

基于用户收视行为与评论情感分析的 收视预测研究

冯小丽,吴肇良,殷复莲

(媒体融合与传播国家重点实验室,中国传媒大学信息与通信工程学院,北京 100024)

摘要: 节目收视预测对提高用户体验起到越来越重要的作用,而针对现有收视预测往往仅考虑用户收视行为忽略了用户评论情感因素,以及要求数据量丰富、易受“奇异点”影响、存在过拟合、欠拟合、参数设置困难等问题。本文提出基于粒子群优化算法的混合核最小二乘支持向量机模型,综合考虑了用户收视行为、评论情感两类因素,并结合时间序列及最小二乘支持向量机模型在预测上的优势对节目收视进行预测。本文采用自适应迭代预测方式,以20天为滑动窗口步长,对用户收视序列进行拟合训练,验证了该模型在收视预测上的有效性及适用性。

关键词: 用户行为;评论情感;收视预测;混合核最小二乘支持向量机;粒子群优化算法

中图分类号:O422 文献标识码:A 文章编号:1673-4793(2022)01-0045-07

Rating prediction based on user viewing behavior and comment emotion analysis

FENG Xiaoli, WU Zhaoliang, YIN Fulian

(State Key Laboratory of Media Convergence and Communication School of Information and Communication Engineering, Communication University of China, Beijing 100024, China)

Abstract: Program rating prediction plays an increasingly important role in improving the user experience. However, existing rating prediction only considers user rating behavior and ignores the emotional factors from user comments. Moreover, it requires abundant data, is susceptible to the influence of “singularity”, and has problems such as over-fitting, under-fitting and difficult parameter setting. In this paper, a hybrid kernel least squares support vector machine model based on particle swarm optimization algorithm is proposed, which comprehensively considers the two factors of user viewing behavior and comment emotion, and combines the advantages of time series and least-squares support vector machine model in forecasting to predict program viewing. In this paper, the adaptive iterative forecasting method is adopted, and the 20-day sliding window step is used to fit the user viewing sequence, which verifies the effectiveness and applicability of the model in viewing forecast.

Key words: user behavior; comment emotion; rating prediction; hybrid kernel least squares support vector machine; particle swarm optimization

1 引言

收视率主要指研究一个地区在同一个时间段内某一频道或某一档节目的收视观众占总收视人数的比率^[1],可以反映节目的口碑与影响度,是评估节目价值的重要指标之一。近年来,由于社交网络的发展和兴盛,受众的情感偏好、网络舆论造势等成为干扰传统电视收视的关键因素^[2];同时,用户在线、及时发表的影评等对节目收视也会产生一定影响。这就使得如何把握电视节目数据、用户行为数据及评论文本数据的影响,对已上映的节目进行科学、有效的收视预测显得尤为重要。

近年来,对收视率的预测方法与技术日益成为相关电视机构、研究者关注的重点内容。收视率预测主要研究如何将影响收视的各种因素转换为一些相关指数,并以相对精确的数学形式表示,以消除主观判断的偏差。早期,有学者利用统计学的线性回归方法来评判收视规律,以刻画节目的收视走向;但这一方法不能深入地把握收视变换规律及建立有效的收视预测模型^[3]。在21世纪初期,基于数据挖掘技术的收视率预测方法开始被引入收视预测领域,为该领域的研究提供了新思路,其以收视率的影响因素为依据对收视率进行预测^[4]。张晶等^[5]将收视率的影响因素作为贝叶斯网络节点,依据先验知识选取分析变量,采集样本数据,通过贝叶斯网络的结构学习和参数学习,提出了新的收视率预测方法。梁招娣等^[5]提出了一种RBF(Radial Basis Function)神经网络预测方法,其通过对收视率数据进行非线性定阶,对模型进行重构,然后利用RBF神经网络寻找最优参数进行预测。汪洋等^[6]采用BP(Back propagation)神经网络进行了收视预测,这一方法更加注重数据之间的稳定规律,因此取得了相对较高的预测精度。程杨^[7]提出采用机器学习中的梯度下降算法对收视率数据进行分析预测,但因数据量的限制,预测误差较大。虽然上述模型在收视率预测上取得了良好效果,但这些方法往往存在易受个别“奇异点”数据影响、相关参数多、设定困难、数据量限制等问题。此外,基于时间序列的方法即依据序列过去变化的特征来预测未来的变化方式,也被用于收视预测。Zheng等^[8]采用基于时间序列的灰色系统模型(Grey Model(1,1), GM(1,1)),在数据少、信息少的情况下,对新一期电视节目的收视进行了预测研究,最终预测误差小于5%。姚芳^[9]等基于时间

