



(12) 发明专利

(10) 授权公告号 CN 115018553 B

(45) 授权公告日 2024. 05. 07

(21) 申请号 202210756922.X

(22) 申请日 2022.06.30

(65) 同一申请的已公布的文献号

申请公布号 CN 115018553 A

(43) 申请公布日 2022.09.06

(73) 专利权人 东南大学

地址 210096 江苏省南京市玄武区四牌楼2号

(72) 发明人 王帅 梅洛瑜 刘尧畅 张心睿

(74) 专利代理机构 南京众联专利代理有限公司  
32206

专利代理师 叶倩

(51) Int. Cl.

G06Q 10/083 (2024.01)

G06Q 10/0631 (2023.01)

G06Q 10/04 (2023.01)

G06Q 30/0204 (2023.01)

G06Q 30/0601 (2023.01)

G06N 3/0464 (2023.01)

G06N 3/042 (2023.01)

G06N 3/0455 (2023.01)

G06F 18/22 (2023.01)

(56) 对比文件

CN 110991713 A, 2020.04.10

CN 113724504 A, 2021.11.30

WO 2021082811 A1, 2021.05.06

WO 2022007526 A1, 2022.01.13

审查员 申冰冰

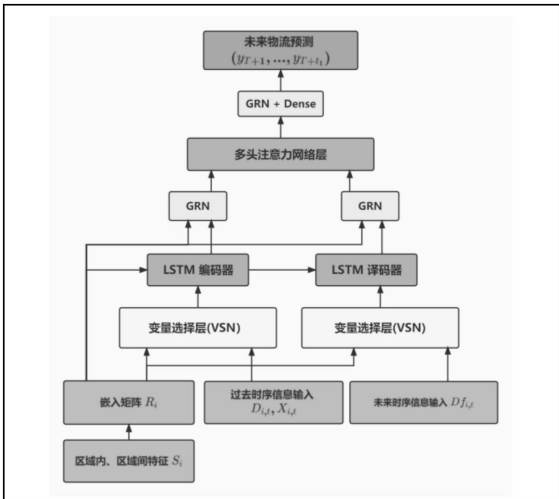
权利要求书4页 说明书10页 附图3页

(54) 发明名称

基于深度学习的区域物流单量预测系统及方法

(57) 摘要

本发明公开了基于深度学习的区域物流单量预测系统及方法,包括特征建模模块、嵌入模块和物流预测模块,特征建模模块用于提取和构建不同物流区域内、区域间的特征并进行特征建模,将每一个区域视为一个节点,通过各个区域之间的相似程度构建图;区域节点嵌入模块用于对每一种区域内、区域间特征进行各区域节点类别的划分,通过图卷积神经网络为每一个区域节点进行特征嵌入,训练完成后每一个特征图均可生成 $E \times N$ 的矩阵,作为预测模型的特征输入;物流预测模块根据区域节点嵌入模块得到的特征向量和时序特征进行长期物流预测。本方法综合考虑区域本身静态特征和相似区域的静态特征上下文信息,对区域级别物流量进行更为显著的长期且精准的预测。



1. 基于深度学习的区域物流单量预测系统,其特征在于,包括特征建模模块、嵌入模块和物流预测模块:

特征建模模块:所述模块用于提取和构建不同物流区区域内、区域间的特征并进行特征建模,将每一个区域视为一个节点,通过各个区域之间的相似程度构建图;

所述构建图的过程具体为:针对每一种特征使用Top-K算法筛选出与每个区域节点最相似的K个节点,并将这些节点用边相连接,遍历每一个区域节点并添加边后,完成了图的构建;其中,提取的特征至少包括区域人口、性别比例、购买力、区域类型、年龄结构和区域级别,并依次生成特征对应的图;区域人口、性别比例和购买力特征下的区域之间相似程度计算方法为:

$$S_{cn} = 1 - \frac{|c_i - c_j|}{c_{max} - c_{min}}$$

其中, $c_i$ 和 $c_j$ 分别代表两个区域的特征数值, $c_{max}$ 和 $c_{min}$ 分别代表所有区域中的该特征下的最大值和最小值;

所述区域级别特征下的区域之间相似程度计算方法为:

$$S_{dn} = \begin{cases} 1, & \text{if } d_i = d_j \text{ or } d_i = d_j \pm 1 \\ 0, & \text{else} \end{cases}$$

其中, $d_i$ , $d_j$ 分别代表两个区域之间的离散特征数值;

所述年龄结构特征下的区域之间相似度计算方法为:将两个区域之间的年龄结构视作两个集合,通过计算二者之间的交并比来定义;

所述区域类型特征下的区域之间相似度计算方法为:通过一个9维向量来描述每个区域的各路区类型组成情况,计算两个区域之间的余弦相似度来定义;

区域节点嵌入模块:所述模块用于对每一种区域内、区域间特征进行各区域节点类别的划分,最后通过图卷积神经网络为每一个区域节点进行特征嵌入,训练完成后每一个特征图均可生成 $E \times N$ 的矩阵,作为物流预测模块模型的特征输入;所述卷积神经网络的公式如下:

$$H^{l+1} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^l W^l)$$

其中, $A \sim = I + A$ , $A$ 是图的邻接矩阵, $I$ 为单位矩阵, $D \sim$ 代表 $A \sim$ 的度矩阵, $\sigma$ 为激活函数,而 $H^1$ 则为第1层网络的特征向量, $W^1$ 为参数;

物流预测模块:所述模块根据区域节点嵌入模块得到的特征向量和时序特征进行长期物流预测;所述模块至少包含一组编码、解码器和特征选择子模块,特征选择子模块通过训练得到每个特征的重要程度并进行软性的特征选择,随后将处理过的特征输入编码器、解码器,编解码器根据特征和过去物流量信息给出未来长期的物流单量预测。

2. 如权利要求1所述的基于深度学习的区域物流单量预测系统,其特征在于:所述区域节点嵌入模块中,在输入图卷积网络进行特征嵌入之前,利用监督学习训练模型对特征图优化,进行各区域节点类别的划分;对于区域人口、性别比例、购买力和年龄结构四种特征,通过取平均值的方式即可进行区域节点类别的划分。

3. 基于深度学习的区域物流单量预测方法,该方法应用于如权利要求1所述的系统,其特征在于,包括如下步骤:

S1,提取和构建不同物流区区域内、区域间的特征并进行特征建模,将每一个区域视为一个节点,通过各个区域之间的相似程度构建图;所述提取的特征至少包括区域人口、性别比例、购买力、区域类型、年龄结构和区域级别,并依次生成特征对应的图;

S2,在完成了特征图的构建之后,通过图卷积神经网络对每一个特征图内的区域节点进行节点嵌入,即得到每个区域对应特征的特征向量;

S3,将区域内、区域间嵌入特征向量和未来时序特征向量输入物流预测模块的模型中,给出一个区域内过去时间的物流变化数据,即可得到预测得到的未来物流量数据。

4. 如权利要求3所述的基于深度学习的区域物流单量预测方法,其特征在于:所述步骤S1具体包括:

S11,特征选取:选择能够影响一个城市物流单量的特征进行静态特征建模,所述特征包括区域人口、性别比例、购买力、区域类型、年龄结构和区域级别;

