

基于 LSTM-Att 方法的音乐流行趋势预测

王振业¹, 叶成绪^{1*}, 王文韬¹, 杨萍²

(1. 青海师范大学 计算机学院, 青海 西宁 810008;

2. 青海师范大学 地理科学学院, 青海 西宁 810008)

摘要:利用循环神经网络的分支长短时记忆网络与注意力机制结合的模型进行音乐流行趋势的预测。首先,分析了传统的支持向量机以及循环神经网络等方法在预测时间序列数据上不能捕捉长时间序列信息等不足之处;其次,基于以上分析建立了长短时记忆网络加注意力机制结合的预测模型,针对所要预测的未来两个月歌手歌曲播放量,对数据集进行分析及相关属性选取、归一化等预处理,选取组合相应的歌曲日播放量、连续3天播放均值作为相应时间点的样本构建神经网络训练集;最后,设计实现了基于长短时记忆网络加注意力机制相结合的预测模型实验。实验结果表明,所使用的预测模型较传统的机器学习方法支持向量机以及长短时记忆网络等在均方根误差和平均绝对误差两个指标上取得了较为明显的提升。

关键词:音乐流行趋势;时间序列;长短时记忆网络;注意力机制;支持向量机

中图分类号:TP181

文献标识码:A

文章编号:1673-629X(2020)09-0188-06

doi:10.3969/j.issn.1673-629X.2020.09.034

Music Trend Forecast Based on LSTM-Att Method

WANG Zhen-ye¹, YE Cheng-xu^{1*}, WANG Wen-tao¹, YANG Ping²

(1. School of Computer, Qinghai Normal University, Xining 810008, China;

2. School of Geography, Qinghai Normal University, Xining 810008, China)

Abstract: Prediction of music popularity trend is carried out by using a model that combines the long short-time memory network of the cyclic neural network and the attention mechanism. Firstly, the shortcomings of traditional support vector machines and cyclic neural networks in predicting time series data like not capturing long time series information are analyzed. Secondly, based on the above analysis, the prediction model of the combination of long short-time memory network and attention mechanism is established. According to the predicted amount of songs played by artists in the next two months, the data set is analyzed and pre-processed such as relevant attribute selection and normalization, and the daily amount of songs played in combination and the average amount of songs played in 3 consecutive days are selected as samples at corresponding time points to construct the neural network training set. Finally, the prediction model experiment based on the combination of long short-time memory network and attention mechanism is designed, which shows that the proposed prediction model has achieved significant improvement over the traditional machine learning method support vector machine and long short-term memory network in terms of root mean square error and average absolute error.

Key words: music trend; time series; long short-term memory network; attention mechanism; support vector machine

0 引言

流行音乐作为当下人们文娱消费的产品之一,越来越受到人们的关注。根据速图研究院研究报告指出,2013年到2018年中国的移动音乐市场呈现井喷的模式。在当今大数据的背景下,众多的流行音乐听众在很大程度上能够决定流行音乐的发展趋势,这主要体现在众多听众在各大音乐平台上的诸如试听、收藏、下载等行为对流行音乐的影响,这些影响反映出了

听众对相应音乐的好感和偏爱。针对流行音乐趋势的预测,可以在海量用户数据的基础上,对听众的好感和偏爱进行深层次的挖掘和分析来获得。收集整理各大平台的海量曲库资源以及用户行为,从而形成流行音乐大数据集,对该数据集进行特定属性的精准分析,最终可以实现对流行音乐趋势走向的精准把控,也实现了用户喜好偏爱的聚合决定流行音乐的趋势走向。

针对流行音乐趋势走向的预测,国内外现在还鲜

收稿日期:2019-10-20

修回日期:2020-02-25

基金项目:国家自然科学基金(61762075);青海省应用基础研究(2018-ZJ-787,2016-ZJ-739)

