重项, 用来表示 2 个节点之间边连接的重要程度. 根据式(1), 迭代传播各节

点的权重直至收敛, 再对节点权重进行倒序排序, 从而得到最重要的 N 个 ICT 系统故障关键词, 作为候选故障类别关键词, 最终构建 ICT 系统故障多标签类别集.

2 基于类别嵌入的层次化分类方法

2.1 层次化多标签文本分类

目前层次化分类算法可以分为局部算法和全局算法^[7] 2 种, 如 Kowsari 等^[8] 提出的层次化深度学习文本分类 (HDLTex, hierarchical deep learning for text classification) 算法是一种局部算法, 其策略是为每个父节点训练一个分类器. 常见局部算法有 3 种: ① 逐节点地为每个节点训练一个二元分类器, 判断样本是否属于该节点代表的标签, 该方法在遇到数量庞大的标签时, 要给每个节点训练一个分类器, 所以模型参数会增加, 计算成本也会随之增大; ② 为每个父节点训练一个多分类器, 然后下一层的分类结果就从被选中父节点的子类别中产生, 该分类器的候选结果都是父类别的子类别, 该方法过于依赖层级结构, 如果 $l-1$ 层的分类结果是错的, 会直接影响到 l 层的分类结果^[9]; ③ 为每一层训练一个多分类器^[10], 该分类器的候选结果是当前层中的所有类别都参与分类. 该方法虽然受层级约束的影响比较小, 但存在错误传递问题, 当层次数量比较多时, 不适合局部的层次化分类. 基于全局结构的层次化分类算法只用一个分类器就能得出样本在类别层次中的分类结果, 优点是分类不会出现错误传递. 因为只有一个分类器, 所以模型参数的数量也比局部算法少很多, 但全局算法在捕捉类别层次结构中的局部信息时不如局部算法, 其模型往往比较复杂^[11], 而 HDPCNN 和 HDRNN 算法是逐层基于类别嵌入的一种层次化文本分类局部算法.

2.2 类别嵌入

构建层次化结构的故障类别 (标签) 体系时, 考虑到层次化分类最大的特点是在进行第 l 层分类时, 必须考虑 l 层之前的分类情况, 为此, 采用层次化类别嵌入的方法解决层次结构之间有联系的问题. 将文本表示 t 与其第 $l-1$ 层的分类结果 p_{l-1} 进行拼接, 将拼接的结果 d_l 用于第 l 层的分类, 以便在进行 l 层分类时利用到上层的分类信息. 由于第 1 层是首次分类, 所以 $d_l = t$, 没有上层可用的类别嵌入, 有

$$d_l = p_{l-1} \oplus t \quad (2)$$

其中符号 \oplus 代表向量的拼接运算。

最后使用一个全连接层和一个 softmax 层对第 l 层进行分类,有

$$\left. \begin{aligned} d'_l &= \tanh(W_1 d + b_1) \\ p_l &= \text{softmax}(W_2 d'_l + b_2) \end{aligned} \right\} \quad (3)$$

其中:softmax 是在第 l 层进行所有类别的计算,该分类结果将用于下一层的分类;激活函数为 \tanh , W 为权重; b 为偏置项; d'_l 是经过 \tanh 线性变换的结果。HDPCNN 和 HDRNN 算法对每层分类进行如上所述的操作,直到得到最后一层的结果。

2.2.1 基于层次化类别嵌入的多标签文本分类

用于 ICT 客服系统电力故障文本第 l 层分类的 HDPCNN 和 HDRNN 算法模型如图 1 所示,利用 HDPCNN 或 HDRNN 算法对电力故障文本进行

层次化多标签分类,基于深层金字塔卷积神经网络 (DPCNN, deep pyramid convolutional neural networks)^[12] 或 中断循环神经网络 (DRNN, disconnected recurrent neural networks)^[13] 算法分别对第 l 层和第 $l-1$ 层进行特征提取,得到文本表示 t , 输入中的 $w_{[n]}$ 表示文本的第 n 个词向量。根据文本第 $l-2$ 层的分类结果 p_{l-2} , 将第 $l-1$ 层得到的文本表示 t 与 p_{l-2} 进行拼接,得到新的文本表示 d_{l-1} 。同样,将第 l 层得到的文本表示 t 与 p_{l-1} 进行拼接,得到新的文本表示 d_l , 最后全连接层和 softmax 层基于文本表示 d_l 对文本进行分类,输出其第 l 层的分类结果 p_l 。

