

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

武 维¹, 李泽平^{1*}, 杨华蔚², 林 川¹, 王忠德¹

(1. 贵州大学 计算机科学与技术学院, 贵阳 550025; 2. 贵州财经大学 大数据应用与经济学院, 贵阳 550025)

(* 通信作者电子邮箱 zpli1@gzu.edu.cn)

摘 要: 针对视频流行度动态变化过程中的时序信息难以捕捉的问题, 提出一种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(DAFCT)。首先, 根据用户的反馈信息, 构建基于注意力机制的长短期记忆网络(Attention-LSTM)模型来捕捉流行趋势并挖掘时序信息; 然后, 采用神经网络因子分解机(NFM)处理多模态的内容特征, 并采用嵌入技术对稀疏的高维特征进行降维处理, 从而降低模型的计算复杂性; 最后, 采用 concatenate 方法融合时序信息和内容特征, 并设计了一种深度注意力视频流行度预测(DAVPP)算法来求解 DAFCT。实验结果表明, 与 Attention-LSTM 模型和 NFM 模型相比, DAFCT 的召回率分别提高了 10.82 和 3.31 个百分点, F1 分数分别提高了 9.80 和 3.07 个百分点。

关键词: 流行度预测; 内容特征; 时序信息; 注意力机制; 神经网络因子分解机

中图分类号: TP391.1 **文献标志码:** A

Deep attention video popularity prediction model fusing content features and temporal information

WU Wei¹, LI Zeping^{1*}, YANG Huawei², LIN Chuan¹, WANG Zhongde¹

(1. College of Computer Science and Technology, Guizhou University, Guiyang Guizhou 550025, China;

2. College of Big Data Application and Economics, Guizhou University of Finance and Economics, Guiyang Guizhou 550025, China)

Abstract: Aiming at the problem that it is difficult to capture the temporal information during the dynamic change of video popularity, a Deep Attention video popularity prediction model Fusing Content and Temporal information (DAFCT) was proposed. Firstly, according to the users' feedback information, an Attention mechanism based Long Short-Term Memory network (Attention-LSTM) model was constructed to capture the popular trend and mine the temporal information. Secondly, Neural Factorization Machine (NFM) was used to process multi-modal content features and embedding techniques were adopted to reduce the computational complexity of the model by reducing the dimension of sparse high-dimensional features. Finally, the concatenate method was employed to fuse the temporal information and content features, and a Deep Attention Video Popularity Prediction (DAVPP) algorithm was designed to solve the proposed DAFCT. Experimental results show that compared with Attention-LSTM model and NFM model, the recall of DAFCT is improved by 10.82 and 3.31 percentage points, and the F1 score was improved by 9.80 and 3.07 percentage points, respectively.

Key words: popularity prediction; content feature; time-series information; attention mechanism; Neural Factorization Machine (NFM)

0 引言

提前预测视频的受欢迎程度是许多应用的重要部分,如推荐、广告和信息检索等^[1]。通过对 YouTube 视频网站上的大量用户反馈行为观察发现,部分视频在发布后的一段时间内,经用户反馈后该视频的流行度呈增长趋势。为了捕获视频流行度的动态变化过程,本文首先采用循环神经网络(Recurrent Neural Network, RNN)建模并计算出视频的流行趋势指数,同时离散化视频的点赞量和点踩量,把流行度预测任务转化为分类问题;将视频的流行度分为“受欢迎”和“不受欢迎”两类,采用神经网络模型对视频的内容特征进行建模;

最后融合视频的流行趋势和内容特征以预测视频的流行度。

基于用户反馈量的宏观积累过程来预测流行度有很大的实用价值,长短期记忆(Long Short-Term Memory, LSTM)网络能够有效地捕捉事件的变化过程^[2],被广泛应用于股票走势预测^[3]、温度变化趋势预测^[4]以及医学研究中的抑郁趋势预测^[5]等。LSTM 网络能够有效地捕捉视频的流行趋势,目前已有研究者采用 LSTM 网络对视频的流行动态进行建模和预测流行度,且取得了较好的效果^[6]。受到这些研究工作的启发,本文采用 LSTM 网络对视频的流行趋势进行建模,捕获视频流行度的变化趋势。

收稿日期: 2020-10-19; 修回日期: 2021-01-17; 录用日期: 2021-01-18。

基金项目: 国家自然科学基金资助项目(61462014); 贵州省科学技术基金资助项目(S2017JP00401185093)。

作者简介: 武维(1995—),女,贵州毕节人,硕士研究生,主要研究方向: 流行度预测、流媒体; 李泽平(1964—),男,贵州贵阳人,教授,博士,主要研究方向: 流行度预测、计算机网络、流媒体; 杨华蔚(1965—),女,贵州贵阳人,教授,博士,主要研究方向: 金融风险管控; 林川(1975—),男,四川自贡人,副教授,硕士,主要研究方向: 计算机网络、云计算; 王忠德(1963—),男,山东青岛人,实验师,主要研究方向: 流媒体。

最近,基于深度学习的模型被应用于用户的偏好预测^[7]、App的流行度预测^[8]、电影的受欢迎程度预测^[9]。部分深度学习模型主要分析了内容特征对预测性能的影响^[10],而另一部分较为新颖的研究工作则重点关注分类模型的性能^[11],其中较为著名的基于深度学习的分类模型是神经网络因子分解机(Neural Factorization Machine, NFM)^[12],它比传统的深度神经网络模型具有更优秀的特征表达能力。NFM是一种新颖的深度学习模型,结合了线性的二阶特征交互和非线性的高阶特征交互,能有效地学习稀疏特征。实际情况下的流行度受外部因素影响容易波动,难以捕捉,但是用户反馈事件的变化趋势在很大程度上影响了视频的流行度,NFM能够有效地学习内容特征,但它不能捕捉内容的流行趋势,因此如何结合视频流行度变化过程和内容特征建模是流行度预测研究工作的难题。

