

基于生成式对抗网络的股价预测研究

许飞飞,胡月,汪召兵

(浙江科技学院 理学院,杭州 310023)

摘要: 为了降低股票市场中噪声信息和投资者情绪对股票价格的影响,以便给投资者带来较高的投资回报并降低交易风险,特提出一种基于金融双向编码器表征和瓦瑟斯坦距离的生成式对抗网络(financial bidirectional encoder representation from transformers and Wasserstein generative adversarial networks,FWGAN)股价预测模型。本模型首先采集东方财富网股评数据,并利用自然语言处理预训练模型将股评数据量化为情绪值,然后将情绪值连同历史股票交易数据、技术指标数据输入由长短期记忆网络(long-short-term memory,LSTM)为生成器和卷积神经网络(convolutional neural network,CNN)为判别器组成的FWGAN模型中进行训练。对比LSTM模型、门控神经网络(gated recurrent units,GRU)模型和生成式对抗网络(generative adversarial networks,GAN)模型对山西汾酒股价的预测性能,结果表明,FWGAN模型的均方根误差为2.572,达到最低,预测效果最好。试验结果验证了本模型对股票时间序列预测的有效性和优越性,可以为投资者进行股价预测提供参考。

关键词: 股价预测;股民情绪;生成式对抗网络股价预测模型;时间序列

中图分类号: F830.91 **文献标志码:** A **文章编号:** 1671-8798(2022)03-0207-09

Research on stock price forecasting based on FWGAN model

XU Feifei, HU Yue, WANG Zhaobing

(School of Sciences, Zhejiang University of Science and Technology, Hangzhou 310023, Zhejiang, China)

Abstract: In order to reduce the impact of noise information and investor sentiment on stock prices, bring investors higher investment returns and lower transaction risks, a generative adversarial networks stock price forecasting model was proposed on the basis of financial bidirectional encoder representation from transformers and Wasserstein distance (FWGAN). This model firstly collected the stock evaluation data of eastmoney. com, and then quantified those data into sentiment values by virtue of the natural language processing pre-training model. Finally, the sentiment values, together with the historical stock transaction data and technical index data, were input into the long-short-term memory network (LSTM) to conduct training in the FWGAN model consisting of a generator and a convolutional neural network (CNN) as a discriminator. Comparing the forecasting performance of the LSTM model, the gated recurrent units (GRU) model and the generative adversarial networks (GAN)

收稿日期: 2021-03-26

基金项目: 浙江省科技计划项目(2015C33088)

通信作者: 胡月(1964—),男,河南省西峡人,教授,硕士,主要从事概率论极限理论和金融数学研究。E-mail: huyue@zust.edu.cn。

model on Shanxi Fenjiu stock price, the results show that the root mean square error of FWGAN model is 2.572, which is the lowest, achieving the best forecasting effect. Experimental results have verified the effectiveness and superiority of this model for stock time series forecasting, which can provide reference for investors to forecast stock prices.

Keywords: stock price forecasting; shareholder sentiment; generative adversarial networks
stock price forecasting model; time series

股票市场预测一直是量化金融投资的一个重要研究课题,然而股票市场是一个十分复杂的系统,其非线性、不平稳性、复杂性等特点导致准确预测股票价格十分困难。来自不同领域的研究人员对金融时间序列数据进行了大量研究,希望能从海量数据中找到规律,从而对股票价格进行精确预测。随着人工智能技术的不断深入,机器学习和深度学习等技术被引入金融预测。深度学习的非参数性及非线性能够为拟合复杂的、动态的、不确定性的股票市场提供很大帮助,成了当下热门的研究方法。张晨希等^[1]利用支持向量机(supper vector machine, SVM)对上市公司股票未来趋势进行预测,结果发现该模型预测精度高于传统神经网络分类方法。Akita 等^[2]对长短期记忆网络(long-short-term memory, LSTM)和传统神经网络做比较,发现 LSTM 对股票序列数据具有良好的拟合性能。彭燕等^[3]在 LSTM 网络的基础上,在数据预处理时进行小波降噪,预测效果得到提升。史建楠等^[4]提出了基于动态模态分解的 LSTM 股价预测方法,此方法较传统预测方法在特定市场背景下价格预测精度更高。赵红蕊等^[5]使用 LSTM 提取数据时序特征,并利用卷积神经网络(convolutional neural network, CNN)挖掘深层特征,发现该联合模型特征提取性能优于单一模型。为了筛选影响股市变化规律的相关特征, Rahman 等^[6-8]从媒体新闻中提取投资者情绪进行建模预测,发现将投资者情绪量化可以帮助模型更好地预测金融数据变化趋势。