序列构建了四种拟合模型,对全国30家电视台卫星频道的收视率进行了分析预测,得出基于时间序列的方法可利用较少的数据信息构建较好的拟合模型并达到较高精度的预测效果。车睿佳^[10]基于时间序列建立了二阶自回归模型,为降低异常值带来的影响,对模型进行了平滑滤波改进;其以节目热播期前29天的收视数据作为训练集,最后3天的数据作为测试集,对电视剧的收视率变化趋势进行预测,取得了不错的预测效果。可见,基于时间序列的方法可在一定程度上解决收视数据量少、信息少等问题,达到对节目未来的收视率有较佳的预测效果。近年来,也有学者对贝叶斯证据框架下的一类支持向量的估计方法进行了研究^[11],引入贝叶斯证据框架对其参数和样本隶属度取值进行优化,将收视率数据的前期认识加入到收视预测模型中,并针对收视率数据的特点给出其参数调整方法,为收视率分析预测提供了一种新思路及方法,且相较于传统的方法取得了更好的预测精度。支持向量机(Support Vector Machine, SVM)被广泛应用于小样本、非线性数据拟合上,较传统的神经网络算法要求数据量丰富、对数据量依赖性强等方面具有较好的优势^[12]。而由SVM扩展的最小二乘支持向量机(Least Squares Support Vector Machine, LSSVM)技术具有算法简单、易于实现、计算速度快等优点^[13],在物流需求量、网络流量、降雨量、就业率等各领域的预测中,都展现出了其的良好优势。

因本文对节目播放期间短期内的收视率做预测研究,基于时间序列即以每天的收视数据为实验样本时,从样本量上属于小样本;从样本值上,因收视用户、收视时长均处于不断变化中,即收视数据也满足小样本,故本文将支持向量机技术引入到收视预测领域以验证其适用性。而针对传统预测技术存在的问题,以及最小二乘支持向量机的参数确定、核函数选择等问题,本文提出基于混合核的PSO_LSSVM(Particle Swarm Optimization, 粒子群优化算法)模型,并综合考虑节目在播期间的用户收视行为和评论情感两大因素进行电视节目收视预测。本文第一节介绍本文研究的背景、意义及发展现状;第二节主要介绍本文构建的混合核PSO_LSSVM收视预测模型的研究框架及模型采用PSO算法^[14]进行参数优化的训练步骤;第三节主要介绍本文的实验与结果分析,并给出对比实验结果;第四节为结论部分并进行了展望。

