

基于优化 LSTM 模型的股票预测

胡聿文

江西财经大学统计学院 南昌 330013

摘要 股票预测研究一直是困扰投资者的难题。以往,投资者采用传统分析方法如 K 线图、十字线等方法来预测股票走势,但随着科技的进步和经济市场的发展,以及经济政策的变动,股票的价格走势受到越来越多方面因素的干扰,仅靠传统的分析方法远远不能解析出股票价格波动中隐藏着的重要信息,因此预测精度大打折扣。为了提高股票价格的预测精度,提出一种基于 PCA 和 LASSO 的 LSTM 神经网络股票价格预测模型。采用 2015—2019 年平安银行(000001)五大类技术指标数据,通过 PCA 和 LASSO 方法对五大类技术分析指标进行降维筛选,再使用 LSTM 模型进行平安银行股票收盘价预测,对比前两种模型和单纯使用 LSTM 模型的预测效果稳定性及准确性。结果表明,相比于 LASSO-LSTM 模型和 LSTM 模型,PCA-LSTM 模型能够大幅削减数据冗余,并且获得了更优异的预测精度。

关键词: 平安银行;技术指标;LSTM;PCA;LASSO

中图分类号 TP183

Stock Forecast Based on Optimized LSTM Model

HU Yu-wen

College of Statistics, Jiangxi University of Finance and Economics, Nanchang 330013, China

Abstract Stock forecasting research has always been a problem that plagued investors. In the past, investors used traditional analysis methods such as K-line diagrams and Yin-Yang lines to predict stock trends. However, with the advancement of science and technology, the development of economic markets, and changes in economic policies, the price trend of a stock is disturbed by various factors. Traditional analysis methods are far from being able to analyze the information in volatility of a stock. So prediction accuracy is greatly reduced. In order to improve the prediction accuracy of stock prices, this paper proposes a stock price prediction model based on PCA, LASSO, and LSTM neural networks. Based on the data of the five major categories of technical indicators of Ping An Bank (000001) from 2015 to 2019, the five major categories of indicators are reduced and screened using the PCA and LASSO methods, and the LSTM model is used to predict the closing price of Ping An Bank's stock, compared with the stability and accuracy of the previous two methods and using LSTM alone. The experimental results show that the PCA-LSTM model significantly reduces data redundancy and obtains better prediction accuracy than the LASSO-LSTM model and LSTM model.

Keywords Ping An bank, Technical index, LSTM, PCA, LASSO

金融股票的预测一直是金融领域的研究热点,在方法上,大致可分为线性预测模型和非线性预测模型。其中,线性预测模型主要有移动平均自回归模型(ARIMA)、广义自回归条件异方差模型(GARCH)^[1]、指数广义自回归条件异方差(EGARCH)^[2]和整积的广义自回归条件异方差(IGARCH)^[3]等。上述线性预测模型作为早期的股票预测模型,对整个金融股票的预测发展起到了举足轻重的推动作用。但鉴于金融时间序列的高噪声和非线性特征,想要通过线性预测模型去精准预测金融股票价格依旧十分困难。随着计算机科技的高速发展以及深度学习研究的日益精进,机器学习领域的神经网络被日益广泛地应用于股票预测并且已经取得了比线性预测模型更加高效和准确的预测结果。Cao^[4]使用 BP 神经网络和灰色 GARCH-BP 模型所预测出的精度要明显优于 GARCH 模型。由于神经网络预测模型具有显著的非线性,我们把神经网络模型归为非线性预测模型。神经网络分为两大类。第一类人是工神经网络(ANN),如 MP 神经网络和

BP 神经网络。ANN 作为早期的神经网络模型,在股票预测方面亦做出了杰出贡献;Deng^[5]提出 DAE-BP 模型对股票先进行 DAE 降维,再使用 BP 神经网络进行股价预测,取得了不错的预测结果。可惜的是,ANN 模型结构过于单一,存在以下问题:1)过拟合,导致模型的泛化能力大大减弱;2)存在局部极值问题,即梯度下降过程中达到局部极小值点就停止工作,不能精准下降至全局极小值点,导致模型预测能力大大减弱;3)优化过程中容易因为神经元权重过多、过繁,导致梯度消失或者梯度爆炸问题,最终使神经网络模型预测失效。第二类神经网络则是相对于 ANN 来说,更深层次、更高效的深度神经网络模型(DNN),如卷积神经网络(CNN)、循环神经网络(RNN)和长短期记忆神经网络(LSTM)。这一类神经网络模型是当前研究金融预测领域最高效、前沿的预测模型,其具有多方面优势。1)对输入变量的形式没有限制,与预测问题可能相关的信息均可被作为模型输入,这一点极大满足了股票市场容易被各式各样的信息所干扰影响的特征。2)有

