

基于图卷积神经网络和 RoBERTa 的物流订单分类

王建兵, 杨超*, 刘方方, 黄 晓, 项 勇

(安徽港口物流有限公司, 安徽 铜陵 244000)

摘 要: 订单信息贯穿于物流供应链的所有环节, 高效的订单处理是保障物流服务质量和运营效率的关键。面对日益增长的差异化客户物流订单, 人工对订单分类费时、低效, 难以满足现代物流要求的效率标准。为了提升物流订单分类的性能, 该文提出了一种基于图卷积神经网络 (graph convolution network, GCN) 和 RoBERTa 预训练语言模型的订单分类方法。首先, 基于物流订单文本的抽象语义表示 (abstract meaning representation, AMR) 结果和关键词构建全局 AMR 图, 并使用图卷积神经网络对全局 AMR 图进行特征提取, 获取订单文本的全局 AMR 图表示向量; 其次, 基于 AMR 算法构建物流订单文本分句的局部 AMR 图集合, 然后使用堆叠 GCN 处理图集合得到订单文本局部 AMR 图表示向量; 再次, 使用 RoBERTa 模型处理物流订单文本, 得到文本语义表示向量; 最后, 融合三种类型的文本表示向量完成物流订单分类。实验结果表明, 该方法在多项评价指标上优于其他基线方法。消融实验结果也验证了该分类方法各模块的有效性。

关键词: 订单分类; 图卷积神经网络; 抽象语义表示; RoBERTa 模型; 特征提取

中图分类号: TP391

文献标识码: A

文章编号: 1673-629X(2023)10-0195-07

doi: 10.3969/j.issn.1673-629X.2023.10.030

A Logistics Order Classification Method Based on Graph Convolution Network and RoBERTa

WANG Jian-bing, YANG Chao*, LIU Fang-fang, HUANG Jian, XIANG Yong

(Anhui Port Logistics Co., Ltd., Tongling 244000, China)

Abstract: Order information runs through all links of the logistics supply chain. Efficient order processing is the key to ensure the quality of logistics services and operational efficiency. With the trend of growing and differentiated customer logistics orders, manual order classification is time-consuming and inefficient, which is difficult to meet the efficiency standards required by modern logistics. A classification method is proposed to improve the performance of logistics order classification. Firstly, the global AMR graph is constructed based on the abstract meaning representation result and keywords of the logistics order text, and the feature extraction of the global AMR graph is carried out using the graph convolution network to obtain the global AMR graph representation vector. Secondly, the local AMR graph set of the logistics order text is built based on the AMR algorithm, and the stacked-GCN is used to process the graph to obtain the representation vector of local AMR graph. Thirdly, the RoBERTa model is used to process the logistics order text with the purpose of obtaining the text semantic representation vector. Finally, three types of the logistics order text representation vectors are fused to finish the classification task. Experimental results show that the proposed method is superior to other baseline methods in many evaluation indexes. Ablation experiments also verify the effectiveness of each proposed module in the classification method.

Key words: order classification; graph convolution neural network; abstract meaning representation; RoBERTa model; feature extraction

0 引言

近年来, 物流业运营模式逐渐从外包型向综合型转变, 信息系统在物流企业的运营管理中扮演了关键角色。作为物流信息系统的中枢神经, 订单子系统在提升企业运营效率、改善客户服务质量等方面发挥了积极作用^[1]。个性化物流服务对订单操作流程提出了

更高要求, 物流企业亟需提高差异化订单的处理效率。通常而言, 大型物流企业每天要受理成千上万个订单。对于每一个物流订单, 客服人员需要根据不同的起运地、目的地、货物清单、服务要求等信息对订单进行拆解与分类, 工作量巨大且容易出错^[2]。依靠人工处理海量差异化订单, 难以达到现代物流服务的效率标准。

收稿日期: 2022-12-19

修回日期: 2023-04-20

基金项目: 安徽省港航集团立项科技项目 (GHKJ2022-W-1.3.3)

作者简介: 王建兵 (1970-), 男, 正高级工程师, 硕士, 研究方向为大数据、交通运输; 通信作者: 杨超 (1989-), 男, 工程师, 研究方向为物流与供应链管理。

因此,研究智能化物流订单分类对于降低物流服务周期、提升客户满意度具有重要意义。物流供应链服务一般包括公路运输、水路运输、多式联运等业务类型,而多式联运业务又包括公水、铁水、公铁等类型^[3]。客户通过物流平台提供的微信小程序以文本形式提交一站式委托订单,平台客服首先对订单文本进行分解,然后结合历史订单路线、最优路线完成物流订单分类,最后根据订单类型分拨到对应的业务系统以完成服务受理。为高效完成订单分类,平台客服不但需要对客户委托有深入的理解,还需要熟悉各种物流业务类型。然而,平台客服往往缺乏系统的业务培训,使得物流订单分类往往存在错误,造成了不必要的二次分拨。