作者简介:王振业(1992-),男,硕士,研究方向为机器学习及其应用;叶成绪,博士,教授,研究方向为机器学习及其应用。

有研究,2016年阿里举办了音乐流行趋势预测大赛开启了该研究领域的大门。阿里音乐经过将近7年的发展和沉淀,如今已经拥有了将近百万首的曲库资源和数以亿次的用户试听、收藏、下载等行为。通过对历史用户行为数据的分析,预测下一阶段某些艺人或歌曲的播放量,从海量的数据中挖掘出未来的黑马歌手,从根本上精准把控流行音乐的发展趋势,是众多音乐平台的关注焦点。文中在阿里提供的用户音乐数据2015年3月到2015年8月的历史播放数据上,组合长短时记忆网络(long short-term memory, LSTM)和注意力机制(attention mechanism)方法,构建 LSTM-Attention 预测模型,实现对2015年9月到2015年10月艺人播放量较为精准的预测。

1 相关研究

1.1 传统预测方法

对时间序列数据进行预测时,传统的方法包括:支持向量机(support vector machine, SVM)、BP(back propagation)神经网络以及循环神经网络(recurrent neural network, RNN)等^[1]。支持向量机最早是在20世纪90年代中期,由Cortes和Vapnik所提出的。当样本规模较小、维度较高且样本间呈非线性状态时,SVM能够充分发挥其自身优势来解决这些问题。

以VC(Vapnik-Chervonenkis dimension, VC)维理论与结构风险最小原理为基础,SVM模型被建立起来^[2]。通过有限的信息,SVM模型以期在模型的复杂程度和学习技巧、方法之间找到最佳平衡点,从而获得极佳的泛化能力。

SVM的核心思想就是基于训练集在样本空间中找到一个超平面将不同类别的样本分开。基于样本空间E,SVM在寻找超平面时采用如下的方式来表述:

$$\mathbf{V}^T \mathbf{X} + a = 0 \quad (1)$$

其中, $\mathbf{V} = (V_1, V_2, \dots, V_n)$ 为法向量,决定了超平面的方向; a 代表位移发生的程度,SVM所寻求的超平面与 $(0,0)$ 之间的距离由 a 所确定。经过推导后的SVM基本模型如下:

$$\min_{\mathbf{V}, a} \frac{1}{2} \|\mathbf{V}\|^2 \quad (2)$$

$$\text{s. t. } y_i(\mathbf{V}^T \mathbf{X} + a) \geq 1, i = 1, 2, \dots, m$$

SVM本身可以通过决策边界的边缘来控制模型的能力^[3],同时,它在小样本训练集上能够得到比其他算法好很多的结果。相较于其他模型,在对目标进行优化时,SVM在确保分类准确率的基础上,对模型的VC维进行降维,从而使得模型在整个样本数据上的期望误差得到了很好的控制,这种方法使得SVM对新鲜样本具有很强的适应能力。然而不可忽视的是

SVM本身所存在的缺陷:(1)SVM算法无法对大规模的样本进行训练;(2)SVM算法对于解决多分类问题存在困难;(3)SVM算法对缺失的数据敏感,对参数和核函数的选择也很敏感。近年来有学者指出,在进行大规模时间序列预测问题上,SVM存在较为严重的过拟合现象。

循环神经网络是一类用于处理序列数据的神经网络。相对于深度神经网络(deep neural network, DNN)、卷积神经网络(convolutional neural networks, CNN),RNN的不同之处在于:它不仅考虑前一刻的输入,并且赋予了网络对前面信息的一种“记忆”功能^[4]。具体的表现形式为网络会对前面的信息进行记忆并应用于当前输出的计算中,即隐含层之间的节点不再是无连接的而是有连接,并且包括上一时刻隐含层的输出。

RNN结构与传统的前馈神经网络不同,其存在记忆效应,正是这种记忆使得RNN网络可以对上下文进行分析,具体到时间序列数据中,它可以做到对前后文数据的连接。然而,在实践中RNN也存在不可避免的缺陷:无法实现对较长序列数据的记忆^[5]。一个主要的原因是,RNN在训练中很容易发生梯度爆炸和梯度消失。文献[3]指出,利用RNN在进行时间序列数据的预测时,极易发生梯度消失现象,这导致训练时梯度不能在较长的序列汇中一直传递下去,从而使得RNN无法捕捉到长距离的影响。