效拟合输入变量间的非线性复杂关系,提高样本拟合程度,同时通过神经元权重循环使用原理,大大减少了神经元权重的数量,有效防止过拟合现象。3)通过 DNN 中 tanh 激活函数,能够显著解决 ANN 中的梯度爆炸和梯度消失问题。影响股票价格波动的因素有公司财务面、基本面和相关政策以及投资者心理情绪。投资界有一句行话:关于股市的一切影响因素都可以在股价中体现,但是投资者心理情绪对股票在一段时间内的波动具有举足轻重的影响。Zheng 在各自研究结果中明确表示投资者的投资情绪与股票收益存在显著的关联性^[6-7]。因此,本文在股票价格影响因素选取中创新性地将消费者情绪和财务数据、基本面数据等多种数据融合,通过深度学习 LASSO 方法和 PCA 分析法对影响股票价格的多种因素进行降维筛选^[8-9],使各输入数据之间的相关性最大化,再分别导入目前最前沿的 LSTM 神经网络模型^[10-14]进行预测,并进行准确性和稳定性的对比,发现其中最高效的股票模型预测方法。

1 研究方法

1.1 LASSO

实证分析中,通常会设置尽可能多的自变量,选取自变量时容易出现偏主观意愿的疏漏,从而导致实证分析失真。而 LASSO 方法是一个能够客观筛选有效变量并且解决多重共线性等问题的估计方法。它是 1997 年由 Tibshirani 提出的一种压缩估计方法,通过构造一个惩罚函数,让回归系数的绝对值之和在小于一个常数的约束条件下,使得回归模型残差平方和最小,产生严格等于零的回归系数,从而有效解决回归模型中的多重共线性问题^[15]。

LASSO 方法是在普通线性回归模型中增加 L_1 惩罚项,普通线性模型的 LASSO 估计为:

$$\hat{\beta}_{Lasso} = \arg \min_{\beta \in R^d} \|Y - X\beta\|^2$$

$$\text{s. t. } \sum_{j=1}^d |\beta_j| \leq t, t \geq 0$$

等价于:

$$\hat{\beta}_{Lasso} = \arg \min_{\beta \in R^d} (\|Y - X\beta\|^2 + \lambda \sum_{j=1}^d |\beta_j|)$$

其中, t 与 λ 一一对应,为调节系数。

令 $t_0 = \sum_{j=1}^d \hat{\beta}_j(OLS)$, 当 $t < t_0$ 时,一部分系数就会被压缩至 0,从而降低 X 的维度,达到减小模型复杂度的目的。例如:如果 $t = t_0/2$,粗略地说,模型中非零系数个数就会由 d 大约减少至 $d/2$ 。

最后通过对调节系数 λ 的控制,可实现变量筛选。

1.2 PCA

主成分分析法是一种降维的统计方法,它借助于一个正交变换,将其分量相关的原随机向量转化成其分量不相关的新随机向量,这在代数上表现为将原随机向量的协方差阵变换成对角形阵,在几何上表现为将原坐标系变换成新的正交坐标系,使之指向样本点散布最开的 p 个正交方向;然后对多维变量系统进行降维处理,使之能以一个较高的精度转换成低维变量系统,再通过构造适当的价值函数,进一步把低维系统转化成一维系统。

(1)原始指标数据的标准化采集 p 维随机向量 $x = (x_1, x_2, x_3, \dots, x_p)^T$, n 个样品 $x_i = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{ip})^T, i = 1,$

$2, \dots, n (n > p)$, 构造样本阵,对样本阵元进行如下标准化变换: $Z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}, i = 1, 2, \dots, n; j = 1, 2, \dots, p$ 。其中, $\bar{x}_j =$

$$\frac{\sum_{i=1}^n x_{ij}}{n}, s_j^2 = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n-1}$$
。从而得标准化阵 Z 。

(2)对标准化阵 Z 求相关系数矩阵:

$$R = [r_{ij}]_{p \times p} = \frac{Z^T Z}{n-1}$$

$$\text{其中, } r_{ij} = \frac{\sum z_{kj} \cdot z_{ki}}{n-1}; i, j = 1, 2, \dots, p。$$

(3)解样本相关矩阵 R 的特征方程:

$$|R - \lambda I_p| = 0 \text{ 得 } p \text{ 个特征根, 确定主成分。按 } \sum_{j=1}^m \lambda_j \geq 0。$$

85 确定 m 值,使信息的利用率达 85% 以上。对每个 $\lambda_j, j = 1, 2, \dots, m$,解方程组 $Rb = \lambda_j b$ 得单位特征向量 b_j^0 。

(4)将标准化后的指标变数转换为主成分 $U_{ij} = z_{ij}^T b_j^0, j = 1, 2, \dots, m$ 。其中, U_1 称为第一主成分, U_2 称为第二主成分,以此类推, U_p 称为第 p 主成分。

(5)对 m 个主成分进行综合评价。

对 m 个主成分进行加权求和,即得最终评价值,权数为每个主成分的方差贡献率。

1.3 LSTM

长短时记忆神经网络(Long Short-term Memory Networks, LSTM)是一种特殊的 RNN 类型,可以学习长期依赖信息。RNN 神经网络模型一直被广泛用于语言识别和文本分类等多个研究领域^[16]。相比于人工神经网络模型(ANN)而言,RNN 神经网络模型可以循环利用神经元的权重参数,能够很好地将历史数据相关信息应用到预测中去。然而,RNN 神经网络模型的误差反向传播算法只是像 ANN 神经网络模型中一样简单,权重的重复利用能够带来好处,也会带来很大弊端,例如梯度爆炸和梯度消失问题,即对历史数据的长期依赖性问题无法有效解决。为解决这两大难题,机器学习科研工作者们研究出长短时记忆神经网络模型(LSTM)^[17],如图 1 所示。