综合分析现有的研究成果可知,流行度动态变化过程难以捕捉,但视频的内容特征对流行度预测模型的性能有很大的影响。目前,联合流行度变化过程和内容特征建模的研究工作较为少见。LSTM网络应用于流行度的动态变化过程建模具有高效性^[13],能够有效捕捉流行度的变化趋势。以NFM为例的基于深度学习的模型有效地结合了线性的二阶特征交互和非线性的高阶特征交互,具有优秀的模型表达和泛化能力,但无法捕获视频的流行度变化趋势。对此,本文提出一种融合内容特征和时序信息的深度注意力视频流行度预测(Deep Attention video popularity prediction model Fusing Content and Temporal information, DAFCT)。该模型融合了视频的内容特征和时序信息,具有优秀的特征表达及泛化能力,并且能够捕获视频的流行趋势。

1 相关工作

1.1 基于LSTM网络的流行趋势预测

基于注意力机制的LSTM网络减少了对外部信息的依赖^[14],已经被广泛应用于文本分类、情感分析、点击率预测等领域。2017年,Wu等^[15]提出了一个深度时间上下文网络(Deep Temporal Context Network, DTCN),通过对facebook推文的流行度动态变化过程建模从而预测帖子的流行度,预测结果表明该模型对预测长期的流行动态有显著的能力。2018年,Yuan等^[16]通过将注意力引入LSTM网络对引文的流行趋势建模,从而预测引文的流行度,在computer science引文数据集上准确率可达85%。2019年,Varuna等^[17]通过分析Github上项目的Fork,Star等反馈事件,构建了时间序列信息,采用LSTM网络对Github上的时间序列信息建模并预测流行趋势。同年7月,在国际人工智能大会上,Liao等^[18]分析了在线文章的点击量、浏览量、转发量、点赞量等宏观事件发生的次数,将注意力机制引入LSTM网络对微信文章随时间的变化过程进行建模以预测流行趋势,进一步融合了内容特征,从而预测出微信文章的流行度。该文的流行度趋势建模过程如下,其中, c 是当前的视频, α_i^c 为注意力权重, V_i^c 为序列向量, h_i^c 为隐层输出, W_j^c 为权重矩阵。

$$\alpha_i^c = V_i^c \tanh\left(\sum_{j=1}^k W_j^c c + b^c\right) \quad (1)$$

$$\alpha_i^c = \frac{e^{\alpha_i^c}}{\sum_{j=1}^k \exp(\alpha_j^c)} \quad (2)$$

$$h_i^c = \sum_{j=1}^k \alpha_j^c c \quad (3)$$

从以上对流行趋势预测的最新研究成果的分析总结可知,LSTM网络能够很好地捕获视频的流行趋势,因此,本文采用了基于注意力机制的LSTM网络来捕捉视频的流行趋势。

1.2 基于深度学习的流行度预测模型

基于深度学习的流行度预测模型自2016年以来便一直是学术界和工业界的研究热点,对深度学习的流行度预测模型的关注始于2016年Chen等^[19]提出的基于深度神经网络(Deep Neural Network, DNN)的模型,该模型基于图像像素以及图像描述等内容特征建模,最初应用于广告的点击率预测,后来被广泛应用于流行度预测研究当中。为了提高预测性能,2017年He等^[20]对深度学习模型进行改进,提出了一个基于神经网络的协同过滤(Neural Collaborative Filtering, NCF)模型,经实验验证NCF模型的预测准确率达到87.30%。2019年,Luo等^[21]提取了新媒体中的政策信息,分析了政策信息中的内容特征,采用DNN建立了一个基于政策信息的流行度预测模型,实验结果显示该模型优于梯度提升决策树(Gradient Boost Decision Tree, GBDT)模型。基于深度学习的流行度预测模型大多都是基于DNN改进的,计算公式如下:

$$\hat{y} = \frac{1}{1 + e^{-z}} \quad (4)$$

$$z = f(x) \quad (5)$$

$$f(x) = \text{sigmoid}[W(y + b)] \quad (6)$$

其中: \hat{y} 为预估的流行度, W 为权重矩阵, b 为偏置向量。

以上研究成果只基于单个项目的内容特征进行建模,忽略了流行趋势对流行度预测模型性能的影响,在实际情况下视频的流行度常常受到外界因素的干扰不容易建模,因此如何联合内容特征和流行趋势建模是流行度预测研究亟待解决的问题。

2 本文模型

2.1 Attention-LSTM模型

本文采用LSTM网络捕捉视频的流行趋势,将注意力机制引入LSTM网络以降低外界因素的干扰,构建了基于注意力机制的Attention-LSTM模型。为了更形象地表征流行度的增长趋势,本文使用 O_{A-L} 表示视频流行度的趋势指数,由Attention-LSTM模型计算可得,Attention-LSTM模型如图1所示。通过分析用户的点赞、浏览、评论等反馈信息,给定时间间隔 t ,将随时间变化的用户反馈量构建为时间序列 d_t ,则得到反馈序列 $\{d_1, d_2, \dots, d_t\}$,LSTM网络的计算公式改写为:

$$G_i^i = \sigma(W^i x_{t-1} + U^i c_{t-1} + D^i h_{t-1} + b^i) \quad (7)$$

$$G_i^f = \sigma(W^f x_{t-1} + U^f c_{t-1} + D^f h_{t-1} + b^f) \quad (8)$$

$$c_t = G_i^i * h_{t-1} + G_i^f * \tanh(W^c x_{t-1} + D^c h_{t-1} + b^c) \quad (9)$$

$$G_i^o = \sigma(W^o x_{t-1} + U^o c_{t-1} + D^o h_{t-1} + b^o) \quad (10)$$

$$h_t = G_i^o * \tanh(c_t) \quad (11)$$

其中: G_i^i 、 G_i^f 、 G_i^o 分别为在给定时间间隔 t 时刻的输入门、遗忘门和输出门, h_{t-1} 为 $t-1$ 时刻的隐层输出, x_{t-1} 为 $t-1$ 时刻的输入; c_{t-1} 为 $t-1$ 时刻的细胞状态; σ 为sigmoid激活函数。

Attention-LSTM模型计算过程如下:

给定某一时间间隔 t ,将 n 个隐层输出记为 H_t :

$$H_i = (h_1, h_2, \dots, h_n) \quad (12)$$

这些隐层输出 h_i 经 softmax 层后, 得到注意力权重:

$$a_i = \text{softmax}(h_i) * [(G_i^v) * \tanh(c_i)] \quad (13)$$

将注意力权重记为 A_i :

$$A_i = (a_1, a_2, \dots, a_n) \quad (14)$$

则 Attention-LSTM 模型输出此时的流行趋势:

$$O_{A-L} = a_i \{G_i^o * \tanh(c_i)\} \quad (15)$$

由上述计算可得 O_{A-L} 的值为小于 1 的数, 表征某一视频的流行趋势指数, 接下来通过与 NFM 模型的预测结果结合, 进而计算出视频的流行度。从实验结果来看, Attention-LSTM 模型能够有效地捕获视频的流行趋势, 且对提高流行度预测模型的性能有很大帮助。

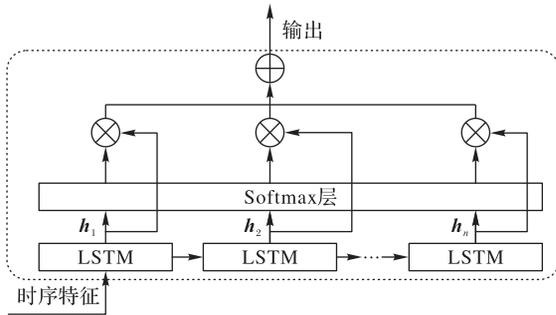


图 1 Attention-LSTM 模型
Fig. 1 Attention-LSTM model

2.2 NFM 模型

视频的内容包括视频类型和数值信息, 通常能为流行度预测提供有用的信息, 是影响视频流行度的关键因素之一。针对不同类型的视频, 不同用户的喜好不同, 用户的反馈表现则截然不同。本文采用 NFM 模型对视频的内容特征建模。NFM 模型首先采用 one-hot 编码技术将类型特征转换为 one-hot 向量, 然后将视频类型的 one-hot 向量输入到 NFM 模型的嵌入 (embedding) 层, 接着视频类型特征和数值特征通过一个二阶特征交互池层组合, 输入隐藏层后由激活函数计算得到输出结果。NFM 模型结合了线性的二阶特征交互和非线性的高阶特征交互, 能够从稀疏数据中学习特征, 有效地提高了特征的表达能。图 2 给出了用于内容特征学习的 NFM 模型。

例如, 给定视频类型的集合为 $M = \{m_1, m_2, \dots, m_k\}$, 对于第 $i (i = 1, 2, \dots, k)$ 个视频 m_i 的类型, 将视频类型的 one-hot 特征向量 x 使用嵌入技术进行降维后得到视频类型的 embedding 向量表示:

$$v_x = \{x_1 \lambda_1, x_2 \lambda_2, \dots, x_n \lambda_n\} \quad (16)$$

其中: λ_i 为第 i 个视频类型的 embedding 向量, x_i 为第 i 个视频类型的 one-hot 向量。

将 embedding 向量 v_x 输入到二阶交互层:

$$z_{BI} (v_x) = \sum_{i=1}^n \sum_{j=i+1}^n x_i \lambda_i \odot x_j \lambda_j \quad (17)$$

其中: \odot 表示 $(\lambda_i \odot \lambda_j)_k = \lambda_{ik} \lambda_{jk}$, 二阶交互层是一个池化操作, 它将文本特征的 embedding 矩阵转换成一个向量。二阶交互层的输出 $z_{BI} (v_x)$ 输入到隐藏层后的计算过程为:

$$\begin{cases} y_1 = \sigma(W_1 z_{BI}(v_x) + b_1) \\ y_2 = \sigma(W_2 y_1 + b_2) \\ \vdots \\ y_L = \sigma(W_L y_{L-1} + b_L) \end{cases} \quad (18)$$

其中: σ, W_L, b_L 分别为 sigmoid 函数、隐藏层的权重矩阵和偏置向量, L 为隐藏层的层数。

将隐藏层的输出 y_L 输入到全连接层后得到 NFM 模型的输出:

$$O_{NFM} = w_0 + \sum_{i=1}^n w_i x_i + q^T (\text{softmax}(W_L y_{L-1} + b_L)) \quad (19)$$

其中: $w_0 + \sum_{i=1}^n w_i x_i$ 为 NFM 模型的线性部分, w_0 为初始化的权重, w_i 表示第 i 个特征的权重; $q^T (\text{softmax}(W_L y_{L-1} + b_L))$ 为非线性部分, 即 DNN, q 为输出层的权重矩阵。

