

一种基于图注意力机制的气温预测模型

韩忠明^{1,2}, 周朋飞¹, 段大高^{1,2}, 张珣^{1,2*}

1. 北京工商大学计算机学院, 北京 100048

2. 食品安全大数据技术北京市重点实验室, 北京 100048

摘要 积极应对气候变化是可持续发展的目标之一。针对气温准确预测任务, 提出了一种基于图注意力机制的气温预测模型。该模型在气温站点组成的拓扑结构上使用了注意力机制, 选择性地聚合周围区域的气温特征, 再使用神经网络拟合复杂的气温变化规律, 得到预测结果。实验使用了2000—2010年京津冀地区的气温数据, 经大量实验验证, 在极少依赖历史气温数据的情况下, 模型能够得到更准确的预测值。模型能够为气候预测和气候灾害预防提供决策支持。

关键词 气温预测; 图神经网络; 注意力机制

水资源和能源的缺乏是可持续发展面临的两大问题。生态环境的改变, 尤其是气候的变化带来的降水变化, 不仅对农业作物的生产造成影响, 石油、核能等能源产业也需要大量的水资源。为了提高能源利用率、降低自然灾害的负面影响, 大量机器学习、深度学习技术应用到实际场景中, 如森林火灾预测^[1-2]、发电量预估^[3]、煤矿安全预警^[4]、石油储藏量预测^[5]等。

自世界气象组织(World Meteorological Organization, WMO) 1976年首次提出气候变化可对人类社

会带来严重威胁以来, 全球气候变化及其不利影响逐渐成为人类共同关心的问题^[6]。工业革命以来的人类活动, 尤其是发达国家在工业化过程中大量消耗能源资源, 导致大气中温室气体浓度增加, 引起全球气候近百年来以变暖为主要特征的显著变化, 波及到世界每个国家, 对全球自然生态系统产生了明显影响, 对人类社会的生存和发展带来严重挑战。2015年, 联合国可持续发展峰会正式通过“采取紧急行动应对气候变化及其影响”为可持续发展目标之一, 趋利避害、减少气候变化对人类社会可

收稿日期: 2019-12-10; 修回日期: 2020-03-30

基金项目: 国家重点研发计划项目(2019YFC0507800); “十三五”时期北京市属高校高水平教师队伍建设支持计划项目(CIT&TCD201904037); 中国博士后科学基金项目(2017M620885)

作者简介: 韩忠明, 教授, 研究方向为互联网数据挖掘、大数据分析与信息检索, 电子信箱: hanzm@th.btbu.edu.cn; 张珣(通信作者), 副教授, 研究方向为GIS软件技术、地理空间智能, 电子信箱: zhangxun@btbu.edu.cn

引用格式: 韩忠明, 周朋飞, 段大高, 等. 一种基于图注意力机制的气温预测模型[J]. 科技导报, 2020, 38(11): 115-121; doi: 10.3981/j.issn.1000-7857.2020.11.013

持续发展的危害已经成为全人类需要共同努力解决的问题。

气候变化会导致全球气候变暖、冰川冻土消融、海平面上升、极端气候频繁、自然灾害加重、生态系统濒危、疾病传播加剧等问题^[7]。中国人口众多、地域辽阔、气候条件复杂、生态环境脆弱,易受气候变化的不利影响,且近年来气候变化面临异常状况,这将对农牧业、林业、自然生态系统、水资源等领域以及沿海和生态脆弱地区造成严重影响,适应气候变化已成为中国的迫切任务。同时气候变化对人类健康产生不利影响^[8],气温升高有利于疾病的传播,还会促使身体机能的紊乱^[9],增加心血管疾病、呼吸系统疾病的发病率和死亡率。因此,深入分析气温的变化规律,提高气温预测的科学性和准确性,具有重要的科学意义和实际意义。

气温预测的传统做法是通过建立大气模型模拟气候变化,构建专门气温预测系统^[10-13],这种方法具有较高的复杂度,建立的模型也难以准确模拟

真实世界中的复杂气象变化。目前很多学者使用机器学习方法,建立数学模型预测未来的气温,例如多元线性回归^[14-16]、自回归滑动平均模型^[17]、整合移动平均自回归^[18]随机森林^[19]等,但这类方法难以拟合复杂的气温。为了提升模型的准确性,有学者提出了基于神经网络的气温预测模型^[20-22]。基于机器学习或者基于神经网络的气温预测模型通常只针对一个小区域,建模时只考虑单点气温值,忽略了不同区域间的相互影响^[23-24]。