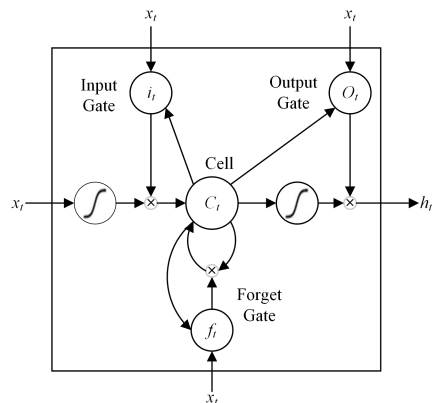


图 1 LSTM 模型神经元结构示意图

Fig. 1 Neuron structure of LSTM model

LSTM 模型相较于 RNN 模型最明显的改进是增加了 1 个细胞状态 C 和 3 个阀门,3 个阀门分别是遗忘门 f 、输出门 o 和输入门 i 。在 LSTM 模型误差反向传播校正权重时,有些误差可以直接通过输入门传递给下一层神经元,有些误差则

可以通过遗忘门去进行数据遗忘,这样就解决了梯度爆炸与消失的难题,即有效地处理历史数据中相关信息的冗余等问题^[18-19]。本文研究的股票价格预测是典型的时序问题,且某一个时刻的价格受前一时刻和历史多时刻价格影响,所以选择 LSTM 模型进行股票价格预测。

2 股票价格预测的实证分析

2.1 数据来源及指标选取

文章所采用的数据为 2015 年 1 月 5 日至 2020 年 2 月 7 日的平安银行(0000001)股票数据(数据来源于通达信金融终端),共 1240 条数据。其中 80% 作为训练集用于训练模型,其余 20% 作为测试集中来验证模型的泛化能力。

在指标选取的过程中,应尽可能全方面地考虑影响因素,全方位地对问题进行分析,尤其对股票价格波动这种影响因素较多且各因素之间并不呈明显线性关系的难题,在指标选择过程中更应该精准筛选。相较于其他研究,实验选用股价的开盘价、最高最低价、成交量以及一般技术指标 OBV、KDJ、BIAS 等常见的技术指标,本文创新性地添加了最前沿的 CCI, MFI, MTM 等若干股价判断技术指标以及准确反映投资者心理情绪的 PSY 指标。这些技术指标能够多方位地涵盖股价波动的潜在信息,具有很强的股价解释性。为了更清楚地对这 57 个技术指标进行理解,表 1 进行了详细说明。

表 1 技术指标变量解释

Table 1 Explanation of technical index variables

一级指标	二级指标	三级指标	变量	解释
能量型	BRAR	AR	Z1	判断行情热点
		BR	Z2	
		CR, CR	Z3	
	CR	CR, MA1	Z4	买入卖出信号
		CR, MA2	Z5	
		CR, MA3	Z6	
		CR, MA4	Z7	
	VR	VR, VR	Z8	买入卖出信号
		VR, MAVR	Z9	
	MASS	MASS, MASS	Z10	买入卖出信号
		MASS, MAMASS	Z11	
PSY	PSY, PSY	Z12	研究投资者对股市涨跌产生心理波动	
	PSY, PSYMA	Z13		
NVI	NVI, NVI	Z14	追踪大户资金流向	
	NVI, MA	Z15		
PVI	PVI, PVI	Z16	追踪散户资金流向	
	PVI, MA	Z17		
超买超卖型	CCI	CCI, CCI	Z18	多/空头市场判断信号
		KDJ, K	Z19	
	KDJ	KDJ, D	Z20	买进卖出信号
		KDJ, J	Z21	
	MFI	MFI, MFI	Z22	短线买进卖出信号
		MTM, MTM	Z23	
	MTM	MTM, MTMMA	Z24	买入卖出信号
		OSC, OSC	Z25	
	OSC	OSC, MAOSC	Z26	多/空头市场判断信号
		ROC, ROC	Z27	
	SKDJ	ROC, MAROC	Z28	超买超卖界定信号
SKDJ, K		Z29		
BIAS	SKDJ, D	Z30	买入卖出信号	
	BIAS, BIAS1	Z31		
	BIAS, BIAS2	Z32		
WR	BIAS, BIAS3	Z33	判断股价拉升拉回信号	
	WR, WR1	Z34		
	WR, WR2	Z35		

(续表)

一级指标	二级指标	三级指标	变量	解释
趋势型	DMI	DMI, PDI	Z36	多/空头市场判断信号
		DMI, MDI	Z37	
		DMI, ADX	Z38	
		DMI, ADXR	Z39	
	DPO	DPO, DPO	Z40	多/空头市场判断信号
		DPO, MADPO	Z41	
	MACD	MACD, DIF	Z42	买入卖出信号
		MACD, DEA	Z43	
		MACD, MACD	Z44	
	AMO	AMO, AMOW	Z45	判断热门/非热门股票
		AMO, AMO1	Z46	
AMO, AMO2		Z47		
OBV	OBV, OBV	Z48	短线买入卖出信号	
	OBV, MAOBV	Z49		
	VOL	VOL, MAVOL1		Z50
VOL	VOL, MAVOL2	Z51		
股价基本型	开盘价	开盘价	Z52	开盘时股票价格
	最高价	最高价	Z53	股票当日最高价格
	最低价	最低价	Z54	股票当日最低价格
	成交量	成交量	Z55	股票当日成交量
	MA	MA, MA1	Z56	股价 10 日均线
		MA, MA5	Z57	股价 120 日均线