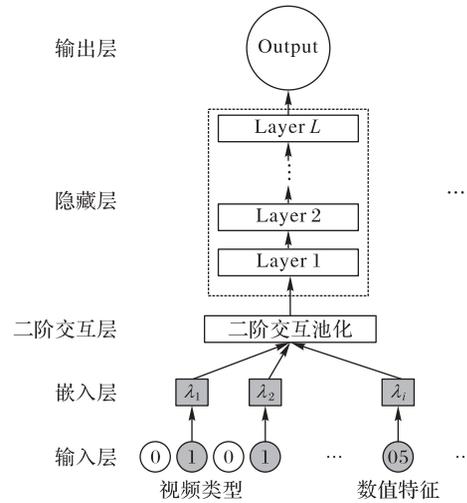


图 2 NFM 模型
Fig. 2 NFM model

2.3 DAFCT

流行度变化过程中的时序信息难以捕获, 而视频的内容特征很大程度上决定了视频的流行度, 是流行度预测任务必不可少的条件。本文的 DAFCT 首先采用 RNN 挖掘时序信息以捕获视频的流行趋势, 引入注意力机制排除外界因素的干扰; 然后采用深度神经网络处理内容特征, 针对稀疏的高维特征则采用嵌入技术进行降维以降低模型的计算复杂性; 最后, 使用 concatenate 方法融合时序信息和内容特征。本文 DAFCT 如图 3 所示。

如图 3 给定 n 个视频, 用 p_i 表示视频的流行度, 将所有的 n 个流行度 p_i 表示为 P :

$$P = \{p_1, p_2, \dots, p_n\} \quad (20)$$

$$p_i = \frac{1}{1 + e^{-z}} \quad (21)$$

式(21)为视频的点赞概率, 即为视频的受欢迎程度。结合 2.1 节 Attention-LSTM 模型的流行趋势和 2.2 节 NFM 模型的输出, 经一个全连接层计算:

$$z = W [O_{A-L}, O_{NFM}] + b \quad (22)$$

其中: O_{A-L} 为视频的流行趋势指数, O_{NFM} 是 NFM 模型的预测结果。代入式(21)后得到流行度:

$$p_i = \frac{1}{1 + e^{-W[O_{A-L}, O_{NFM}] + b}} \quad (23)$$

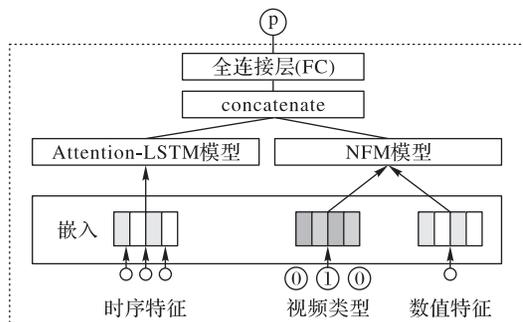


图3 DAFCT结构

Fig. 3 DAFCT structure

下面说明 DAFCT 的信息融合过程及应用:

1) 内容特征和时序信息嵌入。本文的研究工作将视频的统计信息分为时序信息和视频的内容,其中时序信息为给定时间间隔 t 内的用户反馈序列,视频的内容包括视频类型等信息。进行实验验证时,模型训练的时序信息对应的特征为时序特征,视频的内容则对应内容特征,视频的内容包括视频类型等特征。由图3可知,DAFCT使用了嵌入技术对高维的特征降维,从而得到低维的稠密特征。假设当前视频有7种类型,这7种视频类型表示为 $\{0, 1, 2, 3, 4, 5, 6\}$,如果第 i 个视频的类型为5,在数据预处理阶段使用 one-hot 编码技术得到该视频的类型 $\mathbf{x} = [0, 0, 0, 0, 0, 1, 0]$,由此得到一个稀疏的向量,而在实际情况中的数据维度往往是非常巨大的,直接输入模型会影响预测性能。针对这种情况,定义一个嵌入(embedding)向量 $\mathbf{em} = [0, 0.1, 0.25, 0, 0, 1, 0]$,则 $\mathbf{m} \cdot (\mathbf{em})^T = [0.5]$,可以看到这个值表征了该视频所属的类型。本文只是做了简单的假设,实际情况下的视频类型远远多于这个设定,数据维度往往大得多,计算复杂度高,嵌入向量是根据实验而设定的,各不相同。由此可知使用嵌入技术能有效降低特征向量的维度。

2) 信息提取。由图3可知,DAFCT的流行度趋势捕捉过程和视频的内容特征处理过程是并行的。Attention-LSTM模型首先根据输入的时序信息提取时序特征,输入的时序特征经遗忘门 G_i^f 后通过一个 sigmoid 函数计算出要丢弃的时序信息,然后根据要丢弃的信息更新旧的细胞状态;同时通过输入门 G_i^i 将细胞 \tilde{c}_i 存储的候选信息添加到新的细胞 c_i 中将时序信息存储下来;最后通过输出门 G_i^o 输出视频的流行趋势,计算过程如2.1节所示。NFM模型首先通过一个特征的二阶交互层,使得视频类型与其他内容特征如视频id等特征之间进行了交互组合,然后将这些组合的特征输入到隐藏层学习内容特征,最后通过一个激活函数输出预测结果,计算过程如2.2节所示。NFM主要是对组合的特征学习,而不是单一的特征,使用了2层隐藏层,相对于其他基于神经网络的模型更浅,有效地提高了模型的预测性能。