本研究提出一种基于图注意力机制的气温预测模型,使用图神经网络对2000—2010年京津冀地区的171个站点记录的数据进行建模,同时考虑站点本身的历史数据和相邻站点的历史数据。模型技术框架如图1所示,区别于传统机器学习预测模型,该模型增加了站点间的网络构建步骤。实验表明,使用图注意力模型进行气温预测时,能够从周围站点聚合有效信息,提升气温预测的精度,同时减少对历史气温数据的依赖。

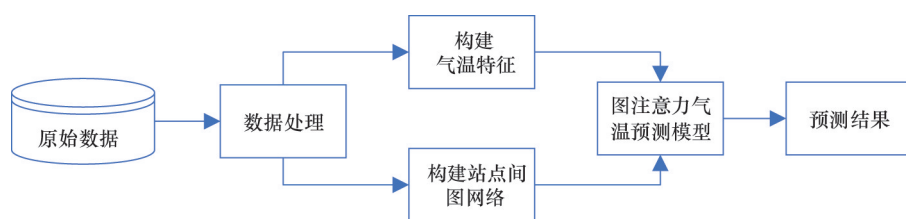


图1 模型技术框架

1 图注意力模型

神经网络是一种模仿生物神经结构的数学模型,通常以层为单元,使用非线性激活函数将各层组成链式结构。神经网络每层的结构没有固定大小,其构建的模型灵活多变,且根据万能近似定理^[25],神经网络可以任意精度来近似任何从一个有限维空间到另一个有限维空间的Borel可测函数。

神经网络通常使用反向传播算法进行学习,首先在训练集上计算预测误差,然后对模型中的可学习参数计算误差的梯度,再根据梯度调整网络中的参数,使整个模型向着预测误差最小的方向学习。

图神经网络是神经网络的一个分支,最早于2005被提出^[26]。其基本思想是,网络中一个结点的性质由其自身的属性和邻居结点的属性决定^[27],目标是为每一个结点学习其表示向量 $\mathbf{h}_v \in \mathbb{R}^s$,该向量长度为 s ,包含了自身和邻居的信息,可以用于节点预测、图分类等任务。

图注意力(graph attention networks, GAT)^[28]模型,在图神经网络中引入注意力机制^[29],赋予邻居结点不同的重要性。

如图2所示,以结点1为研究对象,把结点自身也视为一阶邻居,根据结点的特征向量 $\mathbf{h}$ 计算每条边的权重 α 。在更新结点的特征向量时,根据权重

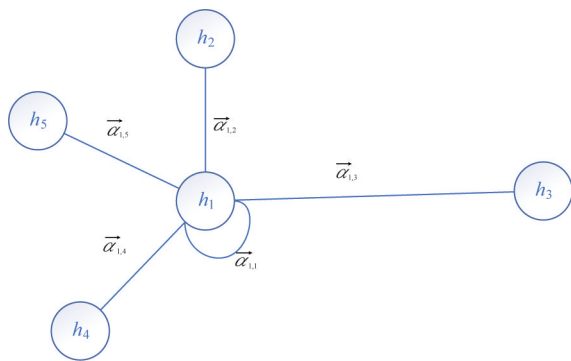


图2 图注意力模型示意

对邻居信息进行聚合,得到包含邻居信息的结点向量

$$\alpha_{ij} = \frac{\exp\left(\text{Leaky ReLU}\left(\mathbf{a}^T[\mathbf{W}\mathbf{h}_i \parallel \mathbf{W}\mathbf{h}_j]\right)\right)}{\sum_{k \in \mathcal{N}(i)} \exp\left(\text{Leaky ReLU}\left(\mathbf{a}^T[\mathbf{W}\mathbf{h}_i \parallel \mathbf{W}\mathbf{h}_k]\right)\right)} \quad (1)$$

$$\mathbf{h}'_i = \sigma\left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij} \mathbf{W}\mathbf{h}_j\right) \quad (2)$$