2.2 数据的筛选

表 1 所列 5 个一级指标、28 个二级指标和 57 个三级指标描述了股价的波动影响因素。由于指标个数较多,为了不给网络运行带来负担,提高 LSTM 神经网络的预测能力,本文分别采用主成分分析法和 LASSO 回归法对 57 个指标进行筛选。

(1) LASSO 回归法

对数据进行标准化处理,接下来利用 LASSO 模型对 57 个指标数据进行分析,以进行变量筛选。运行结果显示,模型 91 步可以得到 LASSO 的全部解。根据最小角回归原理,选择 C_p 统计量最小值对应的拟合方程,第 80 步的 C_p 值最小为 45.04705,模型最优。根据 LASSO 结果,剔除变量为 0 的 MACD, DEA, OBV, OBV, BRAR, BR, BRAR, AR, NVI, MA, PVI, PVI, KDJ, K, ROC, ROC 和 SKDJ, K 9 个变量,选取 50 个变量进行后续的神经网络模型预测。

图 2 为 LASSO 模型筛选结果。图中右侧数据分别代表相应自变量。

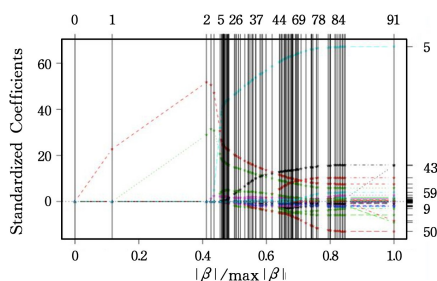


图 2 LASSO 模型筛选结果

Fig. 2 Screening results of lasso model

(2) 主成分分析法

通过 RStudio 软件对平安银行 57 个指标数据进行标准化处理,然后画出碎石图。由图 3 可知,特征值大于 1 的主成分个数有 9 个。

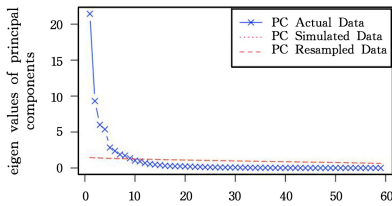


图3 碎石图

Fig. 3 Scree plot

对数据进行提取主成分,并对提取出的主成分结果取85%贡献率,我们得到8个主成分。为了使主成分分析法解释更高效并且最小化LSTM神经网络模型的数据冗余,我们保留数值大于0.5的参数数据。处理后的8个主成分如下:

$$\alpha_1 = 0.777 Z_7 + 0.590 Z_8 + 0.691 Z_9 + \dots + 0.826 Z_{59}$$

$$\alpha_2 = 0.544 Z_1 + 0.565 Z_2 + 0.523 Z_3 + \dots + 0.923 Z_{22}$$

$$\alpha_3 = 0.635 Z_{32} + 0.572 Z_{33} + 0.577 Z_{42} + \dots + 0.632 Z_{55}$$

$$\alpha_4 = 0.800 Z_{36} + 0.917 Z_{37} + 0.584 Z_{42} + \dots + 0.691 Z_{58}$$

...

$$\alpha_8 = 0.751 Z_1 + 0.744 Z_2 + 0.761 Z_3 + \dots + 0.803 Z_6$$

2.3 预测方法及思路

为客观比较LASSO-LSTM和PCA-LSTM之间的预测效率,我们加入单纯LSTM模型的预测结果,进行三者对比。

(1)LSTM神经网络方法及预测思路

LSTM神经网络用历史1240个交易日的数据信息对股票价格走势进行预测。输入数据为未做数据筛选的所有57个参数变量,输出数据为历史股价下一日的收盘价预测值。

(2)LASSO-LSTM模型方法及预测思路

通过LASSO回归法构造惩罚函数,将历史1240个交易日内的57个参数变量进行去共线性筛选,留下的50个参数变量具有低共线性、高相关度等特征,再将其当作输入变量输入LSTM神经网络模型中,输出变量是当日历史数据的下一日收盘价预测值。

(3)PCA-LSTM模型方法及预测思路

通过PCA分析法,从57个原始数据提取出8个主成分用于LSTM模型输入,这8个主成分因子分别以不同的参数系数囊括了57个历史数据的信息,显著地精简了神经网络模型输入端,同时又不丢失重要数据信息。输出变量是当日历史数据的下一日收盘价预测值。

2.4 实证结果对比

现有运用神经网络进行股票预测的研究实验较多是将历史数据全部放入训练集中进行模型训练,将最新一天的股价作为预测数据^[20-21]。以该方式预测可以得出非常拟合的预测结果图,是因为神经网络模型独特的多次校正权重原理能够让神经网络模型最大程度地拟合出与历史股价重合的预测股价,但这在实际炒股应用中的失真率非常大,毫无实际应用价值。因此,本文采用历史股价80%的数据即前960日的历史股价作为训练集,剩余20%的数据即第960日至1240日的历史股价作为预测集。不同的LSTM模型超参数,对LSTM模型的预测能力有着显著的影响,例如神经元层数及每层神经元中的神经元个数的改变会使模型运算繁琐度指数增长且影响最后的预测精度,学习率的千分位改变会显著影响模型在梯度下降时的效率与准确率^[22]。