3) 信息融合。如图3所示,DAFCT使用了 concatenate 方法将 Attention-LSTM 模型捕捉的流行趋势和 NFM 模型挖掘的内容特征融合起来,计算公式为 $\mathbf{z} = \mathbf{W}[O_{A-L}, O_{NFM}] + \mathbf{b}$,其中 O_{A-L} 为 Attention-LSTM 模型的输出, O_{NFM} 为 NFM 模型的输出。

4) DAFCT 的应用。为了验证 DAFCT,本文设计了深度

注意力视频流行度预测(Deep Attention Video Popularity Prediction, DAVPP)算法求解该模型,并应用于视频的流行度预测。

算法1 DAVPP算法。

输入 历史反馈序列 d_i , $t-1$ 时刻的时序特征 x_{t-1} , 经 one-hot 技术编码后的内容特征向量 \mathbf{x} , 权重矩阵 \mathbf{W} 和 \mathbf{U} , 偏置向量 \mathbf{b} , 视频类型 m_i ;

输出 流行度 p_i 。

- 1) 初始化权重矩阵 \mathbf{U} 和 \mathbf{W} , 偏置向量 \mathbf{b} ; 嵌入矩阵 λ , 输出层的权重矩阵 \mathbf{q}
- 2) 利用式(7)、(8)计算输入门 G_i^i , 遗忘门 G_i^f
- 3) 利用 G_i^i, G_i^f 计算当前细胞状态 c_i
- 4) 利用式(10)计算输出门的值 G_i^o
- 5) 利用 G_i^o, c_i 计算 h_i
- 6) for all $h_i \in H_i$ and $a_i \in A_i$ do
- 7) $a_i = \text{softmax}(h_i) * [(G_i^o) * \tanh(c_i)]$
- 8) end for
- 9) 利用 a_i 和 h_i 计算 O_{A-L}
- 10) for all $i \in \{1, 2, \dots, n\}$ do
- 11) $\mathbf{v}_x \leftarrow \lambda_i x_i$
- 12) end for
- 13) for all $i \in \{1, 2, \dots, n\}$ do
- 14) while $j < i$ do
- 15) $z_{BI}(\mathbf{v}_x) \leftarrow x_i \lambda_i \odot x_j \lambda_j$
- 16) end while
- 17) end for
- 18) 利用 $z_{BI}(\mathbf{v}_x)$ 计算 y_1
- 19) for all $i \in \{1, 2, \dots, L\}$ do
- 20) $\mathbf{y}_L \leftarrow \mathbf{y}_{L-1}$
- 21) if $i = 1$ then
- 22) $\mathbf{y}_L = \mathbf{y}_1$
- 23) end for
- 24) 利用 \mathbf{y}_L 计算 O_{NFM}
- 25) 利用 O_{A-L}, O_{NFM} 计算 \mathbf{z}
- 26) while $i < n$ do
- 27) $p_i = \frac{1}{1 + e^{-\mathbf{z}}}$
- 28) end while
- 29) return p_i

3 实验与结果分析

实验的硬件平台: Dell服务器6个CPU核心, 1颗Nvidia Geforce Raytracing 2080共享GPU核, 内存32 GB, 显存11 GB。软件环境: 服务系统Ubuntu16.04, 终端系统Windows 10, 编译平台Pycharm Profession, Python 3.7, TensorFlow 1.9.0。

3.1 数据预处理

实验采用 kaggle (<https://www.kaggle.com>) 平台的 YouTube 视频数据集, 该数据集是情感分析、文本分类、流行度分析、时间序列变化分析以及训练机器学习算法研究的国际常用数据集之一。本文选取了数据集中2018年5月13日到8月26日的历史数据作为数据集, 数据预处理分为以下三个部分:

1) 时间序列特征构建。

本文的 Attention-LSTM 模型主要用于捕获视频的流行趋势, 因此在构建时间序列时以 view、like、dislike、video_id 等特征构建了时间序列, 没有考虑这些操作数的变化规律是否具

有周期性。

构建的宏观时间序列为:从数据集中抽出每隔一天的 view、like、dislike、video_id 操作数作为宏观时间序列来计算,并将这些数据按照时间顺序进行排序,加上 video_id 后得到时间序列数据集。

2) 内容特征。

为了保证实验的有效性,内容特征的选取在时间上与时序特征同步,内容特征包括 publish_time、video_type、channel、video_id 和 describe,内容特征的预处理步骤如下:

步骤 1 对每个样本计算点赞量与点踩量之和的比值,得到的结果若大于 0.5,则认为该视频内容是受用户欢迎的;反之则认为该视频不受欢迎。

步骤 2 将视频的受欢迎程度分为两类:“受欢迎(1)”和“不受欢迎(0)”。

步骤 3 去除数据集中特征取值相同的字段,这些字段会对流行度预测效果产生干扰。

最后,为了使得正负样本平衡,分别采用了上采样和降采样的方式对数据集做了处理,但从实验效果来看降采样方式最佳,于是最终本文采用了降采样方式对数据集做了处理,得到了 20 000 个正样本和 20 000 个负样本。最后选取数据集的前 80% 作为训练集,剩下的 20% 作为测试集。

3.2 评估指标

本文使用准确率(Acc)、召回率(re)和 F1 分数(F_1) 3 个指标评估 DAFCT 的性能。给定训练集 $V = \{v_1, v_2, \dots, v_n\}$, 正样本集为 $TR = \{TR_1, TR_2, \dots, TR_k\}$, 负样本集为 $FA = \{FA_1, FA_2, \dots, FA_{n-k}\}$ 。准确率是评估模型整体性能的指标,表示模型预测正确的样本数占所有样本数 $|V|$ 的比例:

$$Acc = (X + Y) / |V| \quad (24)$$

其中: X 表示实际 $v_i \in TR$, 且 DAFCT 算法预测所得为 $v_i \in TR$ 的样本数量; Y 表示实际 $v_i \in FA$, 且 DAFCT 算法预测为 $v_i \in FA$ 的样本数量。