S12,特征相似度定义与计算:在完成区域静态特征选择后,针对每一个特征的各个区域之间的相似度进行计算,其中,

区域人口、性别比例和购买力特征下的区域之间相似程度计算方法为:

$$S_{cn} = 1 - \frac{|c_i - c_j|}{c_{\max} - c_{\min}}$$

其中, $c_i$ 和 $c_j$ 分别代表两个区域的特征数值, $c_{\max}$ 和 $c_{\min}$ 分别代表所有区域中的该特征下的最大值和最小值;

所述区域级别特征下的区域之间相似程度计算方法为:

$$S_{dn} = \begin{cases} 1, & \text{if } d_i = d_j \text{ or } d_i = d_j \pm 1 \\ 0, & \text{else} \end{cases}$$

其中, $d_i$ , $d_j$ 分别代表两个区域之间的离散特征数值;

所述年龄结构特征下的区域之间相似度计算方法为:将两个区域之间的年龄结构视作两个集合,通过计算二者之间的交并比来定义;

所述区域类型特征下的区域之间相似度计算方法为:通过一个9维向量来描述每个区域的各路区类型组成情况,如果一个区拥有向量对应维度所代表的路区类型,则在向量的该维度上加1,然后通过计算两个区域类型向量之间的余弦相似度来定义区域类型特征相似度;

S13,特征图的构建:构建方式为将每个区域视作一个节点,针对每一类特征使用Top-K算法将与每个节点相似度最高的K个节点用边连接在一起,最终完成特征图的构建。

5. 如权利要求3所述的基于深度学习的区域物流单量预测方法,其特征在于:所述步骤S2具体包括:

S21,图节点分类:针对每一类特征进行节点分类,其中,针对区域人口、性别比例、购买力、年龄结构和区域级别这五类数值型特征,直接根据其数值大小将其分为8类;针对区域类型特征,使用无监督分类算法进行分类;

S22,图卷积神经网络训练:通过图神经网络的监督学习来优化每个区域节点最终的特征嵌入向量,所述图神经网络训练公式如下:

$$H^{l+1} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^l W^l)$$

其中, $A \sim = I + A$ , $A$ 是图的邻接矩阵, $I$ 为单位矩阵, $D \sim$ 代表 $A \sim$ 的度矩阵, $\sigma$ 为激活函数,而 $H^1$ 则为第1层网络的特征向量, $W^1$ 为参数;

S23,特征向量生成:实现针对每一个特征图的图神经网络训练后,将图神经网络的输出层输出的8维向量视作代表一个区域本身特征和区域间关联性特征的特征向量,该向量用作步骤S3中物流预测模块的输入。

6.如权利要求3所述的基于深度学习的区域物流单量预测方法,其特征在于:所述步骤S3预测物流模块中的模型运行具体包括:

S31,特征变量选择:将用于未来物流预测的过去、未来时序信息和步骤S2得到的特征向量输入物流预测模块的变量选择层中,

设定 $t$ 时间步的特征输入为 $F_t = [\xi_t^{(1)}, \dots, \xi_t^{(m)}]$ , $c_s$ 代表区域内、区域间特征嵌入,变量选择层的计算如下所示:

$$\tilde{\xi}_t^{(j)} = GRN_{\xi_t^{(j)}}(\xi_t^{(j)})$$

$$v_t = Softmax(GRN_{v_t}(F_1, c_s))$$

$$\tilde{\xi}_t = \sum_{i=0}^m v_t^{(j)} \tilde{\xi}_t^{(j)}$$

对于变量选择层中的GRN模块,该模块的输入为一个时序特征输入 $a$ 和区域内、区域间特征嵌入 $c$ ,该模块所做的计算如下所示:

$$GRN(a, c) = LayerNorm(a + GLU(\eta_1))$$

$$GLU(\eta_1) = \sigma(W_1 \eta_1 + b_1) \odot (W_2 \eta_1 + b_2)$$

$$\eta_1 = W_3 \eta_2 + b_3$$

$$\eta_2 = ELU(W_4 a + W_5 c + b_4)$$

其中, $ELU$ 为激活函数; $GLU$ 通过为每一个输入特征加权的方式来实现特征选择; $\sigma$ 是sigmoid函数; $W_1$ 到 $W_5$ 和 $b_1$ 到 $b_4$ 均为需要学习的网路参数;

S32,编码与译码:将经过特征选择层选择的特征过去时序信息和特征向量输入LSTM编码器中,将未来时序信息输入LSTM译码器中,进而得到LSTM编码器、译码器的初步输出;

S33,特征变换:将编码器、译码器的输出输入GRN模块中,进一步通过控制输入特征的非线性变换程度来增强模型的鲁棒性;

S34,多头注意力网络进行长期时序特征捕捉:通过GRN处理后输出输入到一个多头注意力网络层中,使用多头注意力网络共享 $V$ 参数来进一步强调关键特征;

S35,模型输出:模型最终输出形式为概率预测,即预测一个置信区间,模型损失函数定义如下:

$$L(q, \hat{y}, y) = \max(q * (y - \hat{y}), (1 - q) * (\hat{y} - y))$$

其中, $\hat{y}$ 和 $y$ 分别代表预测值和真实值, $q$ 为一个人为定义的分位数,根据分位数的不同

模型训练出不同的参数,进而导致预测值差异。

## 基于深度学习的区域物流单量预测系统及方法

### 技术领域

[0001] 本发明属于计算机计算技术领域,主要涉及了一种基于深度学习的区域物流单量预测系统及方法。

### 背景技术

[0002] 由于电子商务的蓬勃发展,订单投递量激增,对物流服务的承载量、质量、时效性提了更高的要求。目前包裹投递最常用的方法是通过配送员从配送网点接单后上门交付给消费者。但快递包裹的交付过程中往往存在各种问题,例如:由于无法精准预测下一周的营业部区域配送包裹数,进而无法对配送人员做出合理安排调度,使得配送效率下降,相关企业也会损失经济利益。相关研究表明,当某个区域的物流量短时间内激增后,往往会导致该区域的包裹运送时间大大延长,这使得特殊时期的物流服务往往会让消费者等待时间过长、无法按时完成配送任务。

[0003] 当下的电商物流预测的现有解决方案侧重于宏观平台预测,倾向于利用单源数据(例如,历史销售数据和日期类型)来模拟与时间相关的物流变化模型并预测物流量。这一类方法并没有考虑到区域本身的特征,例如区域人口、性别比例、购买力等静态特征。静态特征是指一些能够显著影响区域物流单量且长时间内不随时间变化而变化的特征。另一方面,对于区域需求预测场景下的问题,一些解决方案将区域本身的特征纳入预测,但是这一类研究往往将每个区域视作独立的个体,没有显式建模这类特征带来的区域之间的相似性,进而在研究中遗漏了区域关联性的信息。考虑到当今的区域细分物流场景下越来越多的线上和线下区域关联性,上述这两类研究显然不适合于进行多区域的物流需求预测。

[0004] 因此,末端物流的智慧化、便捷化、规范化是我国物流行业发展迫在眉睫的需求,鉴于更精准预测长时间物流变化而更好调度物流资源而带来的经济效益,通过计算机技术对物流单量的预测是十分有必要的。

