

DOI:10.3969/j.issn.1003-5060.2023.11.006

基于注意力与神经网络的视频流行度预测模型

马学森^{1,2}, 杨智捷^{1,2}, 储昭坤^{1,2}, 周天保^{1,2}

(1. 合肥工业大学 计算机与信息学院, 安徽 合肥 230601; 2. 安全关键工业测控技术教育部工程研究中心, 安徽 合肥 230601)

摘要:针对传统预测算法预测精度低及难以处理多变量的时序数据等缺点,文章提出一种采用双向长短期记忆(bi-directional long short-term memory, BiLSTM)网络和时间模式注意力(temporal pattern attention, TPA)机制相结合的视频流行度预测模型。双向长短期记忆网络从视频流行度时间序列的正向和反向提取时间特征,时间模式注意力机制从双向长短期记忆网络输出状态的深层特征提取时间模式,有利于视频流行度预测。真实视频数据的实验结果表明,与经典时序预测方法相比,TPA-BiLSTM 模型能够有效地降低预测的误差,提高预测的准确性。

关键词:流行度预测;多元时间序列;双向长短期记忆(BiLSTM)网络;注意力机制;卷积神经网络(CNN)

中图分类号:TP389.1 **文献标志码:**A **文章编号:**1003-5060(2023)11-1472-07

Video popularity prediction model based on attention and neural network

MA Xuesen^{1,2}, YANG Zhijie^{1,2}, CHU Zhaokun^{1,2}, ZHOU Tianbao^{1,2}

(1. School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230601, China; 2. Engineering Research Center of Safety Critical Industrial Measurement and Control Technology of Ministry of Education, Hefei 230601, China)

Abstract: In view of the shortcomings of traditional prediction algorithms such as low prediction accuracy and difficulty in processing multivariate time series data, a video popularity prediction model using a combination of bi-directional long short-term memory (BiLSTM) and temporal pattern attention (TPA) is proposed. BiLSTM extracts temporal features from the forward and reverse of the video popularity time series, while TPA extracts temporal patterns from the deeper features of the output states of BiLSTM, which is beneficial for video popularity prediction. Experimental results on real video data show that the proposed TPA-BiLSTM combined model effectively reduces the prediction error and improves the prediction accuracy compared with classical time series prediction methods.

Key words: popularity prediction; multivariate time series; bi-directional long short-term memory (BiLSTM); attention mechanism; convolutional neural network (CNN)

根据思科统计^[1],2021 年全球上传至网络的视频内容激增,每月全球网络上新增的视频需要耗费每个人超过 500 万年的时间才能观看完。因为用户的时间和注意力是有限的,所以大多数视频几乎并没有被注意到,只有少数受欢迎的视频获得大多数人的观看。文献[2]指出,许多应用程序中使用着预测视频受欢迎程度的功能,此项功

能是广告和推荐的关键一环。通过观察 YouTube 视频网站上很多用户的反馈行为,一部分视频在公开后的一段时间内,收到用户的正向反馈后,视频的流行度有显著增加。文献[3]指出,由于发布在线视频的形式复杂,内容多样,巨大的数据量给人们造成信息过量的困扰,研究者在视频流行度预测的研究中侧重于从海量信息中预测筛

收稿日期:2022-01-21;修回日期:2022-03-25

基金项目:国家重点研发计划资助项目(2020YFC1512601);广东省科技发展专项基金资助项目(2017A010101001);安徽省自然科学基金联合基金资助项目(2008085UD08)和中央高校基本科研业务费专项资金资助项目(PA2021GDGP0061)

作者简介:马学森(1976—),男,安徽合肥人,博士,合肥工业大学教授,硕士生导师,通信作者,E-mail:mxs@hfut.edu.cn.

选出将来热门的信息。

目前主流的视频流行度预测模型有基于统计学和基于机器学习的预测方法。基于统计学的预测方法^[4]在不同时间段,使用累计浏览量的序列间关系进行建模,未引入外部因素辅助预测,准确度仍不足。文献[5]提出一种新的算法 K -ARMA,找出 K 个最近的样本并给予一定的权重值,根据 K 个样本的流行度和自回归滑动平均(auto regressive moving average, ARMA)模型相结合开展预测;该模型相比传统 ARMA 模型在预测准确度上有明显的提升,但是未结合用户行为数据,未设置多组特征值。文献[6]提出一种融合搜索引擎数据与社交网络数据的多元线性回归模型预测点播量排名,该模型比单独用搜索引擎数据或社交网络的预测模型预测效果更佳。