式(1)为GAT模型计算权重的步骤,其中 $\mathcal{N}(i)$ 为结点 i 的一阶邻居集合, $\mathbf{W}$ 和 $\mathbf{a}$ 均为可学习的参数, $\mathbf{a} \in \mathbb{R}^{2C \times 1}$, $\mathbf{W} \in \mathbb{R}^{C \times C'}$, C 为每个节点输入时特征向量的长度, C' 为输出的结点特征维度。

式(1)所得到的 α_{ij} 为将结点 j 的特征聚合到结点 i 时所参考的重要性权重。式(2)为GAT模型的聚合步骤,通过权重对邻居的特征向量进行加权求和, $\mathbf{h}'_i$ 为结点 i 更新后的表示向量。

单层的GAT模型能够聚合一阶邻居,通过堆叠多层的GAT模型不但能够提升模型整体的表示能力,还能够获取高阶的邻居信息,能够对复杂的图网络进行建模。

2 基于图注意力机制的气温预测模型

2.1 数据分析与处理

本研究数据来自中国国家级地面气象站,为中国地面气候资料日值数据集(V3.0)。考虑数据的完整性,以京津冀地区171个气象观测点在2000年至2010年所记录的历史气温为研究对象,包含每

日的最低、最高和平均气温。使用大量气温数据训练图神经网络模型,进行气温预测任务。京津冀地区占地218000 km²,包含北京市、天津市、河北省的13个城市,属于暖温带大陆性季风型气候,夏季高温多雨,冬季寒冷干燥。

图3为京津冀地区天津市武清区11年的平均气温变化情况。若以年为单位,能看出该地区的气温呈现周期性变化,年最高气温在30℃左右,最低气温为-10℃左右。但如果以天为单位观察气温变化(图4),短期内气温频繁地上升和下降,很难观察到变化的规律。气温在时间维度上变化较大,但在空间范围上气温的变化是连续的、有规律的。图5为2000年1月1日京津冀地区各气象站点所测量的日平均气温,可以看出,局部范围、相邻站点的日平均气温差异较小,总体上,该区域的气温自西北到东南逐渐升高。