1.2 文中使用的模型

(1) 长短时记忆网络模型。

利用长短时记忆网络来进行音乐时序数据的相关预测。长短时记忆(long short-term memory, LSTM)神经网络模型是RNN的一种特殊变形,由瑞士人工智能实验室Hochreiter & Schmidhuber(1997)提出,并在近期由Alex Grave进行了改良和推广^[6]。此处主要介绍2000年Gers改进后的LSTM算法。在传统的RNN中,使用反向传播算法(back propagation through time, BPTT)训练RNN,存在以下问题:(1)网络中输出步长的梯度计算不仅取决于当前步长,还依赖于前一或许多步长的计算;(2)由于存在长期的依赖,会造成梯度消失或梯度爆炸问题^[7]。LSTM模型就是在此基础上被提出的。

LSTM跟RNN最大的不同之处在于,为了判断网络信息流中相应的信息是否有效^[8],LSTM网络中加入了“处理器”,该处理器的作用就是甄别当前输入和上一时刻的输出是否对接下来的信息修整有用,这个“处理器”通常被称为“细胞”^[9]。在LSTM中,每个“细胞”中又被放入了三个“门”,分别是:输入门、遗忘门和输出门。正是通过这三个门,LSTM可以处理相

应的信息。

(2)注意力机制。

2014 年,在人类视觉注意力机制的基础上,Bahdanau 等人首次将注意力模型应用到了自然语言处理领域。同时,注意力模型被应用到基于神经网络的机器翻译中并取得了当时最好的效果。随着深度学习的发展,注意力机制因其优异的表现而得到广泛应用,以时序序列为代表的领域也开始引入注意力机制进行相关的研究^[10]。以文中对阿里音乐平台音乐流行预测研究为例,由 LSTM 层得到相应的输出后,可以对这些输出信息进行计算,进一步得到相关信息的注意力概率分布,这为掌握时序序列中每一时刻 LSTM 细胞单元的输出状态对预测未来歌手歌曲流行

趋势提供了重要依据。

(3)LSTM-Attention 模型。

LSTM-Attention 模型就是在 LSTM 结构后加上注意力层。虽然 LSTM 与 Attention 的组合模型早被用于关系分类中并取得了较好的效果^[11],但是该组合模型在时序序列数据尤其是公共数据中的预测研究上还处于不断探索改进的地步。文中利用 LSTM 与 Attention 相结合的机制,对阿里云音乐平台上的相应歌手及歌曲进行预测研究,利用过往历史数据选取组合“日播放量”、“连续 3 天播放均值”2 个维度来预测未来 60 天的歌手歌曲播放量。所采用的整体模型结构如图 1 所示。

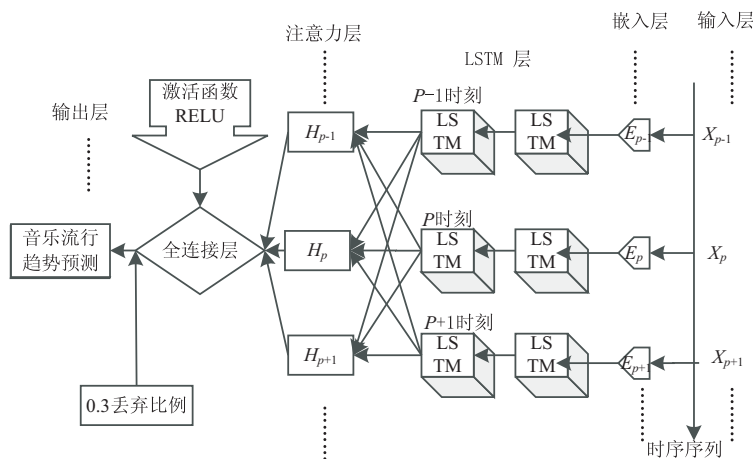


图 1 所采用的 LSTM-Attention 模型结构

2 相关实验及分析

2.1 样本数据描述

2.1.1 样本数据来源

文中所用到的数据集来自于阿里云音乐平台^[12]。阿里云音乐是阿里巴巴旗下的一个大型音乐平台,该平台经过多年的发展积累了大量的用户以及曲库资源,不论是歌手数量还是歌曲种类,该平台都在同类平台中名列前茅。更重要的是,在平台下,存在丰富的用户行为:上传歌曲、下载专辑等。无论是丰富的资源还是用户的活跃程度都对把握音乐流行趋势起到了极其重要的指引作用^[13]。