召回率表示在所有的实际正例中,被正确预测的样本有多少:

$$re = X / (X + Z) \quad (25)$$

其中: Z 表示实际 $v_i \in FA$, 而 DAFCT 算法预测为 $v_i \in TR$ 的样本数量。 F_1 分数表示综合了准确率和召回率之后的性能指标:

$$F_1 = 2 * pre * re / (pre + re) \quad (26)$$

其中 pre 为精确度,表示被分为正例的样本中,实际为正例的比例:

$$pre = X / (X + U) \quad (27)$$

其中: U 表示实际 $v_i \in TR$, 而 DAFCT 算法预测为 $v_i \in FA$ 的样本数量。

3.3 实验设置

在使用 Attention-LSTM 模型挖掘时序信息以捕捉视频流行趋势时,采用了神经元为 128 的单层 LSTM 网络;使用 NFM 模型学习内容特征时,DNN 隐藏层的神经元个数为 128,激活函数使用的是 ReLU,同时利用 L2 正则化对模型参数进行约束;DAFCT 训练过程中的 batch size 取 1 024,模型在训练过程中训练了 60 个 epoch,输出层的激活函数采用 sigmoid 函数。实验流程如图 4 所示。

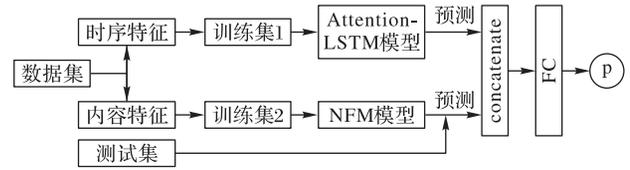


图 4 实验流程

Fig. 4 Experimental flow

3.4 实验与结果分析

为了便于观察模型的预测效果,实验分别与 Attention-LSTM 模型和 NFM 模型做了对比。

图 5 是 DAFCT 与 Attention-LSTM 模型和 NFM 模型的准确率、召回率和 F1 分数对比。其中:DAFCT 的准确率、召回率、F1 分数的均值分别为 0.892 6、0.780 3、0.765 3; Attention-LSTM 模型的准确率、召回率、F1 分数的均值分别为 0.756 9、0.672 1、0.667 3; NFM 模型的准确率、召回率、F1 分数的均值分别为 0.844 7、0.747 2、0.734 6。与 Attention-LSTM 模型相比,DAFCT 的召回率和 F1 分数分别提高了 10.82、9.80 个百分点;与 NFM 模型相比,召回率和 F1 分数分别提高了 3.31、3.07 个百分点。由此可见,Attention-LSTM 模型的整体性能不如 NFM 模型和 DAFCT,虽然 Attention-LSTM 模型能够有效地捕获视频的流行趋势,但对时间序列数据要求严格,而且特征的表达能力有限;而 NFM 模型虽然特征的表达能力强,但不能有效地挖掘视频的流行趋势,也存在一定缺陷。从实验结果来看,结合二者能够有效地提高流行度预测模型的性能。

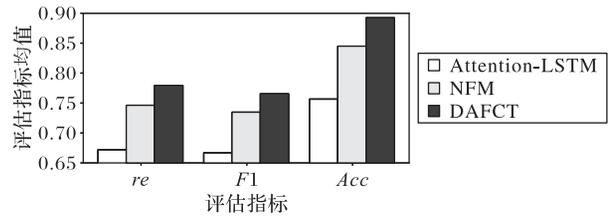


图 5 三个对比模型的召回率、F1 分数和准确率对比

Fig. 5 Comparison of recall, F1 score and accuracy of three comparison models

从图 6 可以观察到,整体上 Attention-LSTM 模型的 F1 分数最小,DAFCT 的 F1 分数最大。Attention-LSTM 模型在第 17 个 epoch 时 F1 分数突然增长,第 19 个 epoch 时突然下降,而后便一直增长,但增速不高,在第 56 个 epoch 达到平稳,可以看出此时 Attention-LSTM 模型开始收敛;NFM 模型在第 9 个 epoch 时 F1 分数突增到 0.734 0,在第 10 个 epoch 时增长至最大值 0.738 2 后在第 11 个 epoch 时降低,而后的几个 epoch 内都呈增长趋势,到了大约第 17 个 epoch 时突然下降并在接下来的几次训练过程中有轻微波动,在第 31 个 epoch 时趋于稳定,逐步收敛;而 DAFCT 在第 7 个 epoch 前的 F1 分数波动较大,由图 6 可清晰观察到从第 7 到 16 个 epoch 过程中的 F1 分数的增速最快,第 17 个 epoch 之后增速平稳,约第 30 到 48 个 epoch 过程中出现轻微的起伏,而后稳步收敛。由以上分析可知,从评估指标 F1 分数来看,与 Attention-LSTM 模型相比,本文的 DAFCT 的特征学习能力更强,虽然 DAFCT 的学习速度没有 NFM 模型快,但 DAFCT 更稳定,且整体性能更好。