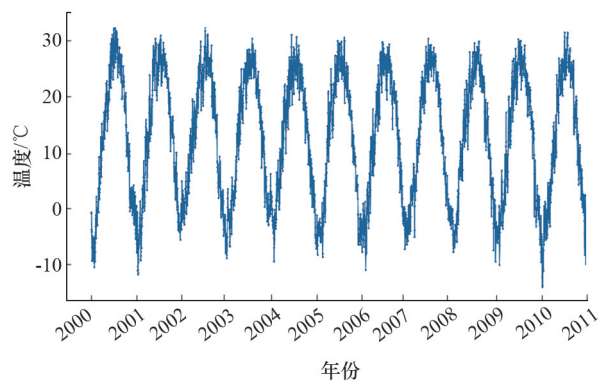


图3 天津市武清区2000—2010年平均气温变化

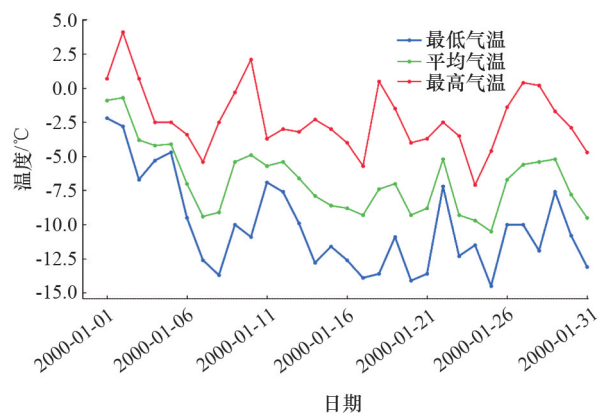


图4 天津市武清区2000年1月气温变化

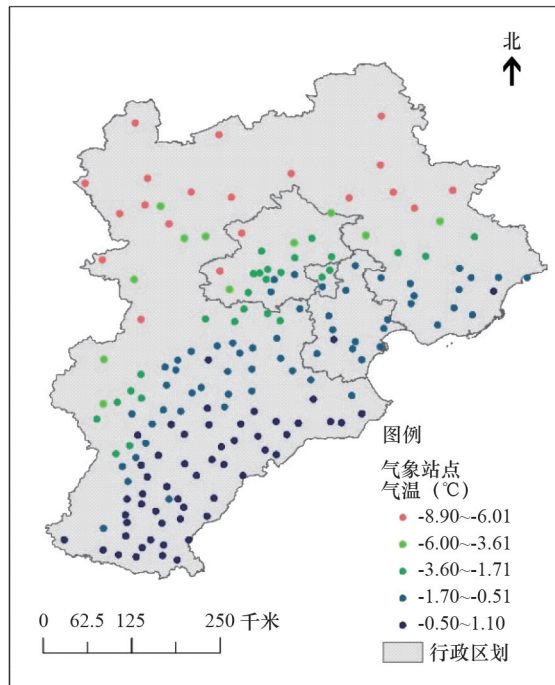


图5 京津冀地区2000年1月1日各地平均气温

气温属于复杂气候系统中的一部分,这样的规律受多种因素影响,难以穷尽且频繁变化。使用图神经网络能够利用各节点间的关系来提升预测任务的准确率。本研究使用GAT模型进行气温预测,将京津冀地区的171个气候站点作为结点,利用神经网络强大的表示能力挖掘各站点间气温的潜在关系,提高气温预测的准确率。

在数据处理过程中,由于观测任务、仪器故障等原因造成的小部分数据缺失,本文使用该站点前一日的数据进行填充。另外,为方便建模和相关指标的计算,对气温数据进行归一化处理,将气温的值域映射到0到1的开区间。

2.2 模型构建

为建立基于图神经网络的气温预测模型,使用双层GAT模型预测日平均气温。模型的输入为所有站点的气温特征矩阵和描述站点间拓扑结构的邻接矩阵,输出为每个站点未来1天平均气温的预测值。

使用图神经网络进行气温预测需要每个站点每日的气温特征和站点间形成的网络。在构建站

点每日气温特征时,使用滑动窗口选取该地区前 K 天的历史气温。对于站点间的图网络,本文根据各个站点的地理坐标,通过半正矢公式(haversine)计算2个站点之间的实际距离,选取直线距离小于阈值 D 的站点作为邻居,最后形成 171×171 邻接矩阵,其中为1的点表示站点相连。气温预测模型结构如图6所示。

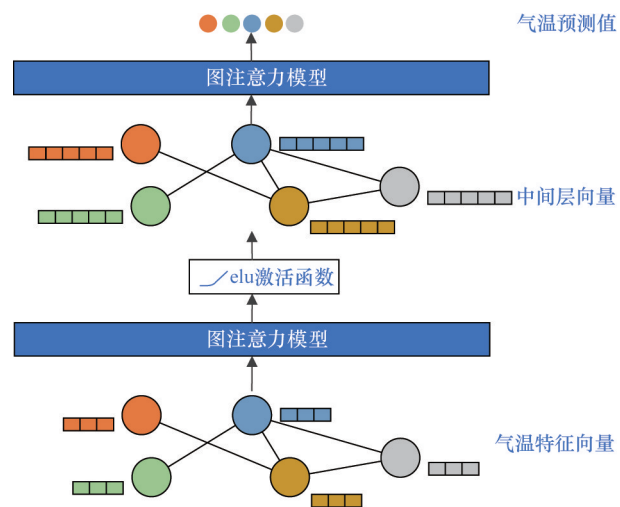


图6 气温预测模型结构

首先将所有站点的气温特征构建为特征矩阵 X , $X \in \mathbb{R}^{N \times F}$, 其中 N 为站点数量, F 为气温特征的维度;按照站点间的距离构建邻接矩阵 A , $A \in \mathbb{R}^{N \times N}$ 。根据式(3),将 X 和 A 输入到第一层的图注意力网络中,经过正向传播,得到各个站点的状态向量矩阵 H_1 , $H_1 \in \mathbb{R}^{N \times F'}$, F' 为隐藏层中每个站点状态向量的维度。

$$H_1 = \text{GAT}_1(X, A) \quad (3)$$

$$\text{elu}(x) = \max(0, x) + \min(0, \alpha \cdot (\exp(x) - 1)) \quad (4)$$

为了让模型能够学习非线性关系,对第一层图注意力网络的输出进行非线性变换。模型中使用非线性激活函数 $\text{elu}^{[30]}$,如式(4)所示,保留大于0的值,将小于0的值进行非线性变换。其中 α 为系数,通常设置为1。

$$H_1' = \text{elu}(H_1) \quad (5)$$

$$Y' = \text{GAT}_2(H_1', A) \quad (6)$$

将状态矩阵 H_1 中的每一个值输入非线性激活

函数 elu 进行变换得到 H_1' 。最后根据式(6)将站点的状态向量 H_1' 和站点间的邻接矩阵 A 输入第二层的图注意力网络中,把长度为 F' 的状态向量映射到 1 维,得到模型最后的输出 Y' 。其中 $Y' \in \mathbb{R}^{N \times 1}$, 每一个值代表该站点的预测温度。