2 混合核 PSO_LSSVM 收视预测模型

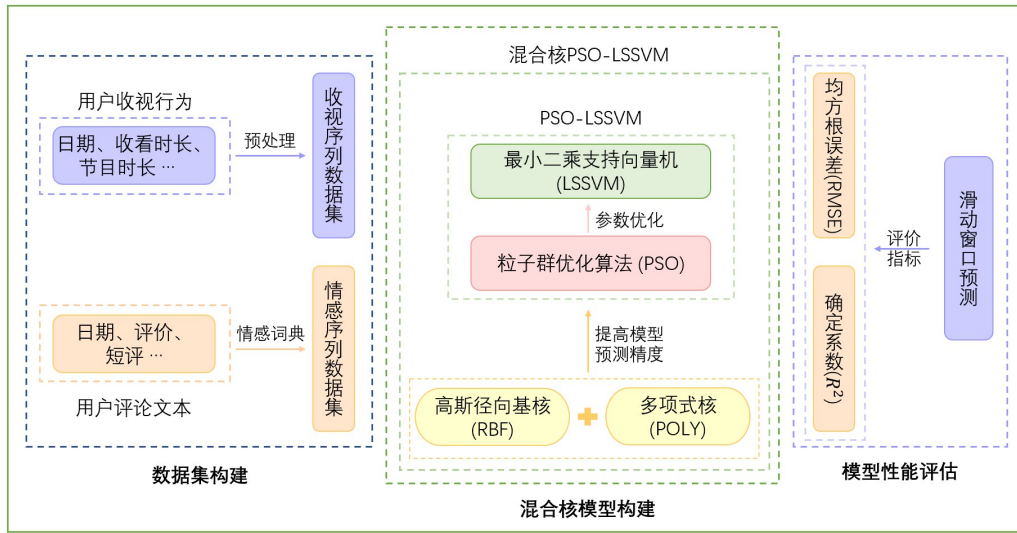


图1 研究架构图

本文基于用户收视行为与评论情感分析的收视预测展开研究,研究架构如图1所示,主要包括数据集构建、混合核模型构建以及模型性能评估三个部分。首先在数据集构建部分获得模型所需的实验数据集,然后在混合核模型构建部分介绍最小二乘支持向量机及粒子群优化算法的基本原理、混合核函数的构建及模型训练优化步骤,最后在模型性能评估部分给出本文采用的两个评价指标及其定义。

2.1 数据集构建

本文采用的用户收视行为数据是从某有线电视网络股份有限公司的机顶盒收视数据中整理获得的,针对其包含的自2015年5月1日至10月31日期间所有用户收视数据记录,并结合分析了2015年下半年国内电视剧热播排行榜,最终选取《琅琊榜》作为本文的收视预测研究对象。每一条用户收视记录数据中包含用户编号、日期、开始收看时间、结束收看时间、节目时长、节目名称、节目标签等要素,其中时长以秒为计算单位。因考虑到用户的收看时长在很大程度上可以反映出其对节目的喜爱程度^[15],故基于每位用户的收看时长、以及当天该节目的总播放时长,利用Python计算了模型需要的每一天的用户收视数据,即构建得到介于[0,1]之间的收视序列数据集。采用的计算方法如公式(1)所示:

$$B = \frac{\sum_{i=1}^n t_i}{n \times T} \quad (1)$$

其中, n 表示当天收看该节目的用户数, t_i 表示第 i 位用户的收看时长, T 表示节目在当天的总播放时长, B 表示用户收视行为数据。则此公式中的分子表示节目播放当天被所有用户收看的总时长,分母表示若每一位用户在节目播放当天完整收看该节目的总时长。

本文采用的用户评论文本数据是从国内最具影响力的电影社区,同时拥有着较大规模用户群体的豆瓣电影网站上基于Python爬虫技术获得的。其中每一条评论数据包含用户昵称、日期、评价、短评等要素,通过选用较为常用的、基础的、基于情感词典的文本情感分析方法经分词、去停用词等数据预处理以及构建情感词典、否定词词典、程度级别词典等操作后利用Python计算出每一条评论文本的情感值。情感值的正值、负值以及零值,分别对应了用户观看节目后所发表短评的积极、消极、中性三种情感态度。同时为避免值过大或过小而影响模型的拟合预测,故将情感值的上下限设置为[-10,10]。最后通过整理计算将每天所有用户的短评情感值均值作为当天的用户评论情感值,即构建得到情感序列数据集。

2.2 混合核模型构建

本文以标准最小二乘支持向量机模型为基础,因该模型的性能在很大程度上取决于模型参数的确定及其核函数的选择,但目前尚未确定统一的理论来指导如何获取有效的模型参数值及核函数。故本文首先针对模型参数值的确定问题,引入了PSO优化算法通过循环迭代的方式得到模型训练拟合时所需的最