### 发明内容

[0005] 本发明正是针对当下电商平台细分区域的物流配送场景,如何充分利用各个区域之间的内部特征、相似关联关系和时序特征,更加精确地预测区域物流需求单量的长期变化情况的问题,提供一种基于深度学习的区域物流单量预测系统及方法,包括特征建模模块、嵌入模块和物流预测模块,特征建模模块用于提取和构建不同物流区域内、区域间的特征并进行特征建模,将每一个区域视为一个节点,通过各个区域之间的相似程度构建图;区域节点嵌入模块用于对每一种区域内、区域间特征进行各区域节点类别的划分,最后通过图卷积神经网络为每一个区域节点进行特征嵌入,训练完成后每一个特征图均可生成 $E \times N$ 的矩阵,作为预测模型的特征输入;物流预测模块根据区域节点嵌入模块得到的特征向量和时序特征进行长期物流预测。本方法综合考虑区域本身静态特征和相似区域的静态特征上下文信息,一方面,考虑区域内与区域间特征提取与处理。在长期物流预测的过程中,区域本身特征会对物流量变化情况产生较大影响,针对这一类基本不随时间变化而变化的

特征,充分考虑如何从中抽象提取出有利于物流预测的信息,并在此基础上通过这些特征描述区域之间的相似关联性进而进一步优化预测精准;另一方面,考虑对物流进行的长期变化预测,部分区域的物流单量会呈现以一周为长度的单量周期性变化,捕捉到这种周期性变化也并以其为依据对物流进行长期变化预测是物流规划的难题之一,本发明对该方面也提出了有效的解决模型,对区域级别物流量进行更为显著的长期且精准的预测。

[0006] 为了实现上述目的,本发明采取的技术方案是:基于深度学习的区域物流单量预测系统,包括特征建模模块、嵌入模块和物流预测模块:

[0007] 特征建模模块:所述模块用于提取和构建不同物流区区域内、区域间的特征并进行特征建模,将每一个区域视为一个节点,通过各个区域之间的相似程度构建图;

[0008] 所述构建图的过程具体为:针对每一种特征使用Top-K算法筛选出与每个区域节点最相似的K个节点,并将这些节点用边相连接,遍历每一个区域节点并添加边后,完成了图的构建;

[0009] 区域节点嵌入模块:所述模块用于对每一种区域内、区域间特征进行各区域节点类别的划分,最后通过图卷积神经网络为每一个区域节点进行特征嵌入,训练完成后每一个特征图均可生成 $E \times N$ 的矩阵,作为物流预测模块模型的特征输入;所述卷积神经网络的公式如下:

$$[0010] \quad H^{l+1} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^l W^l)$$

[0011] 其中, $A \sim = I + A$ , $A$ 是图的邻接矩阵, $I$ 为单位矩阵, $D \sim$ 代表 $A \sim$ 的度矩阵, $\sigma$ 为激活函数,而 $H^1$ 则为第1层网络的特征向量, $W^1$ 为参数;

[0012] 物流预测模块:所述模块根据区域节点嵌入模块得到的特征向量和时序特征进行长期物流预测;所述模块至少包含一组编码、解码器和特征选择子模块,特征选择子模块通过训练得到每个特征的重要程度并进行软性的特征选择,随后将处理过的特征输入编码器、解码器,编解码器根据特征和过去物流量信息给出未来长期的物流单量预测。

[0013] 为了实现上述目的,本发明还采取的技术方案是:基于深度学习的区域物流单量预测方法,包括如下步骤:

[0014] S1,提取和构建不同物流区区域内、区域间的特征并进行特征建模,将每一个区域视为一个节点,通过各个区域之间的相似程度构建图;所述提取的特征至少包括区域人口、性别比例、购买力、区域类型、年龄结构和区域级别,并依次生成特征对应的图;

[0015] S2,在完成了特征图的构建之后,通过图卷积神经网络对每一个特征图内的区域节点进行节点嵌入,即得到每个区域对应特征的特征向量;

[0016] S3,将区域内、区域间嵌入特征向量和未来时序特征向量输入物流预测模块的模型中,给出一个区域内过去时间的物流变化数据,即可得到预测得到的未来物流量数据。

[0017] 与现有技术相比,本发明具有的有益效果:

[0018] (1) 综合考虑区域本身静态特征和相似区域的静态特征上下文信息,对区域级别物流量进行更为显著的长期且精准的预测。

[0019] (2) 通过区域特征,构建系列数据驱动模型,更精准预测各区域的物流量,可以有效改善末端物流的资源分配问题,使相关企业未雨绸缪,更好地调度相关资源,实现企业经济效益的最大化,也能够提高整个物流体系的服务质量。

[0020] (3) 开创性地将图卷积神经网络应用在长期时间序列预测问题中。

[0021] (4) 物流预测模型同时考虑了区域内、区域间特征和未来时序特征,在实现了特征融合的基础上机型未来物流量预测。

## 附图说明

[0022] 图1为本发明特征建模模块中,使用Top-K算法构建图的过程示意图;

[0023] 图2为本发明针对区域类型特征进行节点分类的无监督分类算法图;

[0024] 图3为本发明特征相关图经过训练后转化为特征矩阵的全过程示意图;

[0025] 图4为本发明物流预测模块中模型的架构图;

[0026] 图5为本发明物流预测模块中LSTM神经元结构示意图;

[0027] 图6为本发明实施例2的预测结果示意图。

## 具体实施方式

[0028] 下面结合附图和具体实施方式,进一步阐明本发明,应理解下述具体实施方式仅用于说明本发明而并不用于限制本发明的范围。

[0029] 实施例1

[0030] 基于深度学习的区域物流单量预测系统,包括特征建模模块、嵌入模块和物流预测模块:

[0031] 所述特征建模模块用于提取和构建不同物流区区域内、区域间的特征并进行特征建模,将每一个区域视为一个节点,通过各个区域之间的相似程度构建图;具体方法如下:

[0032] S11,针对选择最能够影响一个城市物流单量的特征进行特征建模,在本发明中,我们选择区域人口、性别比例、购买力、区域类型、年龄结构和区域级别六个区域特征来进行各个区域的静态特征建模;

[0033] S12,在完成区域静态特征选择后我们进行针对每一个特征的各个区域之间的相似度计算,各个特征的相似度计算如下:

[0034] 区域人口、性别比例和购买力特征下的区域之间相似程度计算方法为:

$$[0035] \quad S_{cn} = 1 - \frac{|c_i - c_j|}{c_{\max} - c_{\min}}$$

[0036] 其中, $c_i$ 和 $c_j$ 分别代表两个区域的特征数值, $c_{\max}$ 和 $c_{\min}$ 分别代表所有区域中的该特征下的最大值和最小值;

[0037] 所述区域级别特征下的区域之间相似程度计算方法为:

$$[0038] \quad S_{dn} = \begin{cases} 1, & \text{if } d_i = d_j \text{ or } d_i = d_j \pm 1 \\ 0, & \text{else} \end{cases}$$

[0039] 其中, $d_i$ , $d_j$ 分别代表两个区域之间的离散特征数值;

[0040] 所述年龄结构特征下的区域之间相似度计算方法为:将两个区域之间的年龄结构视作两个集合,通过计算二者之间的交并比来定义;

[0041] 所述区域类型特征下的区域之间相似度计算方法为:通过一个9维向量来描述每



个区域的各路区类型组成情况,如果一个区拥有向量对应维度所代表的路区类型,则在向量的该维度上加1。然后通过计算两个区域类型向量之间的余弦相似度来定义区域类型特征相似度;