针对该音乐平台下的数据信息,平台给出的样例数据是两张表,一张是用户行为表(时间跨度为 2015.03.01 ~ 2015.08.31) p2_mars_tianchi_users_action,该表的作用是对阿里音乐平台下用户的行为进行表征;另一张是歌曲信息表 p2_mars_tianchi_songs,该表所表述信息包含:歌手以及相应歌曲的相关信息,如所属专辑、初始热度等。

文中所用到的数据集 p2_mars_tianchi_user_

actions.csv 和 p2_mars_tianchi_songs.csv 中涵盖了(2015.03.01 ~ 2015.08.30)共计 6 个月的用户行为信息和歌曲信息,经过对数据集的统计,其中所包含歌曲的初始播放总热度(即历史播放人次)为 26 107 万。用户的播放行为记录为 478 万。

2.1.2 数据预分析及预处理

数据分析包括数据质量分析和数据特征分析等。数据质量分析是数据准备过程中重要的环节,是数据预处理的前提。数据质量分析主要是检查原始数据中是否存在脏数据,包括缺失值、异常值(用户表出现新歌曲或用户表行为不在 20150301 ~ 20150831 等)、重复数据以及含有特殊符号的数据。由于数据集中存在着大量对进行预测无意义的数(比如 gmt_create 即用户播放时间),因此要考虑删除。对原始数据集 p2_tianchi_user_actions.csv 进行处理,可得到两个文件:一个是 song_p_d_c.txt,该文件包含 songs_id、每天的播放量、每天的下载量以及每天的收藏量;另一个是 song_fan.txt,包含 songs_id 以及每天不同用户的数目。同理,利用 ifNoArtistTXT() 对 artist 的 songs 进行整理。

(1) 相关属性选取。

为了提高预测模型的准确率、效率及降低算法的复杂度,在进行实验之前需对数据集中的若干个属性进行选择,挑选组合能够决定或者影响到未来歌手歌曲播放量的因素。传统情况下通常采用维度变化、特征转换等处理来达到目的^[14]。文中要对阿里云音乐平台未来 60 天歌手歌曲播放量进行预测,通过相关先验知识可知,所用到的数据集之间很多属性间不存在相关性,例如用户播放时长等。因此,为了降低模型的复杂度,考虑删除不相关属性。文中利用最大信息系数方法来寻找各个属性间的相关性。

最大信息系数(也称互信息方法)是一种具备广泛性和公平性的标准,该方法的主要作用是权衡众多数据中相关两个变量 X 、 Y 之间存在的依赖关系。假设给定两个离散型变量 X 和 Y ,定义两者之间的互信息系数 $R(X, Y)$ 计算方法如下:

$$R(X, Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad (3)$$

其中, $p(x)$ 、 $p(y)$ 分别为变量 X 、 Y 的边缘概率分布, $p(x, y)$ 为两者的联合概率分布。

文中利用以上方法,在删除不相关属性后的数据集中计算以下四个属性:播放量、播放均值、播放量方差和播放时长,最终选取“日播放量”和“连续 3 天播放均值”作为未来歌手歌曲播放量预测的组合指标。

(2) 数据预处理。

在获取的数据集中存在诸如数据缺失、脏数据以及异常数据等不符合要求的数据,因此在模型构建之前需要对这些异常数据进行处理,其处理的步骤如下:

Step1: 数据缺失值处理。

在获取的 470 多万条数据中存在数据缺失现象,文中采用均值填补法进行数据缺失值的填充,具体方法如下:

$$X_t = \frac{1}{2}(X_{t-1} + X_{t+1}) \quad (4)$$

其中, X_t 代表当前时刻缺失的信息, X_{t-1} 代表前一时刻的信息, X_{t+1} 代表后一时刻的信息。

Step2: 数据平滑。

数据平滑的目的是将数据集相关属性中存在的偏度较大的数据进行转化,使其更加服从高斯分布,为取得较好的后续分类结果奠定基础。文中利用最小二乘法对数据进行平滑;

Step3: 数据归一化。

在利用深度学习方法尤其是神经网络进行模型构建时,数据集中数值越大的数据对模型的影响比重越高,这就导致无形中可能丢失那些数值较小的数值属性。因此要对数据集中的数据进行归一化处理,文中

利用最小最大标准化(Min-Max 标准化)方法将数据归一化到 $[0, 1]$:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (5)$$