由于 LSTM 网络无法深层刻画股市复杂的变化规律,研究人员开始聚焦生成式对抗网络(generative adversarial networks, GAN)的应用。GAN 最初被用于图像生成、目标检测等计算机视觉^[9]领域。Hensman 等^[10]随后提出了条件对抗网络,即在 GAN 中添加约束条件,从而有效地避免模型训练崩溃。罗会兰等^[11]结合部分卷积网络与 GAN 提出了一种图像修复模型,提高了图像修复质量。近年来 GAN 在股价预测中也得到了应用。Zhou 等^[12]首次将 GAN 应用到高频量化股票市场中,通过对高频股票市场股价趋势的预测,GAN 的预测效果最好,从而验证了 GAN 网络的优越性。Zhang 等^[13]使用多层感知机作为 GAN 的判别器对标准普尔 500 指数进行预测,结果证明 GAN 模型的预测效果优于其他模型。虽然 GAN 的博弈对抗机制使得股价预测准确率得到进一步提升,但是模型本身还存在训练不稳定、不易收敛等问题,并且在特征工程上并没有考虑到股民情绪对股票价格的影响。

针对上述问题,本研究提出了一种基于金融双向编码器表征和瓦瑟斯坦距离的生成式对抗网络(financial bidirectional encoder representation from transformers and Wasserstein generative adversarial networks, FWGAN)股价预测模型。本模型从股民评论、交易数据和技术指标中动态捕捉股票市场的潜在规律,使用自然语言处理技术中的预训练模型提取文本数据情感特征,并通过优化分布度量距离的瓦瑟斯坦距离生成式对抗网络(Wassersteingenerative adversarial networks, WGAN)进行博弈对抗来生成真实数据,在个股数据集进行对比试验,验证模型的有效性。

1 模型介绍

1.1 LSTM 长短期记忆网络

循环神经网络(recurrent neural network, RNN)是深度神经网络的一种,它通过对过往信息进行不同权重的分配来实现记忆功能,因其能够较好地处理长距离依赖关系而被广泛用于时间序列预测任务中。传统 RNN 在反向传播过程中容易产生“梯度爆炸”和“梯度消失”问题,针对此问题, Hochreiter 等^[14]在 1997 年提出了 LSTM 网络,引入门控方法来避免输入和输出之间的权重冲突,并采用特定的学

习机制来遗忘和更新信息,这更有助于长距离的信息传输。由于 LSTM 有效地解决了 RNN 难以收敛的问题,因此现已被广泛用于处理时间序列数据,如命名实体识别、机器翻译等。

LSTM 网络通过采用门控机制赋予神经单元衡量信息价值的能力。门控机制的 3 个“门”分别为输入门、输出门和遗忘门,其表达式如下:

$$\begin{cases} f_t = \delta(U_f x_t + W_f h_{t-1} + b_f); \\ i_t = \delta(U_i x_t + W_i h_{t-1} + b_i); \\ u_t = \tanh(U_u x_t + W_u h_{t-1} + b_u); \\ c_t = f_t c_{t-1} + i_t u_t; \\ o_t = \delta(U_o x_t + W_o h_{t-1} + b_o); \\ h_t = o_t \tanh(c_t). \end{cases} \quad (1)$$

式(1)中: f_t 为遗忘门输出; δ 为 sigmoid 函数; U_f, U_i, U_u, U_o 和 W_f, W_i, W_u, W_o 均为权重矩阵; x_t 为当前时刻输入; h_{t-1} 为上一时刻隐藏状态; b_f, b_i, b_u, b_o 均为偏置系数; i_t 为输入门输出; u_t 为候选状态; c_t 为内部状态; c_{t-1} 为上一时刻状态; o_t 为输出门输出; h_t 为外部状态。首先利用 h_{t-1} 与 x_t 来计算 f_t, i_t 和 o_t 这 3 个门输出及 u_t 的值;然后结合遗忘门 f_t 和输入门 i_t 更新内部状态 c_t ,这是 LSTM 网络结构的核心部分;最后通过输出门 o_t ,得到当前时刻的外部状态 h_t 。通过以上计算,LSTM 可以有效利用输入数据使其具备长序列的记忆功能。