基于机器学习的流行度预测方法,文献[7]提出一种混合机器学习方法来预测未发布视频内容的流行度,该模型通过视频历史内容的信息预测新内容的受欢迎程度,但未考虑观众的反馈数据(如观众人数、评分等)。随着深度学习理论的出现与发展,长短期记忆(long short-term memory, LSTM)网络广泛运用于如风速^[8]、语音处理^[9]、船舶航迹^[10]等领域,预测效果良好。LSTM 网络在序列建模任务中有很大的优势,它具有长时记忆功能,能够单一方向提取时间序列的时间特征,其变体 GRU(gate recurrent unit)因结构相似,计算简单等特点也被广泛运用于预测方面。双向长短期记忆(bi-directional long short-term memory, BiLSTM)网络由双向 LSTM 组成,主要应用于文本情感分析^[11]、命名实体识别等领域。文献[12]提出将神经网络和马尔可夫相组合的模型(a video popularity prediction algorithm based on a neural network and Markov combined model, Mar-BiLSTM),在使用 Mar-BiLSTM 模型后,精度有所提升的同时仍保持低模型复杂度。文献[13]将自回归与循环神经网络相结合,提出一种基于自回归循环神经网络(auto regressive recurrent neural network, DeepAR),主要应用于零售等领域,该算法可有效融合大量相关时间序列,产生准确的概率预测;此算法运用于视频流行度的预测还需进一步研究。

此外,更多学者将注意力机制引入时间序列预测中。文献[14]提出一种将内容特征与时序信息相融合的注意力视频流行度模型,将内容特征与时间序列分开处理后融合在一起,该模型很好

地捕捉流行趋势。文献[15]研究微信文章的浏览量、点击量等用户反馈事件的出现频次,引入注意力机制作用于 LSTM 网络,将微信文章的变化过程进行建模用于预测流行趋势。文献[16]提出一种多变量时间模式预测的注意力机制(temporal pattern attention, TPA)模型,进行多变量时间序列预测,并有效运用于电力负荷^[16]、股票指数^[17]、泊位占有率^[18]等方面的预测。实验表明,TPA 机制能融合多特征,预测精度较传统方法更佳,但仍需进一步验证在视频流行度预测方向是否适用。

通过以上视频流行度预测方法可以得出视频流行度的预测面临如下难题:

1) 观众点击观看视频的行为具有不确定性,即观众在选择在何时点击何种视频是不确定事件,与个人所面临的环境事件与心理状态有关,这为视频流行度预测增添了不确定性。

2) 视频的流行度与视频播放网站的用户反馈具有相关性,即视频网站会优先推送更热门的视频给观众,观众会通过其他观众反馈的数据(如观看数、评分等)来决定是否观看视频。

基于以上分析,本文提出注意力与神经网络的视频流行度预测算法 TPA-BiLSTM,并验证其在视频流行度预测领域的有效性。本文贡献主要包括以下 3 个方面。

首先,本文考虑不同用户反馈事件^[15]分析视频的观看量、点赞量、转发量、收藏量等宏观事件发生的次数。

其次,本文利用时间模式注意力 TPA 机制,对视频流行度的复杂结构进行建模。TPA 机制能够聚焦最有利于预测的特征,并研究特征之间的关系,提高预测的准确性。注意力机制可以计算每个维度输入特征的权重,当权值大时,意味着该特征更有利于预测;当权值较小时,表明该特征对视频流行度预测的贡献较小。视频流行度的数据集代表具有短期和长期记忆的周期模式。TPA-LSTM 机制包括 LSTM 模块、卷积神经网络(convolutional neural network, CNN)模块和时间注意力模块,适用于各种数据集,甚至是具有弱周期模式的多元时间序列数据。LSTM 组件可以捕获一个相对长期的模式,而 CNN 模块能提取变量之间的局部依赖性和时间维度下的短期模式。此外,时间注意力模块可以选择有助于预测和捕捉时间信息的变量。

最后,通过 LSTM 与双向 CNN 组建的 BiL-

STM 网络,可保留双向时间序列的特征依赖。为使预测网络从视频观看量的时间序列中不仅提取双向时间特征,而且提取多维输入之间的变化规律,结合 BiLSTM 和 TPA 的优点,建立注意力与神经网络的视频流行度预测模型。

1 基于注意力与神经网络模型

1.1 视频流行度的预测问题描述

视频剧集推陈出新,通过视频热度捕捉视频流行度的变化趋势,通常用视频观看量作为流行度的衡量^[5],用序列 pop 来描述视频当前时刻的流行度,分别构造点赞量、转发量、收藏量宏观时间每小时的发次数、点赞量序列 like、转发量序列 forward 和收藏量序列 favorite。利用预测模型 ϕ 预测视频流行度 $\widetilde{\text{pop}}_{t-1+h}$,计算过程为:

$$\widetilde{\text{pop}}_{t-1+h} = \phi \{ \text{pop}_{t-w}, \dots, \text{pop}_{t-1}, \text{like}_{t-w}, \dots, \text{like}_{t-1}, \text{forward}_{t-w}, \dots, \text{forward}_{t-1}, \text{favorite}_{t-w}, \dots, \text{favorite}_{t-1} \} \quad (1)$$