其中, X' 表示归一化完成后的数据, X 表示原始数据, X_{\max} 和 X_{\min} 分别表示原始数据中的最大值和最小值。

经过上述数据预处理后,数据集大致的格式如表 1 所示。

表 1 经预处理后的数据内容

列名	类型	说明	示例
artist_id	String	歌曲属艺人 Id	023406156015e8
Plays	String	当天播放数据	5 000
Ds	String	日期	20150901

2.1.3 数据预分析及预处理

文中的最终目标是预测相应歌手歌曲的播放量,因此在处理数据集时应定位到歌手歌曲播放量这一属性。针对数据集给出的 2015.03.01 ~ 2015.08.30 这六个月数据,统计并制作歌手歌曲播放量的变化趋势图,选取组合日播放量、连续 3 天播放均值,作为某一时间点的样本值,“滑动”构建神经网络的训练集。其模型参数设置如下:

- (1) 输入节点数=2;
- (2) 输出节点数=1;
- (3) 传播步长=64;
- (4) 隐含层节点个数=128;
- (5) 激活函数=ReLU;
- (6) 正则优化(Dropout 丢弃比例)=0.3;
- (7) 学习率=0.001。

2.2 实验结果

图 2 所示为任意一名歌手过去六个月内歌曲的平均播放、下载和收藏量。从中可以看出,某位歌手的歌曲播放量、收藏量及下载量三者中,最适合用来预测未来某段时间内某位歌手是否能流行起来的是播放量。

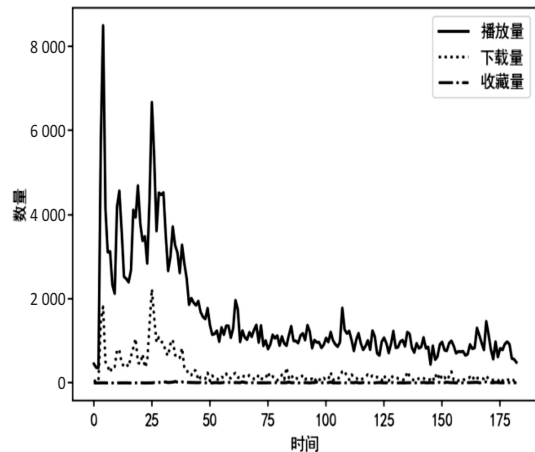


图 2 歌手歌曲的平均播放、下载和收藏量

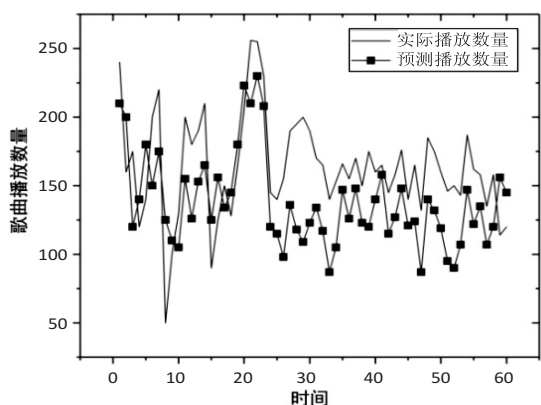


图 3 歌手歌曲预测播放量和真实播放量比对

针对所有歌手的歌曲播放数据,进行去均值、方差归一以及缩放到 $[-1,1]$,同时设置均值滤波的长度为 $aveFilter=4$ 。利用原始的播放量进行预测,能够得到如图 3 所示的预测曲线。从图 3 可以看出,对原始数据进行预测所得的结果中,在某些歌手的预测上存在偏差。

图 4 为利用歌手平均歌曲播放量预测未来 60 天歌手歌曲播放量的实验结果(为缩放后的结果)。由图 4 可以看出,使用平均歌曲播放量进行预测的结果要比使用原始歌曲播放量进行预测的效果好,预测结果基本和未来 60 天歌手的真实歌曲播放量吻合。