HDPCNN 和 HDRNN 是一种逐层进行分类的局部层次化分类算法,在层次化结构的每一层训练一个如图 1 所示的子模型用于该层分类。

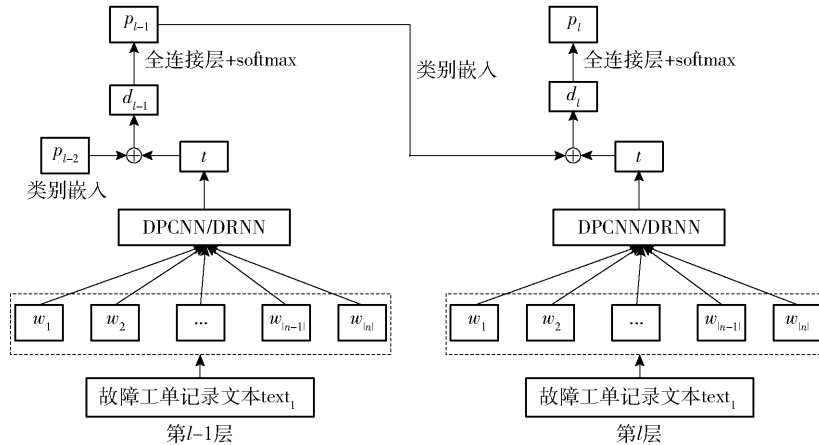


图 1 ICT 系统故障文本第 l 层分类的 HDPCNN/DRNN 算法模型

2.2.2 基于 DPCNN 和 DRNN 的文本分类

HDPCNN 和 HDRNN 分别是 DPCNN^[12] 和 DRNN^[13] 算法层次化多标签分类的扩展模型。DPCNN 算法模型的主要特点是减采样的设计和残差连接的设计,加权平方损失函数为

$$l = \sum_{i,j} \alpha_{i,j} (z_i[j] - p_i[j])^2 \quad (4)$$

其中: z_i 为样本 i 的目标区域,设置权重 $\alpha_{i,j}$ 以达到负采样效果。DRNN 算法模型的主要特点是将位置不变性引入 RNN 模型中,使得 DRNN 算法模型既能捕捉长距离依赖信息,又能抽取关键短语信息。

3 实验

3.1 数据集

实验包含 2 个数据集,一个是根据 ICT 客服系

统的工单故障记录 (WOFR, work order failure record) 构建的多标签数据集;另一个是基准数据集 WOS(web of science)^[8],如表 2 所示。

表 2 数据集各层种类数量信息

数据集	第 1 层	第 2 层	第 3 层	文本数量
WOFR	2	11	35	2 445
WOS	7	134	—	46 985

3.2 实验设置

3.2.1 样本平衡及模型参数

ICT 客服系统工单故障数据集存在样本不平衡的问题,采用数据合成^[14]的方法来处理,利用样本在特征空间的相似性生成新样本,有

$$x_{\text{new}} = x_i + (\hat{x} - x_i)\delta \quad (5)$$

其中: x_i 为小众样本, 从它属于小众类的 K 近邻中随机选取一个样本点 \hat{x} , 生成一个新的小众样本 x_{new} , $\delta \in [0, 1]$ 是随机数.

3.2.2 测评指标

实验需要计算 Overall^[8] 分类准确率和中间各层分类准确率, 其中 Overall 分类准确率指的是利用样本在上层的预测类别来预测样本在最后一层的分类准确率. 中间各层的分类准确率指的是提供样本在上一层的实际类别时该层的分类准确率.

3.2.3 对比算法

层次化多标签文本分类选择流行的 HDLTex^[8] 方法, 平面多标签分类算法包括 TextCNN^[15] 和 BiLSTM^[16]. 另外, 分别采用 Word2vec 和 BERT (bidirectional encoder representation from transformers)^[17] 2 种词嵌入方式作为输入特征对 HDPCNN 和 HDRNN 算法的性能进行实验对比.

3.3 实验结果分析

3.3.1 Overall 分类准确率

如表 3 所示, 只考虑各分类算法中最后一层的分类准确率, 其中 HDPCNN_b 算法在 2 个数据集上 Overall 分类准确率都是所有算法中最高的. HDLTex 和 HDRNN 算法是为了研究层次化分类问题设计的算法, 但与 TextCNN 和 BiLSTM 平面分类算法相比, 其 Overall 分类准确率并不具明显的优势, 这样就失去了层次化的意义, 而 HDPCNN 算法的分类准确率高于这些平面算法.