支持向量机、决策树、朴素贝叶斯等机器学习方法能够完成工单分类,但特征的分析 and 选择使得特征工程较为复杂,可能会出现入模特征和指定任务不相关的情况^[4]。RoBERTa 预训练语言模型在文本特征的特征提取方面具有较好的优势,可以有效地提取文本的上下文信息,从而实现中文文本语义的向量化表示^[5]。然而 RoBERTa 模型对于长文本处理能力较弱,而基于图结构的神经网络模型,如图卷积神经网络(Graph Convolutional Network, GCN),可以根据文本图有效提取文本的全局和局部图特征^[6]。为了克服人工分类效率低下且容易出错的问题,该文提出了一种基于 GCN 和 RoBERTa 模型的物流订单分类方法。该方法使用抽象语义表示(Abstract Meaning Representation, AMR)解析订单文本的每个句子,获取输入文本多个 AMR 图以形成局部 AMR 图;根据输入文本的关键词以及 AMR 图的根节点构建全局 AMR 图,并使用 GCN 网络和堆叠 GCN(stacked-GCN)网络来提取全局和局部 AMR 图的特征;通过 RoBERTa 模型提取订单文本语义特征,最终通过融合特征来完成物流订单分类。

1 相关研究

1.1 工单分类

文献[7]提出了一种基于 ResNet-BiLSTM 的电力客服工单分类模型,该模型利用残差网络学习句内的细节特征,再通过 BiLSTM 学习句间的上下文关联信息,最终得到工单的类别预测结果。文献[8]提出了一种基于事件提取的政务热线工单分类方法,该方法通过由 CNN-BiGRU-Self-Attention 定义的特征提取层获取工单文本的局部特征和全局特征完成工单分类。文献[9]提出一种基于矩阵分解和注意力的多任务学习方法,实现了运营商客服快速准确地对多层级的投诉工单文本进行分类。文献[10]使用 word2vec 模型对银行工单文本进行词嵌入化表示,使用深度学

习下的 TextCNN 模型进行文本分类和工单判定。文献[11]利用 TF-IDF 方法对经过 word2vec 模型处理后的词嵌入向量进行加权,在 TextCNN 模型中进行训练后利用分类器自动完成银行工单类型判断。

1.2 图卷积神经网络

文献[12]针对在线知识社区中回答者用户之间的协作行为,通过构建基于图卷积神经网络的链接预测模型,对在线知识社区中回答者用户的协作行为进行预测。文献[13]根据交互历史构建读者-图书二部图,搭建图卷积神经网络,通过连续的卷积层捕获二部图的高阶连通性来得到读者的邻域偏好信息以实现图书推荐方法。文献[14]提出了一种基于图卷积网络的专利摘要自动生成方法,旨在通过专利的权利要求书及其结构信息来生成专利摘要。文献[15]提出了一种具有替代训练算法的自调优 GCN 方法,实现通过超参数优化来自动化训练 GCN 模型,从而可以通过自动选择参数的方式来缓解传统的 GCN 模型存在过拟合和过平滑的问题。文献[16]通过预训练语言模型生成全文句子之间的注意力矩阵,并将其作为文本全连通图的加权邻接矩阵,将 GCN 应用于文本图对每个节点进行分类,从而在文本中找出突出句子,最终生成文本摘要。

1.3 RoBERTa 模型

文献[17]针对中文任务对 RoBERTa 模型进行了改进,使用了针对中文的 Whole Word Masking(WWM)训练策略,在不改变其他训练策略的基础上,提升了 RoBERTa 模型在中文任务上的实验效果。文献[18]将多目标类别情感分析转换为多个子任务,使用 RoBERTa 模型从文本和目标短语中提取特征信息,并利用交叉注意力机制找出与给定目标类别最相关的特征。文献[19]提出一种民间文学文本预训练模型 MythBERT,并与 BERT、BERT-WWM 和 RoBERTa 等主流中文预训练模型在情感分析、语义相似度、命名实体识别和问答四个自然语言处理任务上进行比较。文献[20]使用 RoBERTa 训练出荷兰语言模型 RobBERT,实验结果表明在一系列荷兰语自然语言处理任务上 RobBERT 语言模型的性能要超越其他语言模型的性能,尤其在小数据集上表现地更为突出。文献[21]在情感识别方面比较了 BERT、RoBERTa、XLNet 等主流预训练语言模型,研究发现采用预训练语言模型学习到的词向量相比以往模型能够获得更多的上下文语义信息。