1.2 CNN 卷积神经网络

CNN 也是深度神经网络的一种,最早被应用于图像处理中,近些年开始被用于自然语言序列处理任务中,并取得了较好的效果。例如,利用膨胀卷积网络进行命名实体识别任务。常见的 CNN 网络是由卷积层、池化层和全连接层构成。卷积层通过卷积运算来提取数据的特征,池化层通过特征选择与过滤来降低数据维度,全连接层是传统的前馈神经网络中的隐藏层,一般是和输出层进行连接,实现最后的分类。

1.3 GAN 生成式对抗网络

GAN 是由 Goodfellow 等^[15]在 2014 年受博弈论中的二人零和博弈最终达到纳什均衡思想的启发而提出的一种生成式神经网络。其网络结构由生成器(generative model)和判别器(discriminative model)两部分组成。生成器的作用是通过对真实数据的学习训练从而产生逼近真实样本数据的分布;判别器的作用是分辨输入数据是真实数据还是生成器的输出数据。生成器想要产生更接近真实的数据,判别器想要更完美地区分真实数据和生成的数据。由此两个网络在对抗中彼此互相提高,在提高中又相互对抗,最终到达纳什均衡。GAN 网络一提出便引起轰动,受到工业界和学术界的重视,并在图像处理、语音识别等领域得到广泛应用,近年也开始被用于金融领域。GAN 的基本结构如图 1 所示。

图 1 中随机噪声一般符合均匀分布或高斯分布。输入数据从高斯分布中采样,通过生成器会输出生成的“伪”数据,随后将真实数据和生成数据一同送入判别器,经过 sigmoid 函数,最后输出判别类型。生成器和判别器不断博弈,提高各自的网络特征提取性能。当生成器通过训练完全掌握到真实数据的分布时,判别器已无法精确辨认数据的来源,这样就达到用生成数据来代替原始数据的效果。GAN 的目标函数如下:

$$\min_{G(z)} \max_{D(x)} V(D(x), G(z)) = E_{x \sim p(x)} \log_2 D(x) + E_{z \sim p_z(z)} \log_2 [1 - D(G(z))]. \quad (2)$$

式(2)中: V 为目标函数; E 为期望; $D(x)$ 为判别器判断真实数据是否真实的概率,判别器的能力越强, $D(x)$ 越大,则目标函数值越大; $D(G(z))$ 为判别器判断生成的数据是否真实的概率,生成器 $G(z)$ 不断接近真实样本数据分布, $D(G(z))$ 会变大,则目标函数则会变小。这样一个极大值、极小值的博弈交替循环来提高生成器和判别器的能力,最终生成器产生的数据与真实数据无限接近,达到纳什均衡。

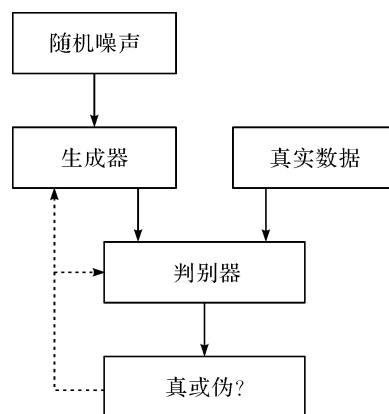


图 1 GAN 的基本结构

Fig. 1 Basic structure of GAN

2 FWGAN 股价预测模型的建立

2.1 WGAN

GAN 在最优判别器下生成器的损失值存在梯度不稳定、梯度消失,或准确性与多样性惩罚不平衡,从而导致模型崩溃的问题。产生这些问题的主要原因是等价优化的距离度量方式不合理。为了有效地解决这些问题,Gulrajani 等^[16]于 2017 年首次提出了 WGAN,用 Wasserstein 距离代替 GAN 计算损失值散度,将其转化为一个线性规划问题来更好地度量 2 个分布之间的距离。Wasserstein 距离又称为地动(earth-mover,EM)距离,常用于衡量 2 个分布之间的距离,其数学表达式为

$$W(P_r, P_g) = \inf_{r \sim \Pi(P_r, P_g)} E_{(x,y) \sim r} \|x - y\|. \quad (3)$$

式(3)中: \inf 为取下界; $\Pi(P_r, P_g)$ 为 P_r 与 P_g 组合起来的所有可能的联合分布集合; $\|x - y\|$ 为样本 x 与样本 y 之间的距离。WGAN 网络的优化函数如下

$$L = E_{x \sim P_r} f_w(x) - E_{z \sim P_g} f_w(g(z)). \quad (4)$$

式(4)中: $f_w(x)$ 和 $f_w(g(z))$ 为满足利普希茨限制的函数,将所需优化目标函数转化为对偶问题进行求解。WGAN 通过最小化距离 L 使生成数据与真实数据之间的距离缩小,可以高效地解决传统 GAN 存在的多样性不足和模型崩溃等问题。