其中:pop 为流行度,已经上线 $t-1$ 天; w 为窗口大小; h 为预测视频流行度的不同预测长度;预测目标为未来几小时流行度的记录 $\{ \widetilde{\text{pop}}_t, \widetilde{\text{pop}}_{t+1}, \dots, \widetilde{\text{pop}}_{t-1+h} \}$ 。

1.2 BiLSTM 网络模型的构建

BiLSTM 是从 LSTM 模型结构改进的双向 LSTM。LSTM 模型提出门控机制,采用 3 种门控机制有效解决梯度爆炸问题,如图 1 所示。

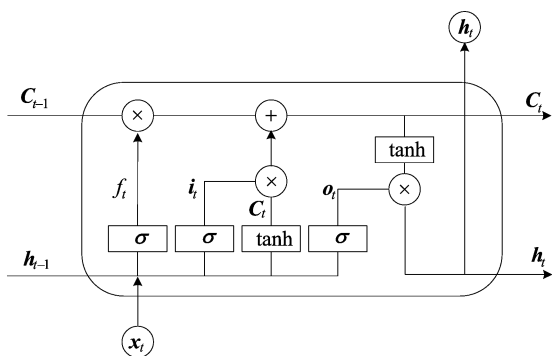


图 1 LSTM 隐藏单元结构

图 1 中, t 时刻输入的特征向量 x_t 以及 $t-1$ 时刻隐藏向量 h_{t-1} 和细胞状态 C_{t-1} ,在结构内计算权值与激活函数获得信息,并输出 t 时刻细胞状态 C_t 和隐藏向量 h_t ,具体计算过程为:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (3)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (4)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (5)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = o_t \times \tanh(C_t) \quad (7)$$

其中: W_f, W_i, W_c, W_o 为权值参数矩阵; b_f, b_i, b_c, b_o 为偏置值; f_t, i_t 分别为 t 时刻的遗忘门、输入门; \tilde{C}_t, C_t, o_t 为输入的细胞状态、细胞状态和输出门; h_t 为最终 t 时刻隐藏单元输出向量。偏置值和权值参数矩阵随机设定。

BiLSTM 用于上下文信息的建模,具有良好的长期信息存储能力。数据的逆传播意味着时间序列以相反顺序被传输到模型,神经元存储 2 个输入方向的信息,使用叠加的多层神经网络结构,完成流行度系列的深层特征挖掘,有效地提高模型的预测效果。BiLSTM 的总体实现流程如图 2 所示。

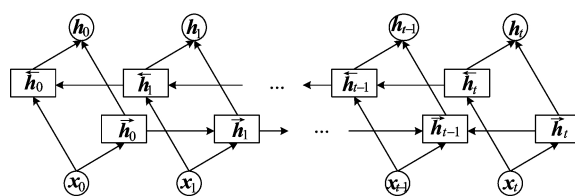


图 2 BiLSTM 总体实现流程

图 2 中,BiLSTM 以 LSTM 作为基础,分别通过正向传播和反向传播来提取历史数据的正向时间特征和反向时间特征,再由此时的 2 个隐藏单元所输出的向量相连组成该时刻输出,其中, h_0, h_1, h_{t-1}, h_t 分别为对应的 0、1、 $t-1$ 、 t 时刻的输出向量。 $t-1$ 时刻 BiLSTM 的计算过程为:

$$\tilde{h}_{t-1} = \text{LSTM}(\tilde{h}_t, x_{t-1}) \quad (8)$$

$$\tilde{h}_{t-1} = \text{LSTM}(\tilde{h}_{t-2}, x_{t-1}) \quad (9)$$

$$h_{t-1} = [\tilde{h}_{t-1}, \tilde{h}_{t-1}] \quad (10)$$

其中: \tilde{h}_{t-1} 为正向时间特征; \tilde{h}_{t-1} 为反向时间特征。

1.3 TPA 机制结构构建

TPA 结构使用 BiLSTM 输出向量 $h_i (i=t-w, \dots, t-1)$,其中每一列代表对不同变量的处理,每一行代表单个特征的时间序列数据。通过卷积计算该特征在卷积核内的时间模式矩阵 $H_{i,j}^C$,使用评分函数计算 $H_{i,j}^C$ 的分值,并使用激活函数归一化分值,用评分函数生成每个状态的注意力权重 a ,通过相加并求和的方式求出信息向量 v_i 。结合信息向量 v_i 及隐藏状态 h_i 求出 h_i' ,最后通过输出层和归一化指数函数来计算输出预测值 $\widetilde{\text{pop}}_{t-1+h}$ 。