优化参数值,即在LSSVM模型的基础上首先构建了基于高斯径向基(RBF)核函数的PSO_LSSVM模型。

而针对模型核函数的选择问题,因模型若只采用单一的RBF核函数拟合训练时,易出现过拟合现象,为保证模型对收视序列数据具有较高拟合效果的同时,也可具有较好的收视预测性能。本文基于具有高学习能力的局部RBF核函数以及具有高泛化能力的全局多项式(POLY)核函数,引入了混合权重系数,实现构建混合核函数 K ,其具体定义公式如下所示:

$$K = a \times K_{RBF} + (1 - a) \times K_{POLY} \quad (2)$$

$$K_{RBF} = \exp\left(-\|x - x_i\|^2 / \sigma^2\right) \quad (3)$$

$$K_{POLY} = (x \times x_i + t^2)^q \quad (4)$$

其中, a 为混合权重系数,表示混合核函数中两个单一RBF核、POLY核所占的比重,其值取值范围为 $[0,1]$ 。当 $a = 1$ 时, K 仅由RBF核函数构成;当 $a = 0$ 时, K 仅由POLY核函数构成。在公式(3)中, σ^2 表示RBF核函数宽度。在公式(4)中, t 表示偏置系数^[16],且 $t \geq 0$,在LSSVM工具箱中其值默认为1; q 表示POLY核函数阶数, $q \geq 1$ 且为整数。在构建得到混合核函数 K 后,将其作为PSO_LSSVM模型的核函数即替换掉原来的RBF核函数,以提高模型的收视预测性能,最终便构建得到基于混合核的PSO_LSSVM收视预测模型。

当构建了基于混合核的PSO_LSSVM收视预测模型后,需由PSO算法实现对两组模型参数组合值即 $[\gamma, \sigma^2], [p, a]$ 的确定,以得到不同训练数据下对应的最优化参数值,模型具体的训练步骤框图如图2所示,主要可分为初始化参数设置、第一次优化确定 $[\gamma, \sigma^2]$ 的最优值、第二次优化确定 $[p, a]$ 的最优值三个步骤。

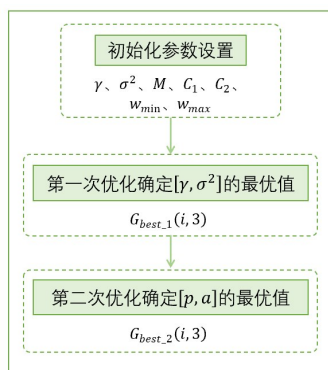


图2 模型训练步骤框图

a) 初始化参数设置

对LSSVM中的两个参数,设其取值范围分别为 $\gamma \in [0.01, 50], \sigma^2 \in [0.01, 20]$;对PSO算法中的几个超参数,设定粒子群个数 $M = 50$,初始化学习因子 $C_1 = C_2 = 1$,初始惯性权值 $w_{min} = 0.4$,终止惯性权值 $w_{max} = 0.95$;最大迭代次数 $K = 500$ 。

b) 第一次优化确定 $[\gamma, \sigma^2]$ 的最优值

令 γ 和 σ^2 的值分别指代粒子群中每个粒子的飞行速度及当前位置,并将模型训练输出的收视拟合值与真实值的均方根误差作为适应度函数计算得到的适应值。并将这三个值存放在3维局部向量 $P_{best_1}(M, 3)$ 中,其中 $P_{best_1}[i, 1]$ 代表第 i 个粒子的 γ 值, $P_{best_1}[i, 2]$ 代表该粒子的 σ^2 值, $P_{best_1}[i, 3]$ 则代表该粒子在当前两个属性取值下的最优适应值。通过循环迭代比较每个粒子找到的局部最优值,当迭代次数达到最大时,确定全局最优参数取值 $G_{best_1}(i, 3)$,即得到模型训练拟合需要 $[\gamma, \sigma^2]$ 的最优参数取值。