2 模型结构

该文提出的物流订单分类模型主要由基于全局 AMR 图的 GCN 特征提取层、基于局部 AMR 图的堆

叠 GCN 特征提取层、基于 RoBERTa 语言模型的语义编码层和订单分类层构成,如图 1 所示。对于给定的物流订单文本,首先采用全局 AMR 算法构建全局 AMR 图,并使用 GCN 网络对全局 AMR 图的节点进行特征提取,获取文本的全局 AMR 图表示向量;其次,采用局部 AMR 算法构建多个 AMR 图,然后使用

堆叠 GCN 网络提取多个 AMR 图的特征,并进行融合得到文本局部 AMR 图表示向量;再次,使用 RoBERTa 预训练语言模型提取订单文本的上下文语义特征,得到订单文本语义表示向量;最后,融合三种类型的文本表示向量,并使用全连接网络(Full Connection, FC)结合 sigmoid 函数完成订单分类。

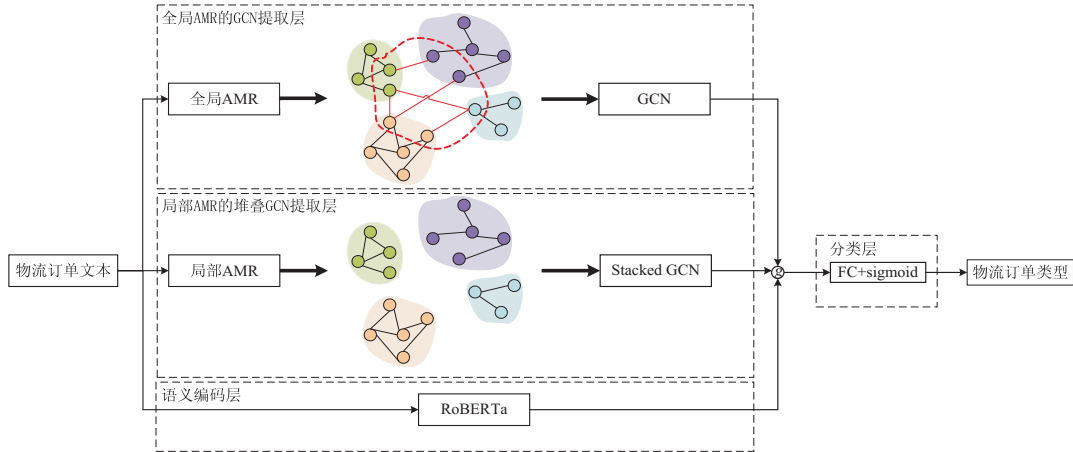


图 1 模型结构

2.1 基于全局 AMR 图的 GCN 特征提取层

从物流订单文本中提取特征信息的一个关键环节是捕捉句内和句间的特征。对于句子之间的相关特征,该文采用全局 AMR 算法处理物流订单文本生成该文本的全局 AMR 图,然后使用 GCN 网络来编码全局 AMR 图,以此得到订单文本句子之间的相关特征和订单文本的主题特征。

2.1.1 全局 AMR 图构建

AMR 是一种全新的领域无关的句子语义表示方法,它将一个句子的语义抽象为一个单根有向无环图^[22]。图 2 给出了一个句子“订单发货地点是铜陵市港口物流公司”的 AMR 图表示的示例,一个自然的句子可以通过 AMR 解析器被解析成一个 AMR 图 $G = (V, E)$ 。 V 表示句子中的实词抽象的概念节点,而边代表一个特定的实词之间的关系(抽象为带有语义关系标签的有向边,且忽略虚词和形态变化体现较虚的语义)。因此,AMR 侧重于语义关系而不是语法的,这种表示方式更有利于理解物流订单,且这种结构更接近订单的“触发词和角色参数(trigger-arguments)”结构。

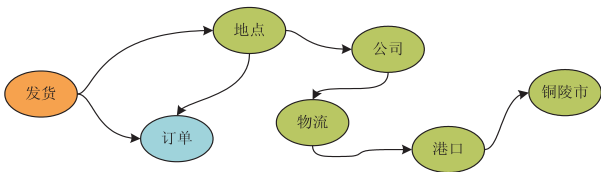


图 2 AMR 图示例

由于物流订单文本存在长短不一问题,且订单文本中的一些句子与订单主题并不相关。因此,该文提

取订单文本的关键词以及 AMR 的根节点作为全局 AMR 图的串节点,构建全局关系图,这样可以提取订单的主题信息和订单文本句子之间的关联关系。句间文本图的构建过程如下:

Step1 对订单文本进行分句处理得到 n 个句子,使用哈工大 NLP 工具包中关键词提取算法提取订单文本每个句子的关键词,得到 n 个句子的关键词集合 $\{K_i\}$;

Step2 对订单文本的 n 个句子进行分词处理,使用 AMR 算法将订单文本的每个句子抽象成 n 个 AMR 图;

Step3 对 n 个句子的 AMR 根节点进行对比,同样语义的根节点进行融合;