TPA 结构在 TPA-BiLSTM 模型中的作用和原理如下。

通过卷积核沿隐藏状态矩阵 H 的行向量使

用滤波器来增强模型学习。卷积核大小为 $1 \times T$ (T 为模型的最大注意长度, 设 $T = \omega$), 卷积配置有 k 个滤波器。将上述卷积核沿隐藏状态矩阵 \mathbf{H} 的行向量计算卷积, 提取该变量在该卷积核范围内的 $\mathbf{H}_{i,j}^C$, 即

$$\mathbf{H}_{i,j}^C = \sum_{l=1}^{\omega} \mathbf{C}_{j,T-\omega+l} * \mathbf{H}_{i,t-\omega-1+l} \quad (11)$$

其中, $\mathbf{H}_{i,j}^C$ 为第 i 行向量和第 j 个卷积核的卷积值。

对时间模式矩阵进行评分, 评分函数 f 为:

$$f(\mathbf{H}_i^C, \mathbf{h}_i) = (\mathbf{H}_i^C)^\top \mathbf{W}_a \mathbf{h}_i \quad (12)$$

其中: \mathbf{H}_i^C 为时间模式矩阵 \mathbf{H}^C 的第 i 行; \mathbf{W}_a 为权值参数; \mathbf{h}_i 为输出的隐藏状态。注意力权重为:

$$a_i = \text{sigmoid}(f(\mathbf{H}_i^C, \mathbf{h}_i)) \quad (13)$$

其中: f 为评分函数; $a_i (i=1, 2, 3, \dots, n)$ 为注意力权重。

1.4 基于 TPA-BiLSTM 的流行度预测模型

基于 1.3 节介绍的 TPA 结构, 构建 TPA-BiLSTM 网络模型, 设计 TPA-BiLSTM 模型求解视频流行度预测问题。

TPA-BiLSTM 模型的视频流行度预测过程, 如图 3 所示。

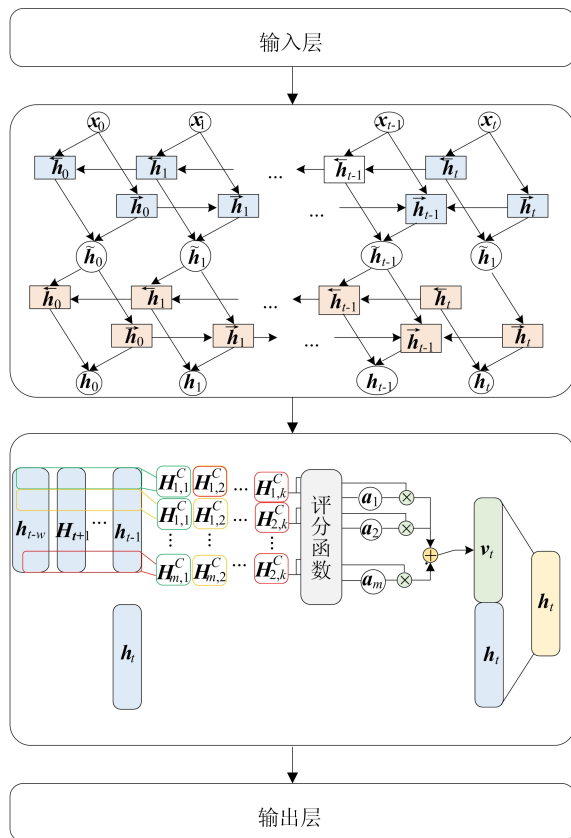


图 3 TPA-BiLSTM 算法模型的预测过程

1) 输入层。输入数据包括视频观看量、点赞

量、转发量、收藏量。输出数据为视频流行度。首先对视频观看量、点赞量、转发量、收藏量的历史数据进行处理, 包括清洗数据、去噪等。

2) BiLSTM 层。预处理后的数据输入 BiLSTM 网络, BiLSTM 模型从正向和反向提取时间特征 h_t 。

3) TPA 层。TPA 使用 BiLSTM 网络的输出特征作为输入, 在所有采样时刻提取单个特征状态所包含的深层信息, 并聚焦有利于预测的时间模式。

4) 输出层。将最后数据相结合生成预测值 $\widetilde{\text{pop}}_{t+h}$ 。

通过评价指数对模型进行评价。对 \mathbf{H}^C 的行加权求和, 得到信息向量 v_t 为:

$$v_t = \sum_{i=1}^m a_i \mathbf{H}_i^C \quad (14)$$

其中, a_i 为时间模式矩阵 \mathbf{H}^C 的第 i 行注意力权重。最后将 v_t 和最后一个时刻的输出 h_t 融合, 通过线性变化, 得到最后的预测输出, 计算过程为:

$$h_t' = \mathbf{W}_h h_t + \mathbf{W}_v v_t \quad (15)$$

$$\widetilde{\text{pop}}_{t+h} = \mathbf{W}_h h_t' \quad (16)$$

其中, $\widetilde{\text{pop}}_{t+h}$ 为流行度的预测值。

2 实验设计与分析

2.1 数据预处理

实验采用 kaggle 网站的 YouTube 视频数据集, 主要用于时间序列变化分析、情感分析和流行度分析等。本文选取 2020 年 6 月 1 日至 9 月 31 日的历史数据作为数据集, 数据信息包括每日内每小时的视频观看量、点赞量、转发量、收藏量等观众反馈事件, 也包含视频 ID、视频播放时间、视频种类。

为了保证实验数据的完整性并提高预测模型的准确性和稳定性。首先, 补充数据集内缺失的值, 保证数据集的完整性; 其次, 去除噪声数据; 最后, 对数据集进行归一化, 将数据进行线性变化。划分每个视频数据集, 训练集的选取为整个数据集的前 80%, 剩余的数据集作为测试集。

2.2 模型的对比分析

选取主流的 5 种时序预测方法与 TPA-BiLSTM 模型进行对比, 这些方法包括: ARMA 模型、LSTM 模型、GRU 模型、DeepAR 模型、Mar-BiLSTM 模型。按照 2.1 节所描述处理时间序列数据, 然后进行数据归一化, 输入模型训练, 最后

对比测试集预测结果。对于 ARMA 模型,本文直接使用文献[5]设定的参数和公开代码进行实验。LSTM 模型与 GRU 模型隐藏层神经元的个数为 24,批处理尺寸为 32,学习率为 0.001,迭代次数为 1 000。DeepAR 模型隐藏层神经元的个数为 40,批处理尺寸为 32,学习率为 0.001,迭代次数为 1 000。对于 Mar-BiLSTM 模型,本文参考文献[12]内的代码与参数设定进行训练模型。经多次调整参数以训练模型后得出 TPA-BiLSTM 流行度预测模型的模型参数如下:迭代次数为 1 000;学习率为 0.001;输入特征数为 4;TPA-BiLSTM 模型训练选取的样本量取 64;BiLSTM 隐藏层神经元的个数为 24,隐藏层的层数为 2。

2.3 评价指标

使用均方根误差 (root mean square error, RMSE) E_{RMS} 和平均绝对误差 (mean absolute error, MAE) E_{MA} 作为评价指标^[3]。其中: E_{RMS} 能够计算得到真实值与预测值的差别; E_{MA} 能够计算得到求平均值后的预测值与真实值之间误差的绝对值。

$$E_{\text{RMS}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (17)$$

$$E_{\text{MA}} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (18)$$

其中: y_i 为真实值; \hat{y}_i 为预测值; n 为采样点总数。

2.4 对比实验与结果分析

将 TPA-BiLSTM 模型与对比模型的实验结果进行对比。实验测试 TPA-BiLSTM 模型与对比模型在不同预测步长 h 的预测精度^[16], h 分别取 6、12、24、48 h 时不同模型的预测性能见表 1 所列。

由表 1 可知,视频流行度趋势变化具有复杂多变的特点,当前发展变化规律与历史规律不完全一致。传统的时间序列预测未考虑外部因素的影响仅强调时间序列预测,更适合在平稳时间序列内进行预测,因此当存在预测误差的缺陷并且外部因素的变化大时,会产生大的偏差。随着 h 逐渐增大时,预测难度增大。LSTM 与 GRU 整体上表现优于 ARMA,说明在视频流行度的预测问题上,基于神经网络的方法能够更好地拟合非线性的时间序列数据。Mar-BiLSTM 模型在预测过程中兼具时间序列的前、后 2 个方向的依赖信息,并通过建立马尔可夫修正模型修正神经网络训练时产生的预测误差,在保持模型复杂度的同时提高了模型的预测精度,该模型相较于其他单一模型的预测表现更优。DeepAR 模型融合自回归与循环神经网络,针对多重时间序列良好的表现,综合表现优于单一模型的使用。本文提出的 TPA-BiLSTM 模型的 RMSE 和 MAE 2 种评价指标均取得较好的实验结果。特别地,在预测未来 6 h 内的视频流行度时,TPA-BiLSTM 模型较 Mar-BiLSTM 模型的 RMSE 降低 12%,较 DeepAR 模型 RMSE 降低 16%。在预测未来 48 h 内的视频流行度时,TPA-BiLSTM 模型较 LSTM 模型 RMSE 降低 5%,较 DeepAR 模型降低 6%。从整体情况来看,TPA-BiLSTM 模型随着 h 的增加,预测指标 RMSE、MAE 随着 h 的增大而增大。经测试 TPA-BiLSTM 模型的表现优于其他模型,在不同预测步长 h 的预测精度上均有良好的表现。