c) 第二次优化确定 $[p, a]$ 的最优值

在第二次PSO优化中,除了改变最大迭代次数为300(该值是经多次试验确定)外;其余参数的初始化设置均与第一次优化的相一致。同时设置混合权重系数 a 的最大值为1,最小值为0;多项式核阶数 p 的取值范围为 $[2, 8]$ 。令 p 和 a 的值分别指代粒子群中每个粒子的飞行速度及当前位置;将第一次优化得到的 $[\gamma, \sigma^2]$ 的全局最优化参数值代入到基于混合核模型构建的新适应度函数中,同样将模型经训练输出的拟合值与真实值的均方根误差作为适应值。并将 p, a 以及新的适应值存放在3维局部向量 $P_{best_2}(M, 3)$ 中,其中 $P_{best_2}[i, 1]$ 表示第 i 个粒子的 p 值, $P_{best_2}[i, 2]$ 表示该粒子的 a 值, $P_{best_2}[i, 3]$ 则表示该粒子在当前两个属性以及第一次优化得到的两个属性取值下的最优适应值。通过循环迭代的方式比较每个粒子找到的局部最优值,当达到最大迭代次数时,可确定第二次的全球最优参数取值 $G_{best_2}(i, 3)$,即确定了 $[p, a]$ 的取值。

最后经两次PSO优化确定了混合核模型拟合训练需要的全部最优化参数后,将两组参数组合值代入到混合核模型中,进行训练,得到训练模型以用于电视节目收视预测。

2.3 模型性能评估

本文采用自适应滑动窗口预测的方式,对标准LSSVM模型、基于单一RBF核的PSO_LSSVM模型、本文所构建的基于混合核的PSO_LSSVM模型以及

常用的数据拟合模型即BP神经网络模型分别做训练拟合及收视预测对比实验。同时为反映各模型的性能优劣,本文选用了在数据拟合预测领域中常用的均方根误差 $RMSE$ 及确定系数 R^2 两个评估指标对模型的收视拟合、预测效果进行客观评价。其中, $RMSE$ 的定义如公式(5)所示:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

R^2 的定义如公式(6)所示:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6)$$

其中, n 表示输入的训练样本个数, y_i 表示训练的真实输出收视样本值, \hat{y}_i 表示经训练模型得到的输出收视预测值, \bar{y} 表示真实输出收视样本值的均值。一般地, $RMSE$ 的值越接近于0,则表示模型的拟合、收视预测效果越好; R^2 的值越接近于1,模型整体的效果越好。

3 实验与结果分析

因用户行为由用户观看日期即时间要素来体现,故本文研究用户收视随时间及评论情感的变化规律构建训练模型进行节目收视预测。而因在预测某一天的收视时,是无法提前获知当天的用户评论文本数据的;故本文提出了先采用基于时间对用户评论情感序列进行二维模型拟合训练,然后基于得到的二维训练模型预测出之后几天的评论情感值,接着将评论情感预测值依次代入到基于时间及评论情感序列对用户收视序列进行三维拟合训练得到的模型中,以预测得到这几天相应的收视值。

此外,考虑到用户的收视习惯、以及节目自身内容对固定用户群体的吸引力,使得用户收视随时间及评论情感的变化也具有一定的规律性。故本文以20天为滑动窗口步长,进行模型自适应迭代预测实验,采用自适应的方式,即依据不同的输入数据得到对应的模型参数,可有效提高模型的拟合、预测性能。

3.1 混合核 PSO_LSSVM 收视预测模型实验

因支持向量机的原理较为复杂,实现起来较为困难,同时针对最小二乘支持向量机,美国的 Suykens 开发的基于 MATLAB 的最小二乘支持向量机工具箱 LSSVMlab 具有编程简单、运行效率高、可计算数据量

大、可扩展性强、算法功能丰富等各大优点^[17],是研究最小二乘支持向量机模型的有力工具。故本文基于 MATLAB2019a 软件安装并添加最小二乘支持向量机工具箱(LSSVMlabv1_8_R2009b_R2011a),通过有效调用其功能函数实现收视模型的训练、预测工作。