Step4 对 n 个句子的关键词集合 $\{K_i\}$ 进行分析处理,不同句子之间存在相同语义的关键词进行融合;

Step5 根据 n 个句子的融合结果构建 AMR 全局图。

2.1.2 GCN 特征提取

在 GCN 特征提取前,需要对构建的物流订单文本关系图进行节点编码。对于给定的物流订单文本 T ,节点编码首先需要对其进行分词,获取订单文本的分词序列 T_q ,然后使用词嵌入的方式对序列 T_q 进行编码,获取编码序列 E_q :

$$E_q = \{\xi_1, \xi_2, \dots, \xi_i^k, \dots, \xi_j^r, \dots\}$$

其中, ξ_i^k 和 ξ_j^r 分别是订单文本 T 中的关键词和 AMR 根节点的词嵌入向量, ξ_m 是其他词的嵌入向量。同时,为了编码关键词和 AMR 根节点在订单文本 T 中

的位置信息,对订单文本进行位置编码,编码结果如下:

$$P = \{p_1, p_2, \dots, p_i^k, \dots, p_j^r, \dots\}$$

其中, p_i^k 和 p_j^r 分别是订单文本 T 中的关键词和 AMR 根节点的位置嵌入向量, p_m 是其他词的嵌入向量。最终的节点编码向量是对词嵌入向量和位置嵌入向量的加权平均,计算公式如下所示:

$$e_m = \lambda_1 \xi_m + \lambda_2 p_m \quad (1)$$

$$e_i^k = \lambda_1 \xi_i^k + \lambda_2 p_i^k \quad (2)$$

$$e_j^r = \lambda_1 \xi_j^r + \lambda_2 p_j^r \quad (3)$$

物流订单文本序列 T_q 的最终编码向量为 $E = \{e_1, e_2, \dots, e_i^k, \dots, e_j^r, \dots\}$, 根据 AMR 图中的节点,在订单文本序列编码对 $\langle T_q, E \rangle$ 中查找对应的节点词编码向量作为全局 AMR 图节点的嵌入向量 $V = \{v_1, v_2, \dots, v_i, \dots\}$ 。该文采用 GCN 网络进行特征提取,从而可以充分提取全局 AMR 图结构特征。使用的图特征编码网络遵循 Kipf and Welling^[23] 提出的 GCN 逐层线性传播规则:

$$X^{(l+1)} = f(X^{(l)}, A) = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} X^{(l)} W^{(l)}) \quad (4)$$

其中, $\tilde{A} = A + I$, $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$, A 是输入图的邻接矩阵, I 是单位矩阵, $X^{(0)} = X$ 是初始输入特征矩阵。因此,对于文中的全局 AMR 图节点的嵌入向量 V ,使用 GCN 网络提取特征,计算得到输入订单文本的全局 AMR 特征表示向量 $V_{\text{global}}^{\text{all}}$,其计算公式如下:

$$V_{\text{global}}^{\text{all}} = \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} \text{ReLU}(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} V W^{(0)}) W^{(1)} \quad (5)$$

其中, $W^{(0)}$ 是输入层到隐藏层的权重矩阵, $W^{(1)}$ 是隐藏层到输出层的权重矩阵。

2.2 基于局部 AMR 图的堆叠 GCN 特征提取层

如图3所示,对于句子内部的相关特征,该文采用局部 AMR 算法处理物流订单文本生成局部 AMR 图,然后使用多层堆叠 GCN 网络来编码局部 AMR 图,并对编码结果进行融合,以得到订单文本句子内部的相关特征的融合结果。

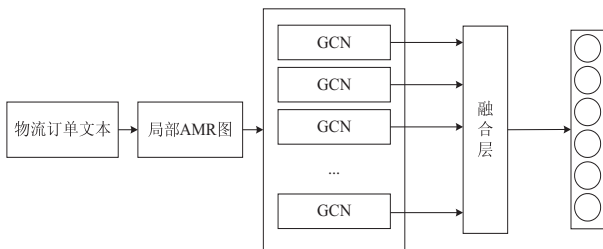


图3 基于局部 AMR 图的堆叠 GCN 特征提取

2.2.1 局部 AMR 图构建

由于订单文本中的各句子对订单分类结果的重要程度不一样,且存在某些句子与订单类型的关联性不

强(如收货联系人信息),因此,该文使用 AMR 算法对订单文本的每个句子进行处理,得到多个句子的 AMR 图。具体步骤如下:

Step1 对订单文本进行分句处理,得到 n 个句子;

Step2 对 n 个句子分别使用 AMR 算法进行处理得到 n 个 AMR 图;