2.2 预训练语言模型

预训练技术通过预先设计网络结构并将编码后的数据输入网络结构中进行训练来提高模型的泛化能力^[17]。预训练技术最初被应用于图像领域,因其效果良好,开始应用于自然语言处理任务中。预训练后的模型可以根据下游具体任务进行微调,从而避免从零开始搭建模型,并且结果证实预训练技术可以显著提高模型性能。双向编码器表征(bidirectional encoder representation from transformers, BERT)^[18]模型是近年提出的一种预训练语言模型,它是由转换器(Transformer)^[19]的编码器堆叠构成。Transformer 是由用于接收文本的编码器和进行结果预测的解码器组成,并且由于在网络中融入了注意力机制,所以可以更好地进行文本上下文关系推理。传统词向量技术(word to vector, Word2Vec)是一种静态的方式,无法充分表征上下文语义信息,有一定不足。而 BERT 通过采用注意力机制可以学习到上下文中的潜在联系,从而提高词向量模型的泛化能力,可以充分表征词意和句意。BERT 模型结构如图 2 所示。

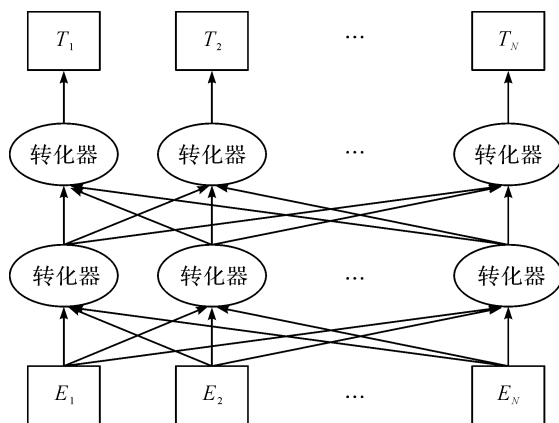


图 2 BERT 模型结构

Fig. 2 BERT model structure

BERT 是针对通用文本语料进行大规模训练的预训练模型,所以针对特定下游任务如金融领域上的训练效果并不理想,随后又有研究者将 BERT 应用于金融领域,提出了金融双向编码器表征模型(financial bidirectional encoder representation from transformers, FinBERT)^[20]。FinBERT 是在大量的金融财经类新闻、上市公司公告和金融类百科词条文本中预训练得到的语言模型,相比 BERT,在金融数据集下游任务的预测准确率中大大提升。本文采用的股民评论是含有大量的金融专业术语的金融文本数据,且上下文间联系紧密,因此使用 FinBERT 预训练模型可以有效提取股民评论数据的情感特征。

2.3 模型框架

为了进一步提升股价预测模型的性能,本研究使用优化分布度量距离的 WGAN 进行对抗博弈,以 LSTM 为生成器, CNN 为判别器,并结合自然语言处理技术构建多特征结合的生成式对抗网络模型进行股价预测。首先爬取股民评论数据,并进行数据清洗,人工标注部分数据用于 FinBERT 预训练模型的微调训练;其次通过微调的 FinBERT 模型将由评论数据转化而来的情绪值作为股价预测的特征之一;最后将股

票交易数据、技术指标数据和获得的情绪值 3 个维度的特征输入 WGAN 进行训练。FWGAN 模型结构如图 3 所示。

3 试验与分析

3.1 数据集

为了提高模型预测的准确率,采用的数据集包含股票交易数据、技术指标数据及股评信息 3 部分。股票交易数据选用 2010 年 7 月 1 日至 2020 年 7 月 1 日山西汾酒开盘价、最高价、最低价、收盘价、成交量、成交额等历史交易数据和与山西汾酒同属酿酒板块具有联动关系的贵州茅台、五粮液、泸州老窖的收盘价^[21]。技术指标数据包括移动平均线、布林线、指数平滑移动平均线(moving average convergence and divergence, MACD)和经过傅里叶变换的趋势线。股票评论信息来自东方财富网股吧中关于山西汾酒股票每日的股评,选取当日阅读量较高的股评,共爬取 70 794 条股评数据。将爬取的股评信息经过处理之后,转化成 0~1 之间的情感值,计算出当日情感均值。将数据集分为 70% 的训练集和 30% 的测试集,训练集用于模型训练,测试集用于模型测试。