表 1 不同模型在不同预测步长下的预测性能对比

模型	RMSE				MAE			
	$h=6$ h	$h=12$ h	$h=24$ h	$h=48$ h	$h=6$ h	$h=12$ h	$h=24$ h	$h=48$ h
ARMA	72 282.126	71 562.409	75 227.104	78 137.408	59 352.832	59 255.331	62 474.289	66 874.689
LSTM	62 533.289	67 583.233	69 544.069	72 911.752	51 323.348	53 369.278	56 739.521	57 969.875
GRU	60 334.227	71 314.532	70 169.243	75 499.651	50 323.101	51 379.201	57 261.321	59 151.861
DeepAR	58 878.437	65 759.855	69 457.001	73 326.605	46 322.625	50 849.267	56 234.465	58 046.645
Mar-BiLSTM	57 150.632	64 039.089	71 548.933	75 142.004	45 518.978	50 094.344	57 689.417	59 636.801
TPA-BiLSTM	50 583.539	59 163.943	66 843.127	69 111.364	40 321.278	42 721.207	53 731.239	56 230.978

综上可得到如下结论:

1) 对比 TPA-BiLSTM 模型和 ARMA 模型的结果,发现 TPA-BiLSTM 模型的效果明显优于基于统计学的模型。传统基于统计学的模型在多变量时序预测场景下,未考虑变量之间的相关性,且对非线性数据无法很好地处理,TPA-BiL-

STM 模型的预测结果更为准确。

2) 对比 TPA-BiLSTM、LSTM、GRU、DeepAR 和 Mar-BiLSTM 模型,发现 TPA-BiLSTM 模型要强于现有基于神经网络方法的相关模型,模型预测在稳定性与准确性有更好地提升。TPA-BiLSTM 模型相较于其他组合模型预测精

度更高,误差更小。这是由于 TPA-BiLSTM 模型捕获多变量的时间序列模式,相比于只获取单一变量的时间序列模式的模型效果更好。

2.5 消融实验及结果分析

采用 YouTube 视频播放数据集进行消融实验,以验证 TPA-BiLSTM 模型各模块设计的作用。每次删去模型中一个组件,并与 TPA-BiLSTM 模型进行对比验证各组件的有效性。将移除组件的模型进行命名区分。即 TPA-BiLSTMw/oCN 模型,即不包含卷积神经网络模块的模型;TPA-BiLSTMw/oBL 模型,即不包含双向隐藏层模块的模型;TPA-BiLSTMw/oAT 模型,即不包含注意力机制模块的模型。3 种模型的预测性能见表 2 所列。

从表 2 可以看出:各组件对视频流行度的预测都有贡献,其中贡献最大的是 CNN 模块,CNN 模块在不同 h 时都起着关键作用,在 h 为 6、12、24、48 h 时,TPA-BiLSTM 和 TPA-BiLSTMw/oCN 模型相比 RMSE 值分别下降 19%、20%、4%、9%。贡献其次的是双向隐藏层模块与注意力机制模块,TPA-BiLSTM 和 TPA-BiLSTMw/oBL 模型相比 RMSE 值也分别降低 16%、11%、3%、6%,TPA-BiLSTM 和 TPA-BiLSTMw/oAT 模型相比 RMSE 值下降 17%、15%、4%、7%。

这意味着缺少任意一个组件都无法发挥该模型的准确与稳定性能,结合实验数据可知,本文提出的 TPA-BiLSTM 模型每个模块的设计都是必要且有效的。

表 2 TPA-BiLSTM 的不同结构变种在不同预测步长下的预测性能对比

模型	RMSE				MAE			
	$h=6$ h	$h=12$ h	$h=24$ h	$h=48$ h	$h=6$ h	$h=12$ h	$h=24$ h	$h=48$ h
TPA-BiLSTMw/oCN	60 334.227	71 314.532	70 169.243	75 499.651	50 323.101	51 379.201	57 261.321	59 151.861
TPA-BiLSTMw/oBL	58 878.437	65 759.855	69 457.001	73 326.605	46 322.625	50 849.267	56 234.465	58 046.645
TPA-BiLSTMw/oAT	59 606.332	68 537.194	69 813.122	74 413.128	48 322.863	51 114.234	56 747.893	58 599.253
TPA-BiLSTM	50 583.539	59 163.943	66 843.127	69 111.364	40 321.278	42 721.207	53 731.239	56 230.978