3 结果及分析

设定滑动窗口大小 K 为 12, 将各个站点前 12 天的最低、最高、平均气温作为特征。用于构建站点间拓扑关系的距离阈值 D 设置为 40 km, 得到 171 个站点中 1637 条相邻关系。在训练模型阶段采用 60% 的数据作为训练集, 验证集和测试集各 20%。

在训练图神经网络时, 模型隐藏层的维度 F' 设置为 1024, 损失函数为平均绝对百分比误差 (MAPE), 优化器使用 SGD, 学习率为 0.001。训练的迭代次数固定为 200, 选取验证集误差最小时的输出作为最终预测结果。

实验选取了整合移动平均自回归模型 (ARIMA) 和 BP 神经网络作为对比模型, 本文模型记为 GAT。ARIMA 模型对原始时间序列进行差分, 将不平稳的序列变为平稳序列, 再结合自回归和移动平均两种方法建立线性的预测模型。BP 神经网络则通过大量神经元相互连接构成的非线性模型, 能够对复杂的关系进行建模, 通常使用反向传播的方法进行训练。

实验结果如表 1 所示, 分别使用均方误差 (MSE)、绝对平均误差 (MAE)、平均绝对百分比误差 (MAPE) 3 种统计方法评估预测精度。

表 1 实验结果

评价指标	ARIMA	BP	GAT
MSE	0.0010	0.0008	0.0006
MAE	0.0237	0.0236	0.0197
MAPE	0.0489	0.0399	0.0346

从表 2 中可知, 线性模型 ARIMA 的预测精度小于两种非线性模型, 在非线性模型中, 加入邻居站点信息的 GAT 模型在 3 种指标下的预测精度均

好于 BP 神经网络。说明 GAT 模型能够有效获取相邻站点的信息, 提升非线性模型的预测精度。

为了探究在气温预测任务中历史气温和邻居站点的选取对预测效果的影响, 实验中修改了模型的 2 个超参数, 滑动窗口大小 K 和相邻站点距离的阈值 D , 并进行了对比试验。

图 7 为 BP 模型和 GAT 模型在选取不同长度的历史气温作为特征时的预测效果。横轴为选取历史气温的天数, 纵轴为 MAPE 指标。从图 7 中可以看出, 当选取的历史气温长度较短时, BP 神经网络的精度较差。在窗口大小为 2 的情况下只考虑 2 天的历史气温信息, BP 神经网络得到信息较少, 模型难以收敛。随着滑动窗口长度的增加, 预测的误差率快速下降。而 GAT 模型在历史气温很少的情况下就有较好的预测效果, 该模型通过获取邻居站点的信息, 减少了对历史气温的依赖。

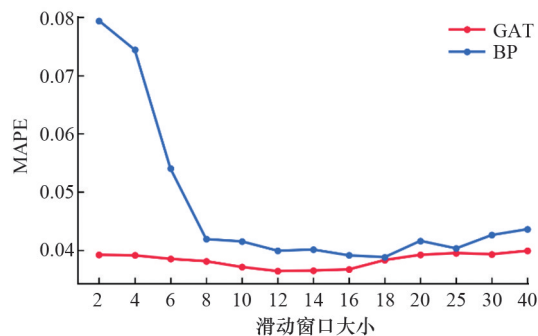


图 7 滑动窗口大小对回归效果的影响

图 8 为 GAT 模型在使用不同邻接矩阵时得到的实验效果。纵轴为 MAPE 指标, 横轴为选取邻居站点时距离的阈值, 单位为 km。从图 8 中可以看出, 邻居选取的范围对结果存在较大影响, 当范围