3.2 数据预处理

从东方财富网爬取的原始股票评论数据进行数据清洗并除噪得到高质量数据,通过微调的 FinBERT 模型将文本数据转化为 0~1 之间的情绪值,随机选取交易日开盘一小时内的部分评论文本,其情绪值见表 1。

傅里叶变换^[22]是通过特殊正交基实现从时域到频域的变换,傅里叶原理揭示任何一个连续时序或信号,都可以表示为不同频率的正弦波信号的无限叠加,其表达式为

$$G(f) = \int_{-\infty}^{+\infty} g(t) e^{-2\pi i f t} dt. \quad (5)$$

式(5)中: $g(t)$ 为时域函数,通过傅里叶变换提取局部和整体趋势后,创建真实股票运动的近似值,并经过降噪作为模型的输入可以使模型预测更准确。使用分别具有 3、6、9 个成分进行变换来提取短期和长期趋势。傅里叶变换趋势如图 4 所示。由图 4 可知,使用的傅里叶变换成分越多,函数值越接近实际值。这表明傅里叶变换可以有效提取股价趋势,从而帮助模型进行预测。

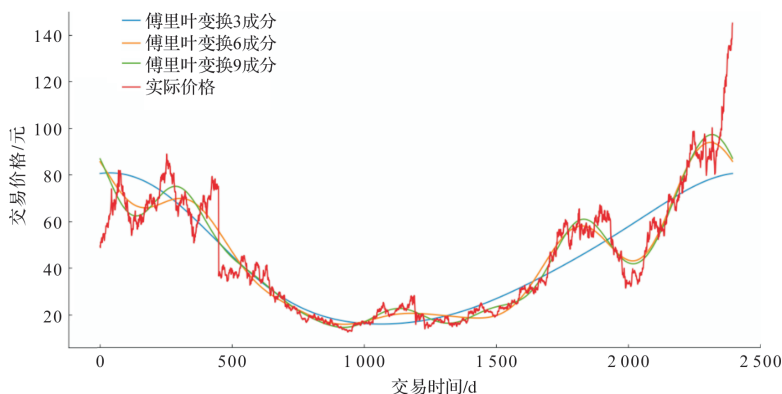


图4 傅里叶变换趋势

Fig. 4 Fourier transform trend



图3 FWGAN 模型结构

Fig. 3 FWGAN model structure

表1 股民评论情绪分析

Table 1 Sentiment analysis on investors' comments

时间	评论	情绪值
2020-05-26 09:35	今天会创新高,130元应该看得见	0.898
2020-05-26 09:31	6个点挂单无效,到底被谁买了又卖了	0.014
2020-05-26 09:27	开什么玩笑,都想要超过五粮液了	0.602
2020-05-26 09:26	只有白酒通货膨胀	0.048
2020-05-26 09:26	哈哈,103买的122卖了,作为山西人的骄傲	0.773

3.3 评价指标

为了评估 FWGAN 模型对股价的预测性能,使用平均绝对误差(mean absolute error,MAE)、平均绝对百分比误差(mean absolute percentage error,MAPE)、均方误差(mean square error,MSE)和均方根误差(root mean square error,RMSE)作为模型评价指标,计算公式如下:

$$\begin{cases} Z_{\text{MSE}} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2; \\ Z_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}; \\ Z_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|; \\ Z_{\text{MAPE}} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\%. \end{cases} \quad (6)$$

式(6)中: \hat{y}_i 为预测值, $i=1,2,\dots,n$; y_i 为真实值; $Z_{\text{MSE}}, Z_{\text{RMSE}}, Z_{\text{MAE}}, Z_{\text{MAPE}}$ 的取值范围为 $[0, +\infty)$ 。取值越趋于 0 则表明预测值越接近真实值,模型的预测性能越好;反之,则表明越差。

3.4 试验设置

试验中计算机操作系统为 64 位 Windows10,处理器为英特尔 i5 2.4 GHz,运行内存为 16 GB,使用框架为 Tensorflow1.14,编程语言为 Python3.6.5,优化器的学习率为 0.000 14。FWGAN 训练损失函数变化如图 5 所示。