Step3 将 n 个 AMR 图构成的集合形成订单文本的局部 AMR 图。

2.2.2 堆叠 GCN 融合网络

对于构建的局部 AMR 图 $G_{\text{local}} = \{G_1, G_2, \dots, G_n\}$, $G_i = \{V_i, E_i\}$, 对每个 G_i 使用词嵌入和位置嵌入的方式进行编码处理,得到对应的词嵌入序列 E_q^i 和位置嵌入序列 P_q^i ,采用加权平均的方式将词嵌入序列 E_q^i 和位置嵌入序列 P_q^i 进行融合,最终得到对应的表示向量序列 E^i 。根据 AMR 图 G_i 中的节点,在输入文本序列编码对中找到对应的节点词编码向量作为 G_i 节点的嵌入向量 $V = \{v_1, v_2, \dots, v_i, \dots\}$,所有 G_i 节点的嵌入向量构成了局部 AMR 图的嵌入向量集合 $V_{\text{local}} = \{V_1, V_2, \dots, V_n\}$ 。对局部 AMR 图的嵌入向量集合 V_{local} 中的每个元素使用 GCN 网络分别进行特征提取:

$$V_i^{\text{GCN}} = \hat{A}_i \text{ReLU}(\hat{A}_i V_i W_i^{(0)}) W_i^{(1)} \quad (6)$$

$$\hat{A}_i = \tilde{D}^{-\frac{1}{2}} \tilde{A}_i \tilde{D}^{-\frac{1}{2}} \quad (7)$$

其中, \hat{A}_i 是局部 AMR 图中第 i 个图的邻接矩阵。多个句子的表示向量构成局部 AMR 图的特征表示向量结合 $V_{\text{local}}^{\text{GCN}} = \{V_1^{\text{GCN}}, V_2^{\text{GCN}}, \dots, V_n^{\text{GCN}}\}$,最后使用门控机制来融合这些表示向量:

$$V_{\text{local}}^{\text{all}} = W_{\text{local}} [V_1^{\text{GCN}}, V_2^{\text{GCN}}, \dots, V_n^{\text{GCN}}] + B \quad (8)$$

其中, W_{local} 和 B 均为可学习参数。融合向量 $V_{\text{local}}^{\text{all}}$ 作为订单文本的局部 AMR 图表示向量。

2.3 基于 RoBERTa 模型的语义编码层

RoBERTa 模型是基于 BERT 的改进模型,相较于 BERT,它拥有更大的模型参数、更大的 batch size、更大规模的训练数据,同时删除了下一个句子预测 (NSP) 任务。这使得 RoBERTa 语言模型能够比 BERT 更好地应用到下游任务,取得更好的模型效果。

该文使用 RoBERTa 预训练语言模型对订单文本进行特征提取,以获取订单文本的上下文语义表征。由于文中任务是中文订单文本,因此使用哈工大讯飞联合实验室发布的 RoBERTa-wwm-ext 作为特征提取模型。如图4所示,对于给定的订单文本序列 $X = \{x_1, x_2, \dots, x_n\}$,其中 x_i 是订单文本中的每个字, x_i 经过字符嵌入、分割嵌入和位置嵌入编码并进行融合得到最终的融合嵌入向量,将 X 的所有融合嵌入向量构成的序列送入 RoBERTa-wwm-ext 模型。RoBERTa-wwm-ext 对融合嵌入向量序列进行特征编码,提取输

入序列 X 的上下文语义信息,最终输出订单文本每个字对应的语义表示向量 $V_{lm} = \{v_1^{lm}, v_2^{lm}, \dots, v_s^{lm}\}$ 。将 s

个语义表示向量进行融合,最终得到订单文本的表示向量 V_{lm}^{text} 。

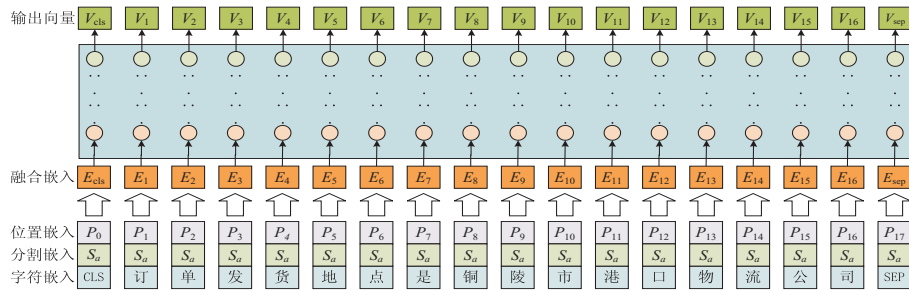


图4 RoBERTa 特征提取

2.4 订单类别预测

三种类型的订单表示向量 V_{global}^{all} 、 V_{local}^{all} 、 V_{lm}^{text} 进行融合得到 V_{concat} 。

$$V_{concat} = \omega_1 V_{global}^{all} + \omega_2 V_{local}^{all} + \omega_3 V_{lm}^{text} \quad (9)$$