首先以第1-20天的评论情感序列经训练优化构建二维拟合模型,预测得到第21、22天的评论情感值为例。采用两次 PSO 优化算法得到混合核模型对应的两个参数组合的最优化取值为:第一次优化结果为 $\gamma = 8.2235, \sigma^2 = 0.3853$;第二次优化结果为 $p = 3, a = 0.7718$,代入模型对应输出的二维拟合曲线图如图3所示。

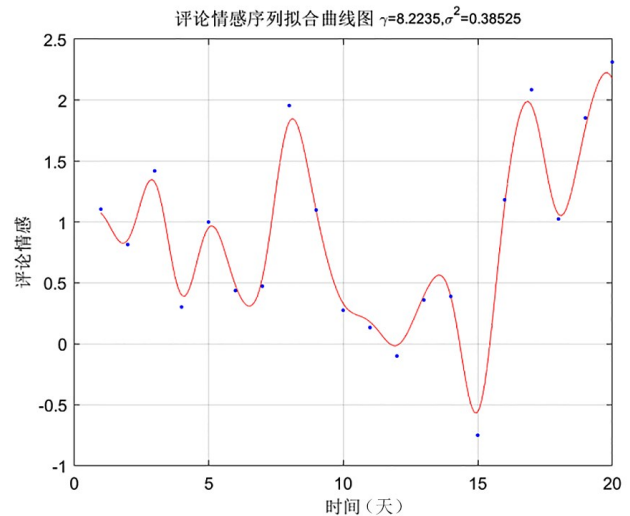


图3 第1-20天评论情感序列拟合曲线图

从图3可以看出,采用基于混合核的模型对用户评论情感序列的拟合与真实情感值的变化是较为接近的,拟合曲线也可以较好地反映出用户观看节目时的情感态度变化。同样,采用20天的滑动窗口步长,以第3-22天的评论情感序列构建二维拟合模型,预测得到第23、24天的评论情感值;以第5-24天的评论情感序列构建二维拟合模型,预测得到第25、26天的评论情感值。

在预测得到评论情感值后,以第1-20天的评论情感序列即用户收视序列经训练优化构建三维混合核拟合模型,来预测第21、22天的收视为例。采用两次 PSO 优化算法得到模型对应的最优化参数值:第一次优化结果为 $\gamma = 40.9683, \sigma^2 = 0.2905$;第二次优化结果为 $p = 4, a = 0.5944$,代入模型对应得到基于 RBF 核及基于 POLY 核的输出三维拟合曲面图分别如图4、图5所示。