表 1 显示了 DAFCT 和其他多种模型的预测性能对比,包括逻辑回归(Logistic Regression, LR)^[22]分类器、NFM 模型、随

机森林(Random Forest, RF)^[23]分类器、支持向量机(Support Vector Machine, SVM)^[24]、Attention-LSTM 模型。基于神经网络的预测模型大多是由 LR 演进而来,LR 通过给特征赋予权重因子来对样本进行分类;NFM 模型通过引入一个二阶交互层使得线性特征和非线性的高阶特征交互来提高模型的特征表达能力;RF 模型是基于决策树的分类器,它结合了多棵树对样本进行训练;SVM 是基于特征空间的线性分类器,通过最大化特征的间隔距离将分类问题转化为凸二次规划问题求解;Attention-LSTM 是基于循环神经网络的一种特殊的神经网络,通过引入注意力机制过滤掉无关信息。本文的 DAFCT 通过引入 Attention-LSTM 网络捕获视频的流行趋势,引入 NFM 模型处理内容特征,此外还引入了注意力机制使得特征的重要程度都能体现,从而达到了更好的预测性能。从实验结果可以看出本文提出的 DAFCT 在性能整体上明显优于对比模型。

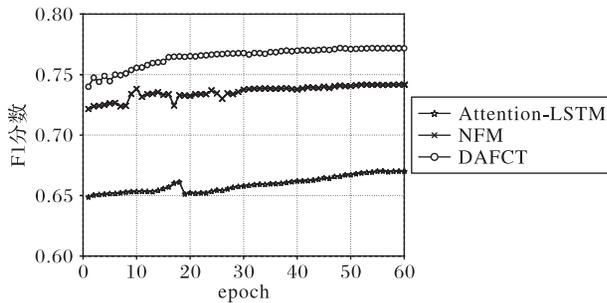


图 6 三个对比模型的 F1 分数对比

Fig. 6 Comparison of F1 scores of three comparison models

表 1 所提模型与其他模型的预测性能对比

Tab. 1 Prediction performance comparison of the proposed model and other models

模型	准确率	召回率	F1 分数
LR	0.7520	0.6531	0.6446
NFM	0.8447	0.7472	0.7346
RF	0.6587	0.4108	0.4230
SVM	0.8352	0.6609	0.6750
Attention-LSTM	0.7569	0.6721	0.6673
DAFCT	0.8926	0.7803	0.7653

为了观察 LR、NFM、RF、SVM、Attention-LSTM 和 DAFCT 的整体分类性能,绘制了各模型的准确率,如图 7 所示。从图 7 可以看出,RF 模型的分类性能较低,而本文的 DAFCT 的性能最好。RF 模型开始的学习速度最快,但很快就收敛且增速呈下降趋势;LR 模型的学习速度相对来说较为缓慢;SVM 和 NFM 模型虽然稳定,但分类的准确率没有 DAFCT 高。由图 7 清晰可见,本文提出的 DAFCT 在训练过程中学习速度快且准确率高。

基于视频内容特征的分类器性能很大程度上依赖于提取的内容特征,并且预测效果受到限制。Attention-LSTM 模型对时间序列要求严格,因此只能捕获视频的流行趋势,模型的特征表达能力较差,从以上实验结果来看存在的缺陷还比较大;基于内容特征的分类器如 NFM、LR、RF 等则无法捕捉视频的流行趋势。与其他分类器相比较,本文的 DAFCT 能够有效地捕获视频的流行趋势,并且能够有效提高流行度预测模型的性能,且特征的学习和表达能力更强。从以上实验结果分析可知,用户反馈事件的积累量能够有效地表达视频的流行趋

势,而且能够帮助提高流行度预测模型的性能,即这些用户反馈的时序信息是影响流行度的重要因素;除此之外,内容特征很大程度上影响了流行度预测模型的效果,是流行度预测研究工作当中必不可少的,因此结合二者才能够使得流行度预测模型的性能更佳。

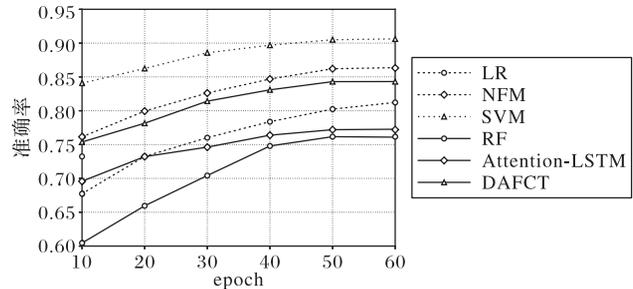


图 7 各对比模型的准确率对比

Fig. 7 Accuracy comparison of models

4 结语

本文提出一种融合内容特征和时序信息的深度注意力流行度预测模型 DAFCT,能够有效地表达视频的流行趋势且提高流行度的预测性能。为了排除外界因素的干扰,增加了注意力机制,采用 Attention-LSTM 模型挖掘时序信息以捕捉流行趋势,利用 NFM 模型处理内容特征,模型成功地捕捉了流行趋势并得到了较好的预测性能。实验结果显示,本文的 DAFCT 的分类性能优于其他分类器,但也可以看出本文设计的 Attention-LSTM 模型存在局限性,在今后的研究工作当中,将会进一步把模型应用在其他领域,探索该模型存在不足的原因并进行改善;另外,还将对 DAFCT 的应用及模型的高效求解算法做进一步的研究。

参考文献 (References)