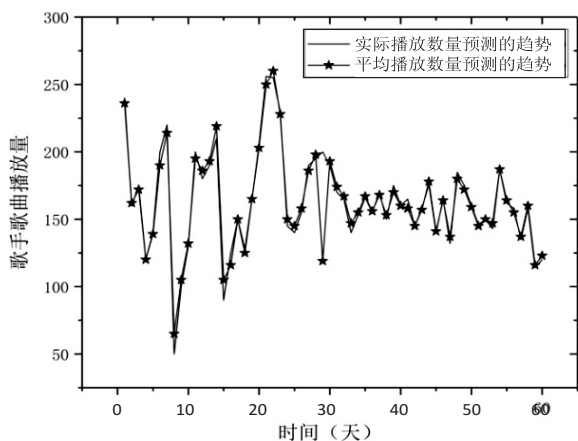


图 4 利用平均播放数据进行预测后歌手歌曲预测播放量和真实播放量比对

对比其他预测方法时,利用了 RMSE、MAE,其具体的计算方法如下:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n [y - \hat{y}]^2} \quad (6)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y - \hat{y}| \quad (7)$$

其中, y 为观测值, \hat{y} 为原始值, n 为观测次数。

文中实验在数据集中任意选取 10 名艺人,分别利用 SVM、LSTM 以及文中用到的 LSTM-Attention 算法预测 2015 年 8 月 1 日到 2015 年 8 月 31 日艺人歌曲播

放量,预测结果—均方根误差、平均绝对误差如图 5 和图 6 所示。

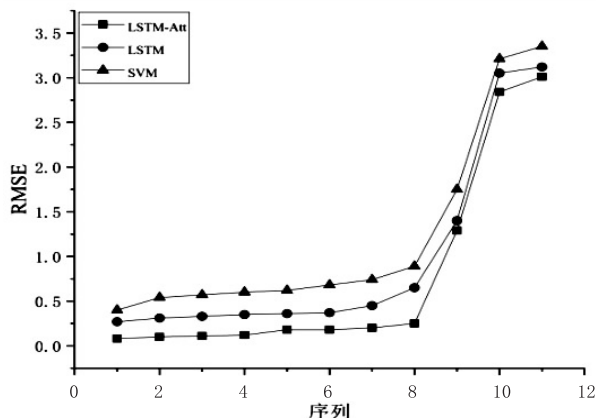


图 5 三种模型在 RMSE 指标上的比对

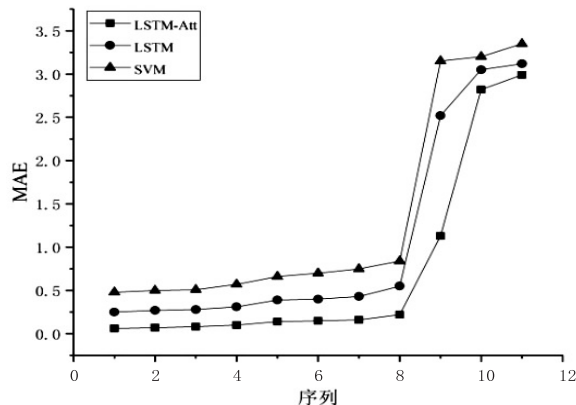


图 6 三种模型在 MAE 指标上的比对

由图 5 和图 6 可以看出,在对 10 名艺人的歌曲播放量预测中,文中提出的预测模型在预测效果 (RMSE、MAE 指标)上,相比于传统方法中的 LSTM 以及 SVM,分别由原来的 0.073、0.046 降低到 0.046、0.033,误差率分别降低了 36.8% 和 28.3%。证明文中所提出的模型更适合对音乐流行趋势进行预测。图 7 显示的是不同的模型参数的 RMSE、MAE 两个指标的影响程度,不难发现,在一定范围内随着模型层数的