从图 5 中可以看出,随着迭代训练次数的增加,损失值减小并趋于稳定,生成样本和真实样本之间的差距逐渐变小。

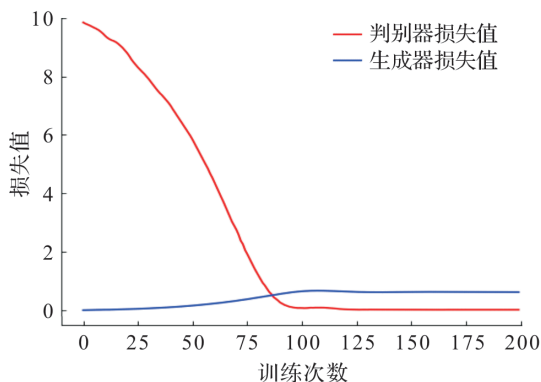


图 5 FWGAN 训练损失函数变化

Fig. 5 FWGAN loss function changes

3.5 试验结果对比分析

为了验证我们提出的多特征结合的生成式对抗网络股价预测模型的性能,对山西汾酒股票每日收盘价进行预测,用前 5 日的指标数据来预测第 6 日收盘价并与 LSTM、门控神经网络(gated recurrent units, GRU)和 GAN 进行对比,观察 4 种模型在训练集和测试集上的表现。各模型的试验结果指标对比见表 2。

表 2 各模型的试验结果指标对比

Table 2 Comparison of experimental indexes of each model

模型	数据	Z_{MSE}	Z_{RMSE}	Z_{MAE}	$Z_{\text{MAPE}}/\%$
LSTM	训练数据	6.718	2.592	1.291	1.375
	测试数据	26.307	5.129	3.358	2.082
GRU	训练数据	4.016	2.004	1.008	0.923
	测试数据	22.915	4.787	2.448	1.206
GAN	训练数据	3.597	1.897	0.813	1.123
	测试数据	15.603	3.950	2.085	3.450
FWGAN	训练数据	3.387	1.946	0.879	1.116
	测试数据	6.615	2.572	1.956	3.368

表2中的试验结果表明,对比 LSTM 模型与 GRU 模型,GRU 模型在训练数据和测试数据上的 MSE、RMSE、MAE、MAPE 的值更小,说明 GRU 模型的预测值更接近真实值,预测性能更好。这主要是因为 GRU 网络是将 LSTM 网络中的更新门和遗忘门合并为重置门,将需要进行训练的参数变小从而防止过拟合,提高预测准确率。因此相比 LSTM 模型,GRU 模型更适合股票时间序列数据。相比 LSTM 模型和 GRU 模型,GAN 模型在训练集和测试集上的 RMSE 值均有所下降,说明经过生成器和判别器的博弈对抗,模型能较好地学习真实数据分布,更好地进行股票价格预测。FWGAN 作为本研究提出的新模型,在融合了历史数据、技术指标和情绪值指标之后,在测试集上的 MSE、RMSE 和 MAE 均达到最小,验证了使用 Wasserstein 距离代替 GAN 计算损失值散度之后,预测能力更强,从而也证明了本研究提出的模型的有效性。为了更清楚地对比 FWGAN 模型与各模型之间的预测效果差异,将不同模型在测试集上的预测结果进行了可视化处理,如图 6~9 所示。