因此,在本次预测实验中设定了不同的LSTM模型超参数进行预测结果比较。在目前的LSTM研究文章中,LSTM模型神经元层数设为2层和3层时效率最高,神经元个数设置为8不变,学习率通常设置为0.001,这里另设对比学习率0.005,迭代次数设置为500和1000次,激活函数设置为LSTM模型常用的tanh函数不变。通过判断实验结果的预测精度以及预测股价和历史股价间走势的拟合程度,来确定何种方式的模型预测结果最有成效。预测精度评估分别采用均方误差函数(MSE)、均方根误差(RMSE)和平均绝对误差(MAE),三者的值越小表明预测结果越精确。表2—表4是3种预测模型的实验结果。

表2 LSTM在LSTM模型参数下的预测结果

Table 2 Prediction results of LSTM under LSTM model parameters

迭代次数	激活函数	学习率	神经元层数	神经元个数	MSE	RMSE	MAE
500	tanh	0.001	3	8	0.0078	0.1344	0.2280
1000	tanh	0.001	3	8	0.0014	0.0313	0.1754
500	tanh	0.005	3	8	0.0036	0.0573	0.1884
1000	tanh	0.005	3	8	0.0151	0.0343	0.1677
500	tanh	0.001	2	8	0.0092	0.0303	0.1822
1000	tanh	0.001	2	8	0.0017	0.0382	0.1467
500	tanh	0.005	2	8	0.0012	0.0351	0.2516
1000	tanh	0.005	2	8	0.0042	0.0490	0.2382

从表2单纯使用LSTM模型对股票的预测结果来看,在LSTM模型设置迭代次数为500次、激活函数为tanh、学习率为0.005、神经层为3层、神经元为8个时,MSE结果最小,为0.0012;在设置迭代次数500次、激活函数tanh、学习率0.001、神经层2层、神经元8个时,RMSE结果最小,为0.0303;在LSTM模型设置迭代次数1000次、激活函数tanh、学习率0.001、神经层3层、神经元8个时,MAE结果最小,为0.1467。图4—图6分别为单纯LSTM模型下,3个误差值最小条件时的预测图。

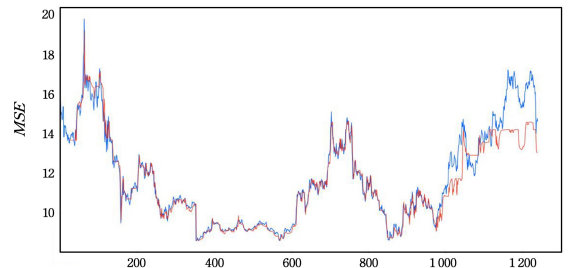


图4 迭代500次、学习率0.005、3层神经元时的MSE
Fig. 4 MSE when iteration=500, learning rate = 0.005 and neurons=3

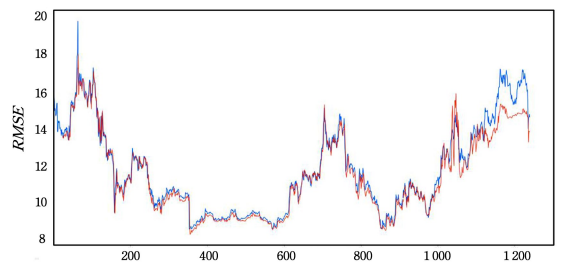


图5 迭代500次、学习率0.001、2层神经元时的RMSE
Fig. 5 RMSE when iteration=500, learning rate=0.001 and neurons=2

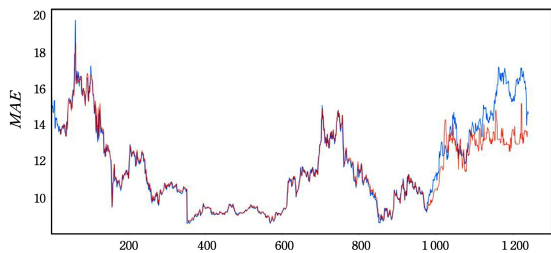


图6 迭代1000次、学习率0.001、3层神经元时的MAE

Fig. 6 MAE when iteration=1000, learning rate=0.001 and neurons=3

从图4—图6可以看出,单纯使用LSTM模型进行股价预测,以MSE和MAE为误差函数做预测时,在预测集一端的误差非常大,预测股价呈现乱序、波动大等负面情况,无法拟合预测集的真实股价的走势和真实股价。相比前两者的预测结果,以RMSE为误差函数做预测时,预测股价则能够相对较好地贴合预测集中真实股价的走势,但是预测结果与预