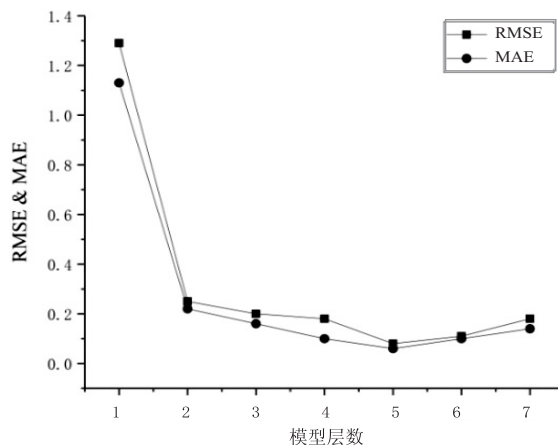


图 7 不同预测模型层数对 RMSE 和 MAE 的影响

增加,两个指标都在下降,但是超过一定的范围后,又呈上升趋势。文中模型所采用的层数为5层。

表2为模型中隐层单元个数以及学习率两个参数对结果的影响,主要体现在训练集以及测试集的均方

根误差和平均绝对误差两个指标中。可以看出,隐层单元个数为64,学习率为0.001时,所得实验结果优于其他参数组合。

表2 不同参数对模型结果的影响

Parametr	Train_rmse	Test_rmse	Train_mae	Test_mae
Cell=64	0.06	0.2	0.04	0.16
Cell=128	0.07	0.12	0.04	0.18
Cell=256	0.61	3.01	0.31	2.99
Lr=0.1	0.61	0.18	0.08	0.15
Lr=0.01	0.07	0.12	0.04	0.18
Lr=0.001	0.06	0.08	0.04	0.06

3 结束语

提出了一种基于 LSTM-Attention 的音乐流行趋势预测模型,即根据过往歌手歌曲的均值播放量来预测未来一段时间内某位歌手的歌曲能否流行起来。利用阿里云音乐上的音乐数据评估了该模型,预测结果超过了利用传统机器学习方法进行的预测,证实了 LSTM-Attention 模型更适合于处理时间序列数据。

未来,在该研究的基础上还可以进行以下工作:文中采用的是基于 LSTM-Attention 模型的预测方法,接下来可以尝试采用 CNN-LSTM-ResNet 的方法进行相关数据的预测,因此后续工作中也可以尝试单模型组合,以期提升预测效果;文中在进行相应特征选择时,影响艺人歌曲播放量的因素考虑略有不足,使得预测结果略差,因此后期可进一步挖掘出对艺人歌曲播放量有影响的其他特征。

参考文献:

- [1] 张蕾,章毅. 大数据分析的无限深度神经网络方法[J]. 计算机研究与发展, 2016, 53(1): 68-79.
- [2] LECUN Y, BENGIO Y, HINTO G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [3] DE JESUS O, HAGAN M T. Backpropagation algorithms for a broad class of dynamic networks[J]. IEEE Transactions on Neural Networks, 2007, 18(1): 14-27.
- [4] CAI Xulin, SHU Jian, AL-KALI M. Link prediction ap-

proach for opportunistic networks based on recurrent neural networks[J]. IEEE Access, 2019, 7: 2017-2025.

- [5] ZHANG Yongzhi, XIONG Rui, HE Hongwen, et al. Long short-term memory recurrent neural networks for remaining useful life prediction of lithium-ion batteries[J]. IEEE Transaction on Vehicular Technology, 2018, 67(7): 5695-5705.
- [6] RAO Guozheng, HUANG Weihang, FENG Zhiyong. LSTM with sentence representations for document-level sentiment classification[J]. Neurocomputing, 2018, 308: 49-57.
- [7] 李大利. 一种时间序列预测方法的应用研究[D]. 长春: 吉林大学, 2008.
- [8] AREL I, ROSE D C, KARNOWSKI T P. Deep machine learning—a new frontier in artificial intelligence research [research frontier][J]. IEEE Computational Intelligence Magazine, 2010, 5(4): 13-18.
- [9] 郭冠呈, 刘书明, 李俊禹, 等. 基于双向长短时神经网络的水量预测方法研究[J]. 给水排水, 2018, 44(3): 123-126.
- [10] 邢吉亮. 结合注意力机制的 Bi-LSTM 循环神经网络对关系分类的研究[D]. 长春: 吉林大学, 2018.
- [11] 焦李成, 杨淑媛, 刘芳, 等. 神经网络七十年: 回顾与展望[J]. 计算机学报, 2016, 39(5): 1-22.
- [12] 吕倩倩. 基于机器学习的音乐流行趋势预测[D]. 兰州: 兰州大学, 2017.
- [13] 颜家康. BP 神经网络算法在音乐流行趋势预测中的应用研究[D]. 成都: 西南交通大学, 2016.
- [14] 彭晓冰, 朱玉全. 基于特征内相关和互信息的加权 SVM 算法[J]. 计算机科学, 2018, 45(12): 182-186.