表 3 Overall 分类准确率

算法分类类型	算法	WOFR	WOS
平面	TextCNN	66.28	62.34
	BiLSTM	71.16	73.08
层次化	HDLTex	70.05	76.58
	HDRNN _{w2v}	70.30	76.29
	HDPCNN _{w2v}	72.18	81.46
	HDRNN _b	75.78	79.58
	HDPCNN_b	76.64	84.37

实验结果显示, HDPCNN_b 算法的分类效果最好. 究其原因, TextCNN 难以获得文本的长距离依赖关系, 而 HDPCNN 算法通过不断加深网络, 可以抽取长距离的文本依赖关系, 从而获得最佳的分类准确率; HDRNN 算法的原型是 DRNN, 保持了循环神经网络和卷积神经网络的一部分优点, 所以

分类的效果较 HDLTex 和 TextCNN 要好; 通过 BiLSTM 可以更好地捕捉双向的语义依赖, 所以 BiLSTM 是除了 HDPCNN 算法外 Overall 分类准确率较优的一个模型; HDPCNN 算法之所以获得最佳的准确率是因为其通过不断加深网络, 可以抽取长距离的文本依赖关系有效地表示文本中的长范围关联, 而 HDRNN 算法分类受到卷积核窗口大小影响较大, 窗口太大或太小都会导致模型的效果不佳, HDRNN 算法模型的稳定性没有 HDPCNN 算法好.

由表 3 可看出, 在 2 个数据集上, 分别用 Word2vec 和 BERT 的词嵌入方式作为输入特征时, 基于 BERT 的算法分类效果都较好, 这是因为相较于 Word2vec, 基于 BERT 的词嵌入可以看成是一个句子级别的上下文的词表示, 并且可以对一词多义进行建模, 进一步增加了词向量模型的泛化能力, 所以分类性能比较好. 实验结果显示, HDPCNN_b 算法获得了最好的分类效果.

3.3.2 分类准确率

表 4 和表 5 所示分别为层次化分类模型在 ICT 客服系统工单故障记录数据集 WOFR 和 WOS 数据集上的分类准确率. 其中 l_1, l_2, l_3 为提供文本在上层的真实类别时本层的分类准确率. 从表 4 和表 5 中可以看出, 各层次化分类模型在 WOFR 数据集的 l_3 分类准确率和 WOS 数据集的 l_2 分类准确率要比对应的 Overall 分类准确率高, 这说明在分类时考虑类别之间的层次关系, 引入上层的分类结果对本层分类准确率有很重要的影响. HDPCNN 算法可有效地提升 ICT 系统故障类别的分类性能, 推动 ICT 客服系统的业务模式从人工慢慢转向自动化, 降低 ICT 客服系统对客服人员的要求, 提高 ICT 客服系统故障研判的质量, 辅助客服人员进行故障定位, 从而保证 ICT 客服系统服务的时效性和准确性.

表 4 层次化分类算法在 WOFR 数据集各层的分类准确率

算法类别	算法	l_1	l_2	l_3	Overall
对比算法	HDLTex	86.79	83.07	70.63	70.05
	HDRNN _{w2v}	86.67	85.25	80.13	70.30
主算法	HDPCNN _{w2v}	92.83	86.75	80.98	72.18
	HDRNN _b	88.32	86.52	80.14	75.78
	HDPCNN_b	96.33	89.74	82.05	76.64

表5 层次化分类算法在 WOS 数据集各层的分类准确率

算法类别	算法	l_1	l_2	Overall
主算法	HDLTex	90.45	84.66	76.58
	HDRNN _{w2v}	89.17	83.22	76.29
	HDPCNN _{w2v}	95.93	86.35	81.46
	HDRNN _b	94.67	84.72	79.58
	HDPCNN _b	96.63	87.39	84.37

从表4和表5所示的层次化分类结果可见, HDPCNN_b算法的效果最好,尤其在 ICT 客服系统故障数据集 WOFR 上达到了90%以上的分类准确率。究其原因, HDLTex 算法采用逐节点的方法进行层次化分类,第 l 层进行分类时,该文本的候选类别是第 $l-1$ 层分类结果类别的子类别。如果第 $l-1$ 层分类出现错误,那么第 l 层的分类必然错误,而 HDPCNN 算法是逐层地进行分类,相比同样逐层进行分类的 HDRNN 算法, HDPCNN 算法在有效抽取长距离的文本依赖关系的同时,还比 HDRNN 算法更加稳定,在数据集 WOFR 上更有明显的优势。

不难发现,基于 BERT 的词嵌入方式可以有效获取文本特征,在一定程度上提高了模型的分类性能。HDPCNN 算法在每一层进行分类时文本的候选类别是该层的所有类别,所以如果上层的分类结果是错误的,该层的分类器依然有可能获得正确的分类结果。