其中, $\omega_1, \omega_2, \omega_3$ 是融合参数, 满足 $\omega_1 + \omega_2 + \omega_3 = 1$ 。将 V_{concat} 输入分类器中完成订单的多标签分类:

$$p = \text{sigmoid}(WV_{concat} + b) \quad (10)$$

其中, W 和 b 是可学习参数, p 是各类别的分类预测概率。使用多标签分类的交叉熵作为训练损失函数:

$$L = - \sum_i (y_i \log y_i + (1 - y_i) \log(1 - y_i)) \quad (11)$$

其中, y_i 为真实标签, y_i' 为预测结果, i 为订单的分类类别。

3 实验结果与分析

3.1 实验环境

该文使用基于 CUDA 11.0 的深度学习框架 pytorch 1.7.1 构建网络模型, 实验平台为内存 64G, 显存 24G 的 Ubuntu 18.04 LTS 系统。

3.2 模型训练过程

该文提出的物流订单分类模型训练流程如图 5 所示。

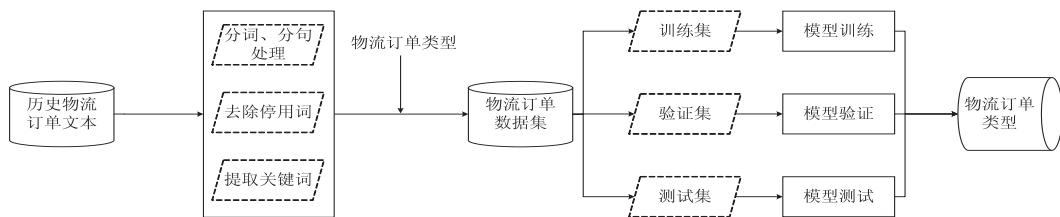


图5 模型训练过程

首先, 通过对历史物流订单文本进行分词、分句处理、去除停用词以及提取关键词后, 结合历史物流订单的处理结果构建物流订单数据集。然后, 将数据集按一定比例划分为训练集、验证集和测试集, 其中训练集用于模型训练, 验证集通过不断迭代更新模型性能, 测试集用来评估模型性能。最后, 使用训练好的模型进行物流订单类型预测。

3.3 数据集

从 2017 年 1 月 1 日-2021 年 12 月 31 日期间安徽港口物流有限公司历史物流订单中挑选了 30 000 条订单构建了实验数据集, 如表 1 所示。数据集包含订单文本数据和对应的订单类型, 其中订单文本是客户物流委托内容, 订单类型是由客服根据订单文本标注所得。同时, 对 30 000 条订单文本进行了统计分析, 这些物流订单文本长度均值为 276 个字。

表1 数据集描述

订单总量	训练集/验证集/测试集	订单文本均长
30 000	20 000/5 000/5 000	276

3.4 超参数设置

在整体网络训练过程中, 文中模型的超参数如表 2 所示。

表2 超参数设置

参数名	参数值
Epochs	10
BatchSize	32
Optimizer	Adam
LearningRate	1e-05
LearningRateDecay	0.5
GCN EmbeddingDim	768

3.5 基线对比实验

该文采用精确率 (Precision)、召回率 (Recall)、F1 值、准确率 (Accuracy) 以及汉明损失 (HammingLoss) 作为物流订单分类性能的评价指标。为了验证文中物流订单分类方法的性能, 与多种基线方法进行了对比。

• TextCNN^[24]: 使用预训练词向量来编码输入文本, 然后采用卷积神经网络提取订单文本的嵌入向量

来获取输入文本的特征,最终采用全连接网络结合 sigmoid 函数完成订单分类。

- HAN^[25]:使用基于词级别的 BiGRU+Attention 和句子级别的 BiGRU+Attention 模型来提取订单文本多层次语义特征,然后采用 sigmoid 函数完成订单分类。

- TextGCN^[26]:首先基于词共现和文档词关系构建语料库的文本图,然后使用 one-hot 编码对构建的文本图进行编码,最终采用 GCN 网络提取编码后的文本图完成订单分类。

- XLNet^[27]:使用哈工大讯飞联合实验室训练的中文 XLNet 预训练语言模型(chinese-xlnet-base)提取订单文本特征,进行 fine-tuning 后应用订单分类任务。

- RoBERTa^[17]:使用哈工大讯飞联合实验室训练的中文 RoBERTa 预训练语言模型(chinese-roberta-wwm-ext)提取输入文本特征,进行 fine-tuning 后应用到订单分类任务。