[0042] S13,在完成了针对各个特征的相似度定义后,我们进行特征图的构建,构建方式为将每个区域视作一个节点,针对每一类特征使用Top-K算法将与每个节点相似度最高的K个节点用边连接在一起,如图1所示,最终完成特征图的构建。

[0043] 本方法中,选择出最能够代表一个区域物流单量需求程度的特征并将其转化为图数据结构。在图数据结构中,一个区域节点既保留了其原本的特征信息,又能够根据其相连的节点来体现出区域之间的相似性和关联性,因而使得预测能够考虑区域静态特征和区域间的相似性特征。

[0044] 区域节点嵌入模块用于对每一种区域内、区域间特征进行各区域节点类别的划分,最后通过图卷积神经网络为每一个区域节点进行特征嵌入,训练完成后每一个特征图均可生成 $E \times N$ 的矩阵,作为预测模型的特征输入;所述模块的工作步骤具体如下:

[0045] S21,针对每一类特征进行节点分类,具体的分类方法为:针对区域人口、性别比例、购买力、年龄结构和区域级别这五类数值型特征,直接根据其数值大小将其分为8类;针对区域类型特征,我们使用如图2所示的无监督分类算法进行其分类。

[0046] S22,在完成了针对各个特征类型的节点分类后,我们进行图神经网络的监督学习来优化每个区域节点最终的特征嵌入向量。图神经网络训练公式如下:

$$[0047] \quad H^{l+1} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^l W^l)$$

[0048] 其中, $A \sim = I + A$ , $A$ 是图的邻接矩阵, $I$ 为单位矩阵, $D \sim$ 代表 $A \sim$ 的度矩阵, $\sigma$ 为激活函数,而 $H^1$ 则为第1层网络的特征向量, $W^1$ 为参数;

[0049] S23,实现针对每一个特征图的图神经网络训练后,将图神经网络的输出层输出的8维向量视作可以代表一个区域本身特征和区域间关联性特征的特征向量,该向量会用作物流预测模块的输入。

[0050] 通过图神经网络将图数据结构的信息转换为向量信息,且该向量信息内同时包含了区域本身特征信息和区域间的关联性信息,该向量能够为下一步的区域物流量最终预测提供重依据。

[0051] 物流预测模块根据区域节点嵌入模块得到的特征向量和时序特征进行长期物流预测;所述模块至少包含一组编码、解码器和特征选择子模块,特征选择子模块通过训练得到每个特征的重要程度并进行软性的特征选择,随后将处理过的特征输入编码器、解码器,编解码器根据特征和过去物流量信息给出未来长期的物流单量预测,其具体工作步骤如下:

[0052] S31,将用于未来物流预测的过去、未来时序信息和S2步骤得到的特征向量输入变量选择层中,这一层所做的计算说明如下:

[0053] 我们设定 $t$ 时间步的特征输入为 $F_t = [\xi_t^{(1)}, \dots, \xi_t^{(m)}]$ , $c_s$ 代表区域内、区域间特征嵌入。变量选择层(Variable Selection Network)的计算如下所示:

$$[0054] \quad \tilde{\xi}_t^{(j)} = GRN_{\xi_t^{(j)}}(\xi_t^{(j)})$$

$$[0055] \quad v_t = \text{Softmax}(\text{GRN}_{v_t}(F_1, c_s))$$

$$[0056] \quad \tilde{\xi}_t = \sum_{i=0}^m v_t^{(i)} \xi_t^{(i)}$$

[0057] 对于变量选择层中的GRN模块(Gated Residual Network),该模块的输入为一个时序特征输入a和区域内、区域间特征嵌入c,该模块所做的计算如下所示:

$$[0058] \quad \text{GRN}(a, c) = \text{LayerNorm}(a + \text{GLU}(\eta_1))$$

$$[0059] \quad \text{GLU}(\eta_1) = \sigma(W_1\eta_1 + b_1) \odot (W_2\eta_1 + b_2)$$

$$[0060] \quad \eta_1 = W_3\eta_2 + b_3$$

$$[0061] \quad \eta_2 = \text{ELU}(W_4a + W_5c + b_4)$$

[0062] 在上述公式中,ELU为激活函数;GLU通过为每一个输入特征加权的方式来实现特征选择; $\sigma$ 是sigmoid函数。 $W_1$ 到 $W_5$ 和 $b_1$ 到 $b_4$ 均为需要学习的网路参数。

[0063] 我们可以将变量选择层模块视作针对时序特征和区域内、区域外特征的一种加权。通过该层神经网络,模型可以分辨出哪些特征对于提高预测准确率有着明显的正面影响,从而进一步关注于这种关键特征;

[0064] S32,将经过特征选择层选择的特征过去时序信息和特征向量输入LSTM编码器中,将未来时序信息输入LSTM译码器中,进而得到LSTM编码器、译码器的初步输出;

[0065] S33,将编码器、译码器的输出输入GRN模块中,进一步通过控制输入特征的非线性变换程度来增强模型的鲁棒性,从而实现软性特征选择,使模型在面对数据集中的噪声特征时有更好的表现;

[0066] S34,为了在长期物流量预测中捕捉到跨多个时间步的长期序列依赖,模型将通过GRN处理后输出输入到一个多头注意力网络层中。与普通多头神经网络的Q,V,K参数所不同的是,本模型使用的多头注意力网络共享V参数来进一步强调关键特征;

[0067] S35,模型最终输出形式为概率预测,即预测一个置信区间,模型损失函数定义如下:

$$[0068] \quad L(q, \hat{y}, y) = \max(q * (y - \hat{y}), (1 - q) * (\hat{y} - y))$$

[0069] 在上述公式中, $\hat{y}$ 和y分别代表预测值和真实值,q为一个人为定义的分位数,根据分位数的不同模型可以训练出不同的参数,进而导致预测值差异。利用该损失函数和历史数据训练整个模型后即可通过所述预测流程实现城市区域的未来物流单量预测。

[0070] 该物流预测模块基于时序特征和区域静态特征进行未来物流单量预测,实现了物流单量预测时的多特征融合。通过GRN模块和变量选择层来处理数据集中的噪声并提取出对物流预测真正重要的特征,实现一种软性的特征选择。在预测模型中,创新性使用基于添加门控机制的LSTM神经元的encoder-decoder结构,该神经元通过在内部加入了记忆遗忘过程,从而在处理时间序列长期依赖时可以获取更多的特征信息。

[0071] 为了进一步加强模型的鲁棒性,增强其对各类噪声信息的处理能力,本模型在编码译码器之后紧跟着加入一层GRN网络结构。

[0072] 最终,为了提取长期时序依赖特征,在模型倒数第二层放入多头注意力网络,该网络增强了模型对于跨时间步序列关系的捕捉,有利于物流量呈现周期性波动的区域的预

测,使得预测的过程中能过考虑到某些区域的物流周期性变化。

[0073] 实施例2

[0074] 本实施例使用京东北京地区共389个营业区域的2021年5月到12月的物流量数据。

[0075] 实验数据预处理和平台配置具体如下:

[0076] 针对原始数据,首先我们利用python语言中的pandas依赖包对4.51GB的物流配送数据进行逐个区域统计,进而得到每个区域2021年5月到12月内各个营业区域的每日物流量数据 $X_{i,t}$ 。然后我们根据脱敏后的用户信息数据统计出各个营业区域内的用户画像特征。