测集中第960日至第1240日的真实历史股价间的重合率过低,无法较好地拟合。

从表3 LASSO-LSTM模型对股票预测结果来看,在LSTM模型设置迭代次数1000次、激活函数tanh、学习率0.001、神经层2层、神经元8个时,MSE结果最小,为0.0009;在设置参数为迭代次数1000次、激活函数tanh、学习率0.001、神经层3层、神经元8个时,RMSE结果最小,为0.0272;在LSTM模型设置迭代次数500次、激活函数tanh、学习率0.001、神经层3层、神经元8个时,MAE结果最小,为0.1398。图7—图9分别为单纯LSTM模型下,3个误差值最小条件时的预测图。从图7—图9可以看出,尽管LASSO-LSTM模型在以MSE, RMSE, MAE三者为误差函数时取得的预测精度有较为可观的结果,但是从预测图的走势来看,LSTM模型所得的预测值呈现杂乱性、无序性并且与历史真实股价严重偏离的现象,因此,LASSO-LSTM模型在股价预测应用上是远远不能付诸于实际中的。

表3 LASSO-LSTM在LSTM模型参数下的预测结果

Table 3 Prediction results of lasso LSTM under LSTM model parameters

迭代次数	激活函数	学习率	神经层数	神经元个数	MSE	RMSE	MAE
500	tanh	0.001	3	8	0.0039	0.0419	0.1398
1000	tanh	0.001	3	8	0.0050	0.0272	0.2209
500	tanh	0.005	3	8	0.0098	0.0535	0.1793
1000	tanh	0.005	3	8	0.0025	0.0423	0.1492
500	tanh	0.001	2	8	0.0013	0.0605	0.1965
1000	tanh	0.001	2	8	0.0009	0.0327	0.2346
500	tanh	0.005	2	8	0.0062	0.0616	0.2441
1000	tanh	0.005	2	8	0.0039	0.0607	0.1856

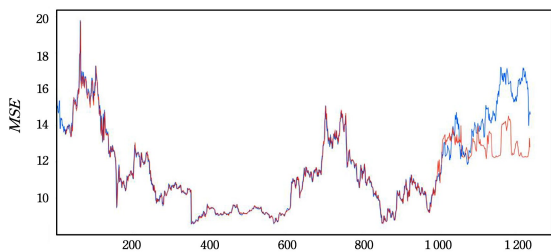


图7 迭代1000次、学习率0.001、2层神经元时的MSE

Fig. 7 MSE when iteration=1000, learning rate=0.001 and neurons=2

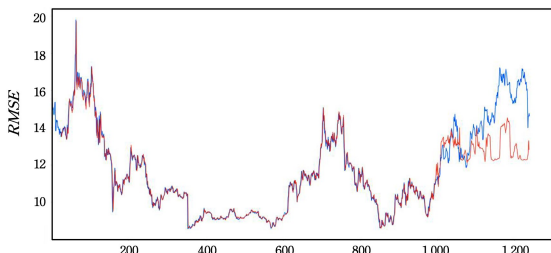


图8 迭代1000次、学习率0.001、3层神经元时的RMSE

Fig. 8 RMSE when iteration=1000, learning rate=0.001 and neurons=3

从表4 PCA-LSTM模型对股票预测结果来看,在LSTM模型设置迭代次数1000次、激活函数tanh、学习率0.001、神经层2层、神经元8个时,MSE结果最小,为0.0005;在LSTM模型设置迭代次数500次、激活函数tanh、学习率0.001、神经层3层、神经元8个时,RMSE结果达到最小值

0.0283;同时,在此条件下,MAE也为最小值0.2101。值得一提的是,MSE的值也很可观,仅比最小0.0005略微高出0.0002。图10—图12分别为PCA-LSTM模型下,3个误差值最小值条件时的预测图。

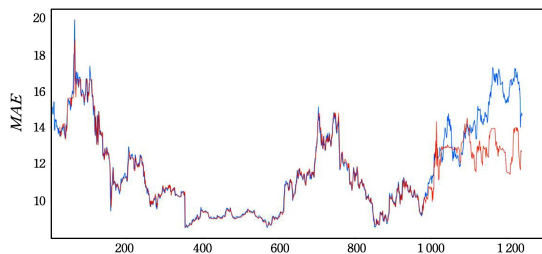


图9 迭代500次、学习率0.001、3层神经元的时MAE

Fig. 9 MAE when iteration=500, learning rate=0.001 and neurons=3

表4 PCA-LSTM在LSTM模型参数下的预测结果

Table 4 Prediction results of PCA-LSTM under LSTM model parameters

迭代次数	激活函数	学习率	神经层数	神经元个数	MSE	RMSE	MAE
500	tanh	0.001	3	8	0.0007	0.0283	0.2101
1000	tanh	0.001	3	8	0.0012	0.0629	0.2221
500	tanh	0.005	3	8	0.0071	0.0501	0.3365
1000	tanh	0.005	3	8	0.0019	0.0423	0.2308
500	tanh	0.001	2	8	0.0017	0.0513	0.2532
1000	tanh	0.001	2	8	0.0005	0.0629	0.2160
500	tanh	0.005	2	8	0.0060	0.0584	0.2355
1000	tanh	0.005	2	8	0.0015	0.0314	0.2814