从表3~5可以看出, HDPCNN 算法模型对电力 ICT 系统故障的识别效果较好,对于提高故障判断以及快速定位,及时作出反应,给出相关的解决方案具有重要意义,从根本上解决了人工研判时遇到的故障误判的问题,大大降低了电力 ICT 客服系统的人工成本,规范了 ICT 客服系统的工单故障记录模式且保证了 ICT 客服系统的服务质量。

4 结束语

面向电力领域构建了一个电力 ICT 系统层次化结构的故障类别标签体系,将 ICT 系统故障类别定位归结为层次化多标签文本分类问题。构建了电力 ICT 系统故障多标签训练集,提出了一种逐层的基于类别嵌入的局部层次化分类方法。实验结果证明, HDPCNN 算法模型可以准确地实现 ICT 系统故障类别的定位,提高层次化多标签分类准确率。

参考文献:

- [1] 傅群超, 王枫. 用于文本分类的多探测任务语言模型微调[J]. 北京邮电大学学报, 2019, 42(6): 76-83.
Fu Qunchao, Wang Cong. Based on multiple probing tasks fine-tuning of language models for text classification[J]. Journal of Beijing University of Posts and Telecommunications, 2019, 42(6): 76-83.
- [2] Xi Ziyue, Chen Xiaona, Almad T, et al. A novel ensemble approach to multi-label classification for electric power fault diagnosis[C]//2019 IEEE 7th International Conference on Computer Science and Network Technology (ICCSNT). Dalian: IEEE Press, 2019: 267-271.
- [3] 汪崔洋, 江全元, 唐雅洁, 等. 基于告警信号文本挖掘的电力调度故障诊断[J]. 电力自动化设备, 2019, 39(4): 126-132.
Wang Cuiyang, Jiang Quanyuan, Tang Yajie, et al. Fault diagnosis of power dispatching based on alarm signal text mining[J]. Electric Power Automation Equipment, 2019, 39(4): 126-132.
- [4] Melamud O, Goldberger J, Dagan I. Context2vec: learning generic context embedding with bidirectional LSTM[C]//Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning. Berlin: Association for Computational Linguistics, 2016: 51-61.
- [5] Peng Hao, Li Jianxin, Wang Senzhang, et al. Hierarchical taxonomy-aware and attentional graph capsule RCNNs for large-scale multi-label text classification[J]. IEEE Transactions on Knowledge and Data Engineering, 2021, 33(6): 2505-2519.
- [6] Mihalcea R, Tarau P. TextRank: bringing order into texts[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Barcelona: Association for Computational Linguistics, 2004: 404-411.
- [7] Silla C N, Freitas A A. A survey of hierarchical classification across different application domains[J]. Data Mining and Knowledge Discovery, 2011, 22(1/2): 31-72.
- [8] Kowsari K, Brown D E, Heidarysafa M, et al. HDLTex: hierarchical deep learning for text classification[C]//2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA). Cancun: IEEE Press, 2017: 364-371.
- [9] Liu Shaohui, Dong Minghai, Zhang Haijun, et al. An approach of multi-hierarchy text classification[C]//2001 International Conferences on Info-Tech and Info-Net. Beijing: IEEE Press, 2001: 95-100.

- [10] Wehrmann J, Barros R C, Dôres S N D, et al. Hierarchical multi-label classification with chained neural networks[C] // Proceedings of the Symposium on Applied Computing. Marrakech: ACM, 2017: 790-795.
- [11] Cerri R, Barros R C, de Carvalho A C P L F, et al. A grammatical evolution algorithm for generation of hierarchical multi-label classification rules[C] // 2013 IEEE Congress on Evolutionary Computation. Cancun: IEEE Press, 2013: 454-461.
- [12] Johnson R, Zhang Tong. Deep pyramid convolutional neural networks for text categorization[C] // Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Vancouver: Association for Computational Linguistics, 2017: 562-570.
- [13] Wang Baoxin. Disconnected recurrent neural networks for text categorization[C] // Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Melbourne: Association for Computational Linguistics, 2018: 2311-2320.
- [14] Xu Zhihui, Yang Lihong. Application of improved SMOTE algorithm in logistic regression credit scoring model[J]. Hans Journal of Data Mining, 2021, 11(2): 50-58.
- [15] Kim Y. Convolutional neural networks for sentence classification[C] // Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha: Association for Computational Linguistics, 2014: 746-751.
- [16] Schuster M, Paliwal K K. Bidirectional recurrent neural networks[J]. IEEE Transactions on Signal Processing, 1997, 45(11): 2673-2681.
- [17] Cai Qing. Research on Chinese naming recognition model based on BERT embedding[C] // 2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS). Beijing: IEEE Press, 2019: 1-4.