- BERT-GCN^[28]:基于词节点与文档节点构建异质图,采用 BERT 预训练模型初始化文档节点,联合训练 BERT 模块和 GCN 模块完成订单分类。

表 3 基线对比结果

分类方法	Precision	Recall	F1	Accuracy	Hamming Loss
TextCNN	0.718 5	0.776 1	0.746 2	0.707 5	0.180 83
HAN	0.738 4	0.798 1	0.767 1	0.731 0	0.166 00
TextGCN	0.765 8	0.820 2	0.792 1	0.763 8	0.147 47
XLNet	0.781 4	0.829 3	0.804 6	0.778 5	0.137 93
RoBERTa	0.797 8	0.844 7	0.820 6	0.797 9	0.126 50
BERT-GCN	0.818 0	0.863 0	0.839 9	0.817 3	0.112 67
文中方法	0.846 6	0.884 9	0.865 3	0.849 5	0.094 37

基线对比实验结果如表 3 所示。从表中可以看出,该文提出的物流订单分类方法在各项指标上均优于其他基线模型。值得注意的是,TextCNN 使用传统的 CNN 网络提取文本特征并进行分类,效果不佳,这是因为 CNN 网络仅仅能提取文本的局部特征,从而无法获取文本上下文信息。HAN 由于是用双向 GRU 并结合 Attention 机制提取上下文语义信息,但是缺乏对文本局部特征的提取同时对于长句的特征提取能力也较差,因此仅取得了比 TextCNN 好的效果。对于 XLNet 和 RoBERTa 模型,可以有效提取订单文本的特征信息,因此取得了较好的效果。TextGCN 使用 GCN 对文本图进行编码,可以有效地获取文本的句法结构信息,更好地解决文本长度问题,然而它对文本上下文信息的提取能力较差。而 BERT-GCN 由于加入了

BERT 模块,从而达到了更优的效果。由于 RoBERTa 模型在预训练阶段充分利用大规模无标注数据,可以更好地掌握通用语言能力,在绝大多数任务上都能表现出超越传统模型对文本上下文语义的提取效果,并且对于长短不一的文本可以使用 GCN 网络来提取文本的主题特征,因此该文提出的基于多种层次的图结构的订单分类方法拥有更好的性能。

3.6 消融实验

为了说明物流订单分类模型各模块的有效性,进行了消融实验:

①移除基于全局 AMR 图的 GCN 特征提取层,仅使用剩余两个模块的特征表示向量,其他部分保持不变。

②移除基于局部 AMR 图的堆叠 GCN 特征提取层,仅使用剩余两个模块的特征表示向量,其他部分保持不变。

③移除基于 RoBERTa 模型的语义编码层,仅使用剩余两个模块的特征表示向量,其他部分保持不变。

物流订单分类模型各模块的消融实验结果如表 4 所示。可以看出,该文提出的分类模型各项评价指标均优于①、②和③(文中模型>②>①>③)。③效果最差,说明了 RoBERTa 模型在订单文本特征提取方面有较大贡献,因而可以取得较好的效果。由②>①可知,基于全局 AMR 图的 GCN 特征提取层比基于局部 AMR 图的堆叠 GCN 特征提取层对订单文本特征提取的效果更好。由此可见,文中模型的各个模块均可以有效提高物流订单分类的性能。

表 4 消融实验结果

分类方法	Precision	Recall	F1	Accuracy	Hamming Loss
①	0.809 6	0.853 4	0.830 9	0.808 0	0.118 93
②	0.818 7	0.860 7	0.839 2	0.816 4	0.113 00
③	0.776 9	0.827 0	0.801 2	0.772 1	0.140 60
文中方法	0.846 6	0.884 9	0.865 3	0.849 5	0.094 37

4 结束语

该文提出了一种基于图卷积神经网络和 RoBERTa 语言模型的物流订单分类方法。该方法通过提取订单文本的全局图、局部图以及文本语义的融合特征来实现物流订单分类。首先,基于全局 AMR 算法结合文本关键词构建全局 AMR 图,并使用 GCN 对提取全局 AMR 图的结构特征,获取订单文本的全局 AMR 图表示向量。其次,基于局部 AMR 算法处理订单文本分句,以生成局部 AMR 图集合,使用 stacked-GCN 处理局部 AMR 图集合,获取的局部 AMR 图表

示向量集,并将向量集进行融合得到订单文本的局部 AMR 图表示向量。再次,使用 RoBERTa 模型提取订单文本的上下文语义特征,得到订单文本的语义表示向量。最后,融合三种类型的订单文本表示向量,并使用全连接网络结合 sigmoid 函数完成订单分类。由于采用了 RoBERTa 模型作为物流订单文本的上下文语义特征提取模型,对于长文本会进行截断,从而丢失语义信息,可能会导致错误的订单分类。未来,将进一步研究如何降低 RoBERTa 模型对物流订单文本截断所带来的性能影响。

参考文献:

- [1] 张羽政. 人工智能技术在现代物流管理模式中的作用探究[J]. 物流科技, 2022, 45(10): 105-107.
- [2] 杨炳康. 物流企业运营管理效率研究[J]. 中国储运, 2021(7): 196-198.
- [3] 马妙明. 多式联运信息集成平台的构建[J]. 微型电脑应用, 2022, 38(3): 52-54.
- [4] NASIR J A, KHAN O S, VARLAMISC I. Fake news detection: a hybrid CNN-RNN based deep learning approach[J]. International Journal of Information Management Data Insights, 2021, 1(1): 1-13.
- [5] LIU Y H, OTT M, GOYAL N, et al. RoBERTa: a robustly optimized BERT pretraining approach[EB/OL]. arXiv: 1907.11692, 2019.
- [6] 吴博, 梁循, 张树森, 等. 图神经网络前沿进展与应用[J]. 计算机学报, 2022, 45(1): 35-68.
- [7] 黄秀彬, 许世辉, 赵阳, 等. 基于 ResNet-BiLSTM 模型的电力客服工单分类研究[J]. 电子设计工程, 2022, 30(22): 179-183.
- [8] 陈钢. 融合 RoBERTa 和特征提取的政务热线工单分类[J]. 计算机与现代化, 2022(6): 21-26.
- [9] 宋勇, 严志伟, 秦玉坤, 等. 基于矩阵分解和注意力多任务学习的客服投诉工单分类[J]. 电信科学, 2022, 38(2): 103-110.
- [10] 薛巍. 基于人工智能技术实现银行工单智能派单[J]. 金融电子化, 2019(4): 56-57.
- [11] 刘佳. 基于改进词嵌入表示和卷积神经网络的银行智能派单系统的研究[J]. 工业控制计算机, 2020, 33(4): 101-104.
- [12] 刘臣, 李秋, 郝宇辰. 基于图卷积神经网络的在线社区行为预测[J]. 计算机技术与发展, 2021, 31(4): 28-33.
- [13] 陈帆, 张文德, 刘田. 基于图卷积神经网络的图书推荐方法研究[J]. 情报探索, 2022(10): 1-5.
- [14] 李健智, 王红玲, 王中卿. 基于图卷积网络的专利摘要自动生成研究[J]. 计算机科学, 2022, 49(S1): 172-177.
- [15] ZHU R, TAO Z, LI Y, et al. Automated graph learning via population based self-tuning GCN[J]. arXiv: 2107.04713, 2021.
- [16] LIN Y C, MA J. On automatic text extractive summarization based on graph and pre-trained language model attention[J]. arXiv: 2110.04878, 2021.
- [17] CUI Y M, CHE X Y, LIU T, et al. Pre-training with whole word masking for Chinese BERT[J]. arXiv: 1906.08101, 2019.
- [18] LIAO W X, ZENG B, YIN X W, et al. An improved aspect-category sentiment analysis model for text sentiment analysis based on RoBERTa[J]. Applied Intelligence, 2021, 51(6): 3522-3533.
- [19] 陶慧丹, 段亮, 王筋辉, 等. 基于 BERT 的民间文学文本预训练模型[J]. 计算机技术与发展, 2022, 32(11): 164-170.
- [20] DELOBELLE P, WINTERS T, BERENDT B. RobBERT: a Dutch RoBERTa-based language model[J]. arXiv: 2001.06286, 2020.
- [21] CORTIZ D. Exploring transformers in emotion recognition: a comparison of BERT, DistillBERT, RoBERTa, XLNet and ELECTRA[J]. arXiv: 2104.02041, 2021.
- [22] XU D, LI J, ZHU M, et al. Improving AMR parsing with sequence-to-sequence pre-training[C]//Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP). Punta Cana: EMNLP, 2020: 2501-2511.
- [23] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[J]. arXiv: 1609.02907, 2016.
- [24] KIM Y. Convolutional neural networks for sentence classification[C]//Proceedings of the 2014 conference on empirical methods in natural language processing. Doha: ACM, 2014: 1746-1751.
- [25] YANG Z, YANG D, DYER C, et al. Hierarchical attention networks for document classification[C]//Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies, association for computational linguistics. San Diego: NAACL HTL, 2016: 1480-1489.
- [26] YAO L, MAO C, LUO Y. Graph convolutional networks for text classification[C]//33rd AAAI conference on artificial intelligence (AAAI-19). Hawaii: AAAI, 2019: 7370-7377.
- [27] YANG Z, DAI Z, YANG Y, et al. XLNet: generalized autoregressive pretraining for language understanding[C]//Proceedings of the 31th conference on advances in neural information processing systems. Long Beach: [s. n.], 2019: 5754-5764.
- [28] LIN Y, MENG Y X, SUN X F, et al. BertGCN: transductive text classification by combining GNN and BERT[J]. arXiv: 2105.05727, 2021.