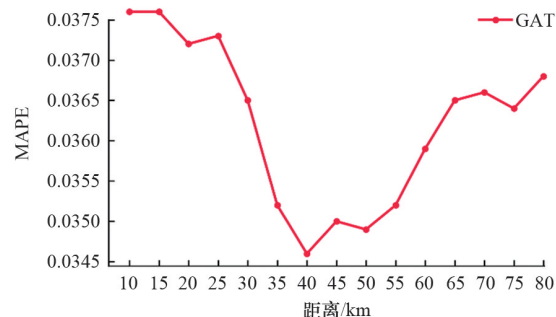


图 8 邻居距离对实验效果的影响

从 10 km 增大到 40 km 时,气温预测的精度逐渐提高,当范围超过 40 km 时,增加邻居并不能提高 GAT 模型的预测精度。这说明使用 GAT 模型对某地区气温进行预测时,使用 40 km 范围内的相邻站点的历史温度信息能提高模型的预测效果,使用超过 50 km 站点的历史温度信息时,带来的噪音会降低模型的预测效果。

4 结论

提出了一种基于图注意力机制的气温预测模型,不同于传统模型单点预测方式,本模型通过图上的注意力机制,融合一个地区内多个站点的气温特征,能够提升气温预测的准确率,在多种指标下均取得较优效果。同时,探究了相关因素对该模型预测效果的影响,在时间尺度上,相较于传统方法,本模型能够通过聚合邻居信息的方法减少对历史气温数据的依赖,在只用 2 天历史数据的情况下,就能够得到较为精确的预测结果;在空间尺度上,通过对比实验,得到了站点邻居选取的范围。

图神经网络能够对复杂的关系进行建模。在实际应用中,气象站点所记录的数据不仅包括气温,还包括其他更多的气象数据。使用图神经网络不仅可以各个站点关联起来提升气温预测的精度,还可以引入其他复杂气象数据,完成对风速、湿度、降水量等指标的预测,从而构建自动化的气象系统。

参考文献(References)

- [1] Sakr G E, Elhajj I H, Mitri G, et al. Artificial intelligence for forest fire prediction[C]//2010 IEEE/ASME International Conference on Advanced Intelligent Mechatronics. Piscataway, NJ: IEEE, 2010: 1311-1316.
- [2] Cheng T, Wang J. Integrated spatio-temporal data mining for forest fire prediction[J]. Transactions in GIS, 2008, 12(5): 591-611.
- [3] Pinson P, Kariniotakis G. Conditional prediction intervals of wind power generation[J]. IEEE Transactions on Power Systems, 2010, 25(4): 1845-1856.
- [4] Zhao X, Gang W, Zhao K K, et al. On-line least squares support vector machine algorithm in gas prediction[J]. Mining Science and Technology (China), 2009, 19(2): 194-198.
- [5] 崔书岳, 黄晓辉, 陈云亮, 等. 基于 HMM 的缝洞型油藏产量预测算法[J]. 西南大学学报: 自然科学版, 2020, 42(2): 137-144.
- [6] 李雪超. 兰州市气温变化趋势及预测[D]. 兰州: 兰州财经大学统计学院, 2019.
- [7] 陈广仁. 气候变化关乎人类未来[J]. 科技导报, 2011, 29(34): 8.
- [8] Boko M, Niang I, Nyong A, et al. Africa Climate Change 2007: Impacts, Adaptation and Vulnerability: Contribution of working group ii to the fourth assessment report of the intergovernmental panel on climate change[R]. Cambridge UK: IPCC, 2018.
- [9] 王颢晨, 路凤, 武继磊, 等. 中国气候变化对人口健康影响研究评述[J]. 科技导报, 2014, 32(28/29): 109-116.
- [10] Xue M, Wang D, Gao J, et al. The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation[J]. Meteorology & Atmospheric Physics, 2003, 82(1): 139-170.
- [11] 杨成荫, 赵苏璇, 程立国, 等. 短期区域气候预测系统及回报试验[J]. 解放军理工大学学报: 自然科学版, 2016, 17(3): 289-295.
- [12] Smith M. Vehicle-centric weather prediction system and method: US, Patent 6,603,405[P]. 2003-08-05.
- [13] 彭京备, 布和朝鲁, 郑飞, 等. 2017 年夏季全国气候趋势展望[J]. 中国科学院院刊, 2017, 32(4): 413-417.
- [14] Abuelia M, Chowdhury B. Solar power probabilistic forecasting by using multiple linear regression analysis[C]//SoutheastCon 2015, Fort Lauderdale, USA: IEEE, 2015: 1-5.
- [15] 王丹, 王建鹏, 白庆梅, 等. 递减平均法与一元线性回归法对 ECMWF 温度预报订正能力对比[J]. 气象, 2019, 45(9): 1310-1321.
- [16] 金秀良, 宋燕, 吴洪, 等. 应用北半球大气环流系统预测中国降水量和气温[J]. 气象科技, 2014, 42(6): 1028-1038.
- [17] 倪淑娜, 唐波, 蔡家辉. 基于 GM-ARMA 组合模型的全球年平均气温预测[J]. 中国新技术新产品, 2008(12): 9-10.
- [18] 陈百硕, 李守伟, 何建敏, 等. 天气衍生品中时变均值回复的气温预测模型研究[J]. 管理工程学报, 2014, 28(2): 145-150.