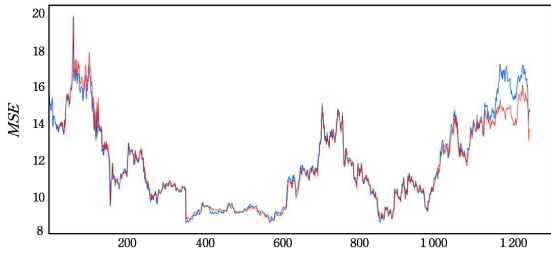


图 10 迭代 1000 次、学习率 0.001、2 层神经元时的 MSE

Fig. 10 MSE when iteration=1000, learning rate=0.001 and neurons=2

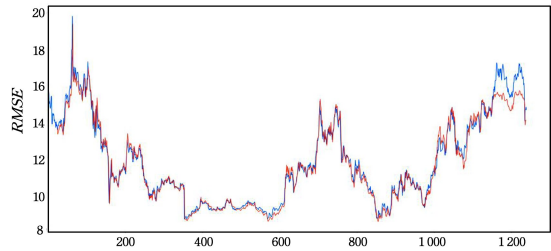


图 11 迭代 500 次、学习率 0.001、3 层神经元时的 RMSE

Fig. 11 RMSE when iteration=500, learning rate=0.001 and neurons=3

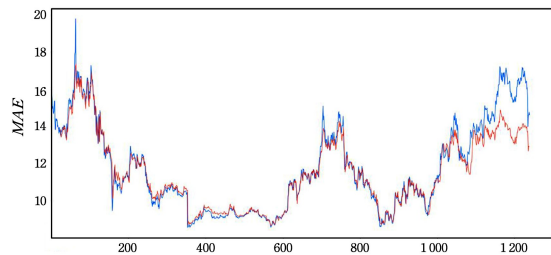


图 12 迭代 500 次、学习率 0.001、3 层神经元时的 MAE

Fig. 12 MAE when iteration=500, learning rate=0.001 and neurons=3

从图 10—图 12 来看, PCA-LSTM 模型无论是 MSE 还是 RMSE 作为误差函数的预测精度都很卓效, 在实际预测中, 预测模型所得的预测股价与第 960 日至第 1160 日之间的真实股价近乎重合, 且与第 1160 日至第 1240 日之间的历史股价走势高度拟合。而 PCA-LSTM 模型在以 MAE 为误差函数的预测实验中, 也有较好的拟合结果。

结束语 本文使用 LSTM 模型、LASSO-LSTM 模型和 PCA-LSTM 模型对上证股市的平安银行股票进行股价预测, 通过对比 3 种预测模型, 发现 PCA-LSTM 模型对平安银行的股票价格预测有很强的泛化能力, 其预测效果也是三者中最为精确高效的。目前市面上较多股市投资团队仅将股票的相关影响因素单纯应用 LSTM 模型进行预测, 影响因素数据繁杂冗余, 导致模型预测精度大打折扣。而 PCA-LSTM 模型在数据预测前将股票的相关影响因素数据进行主成分分析, 对数据进行因子旋转提炼, 提取出以最少主成分数量代表最多影响因素数量的主成分作为 LSTM 神经网络模型的输入变量, 能够让预测模型更加高效, 从而具有较好的预测效果。

参考文献

[1] YANG Q, CAO X B. Analysis and prediction of stock price

based on arma-garch model [J]. Practice and Understanding of Mathematics, 2016, 46(6): 80-86.

- [2] HUANG L J, JIN T X. Stability of Islamic stock market based on egarch-m model [J]. Gansu Academic Journal of theory, 2019 (6): 107-115.
- [3] FANG J. Empirical research on VaR Measurement of China's stock market: semi parametric method based on igarch [J]. Financial Theory and Teaching, 2018(3): 15-18.
- [4] CAO X, SUN H B. Stock price prediction based on Grey GARCH model and BP neural network [J]. Software, 2017, 38(11): 126-131.
- [5] DENG J K, WAN L, HUANG N N. Research on Stock Forecasting Based on dae-bp neural network [J]. Computer Engineering and Application, 2019, 55(3): 126-132.
- [6] ZHENG G J. Stock time series analysis and prediction based on Internet investor sentiment [D]. Hangzhou: Zhejiang University of technology, 2019.
- [7] LIANG X Z. Research on the relationship between investor sentiment and stock returns [D]. Beijing: Beijing Jiaotong University, 2017.
- [8] WANG J Z. Application of improved adaptive lasso method in stock market [J]. Mathematical Statistics and Management, 2019, 38(4): 750-760.
- [9] YU H H, CHEN R D, ZHANG G P. A SVM Stock Selection Model within PCA [J]. Procedia Computer Science, 2014, 31.
- [10] REN T Y. An Empirical Study of Stock Return and Investor Sentiment Based on Text Mining and LSTM [C] // Proceedings of the 2019 4th International Conference on Social Sciences and Economic Development (ICSSSED 2019). 2019.
- [11] JIA M Z, HUANG J, PANG L H, et al. Analysis and Research on Stock Price of LSTM and Bidirectional LSTM Neural Network [C] // Proceedings of the 3rd International Conference on Computer Engineering, Information Science & Application Technology (ICCIA 2019). 2019.
- [12] YANG Q, WANG C W. Global stock index prediction based on deep learning LSTM neural network [J]. Statistical Research, 2019, 36(3): 65-77.
- [13] PENG Y, LIU Y H, ZHANG R F. Modeling and analysis of stock price forecasting based on LSTM [J]. Computer Engineering and Application, 2019, 55(11): 209-212.
- [14] CHEN J, LIU D X, WU D S. Research on stock index prediction method based on feature selection and LSTM model [J]. Computer Engineering and Application, 2019, 55(6): 108-112.
- [15] TIBSHIRANI R. Regression shrinkage and selection via the lasso: A retrospective [J]. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2011, 73(3): 273-282.
- [16] HAN H, LIU G L, SUN T Y, et al. Text sentiment analysis based on multi attention level neural network [J]. Computer Engineering and Application.
- [17] PEI D W, ZHU M. Stock price prediction based on multi factor and multi variable long term and short term memory network [J]. Computer System Applications, 2019, 28(8): 30-38.
- [18] ZENG A, NIE W J. Stock recommendation system based on deep