3 结 论

针对视频流行度预测问题,本文提出一种基于注意力与神经网络的视频流行度预测模型,能够有效地提高流行度的预测性能。该模型从 2 个方向提取视频流行度数据的时间特征,为了提高流行度预测精度,在模型中将点赞量、转发量、收藏量数据特征作为视频流行度的融合特征,使模型在预测精度与泛化能力上更出色,解决现有单一神经网络模型和传统基于统计学方法难以准确预测真实视频流行度、难以融合多特征的问题。为了客观地比较模型的有效性,将真实视频流行度样本分为训练集与测试集进行实验。结果表明,本文所提模型比 ARMA 模型、LSTM 模型、GRU 模型、DeepAR 模型和 Mar-BiLSTM 模型表现出更好的预测效果,并在消融实验中表明每个模块的作用。后续的研究可将不同类型视频的独有特征纳入考虑,设计适用于更为复杂情形的预测模型并进行运用推广。

[参 考 文 献]

[1] CISCO. Cisco visual networking index: forecast and method-

ology, 2016—2021[R]. San Jose, CA: CISCO, 2017: 1-17.

- [2] MONIZ N, TORGO L. A review on web content popularity prediction: issues and open challenges [J]. *Online Social Networks and Media*, 2019, 12: 1-20.
- [3] 曹琦,沈华伟,高金华,等. 基于深度学习的流行度预测研究综述[J]. *中文信息学报*, 2021, 35(2): 1-18, 32.
- [4] TAN Z, WANG Y, ZHANG Y, et al. A novel time series approach for predicting the long-term popularity of online videos [J]. *IEEE Transactions on Broadcasting*, 2016, 62(2): 436-445.
- [5] 赵命燕,李泽平. 一种基于视频特征及历史数据的流行度预测算法[J]. *计算机与现代化*, 2018(2): 49-53.
- [6] 徐晓枫,贺樑,杨静. 融合社交与搜索数据的电视剧点播排名预测研究[J]. *计算机工程*, 2015, 41(8): 6-12, 17.
- [7] JEON H J, SEO W C, PARK E J, et al. Hybrid machine learning approach for popularity prediction of newly released contents of online video streaming services[J]. *Technological Forecasting & Social Change*, 2020, 161: 120303.
- [8] 魏昱洲,许西宁. 基于 LSTM 长短期记忆网络的超短期风速预测[J]. *电子测量与仪器学报*, 2019, 33(2): 64-71.
- [9] SMAGULOVA K, JAMES A P. A survey on LSTM meristic neural network architectures and applications[J]. *The European Physical Journal Special Topics*, 2019, 228(10): 2313-2324.
- [10] 权波,杨博辰,胡可奇,等. 基于 LSTM 的船舶航迹预测模型[J]. *计算机科学*, 2018, 45(增刊 2): 126-131.
- [11] 李洋,董红斌. 基于 CNN 和 BiLSTM 网络特征融合的文本

- 情感分析[J]. 计算机应用, 2018, 38(11): 3075-3080.
- [12] 马学森, 陈树友, 许向东, 等. 基于神经网络与马尔可夫组合模型的视频流行度预测算法[J]. 电信科学, 2021, 37(8): 18-26.
- [13] SALINAS D, FLUNKERT V, GASTHAUS J, et al. DeepAR: probabilistic forecasting with autoregressive recurrent networks[J]. International Journal of Forecasting, 2020, 36(3): 1181-1191.
- [14] 武维, 李泽平, 杨华蔚, 等. 融合内容特征和时序信息的深度注意力视频流行度预测模型[J]. 计算机应用, 2021, 41(7): 1878-1884.
- [15] LIAO D L, XU J, LI G F, et al. Popularity prediction on online articles with deep fusion of temporal process and content features[C]//Proceedings of the 33rd AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2019: 200-207.
- [16] SHIH S Y, SUN F K, LEE H Y. Temporal pattern attention for multivariate time series forecasting[J]. Machine Learning, 2019, 108(8/9): 1421-1441.
- [17] WEI X L, LEI B B, OUYANG H B, et al. Stock index prices prediction via temporal pattern attention and long-short-term memory[J]. Advances in Multimedia, 2020(4): 8831893.
- [18] 王竹荣, 薛伟, 牛亚邦, 等. 基于注意力机制的泊位占有率预测模型研究[J]. 通信学报, 2020, 41(12): 182-192.