[0077] 用户画像特征包含上文中所提到的区域人口特征、性别比例特征、购买力特征和年龄结构特征。这一类用户画像特征从各个角度描述了区域的用户基本面,构成了区域内、区域间特征的主要组成部分。对于实验所需的营业部等级特征和区域类型特征,数据获取方法如下:

[0078] 首先我们根据京东公司对每个营业部的等级划分可以直接得到营业部等级特征。

[0079] 针对区域类型特征,我们可以通过北京市POI地理信息数据进行统计划分,进而将营业区域下的每一个路区都划分成九种区域类型之一,最终得到每个区域的类型数据。

[0080] 将上述用户画像特征、区域类型特征和营业部等级特征结合在一起后即可得到本文中所述的区域内、区间特征 $S_i = (S_{i1}, \dots, S_{iM})$ ,  $S_i \in \mathbb{R}^M$ 。针对于时序特征信息 $D_{i,t}$ 和 $D_{fi,t}$ ,我们根据2021年5月到12月的每一天的日期类型即可得到相关信息,其中购物节信息包含了京东平台内部定义的6.18、双11、双12等促销节。

[0081] 综上所述,本发明所使用的京东营业区域相关数据和各类日期时序数据可以很好地满足实验、模型需求。

[0082] 实验平台为VScode,使用语言python=3.7.9,预测模型基于pytorch=1.10.依赖包构建,图卷积神经网络基于tensorflow=2.4.0依赖包构建。在本实例中,我们设真实值为每个区域统计出的每日物流量数据。为了更好地反映出模型针对不同日期类型下的表现,我们选择5月、6月两个月的数据作为测试集,余下月份的数据作为训练集。

[0083] 对于模型参数设置,我们设置Top-K算法中的 $K=5$ 。图卷积神经网络被设计为4层,每一层的输出向量维度分别为32、16、8和 $x$ 。 $x$ 代表根据各个区域内、区域间特征分类后区域的种类数量,并将倒数第二层的8维向量视为区域节点嵌入向量。对于区域类型特征的分类算法,算法中的相似度阈值 $\tau$ 设置为0.90。在预测物流量所使用的Transformer模型中,实验中设置多头注意力的头数量为2,编码器和译码器长度均为5,分位数为 $q=[0.1, 0.5, 0.9]$ 。基于深度学习的区域物流单量预测方法,包括如下步骤:

[0084] S1:针对对区域内、区域间各类特征的图的构建

[0085] 为了有效地从区域内、区域间特征中提取相关区域关联性上下文特征,将每一个区域视为一个节点,并根据各个特征的区域相似度构建图。每一种区域内、区域间特征都构建一幅图,从而构建多个图 $G_s = \cup \{G_s\} = (V, E_s)$ 来代表这些特征,其中 $V$ 代表着表示各个区域 $r \in R$ 的节点集, $E_s$ 代表着根据不同的特征相似度计算方法构建出的边集。

[0086] 为了基于各个特征的相似度来构建图,首先需要根据每个特征的特点来设计相似度算法。本模型中所选的区域、区域间特征分别为区域人口、性别比例、购买力、区域类型、年龄结构和区域级别,根据不同特征的数据特点,针对性地设计如下区域相似度算法:

[0087] 连续数值特征相似度计算:对于数据特点是连续数值形式的特征,我们设计以下

公式来表达这类特征下两个区域之间的相似程度：

$$[0088] \quad S_{cn} = 1 - \frac{|c_i - c_j|}{c_{\max} - c_{\min}}$$

[0089] 在该公式中,  $c_i$  和  $c_j$  分别代表两个区域的特征数值,  $c_{\max}$  和  $c_{\min}$  分别代表所有区域中的该特征下的最大值和最小值。我们用公式来定义并计算区域人口、性别比例和购买力特征下区域之间的相似程度。

[0090] 离散数值特征相似度计算: 对于数据特点是连续数值形式的特征, 我们设计以下公式来表达这类特征下两个区域之间的相似程度:

$$[0091] \quad S_{dn} = \begin{cases} 1, & \text{if } d_i = d_j \text{ or } d_i = d_j \pm 1 \\ 0, & \text{else} \end{cases}$$

[0092] 此处  $d_i, d_j$  分别代表两个区域之间的离散特征数值。我们用上述公式来定义并计算区域级别特征下区域之间的相似程度。

[0093] 区间特征相似度计算: 由于区域年龄结构特征以区间范围的形式给出, 我们将两个区域之间的年龄结构视作两个集合, 进而通过计算二者之间的交并比来定义年龄结构特征下的区域之间的相似度。

[0094] 向量型特征相似度计算: 由于本文使用一个9维向量来描述每个区域的各路区类型组成情况, 于是我们直接通过计算两个区域之间的余弦相似度来定义区域类型特征下的区域之间的相似度。

[0095] 经过上述的相似度定义, 可以得到每一个区域内、区域间特征的相似度计算方法。这些计算方法均将区域之间的特征相似度转化为一个0到1的数值, 接近1代表两个区域之间的相似度越高。通过各个特征的相似度计算方法, 遍历计算不同特征下任意区域之间的相似度, 进而可以针对每个特征得到一个  $N \times N$  的相似度矩阵, 通过查询对应特征的矩阵元素  $m_{i,j}$  即可得到该特征下区域  $i$  和区域  $j$  之间的相似度。在完成了相似度定义和不同区域内、区域间特征下各区域两两之间的相似度计算后, 我们进行图的构建, 其构建过程如下:

[0096] 将每一个区域视为一个节点, 并通过各个区域之间的相似程度构建图, 具体过程为针对每一种特征使用Top-K算法筛选出与每个区域节点最相似的K个节点, 并将这些节点用边相连接。遍历每一个区域节点并添加边后, 我们就完成了图的构建。图1展示了使Top-K算法构建图的过程。

[0097] 最终, 我们通过将每一个预测区域视作图节点, 然后基于各类特征的相似度使用top-K算法进行节点连接的方式, 可以为区域人口、性别比例、购买力、区域类型、年龄结构和区域级别这六种特征依次生成特征对应的图。

[0098] S2: 区域节点嵌入方法

[0099] 在完成了针对各个区域内、区域间特征的图的构建后, 我们为每一个区域节点进行特征信息的嵌入。在使用图卷积神经网络进行特征嵌入之前, 为了可以通过监督学习训练模型从而不断地优化嵌入结果, 我们首先需要针对每一种区域内、区域间特征进行各区域节点类别的划分。

[0100] 对于数值类型的特征, 由于其数值本身即可代表该区域特征之间是否相似, 我们

通过其数值大小即可进行区域节点类别的划分,举例来讲,针对购买力特征,如果该特征的最大值是4.50,最小值为2.10,共需要分为8个类别,则最大值的区域即对应第8类区域,最小值的区域对应第1类区域,若有一个区域购买力为3.50,则其类别第 $[(3.5-2.1)/4.5-2.1) \times 8] = 5$ 类别。

[0101] 此处为了便于分类,我们将年龄结构特征数值通过取平均值的方式从区间类特征转换为连续型数值特征。因此,我们可以通过这种方法将区域人口特征、性别比例特征、年龄结构特征、购买力特征和区域类别特征进行分类。