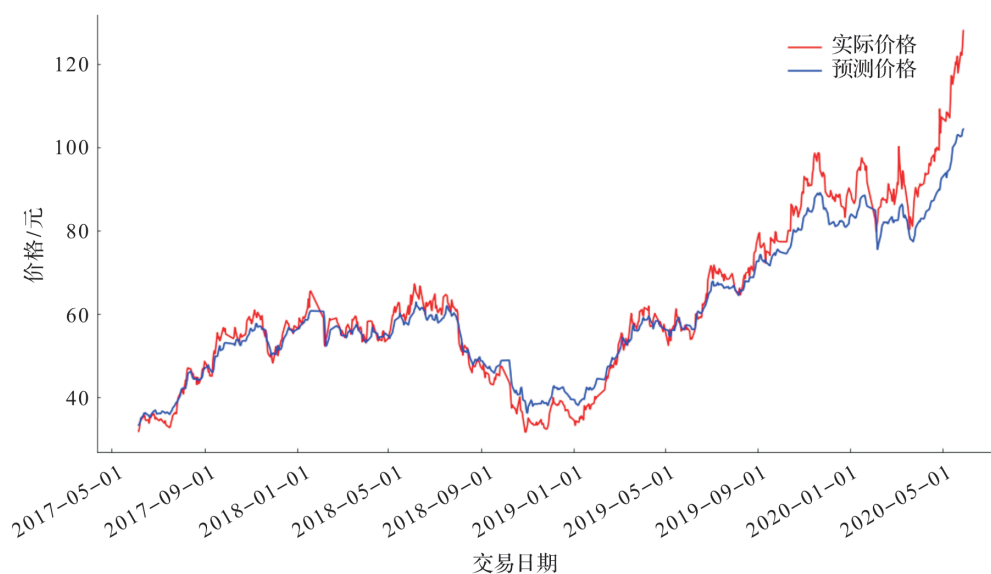


图6 LSTM 模型预测结果

Fig. 6 LSTM model forecasting results

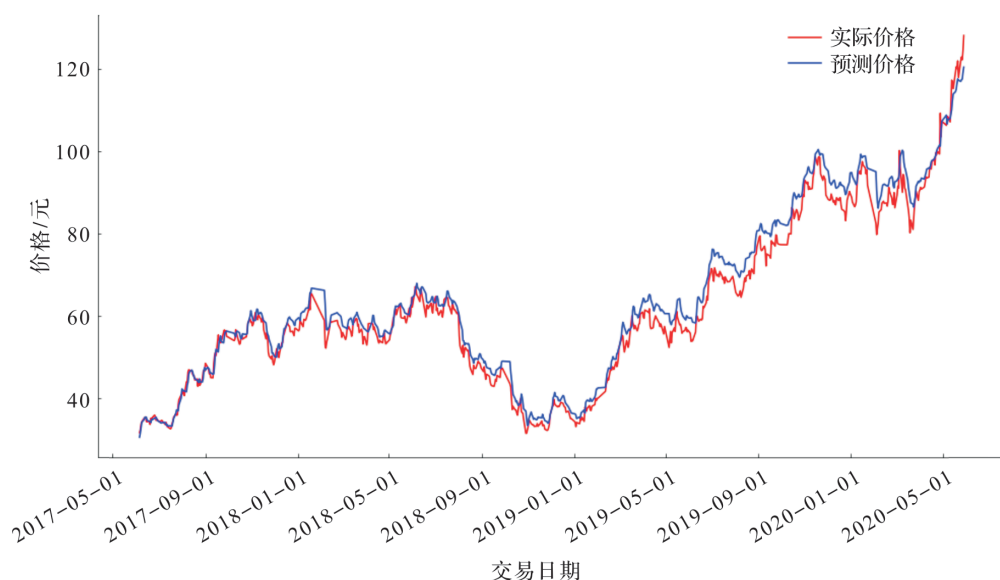


图7 GRU 模型预测结果

Fig. 7 GRU model forecasting results

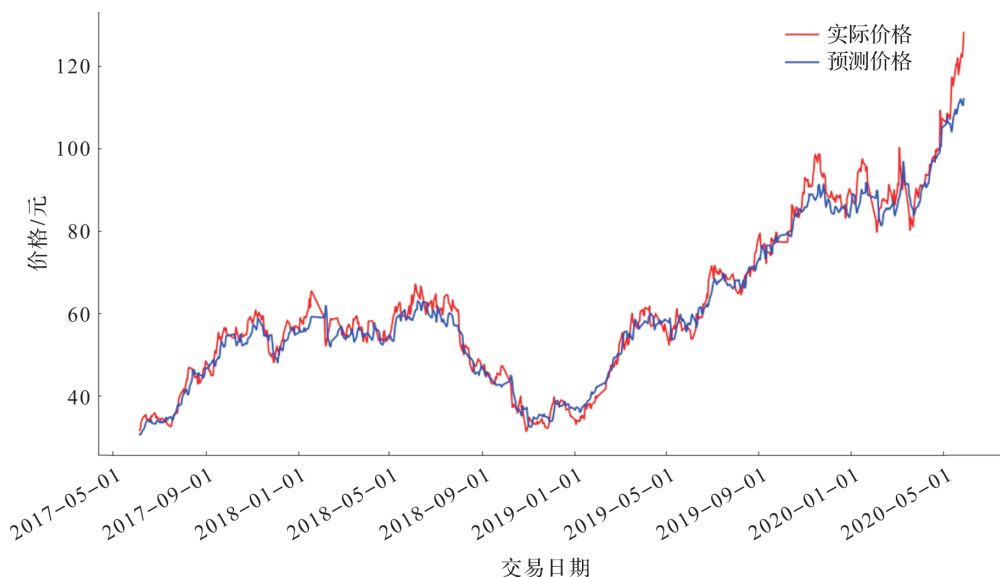


图 8 GAN 模型预测结果

Fig. 8 GAN model forecasting results

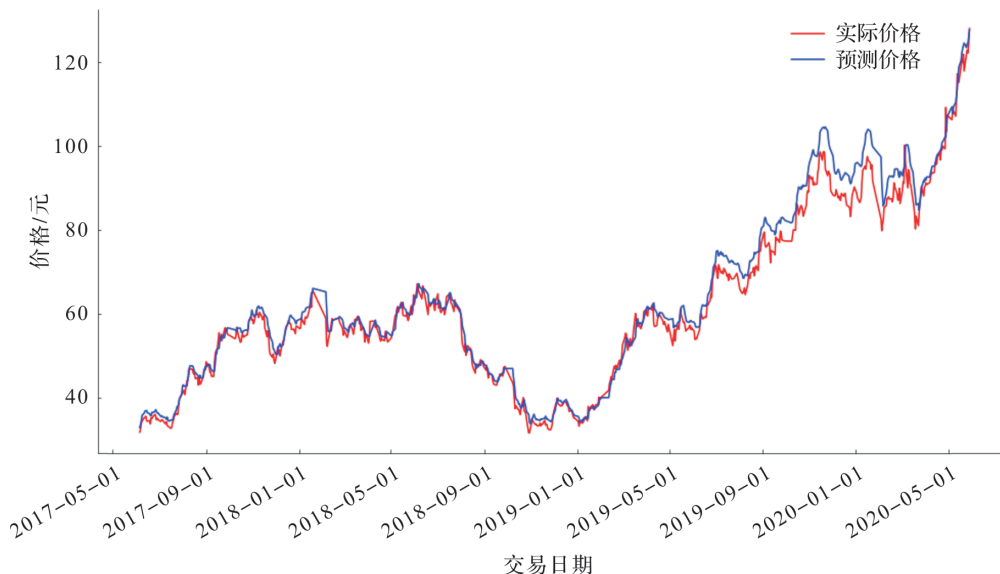


图 9 FWGAN 模型预测结果

Fig. 9 FWGAN model forecasting results

从图 6 中可以看出,LSTM 模型的 RMSE 值为 5.129,该模型的整体预测值低于实际值,预测值与真实值之间的偏差较大,模型可能存在过拟合现象。图 7 GRU 模型的 RMSE 值为 4.787,相比 LSTM 模型,GRU 模型的预测值与实际值差距变小,预测性能上升,但 GRU 的预测值在整体上偏高于实际值。GAN 模型的 RMSE 值为 3.950,预测误差值比 LSTM 和 GRU 更小,但是从图 8 中可以发现,GAN 模型对股票上升趋势的捕捉能力不够强,当股价大幅度上升时,预测值与真实值之间仍有较大的差距。