(责任编辑 李 凯)

(上接第 1453 页)

- [2] FREGO M, BEVILACQUA P, DIVAN S, et al. Minimum time-minimum jerk optimal traffic management for AGVs[J]. IEEE Robotics and Automation Letters, 2020, 5(4): 5307-5314.
- [3] 倪郁东, 李媛媛, 沈吟东, 等. 基于改进人工蜂群算法的机器人路径规划研究[J]. 合肥工业大学学报(自然科学版), 2019, 42(11): 1569-1575.
- [4] LEMERLE P, HÖPPNER O, REBELLE J. Dynamic stability of forklift trucks in cornering situations: parametrical analysis using a driving simulator[J]. Vehicle System Dynamics, 2011, 49(10): 1673-1693.
- [5] 张艳, 张明路, 蒋志宏, 等. 动态环境下移动机器人路径规划的研究[J]. 合肥工业大学学报(自然科学版), 2020, 43(10): 1297-1306.
- [6] 陈虹, 郭露露, 宫洵, 等. 智能时代的汽车控制[J]. 自动化学报, 2020, 46(7): 1313-1332.
- [7] SUN H, ZHANG W, YU R, et al. Motion planning for mobile robots-focusing on deep reinforcement learning: a systematic review[J]. IEEE Access, 2021, 9: 69061-69081.
- [8] 陈秋莲, 郑以君, 蒋环宇, 等. 基于神经网络改进粒子群算法的动态路径规划[J]. 华中科技大学学报(自然科学版), 2021, 49(2): 51-55.
- [9] 王庆, 徐海明, 吕品, 等. 基于改进蚁群算法的多旋翼无人机航迹规划研究[J]. 合肥工业大学学报(自然科学版), 2021, 44(9): 1172-1178.
- [10] 杜明博, 梅涛, 陈佳佳, 等. 复杂环境下基于 RRT 的智能车辆运动规划算法[J]. 机器人, 2015, 37(4): 443-450.
- [11] 刘成菊, 韩俊强, 安康. 基于改进 RRT 算法的 RoboCup 机器人动态路径规划[J]. 机器人, 2017, 39(1): 8-15.
- [12] BAZZI D, PRIORA G, ZANCHETTIN A M, et al. RRT* and goal-driven variable admittance control for obstacle avoidance in manual guidance applications[J]. IEEE Robotics and Automation Letters, 2022, 7(2): 1920-1927.
- [13] ZHAO J, ZHANG Y, MA Z, et al. Improvement and verification of A-star algorithm for AGV path planning[J]. Computer Engineering and Applications, 2018, 54(21): 217-223.
- [14] XIN W, WANLIN L, CHAO F, et al. Path planning research based on an improved A* algorithm for mobile robot[J]. IOP Conference Series: Materials Science and Engineering, 2019, 569(5): 52044.
- [15] 劳彩莲, 李鹏, 冯宇. 基于改进 A* 与 DWA 算法融合的温室机器人路径规划[J]. 农业机械学报, 2021, 52(1): 14-22.
- [16] 宗成星, 陆亮, 雷新宇, 等. 一种基于 A* 算法的空间多自由度机械臂路径规划方法[J]. 合肥工业大学学报(自然科学版), 2017, 40(2): 164-168.
- [17] 吕恩利, 林伟加, 刘妍华, 等. 基于 B 样条曲线的智能叉车托盘拾取路径规划研究[J]. 农业机械学报, 2019, 50(5): 394-402.
- [18] LIN R, HUANG H, LI M. An automated guided logistics robot for pallet transportation[J]. Assembly Automation, 2020, 41(1): 45-54.
- [19] HUANG Y, DING H, ZHANG Y, et al. A motion planning and tracking framework for autonomous vehicles based on artificial potential field elaborated resistance network approach[J]. IEEE Transactions on Industrial Electronics, 2020, 67(2): 1376-1386.
- [20] 辛煜, 梁华为, 杜明博, 等. 一种可搜索无限个邻域的改进 A* 算法[J]. 机器人, 2014, 36(5): 627-633.
- [21] 李冲, 张安, 毕文豪. 单边矩形扩展 A* 算法[J]. 机器人, 2017, 39(1): 46-56.
- [22] 崔宝侠, 王森弛, 段勇. 基于可搜索 24 邻域的 A* 算法路径规划[J]. 沈阳工业大学学报, 2018, 40(2): 180-184.
- [23] 王维, 裴东, 冯璋. 改进 A* 算法的工业机器人最短路径规划[J]. 计算机应用, 2018, 38(5): 1523-1526.

(责任编辑 胡亚敏)