[0102] 然而,上述分类方式并不适用于区域类型这种非数值型的向量类特征,因为我们不可能通过判断向量的模的大小的方式来得到关于区域类型的分类结果,例如,假设区域1的区域类型特征向量为 $\vec{n}_1 = (1, 2, 0, 0, 0, 0, 0, 0, 0)$ ,区域2的区域类型特征向量为 $\vec{n}_2 = (0, 0, 0, 1, 2, 0, 0, 0, 0)$ ,则区域1和区域2的向量模相等,但是由于这两个区域向量维度上的巨大差异,这两个模相等的区域却是完全不同类型的区域。因此,我们设计了一个新的算法来针对向量型特征的区域类别划分,具体的算法如图2所示。该算法详细说明了针对区域类型特征本文所使用的分类方法,该分类方法可以视为一种无监督的分类方法。

[0103] 在算法中,我们首先将第一个区域类型向量放入新建立的集合Set1中,并令Set1的集合特征向量为第一个区域类型向量,随后进行如下循环判断:若当前向量与已有的集合特征向量余弦相似度大于阈值 $\tau$ ,则将该向量放入对应集合,并通过取平均值的方式更新集合的特征向量,这一步对应伪代码的3到12行。如果该区域类型向量与与现有集合的特征向量余弦相似度均小于阈值 $\tau$ ,则创建新集合,这一步对应伪代码的13到15行。

[0104] 在实现了针对各个区域内、区域间的特征划分之后,我们使用图卷积神经网络(GCN)进行个区域节点的特征嵌入。图卷积神经网络的公式如下:

$$[0105] \quad H^{l+1} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^l W^l)$$

[0106] 在上式中, $\tilde{A} = I + A$ ,A是图的邻接矩阵,I为单位矩阵, $\tilde{D}$ 代表 $\tilde{A}$ 的度矩阵, $\sigma$ 为激活函数,而 $H^l$ 则为第l层网络的特征向量,模型最终的目标是通过类别监督学习不断优化 $W^l$ 参数,进而优化神经网络的嵌入效果。

[0107] 通过公式可以看出,虽然图卷积神经网络基于频谱分解方法设计,但是其训练过程中同样也可以体现出非常直观的空间结构关联性,即一个节点的最终嵌入结果不仅仅由其自身的特征决定,而是同时由其自身特征和与之相连接的节点特征共同决定。因此,通过图卷积神经网络进行的区域节点嵌入结果在保留了原有特征本身信息的同时也体现出了跨区域之间的关联性。

[0108] 图3展示了特征相关图在经过训练后最终转化为特征矩阵的全过程,从图3中可以看出,本实施例针对每一个区域内、区域间特征都构建了一个对应的特征图,即针对M个区域内、区域间特征构建了M张特征图。每一个特征图中包含所有实验需要进行未来物流量预测的区域节点,这些节点基于其代表的特征相似度在Top-K算法之后建立关联边生成图。在每一个构建完成的图的基础上,我们基于每一个特征的分类情况进行图卷积神经网络训练,训练完成后每一个特征图均可生成 $E \times N$ 的矩阵,该矩阵接下来将会作为预测模型的特征输入。

[0109] S3:将区域内、区域间嵌入特征向量和未来时序特征向量输入物流预测模块的模型中,给出一个区域内过去时间的物流变化数据,即可得到预测得到的未来物流量数据。

[0110] 在完成了区域节点的特征嵌入后,利用区域内、区域间特征嵌入和时序特征来进行物流单量的Transformer类模型设计。本实施例中的物流预测Transformer模型设计主要参考了Temporal Fusion Transforme的各模块,具体的模型架构如图4所示,其模型设计的主要思路是通过在时序特征和区域内、区域间特征中选择对预测结果真正重要的特征输入编码解码器中进行预测。

[0111] 在该模型中,图卷积神经网络被用于提取区域内、区域间信息的编码器。模型的输入为过去的时序数据 $D_{i,t} = (d_{T-t0}^i, \dots, d_T^i)$ ,  $X_{i,t} = (x_{T-t0}^i, \dots, x_T^i)$ , 区域内、区域间特征的嵌入矩阵 $R_i \in \mathbb{R}^{M \times E}$ , 以及未来已知的时序数据 $Df_{i,t} = (d_{T+1}^i, \dots, d_{T+t1}^i)$ 。

[0112] 模型主要架构为加了多头注意力机制的encoder-decoder结构,此处的神经元由普通的RNN神经元换成了LSTM(Long Short Term Memory)神经元。相对于普通RNN,LSTM神经元通过添加门控机制,在较长的序列中表现更好。LSTM神经元的内部架构如图6所示。

[0113] LSTM神经元内部由左至右所作的计算分别是信息遗忘计算、信息选择和神经元状态更新计算、信息输出计算。首先是信息遗忘计算:

$$[0114] \quad f_t = \sigma(W_f \cdot [h_{t-1}, x_t + b_f])$$

[0115] 在公式上述公式中, $x_t$ 、 $h_{t-1}$ 分别为当前时刻输入和前一时间步的隐状态, $W_f$ 、 $b_f$ 为模型参数。信息遗忘计算决定我们将会从神经元状态中遗忘哪些不重要的信息。接下来是信息选择和神经元状态更新计算:

$$[0116] \quad i_t = \sigma(W_i \cdot [h_{t-1}, x_t + b_i])$$

$$[0117] \quad \tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t + b_c])$$

$$[0118] \quad C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

[0119] 在上述公式中, $C_t$ 代表新的状态量, $W_i$ 、 $b_i$ 、 $W_c$ 、 $b_c$ 为网络参数。通过该公式决定神经元每一个时间步有多少信息需要被更新以及原状态是否应该被遗忘。最终是信息输出计算,定义如下:

$$[0120] \quad o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$[0121] \quad h_t = o_t * \tanh(C_t)$$

[0122] 在上式中, $W_o$ 、 $b_o$ 为网络参数, $h_t$ 为LSTM的最终输出。综合上述计算过程,我们可以看出相对于普通RNN神经元,LSTM神经元在神经元内部加入了记忆遗忘过程,从而在处理时间序列长期依赖时可以获取更多的特征信息。

[0123] 接下来,我们将逐个模块介绍其在模型架构中的具体作用。首先,对于模型中的GRN模块(Gated Residual Network),该模块的输入为一个时序特征输入a和区域内、区域间特征嵌入c,该模块所做的计算如下所示:

$$[0124] \quad \text{GRN}(a, c) = \text{LayerNorm}(a + \text{GLU}(\eta_1))$$

$$[0125] \quad \text{GLU}(\eta_1) = \sigma(W_1 \eta_1 + b_1) \odot (W_2 \eta_1 + b_2)$$

$$[0126] \quad \eta_1 = W_3 \eta_2 + b_3$$

$$[0127] \quad \eta_2 = \text{ELU}(W_4 a + W_5 c + b_4)$$

[0128] 在上述公式中,ELU为激活函数;GLU通过为每一个输入特征加权的方式来实现特

征选择; $\sigma$ 是sigmoid函数。 $W_1$ 到 $W_5$ 和 $b_1$ 到 $b_4$ 均为需要学习的网路参数。GRN模块工作方式类似于LSTM中的门控机制,其通过控制输入特征的非线性变换程度来增强模型的鲁棒性,从而实现软性特征选择,使模型在面对数据集中的噪声特征时有更好的表现。