- bidirectional LSTM [J]. *Computer Science*, 2019, 46(10): 84-89.
- [19] REN J, WANG J H, WANG C M, et al. Stock forecasting system based on elstm-l model [J]. *Statistics and Decision*, 2019, 35(21): 160-164.
- [20] FENG Y X, LI Y M. Research on prediction model of CSI 300 index based on LSTM neural network [J]. *Practice and Understanding of Mathematics*, 2019, 49(7): 308-315.
- [21] LI S S. Securities selection based on long term memory neural network [D]. Zhengzhou: Zhengzhou University, 2019.
- [22] XU T T. Research on stock price rise and fall prediction based on LSTM neural network model [D]. Shanghai: Shanghai Normal University, 2019.



HU Yu-wen, born in 1996, postgraduate. His main research interests include mathematical statistics, financial statistics and deep learning.

(上接第 136 页)

- [8] WANG H, CHEN X, TIAN S Z, et al. SAR image recognition based on few-shot learning [J]. *Computer Science*, 2020, 47(5): 124-128.
- [9] YU Y, WANG B, ZHANG L. Hebbian-based neural networks for bottom-up visual attention and its applications to ship detection in SAR images [J]. *Neurocomputing*, 2011, 74(11): 2008-2017.
- [10] CHEN LF, LIU Y Z, ZHANG P, et al. Road extraction algorithm of multi-feature migh-resolution SAR image based on Multi-Path RefineNet [J]. *Computer Science*, 2020, 47(3): 156-161.
- [11] ZHANG Z F, ZHAO Z, WEI J J, et al. A method for ship detection in SAR imagery based on improved human visual attention system [J]. *Science of Surveying and Mapping*, 2017, 42(4): 108-112.
- [12] ZHANG M C, WU X Q, WANG P W. An adaptive approach for target detection in SAR images [J]. *Computer Engineering and Applications*, 2006(10): 200-203.
- [13] ZHAO Q H, WANG X, LI Y, et al. Ship detection optimization method in SAR imagery based on multi-feature weighting [J]. *Journal on Communication*, 2020, 41(3): 91-101.
- [14] DU L, WANG Z C, WANG Y, et al. Survey of research progress on target detection and discrimination of single-channel SAR images for complex scenes [J]. *Journal of Radars*, 2020, 9(1): 34-54.
- [15] CHEN S, WANG H, XU F, et al. Target classification using the deep convolutional networks for SAR images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(8): 4806-4817.
- [16] BENTES C, VELOTTO D, TINGS B. Ship classification in TerraSAR-X images with convolutional neural networks [J]. *IEEE Journal of Oceanic Engineering*, 2017, PP(99): 1-9.
- [17] NIU L. Sidelobe removal method for ship target based on SAR images [J]. *Radar Science and Technology*, 2018, 16(2): 197-200.
- [18] GAO D, HAN S, VASCONCELOS N. Discriminant saliency, the detection of suspicious coincidences, and applications to visual recognition [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2009, 31(6): 989-1005.
- [19] WANG X H. Moment technique and its applications in image processing and recognition [D]. Xi'an: Northwestern Polytechnical University, 2002.
- [20] XIONG W, XU Y L, CUI Y Q, et al. Geometric feature extraction of ship in high-resolution synthetic aperture Radar images [J]. *Acta Photonica Sinica*, 2018, 47(1): 55-64.
- [21] ZHEN Y, LIU W, CHEN J H, et al. Geometric structure feature extraction of ship target in high-resolution SAR image [J]. *Journal of Signal Processing*, 2016, 32(4): 424-429.
- [22] PAPATSOUMA I, FARMAKIS N. Approximating Symmetric Distributions via Sampling and Coefficient of Variation [J]. *Communications in Statistics*, 2020, 49(1): 61-77.
- [23] AI J, YANG X, SONG J, et al. An adaptively truncated clutter-statistics-based two-parameter CFAR detector in SAR imagery [J]. *IEEE Journal of Oceanic Engineering*, 2017, PP(99): 1-13.
- [24] XIONG W, XU Y L, YAO L B, et al. A new ship target detection algorithm based on SVM in high resolution SAR images [J]. *Remote Sensing Technology and Application*, 2018, 33(1): 119-127.
- [25] ZHAO Z, JI K, XING X, et al. Adaptive CFAR detection of ship targets in high resolution SAR imagery [J]. *Proceedings of Spie the International Society for Optical Engineering*, 2013, 8917: 89170L-89170L-8.



YAN Jun, born in 1962, Ph.D. His main research interests include remote sensing technology and application and so on.



LU Lin-lin, born in 1984, Ph.D, associate professor. Her main research interests include urban remote sensing, remote sensing information extraction and classification.