- [1] 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.
- [2] 王慧健,刘峥,李云,等. 基于神经网络语言模型的时间序列趋势预测方法[J]. 计算机工程, 2019, 45(7): 13-19, 25. (WANG H J, LIU Z, LI Y, et al. Trend prediction method of time series trends based on neural network language model [J]. Computer Engineering, 2019, 45(7): 13-19, 25.)
- [3] JHANG W S, GAO S E, WANG C M, et al. Share price trend prediction using attention with LSTM structure [C]// Proceedings of the 20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing. Piscataway: IEEE, 2019: 208-211.
- [4] WU T, LIU C C, HE C. Prediction of regional temperature change trend based on LSTM algorithm [C]// Proceedings of the IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference. Piscataway: IEEE, 2020: 62-66.
- [5] KUMAR S D, SUBHA D P. Prediction of depression from EEG signal using Long Short Term Memory (LSTM) [C]// Proceedings of the 3rd International Conference on Trends in Electronic and Informatics. Piscataway: IEEE, 2019: 1248-1253.
- [6] DU B W, PENG H, WANG S Z, et al. Deep irregular convolutional residual LSTM for urban traffic passenger flows prediction [J]. IEEE Transaction on Intelligent Transportation Systems, 2020, 21(3): 972-985.

- [7] 任佳智, 田辉, 范绍帅, 等. 基于用户偏好预测的无人机部署和缓存策略[J]. 通信学报, 2020, 41(6): 1-13. (REN J Z, TIAN H, FAN S S, et al. UAV deployment and caching scheme based on user preference prediction[J]. Journal on Communications, 2020, 41(6): 1-13.)
- [8] 张艺璇, 郭彬, 刘佳琪, 等. 基于多级注意力机制网络的 app 流行度预测[J]. 计算机研究与发展, 2020, 57(5): 984-995. (ZHANG Y X, GUO B, LIU J Q, et al. app popularity prediction with multi-level attention networks [J]. Journal of Computer Research and Development, 2020, 57(5): 984-995.)
- [9] ABIDI S M R, XU Y L, NI J Y, et al. Popularity prediction of movies: from statistical modeling to machine learning techniques [J]. Multimedia Tools and Applications, 2020, 79(47/48): 35583-35617.
- [10] CHEN G D, KONG Q C, MAO W J. An attention-based neural popularity prediction model for social media events [C]// Proceedings of the 2017 IEEE International Conference on Intelligence and Security Informatics. Piscataway: IEEE, 2017: 161-163.
- [11] DING K Y, WANG R G, WANG S Q. Social media popularity prediction: a multiple feature fusion approach with deep neural networks [C]// Proceedings of the 27th ACM International Conference on Multimedia. New York: ACM, 2019: 2682-2686.
- [12] HE X N, CHUA T S. Neural factorization machines for sparse predictive analytics [C]// Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2017: 355-364.
- [13] CHEN Q, QI X. Research on trend analysis and prediction algorithm based on time series [C]// Proceedings of the 3rd International Conference on Electronic Information Technology and Computer Engineering. Piscataway: IEEE, 2019: 73-76.
- [14] MOU H L, LIU Y H, WANG L. LSTM for mobility based content popularity prediction in wireless caching networks [C]// Proceedings of the 2019 IEEE Globecom Workshops. Piscataway: IEEE, 2019: 1-6.
- [15] WU B, CHENG W H, ZHANG Y D, et al. Sequential prediction of social media popularity with deep temporal context networks [C]// Proceedings of the 26th International Joint Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2017: 3062-3068.
- [16] YUAN S, ZHANG Y, TANG J, et al. Modeling and predicting popularity dynamics via deep learning attention mechanism [EB/OL]. (2018-11-06) [2020-05-23]. <https://arxiv.org/pdf/1811.02117.pdf>.
- [17] VARUNA T V, MOHAN A. Trend prediction of GitHub using time series analysis [C]// Proceedings of the 10th International Conference on Computing, Communication and Networking Technologies. Piscataway: IEEE, 2019: 1-7.
- [18] 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.
- [19] CHEN J X, SUN B G, LI H, et al. Deep CTR prediction in display advertising [C]// Proceedings of the 24th ACM International Conference on Multimedia. New York: ACM, 2016: 811-820.
- [20] HE X N, LIAO L Z, ZHANG H W, et al. Neural collaborative filtering [C]// Proceedings of the 26th International Conference on World Wide Web. Republic and Canton of Geneva: International World Wide Web Conferences Steering Committee, 2017: 173-182.
- [21] LUO Y, WANG F F, ZHAO F F, et al. A framework for policy information popularity prediction in new media [C]// Proceedings of the 2019 IEEE International Conference on Intelligence and Security Informatics. Piscataway: IEEE, 2019: 209-211.
- [22] WANG Y H, OU Y, DENG X D, et al. The ship collision accidents based on logistic regression and big data [C]// Proceedings of the 2019 Chinese Control and Decision Conference. Piscataway: IEEE, 2019: 4438-4440.
- [23] DEMIDOVA L, LVKINA M. Defining the ranges boundaries of the optimal parameters values for the random forest classifier [C]// Proceedings of the 1st International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency. Piscataway: IEEE, 2019: 518-522.
- [24] DAI H. Research on SVM improved algorithm for large data classification [C]// Proceedings of the IEEE 3rd International Conference on Big Data Analysis. Piscataway: IEEE, 2018: 181-185.

This work is partially supported by the National Natural Science Foundation of China (61462014), the Science and Technology Foundation of Guizhou Province (S2017JP00401185093).

WU Wei, born in 1995, M. S. candidate. Her research interests include popularity prediction, streaming media.

LI Zeping, born in 1964, Ph. D., professor. His research interests include popularity prediction, computer network, streaming media.

YANG Huawei, born in 1965, Ph. D., professor. Her research interests include financial risk management.

LIN Chuan, born in 1975, M. S., associate professor. His interests include computer network, cloud computing.

WANG Zhongde, born in 1963, experimentalist. His interests include streaming media.