[0129] 接下来,我们设定 $t$ 时间步的特征输入为 $F_t = [\xi(1)t, \dots, \xi(m)t]$ ,  $c_s$ 代表区域内、区域间特征嵌入。变量选择层(Variable Selection Network)的计算如下所示:

$$[0130] \quad \tilde{\xi}_t^{(j)} = \text{GRN}_{\xi_t^{(j)}}(\xi_t^{(j)})$$

$$[0131] \quad v_t = \text{Softmax}(\text{GRN}_{v_t}(F_t, c_s))$$

$$[0132] \quad \tilde{\xi}_t = \sum_{i=0}^m v_t^{(i)} \tilde{\xi}_t^{(i)}$$

[0133] 从公式上述可以看出,变量选择层首先使用GRN层来得到每一个输入特征的转换后的特征值,然后将所有特征连接后使用GRN层和softmax层得到特征权重。最后转换后的特征值加权求和后作为变量选择层的最终输出。

[0134] 直观上来说,我们可以将变量选择层模块视作针对时序特征和区域内、区域外特征的一种加权。通过该层神经网络,模型可以分辨出哪些特征对于提高预测准确率有着明显的正面影响,从而进一步关注于这种关键特征。

[0135] 最终,为了在长期物流量预测中捕捉到跨多个时间步的长期序列依赖,模型使用一个多头注意力网络层。与普通多头神经网络的 $Q, V, K$ 参数所不同的是,本模型使用的多头注意力网络共享 $V$ 参数来进一步强调关键特征。

[0136] 模型使用分位数回归,最终的模型输出形式为概率预测,即预测结果为一个置信区间。模型损失函数定义如下:

$$[0137] \quad L(q, \hat{y}, y) = \max(q * (y - \hat{y}), (1 - q) * (\hat{y} - y))$$

[0138] 在上述公式中, $\hat{y}$ 和 $y$ 分别代表预测值和真实值, $q$ 为一个人为定义的分位数,根据分位数的不同模型可以训练出不同的参数,进而导致预测值差异。

[0139] 因而,使用上述方法对北京地区2021年5月到6月的物流单量进行预测,预测结果如图6所示,在结果图6中实线代表实际物流量数据,Prediction\_Px折线代表当分位数 $q$ 等于 $\frac{x}{100}$ 时模型给出的预测结果。

[0140] 将上述预测结果计算以下预测指标:

$$[0141] \quad MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \bar{y}|}{y_i}$$

$$[0142] \quad RMSE = \frac{1}{n} \sum_{i=1}^n (\bar{y} - y_i)^2$$

[0143] 将预测结果进行归一化处理后,计算得到模型预测结果 $MAPE = 0.153$ ,  $RMSE = 0.048$ ,这说明模型预测结果准确可信。

[0144] 需要说明的是,以上内容仅仅说明了本发明的技术思想,不能以此限定本发明的保护范围,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰均落入本发明权利要求书的保护范围之内。