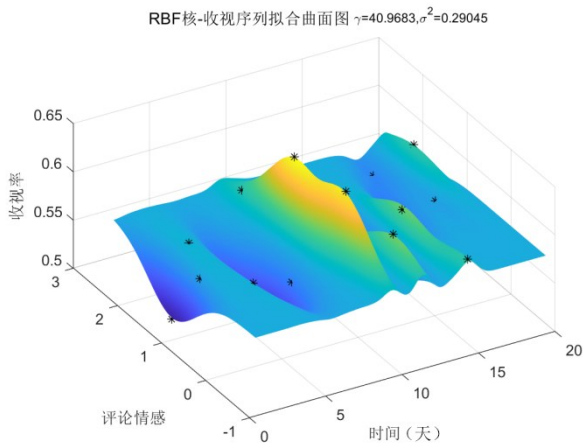


图4 RBF核第1-20天用户收视序列拟合曲面图

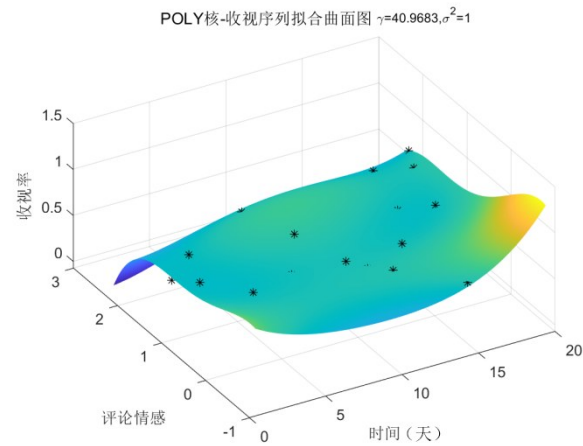


图5 POLY核第1-20天用户收视序列拟合曲面图

从图4、图5中可以看出,由基于局部的RBF核函数得到的拟合曲面图较基于全局的POLY核函数得到的拟合曲面图的趋势随收视的变化更为细腻,即其拟合的结果更接近收视率真实值,反映体现出RBF核函数的高学习能力,也可体现出POLY核函数的泛化性能较RBF核函数的更好。综合两者的优势,便可使得三维用户收视的混合核模型对收视值拟合的整体效果更佳。接着基于得到的三维用户收视混合核模型来预测第21、22天的收视时,代入在二维评论情感混合核模型中预测得到的这两天的评论情感值;同理,基于20天的滑动窗口步长,迭代预测得到第22、23天的收视以及第24、25天的收视值。

3.2 对比实验

当以20天为滑动窗口步长时进行模型收视预测时,得到在各模型下的收视预测结果如表1所示,从表1中可以看出基于混合核模型实验得到的预测值较其

它模型更接近于真实值,并绘制各模型预测效果对比图如图6所示。

表1 各模型收视预测结果

日期	模型	真实值	混合核 PSO_LSSVM	RBF核 PSO_LSSVM	标准 LSSVM	BP神经网络
第21天		0.6241	0.6140	0.5678	0.5614	0.5428
第22天		0.6494	0.6533	0.5704	0.5444	0.5782
第23天		0.6223	0.6069	0.6069	0.6355	0.8943
第24天		0.5710	0.5899	0.5899	0.5949	0.8945
第25天		0.6672	0.6192	0.5804	0.5629	0.6992
第26天		0.6422	0.6486	0.5835	0.5528	0.6899

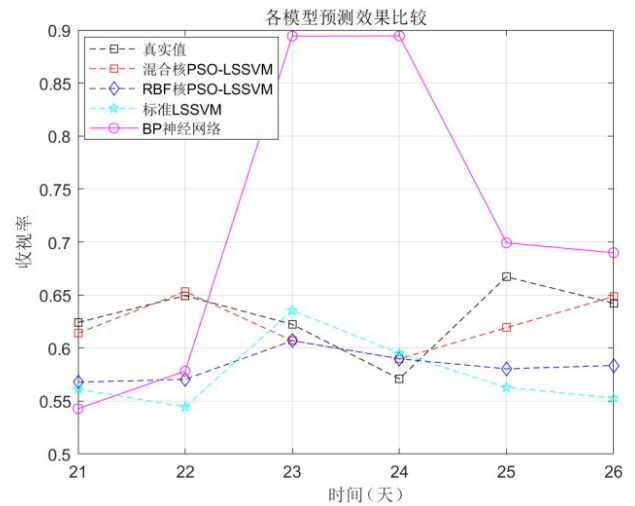


图6 各模型预测效果比较图

从图6中可以看出混合核模型的预测值与真实值的变化趋势在整体上更接近,即模型的泛化性能更好;也可以看出,当采用BP神经网络模型进行收视预测时,因其对样本量的依赖性较强,所以当样本量过少时,整体的预测效果均没有基于LSSVM模型的好。同时,在表2中给出了各模型收视预测性能评估指标值。

表2 各模型收视评估指标

模型	评估指标	均方根误差	确定系数
混合核 PSO_LSSVM		0.0226	0.4425
RBF核 PSO_LSSVM		0.0592	-2.8256
标准 LSSVM		0.0759	-5.3002
BP神经网络		0.1797	-34.2738

从表2中可以以更直观的方式显示出本文所构建基于混合核的 PSO_LSSVM 收视预测模型的预测性能明显优于其他模型。同时也体现出在标准 LSSVM

模型的基础上引入 PSO 优化算法的 RBF 核 PSO_LSSVM 模型,对标准模型做了有效的改进,并在综合了混合核函数后达到了更优的性能,使均方根误差值近似达到 0.02,更接近于 0,说明本文构建的混合核模型取得了较好的收视预测效果及较高的泛化性能。