- [19] Huang C Y, Liu Y W, Tzeng W C, et al. Short term wind speed predictions by using the grey prediction model based forecast method[C]//2011 IEEE Green Technologies Conference (IEEE-Green). Baton Rouge, USA: IEEE, 2011: 1-5.
- [20] 牛志娟, 胡红萍. 基于主成分分析的BP神经网络和RBF神经网络月平均气温预测模型[J]. 高师理科学刊, 2015, 35(11): 6-8.
- [21] 林耿, 郑紫微. 基于灰色-BP神经网络的福州市年平均气温预测模型[J]. 河南工程学院学报: 自然科学版, 2018, 30(4): 71-75.
- [22] 陶晔, 杜景林. 基于随机森林的长短期记忆网络气温预测[J]. 计算机工程与设计, 2019, 40(3): 737-743.
- [23] Miguez-Macho G, Stenchikov G L, Robock A. Regional climate simulations over North America: Interaction of local processes with improved large-scale flow[J]. Journal of Climate, 2005, 18(8): 1227-1246.
- [24] 刘鸿波, 张大林, 王斌. 区域气候模拟研究及其应用进展[J]. 气候与环境研究, 2006(5): 649-668.
- [25] Hornik K. Approximation capabilities of multilayer feed-forward networks[J]. Neural Networks, 1991, 4(2): 251-257.
- [26] Bianchini M, Maggini M, Sarti L, et al. Recursive neural networks for processing graphs with labelled edges: Theory and applications[J]. Neural Networks, 2005, 18(8): 1040-1050.
- [27] Scarselli F, Gori M, Tsoi A C, et al. The graph neural network model[J]. IEEE Transactions on Neural Networks, 2008, 20(1): 61-80.
- [28] Veličković P, Cucurull G, Casanova A, et al. Graph Attention Networks[C]//International Conference on Learning Representations. Vancouver, Canada: International Conference on Learning Representations, 2018: 1-12.
- [29] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Advances in neural information processing systems. Long Beach, CA: Curran Associates, 2017: 5998-6008.
- [30] Clevert D A, Unterthiner T, Hochreiter S. Fast and accurate deep network learning by exponential linear units (elus) [C]//International Conference on Learning Representations. San Juan, Puerto Rico: International Conference on Learning Representations, 2016: 1-14.

A temperature prediction model based on graph attention mechanism

HAN Zhongming^{1,2}, ZHOU Pengfei¹, DUAN Dagao^{1,2}, ZHANG Xun^{1,2*}

1. School of Computer Science and Engineering, Beijing Technology and Business University, Beijing 100048, China

2. Beijing Key Laboratory of Big Data Technology for Food Safety, Beijing 100048, China

Abstract The active response to the climate change is one of the goals of sustainable development. This paper presents a temperature prediction model based on the graph attention mechanism. The attention mechanism on the topology of the temperature sites is used to selectively aggregate the temperature feature of the surrounding area. Then the neural network is used to fit the complex temperature change pattern and forecast the future temperatures. In the experiments, the temperature data of Beijing-Tianjin-Hebei region from 2000 to 2010 are used. A large number of experiments show that with this method more accurate predictions can be made with a small amount of historical temperature data. The model can provide the decision support for the climate prediction and the climate disaster prevention, with an important scientific and practical significance.

Keywords temperature prediction; graph neural network; attention mechanism ●



(责任编辑 傅雪)