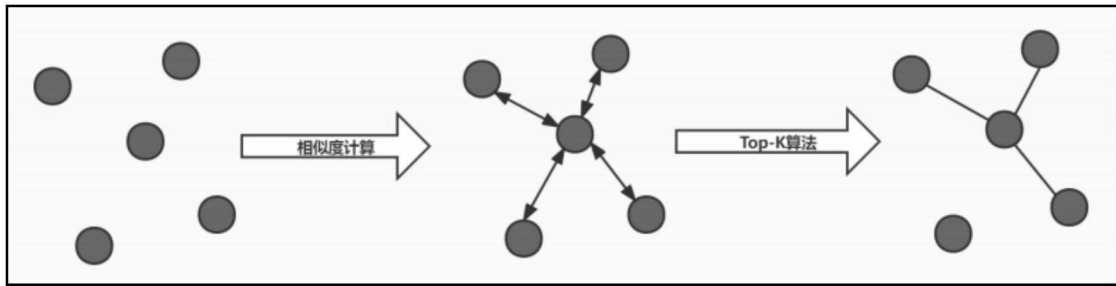


图1

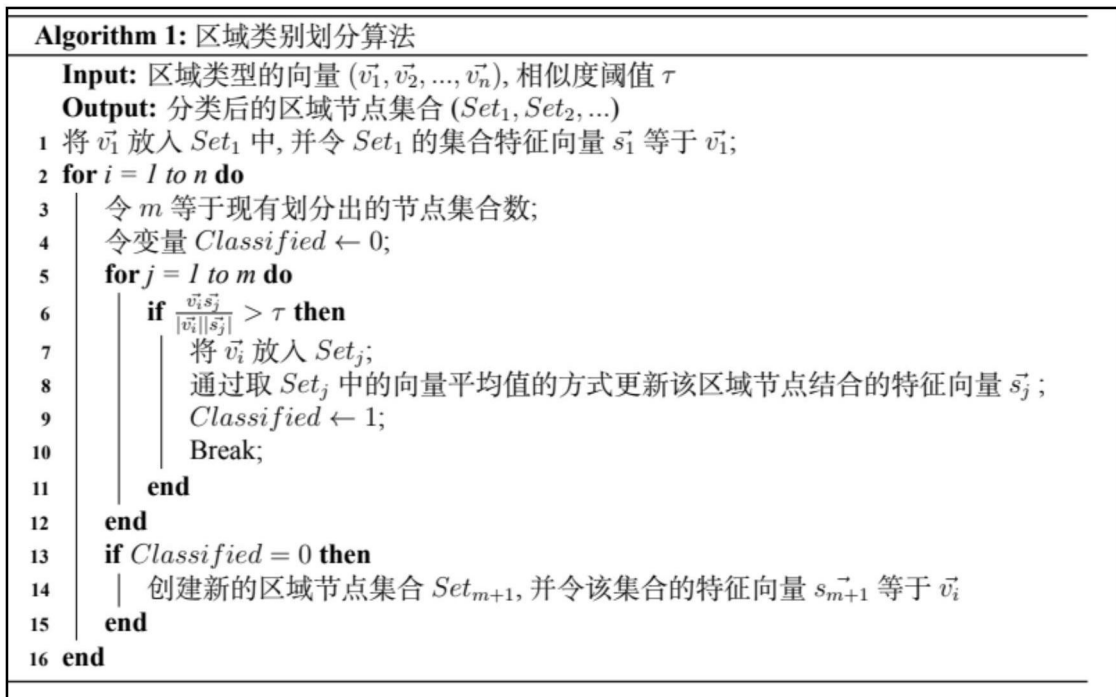


图2



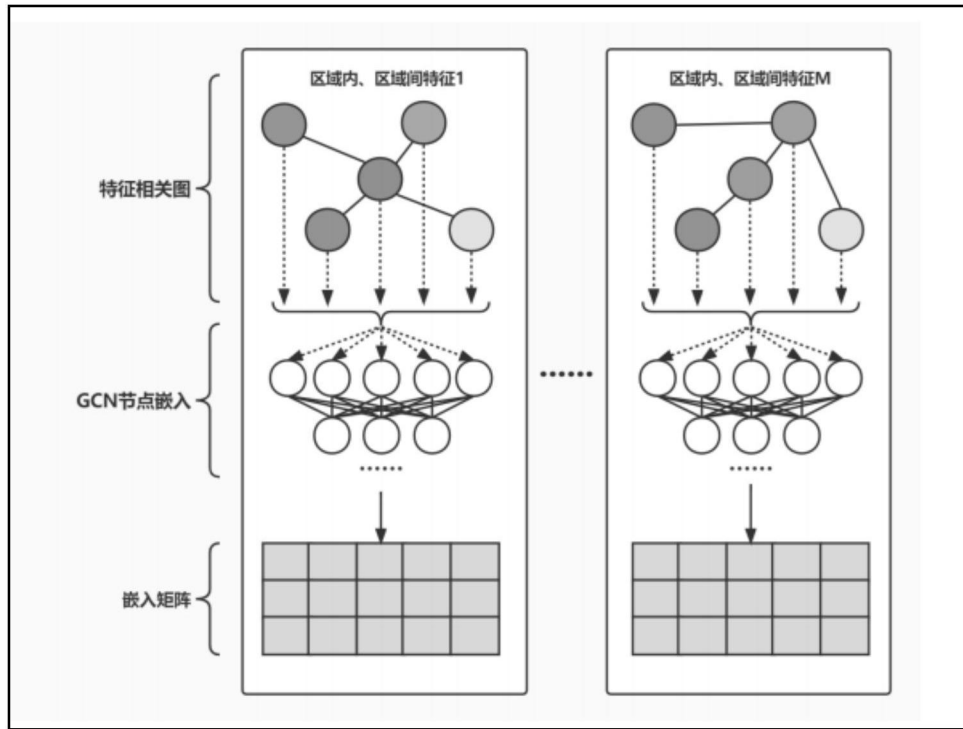


图3

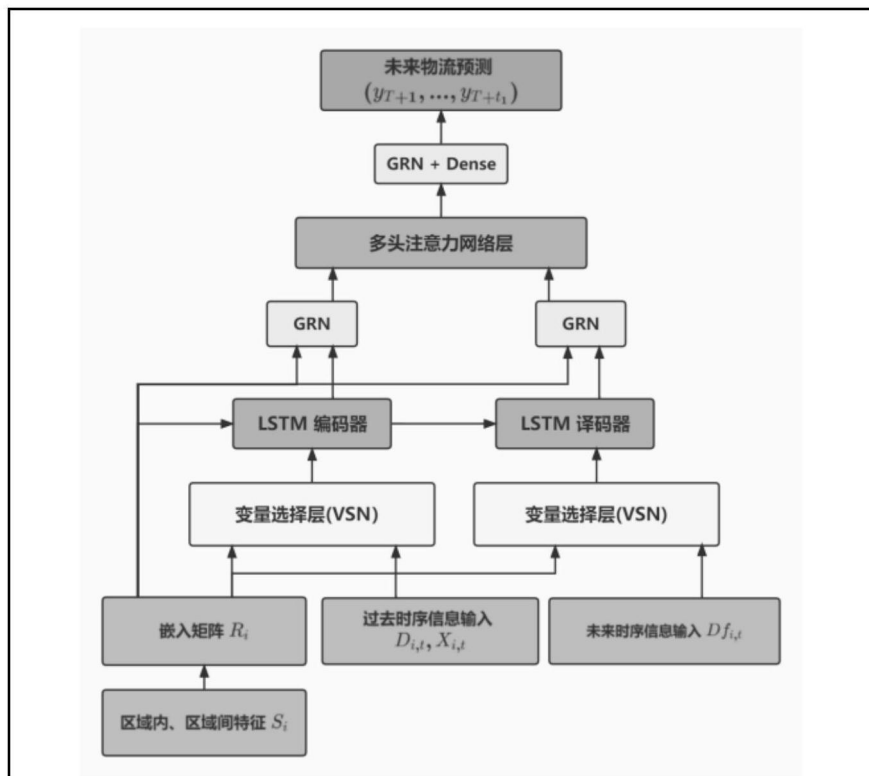


图4

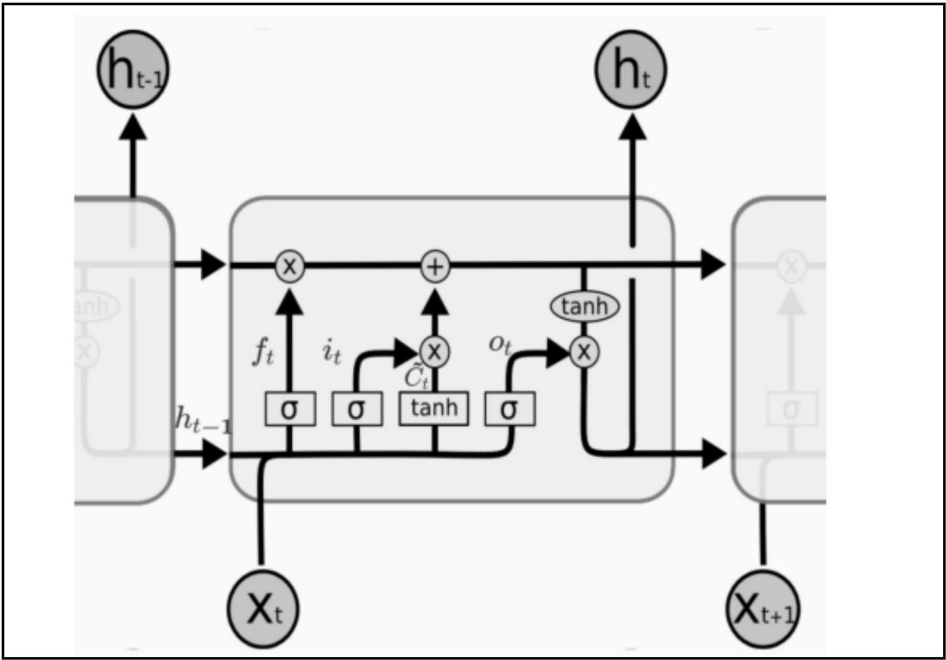


图5

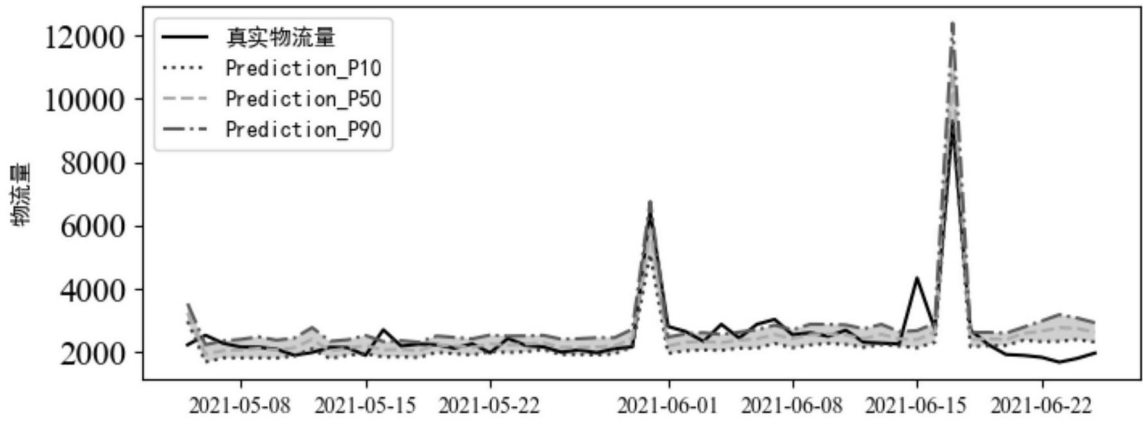


图6