4 结 论

本文基于用户收视行为与评论情感分析的收视预测展开研究,在传统的预测方法未考虑用户观影情感因素对收视的影响上做了进一步的提升。同时构建了具有高拟合、高泛化性能的混合核 PSO_LSSVM 模型,并采用自适应迭代预测的方式以 20 天为滑动窗口步长进行模型拟合训练及收视预测,并通过对比实验分析得到混合核模型下预测均方根误差指标达到最低 0.02 且确定系数在各模型下最高,即验证了本文构建的混合核模型在收视预测领域的有效性及适用性。接下来我们将尝试爬取更多的评论文本数据量并选取其他的滑动窗口步长做进一步的实验于改进工作,提高模型的收视预测泛化性能。

参考文献 (References) :

- [1] 肖启伟. 数字电视收视数据的统计与分析方法的研究 [D]. 福州大学, 2015.
- [2] 王锦慧, 郭美彤. 中国电视剧收视率预测模型研究—基于版权价值评估视角 [J]. 现代传播 (中国传媒大学学报), 2019(8):126-132.
- [3] 张晶, 白冰, 苏勇. 基于贝叶斯网络的电视节目收视率预测研究 [J]. 科学技术与工程, 2007, 7(19):4099-5102.
- [4] 黄玲莉, 刘小龙. 基于 ARIMA 与 BP 神经网络的收视率组合预测模型 [J]. 电视技术, 2015, 9(28):117-121.
- [5] 梁招娣, 刘小龙. 基于 RBF 神经网络的电视收视率预测模型 [J]. 河南科学, 2013(9): 1428-1431.
- [6] 汪洋, 田钢, 温淑鸿. 基于 BP 神经网络的电视节目收视率预测模型 [J]. 电视技术, 2014, 38(6): 94-96.
- [7] 程杨. 机器学习算法用于节目收视率预测之实验探索 [C]. 中国浙江绍兴: 第 18 届全国互联网与音视频广播发展研讨会暨第 27 届中国数字广播电视与网络发展年会论文集, 2019: 262-266.
- [8] Zheng Yilei. Audience rating prediction of new TV programs based on GM(1,1) envelopment model [C]. Nanjing, China: IEEE Press, 2009:388-391.
- [9] 姚芳, 李越, 肖春来. 基于时间序列模型的全国 30 家电台收视率分析 [J]. 数学的实践与认识, 2011, 41(13): 34-39.
- [10] 车睿佳. 基于时序预测和主成分分析的电视剧热度研究 [J]. 现代商贸工业, 2019, 10(29):59-61.
- [11] 陈青, 薛惠锋. 改进 P-SVM 收视率预测方法及其应用研究 [J]. 西安工业大学学报, 2011, 6(10):535-542.
- [12] N Niu, S Nan, J Zhang, Y Wang. LSSVM prediction model based on improved kernel parameter selection method [C]. Jinan, China: 11th International Conference on Prognostics and System Health Management (PHM-2020 Jinan), 2020: 391~394.
- [13] 杨凯帆, 黄婧, 谢泉, 王若昕, 沈丹青, 何丽娜, 陈汝科. 基于最小二乘支持向量机的短期负荷预测 [J]. 电工料, 2021, 172(1): 51-53. DOI: 10.16786/j.cnki.1671-8887.eem.2021.01.015.
- [14] 李爱国, 覃征, 鲍复民, 贺升平. 粒子群优化算法 [J]. 计算机工程与应用, 2002(21):1-3+17.
- [15] 顾军华, 李晓雪, 杨亮. 基于时段-时长耦合 LDA 的用户收视行为挖掘 [J]. 计算机应用与软件, 2020, 37(4):31-39.
- [16] 何岸青. 基于多项式拟合与支持向量机的股票关键拐点预测 [D]. 吉林大学, 2016.
- [17] 李方方, 赵英凯, 颜昕. 基于 Matlab 的最小二乘支持向量机的工具箱及其应用 [J]. 计算机应用, 2006(S2):358-360.

编辑:龙学锋,李树锋