从图 9 中可以看出,经过优化的 FWGAN 模型在测试集上 RMSE 值为 2.572,其预测值与真实值间的差距最小;同时模型能较好地捕捉股价上升、下降地趋势信息,预测性能最好。其原因如下:1) GAN 网络的对抗博弈机制能很好地捕捉原始数据分布信息,使生成的数据能很好地拟合原始分布;2) Wasserstein 距离作为一种新的距离衡量方式,能有效地解决等价距离不合理的问题,从而有利于模型训练中的参数更新。

4 结 语

本研究提出了基于生成式对抗网络的股价预测方法,首先运用自然语言处理中的 FinBERT 预训练

模型进行股民情感分析,将股民情绪值转化为0~1之间的数值;然后将股民情绪值结合股票交易数据和技术指标数据作为模型的输入,采用以LSTM为生成器、CNN为判别器并优化度量距离WGAN网络进行股价预测;最后采用山西汾酒的十年股票数据进行试验,相比LSTM、GRU和GAN模型,优化后的模型预测性能得到了提升。本研究采用的模型虽然具有较好的预测性能,但是对股价数据具有较强的依赖性,下一步我们将考虑在医疗、科技其他板块股票数据上进行试验,在特征工程上考虑加入公司财务指标数据等基本指标数据,并结合更先进的自然语言处理技术,提取细粒度的情感值以帮助模型进行预测。

参考文献:

- [1] 张晨希,张燕平,张迎春,等. 基于支持向量机的股票预测[J]. 计算机技术与发展,2006(6):35.
- [2] AKITA R, YOSHIHARA A, MATSUBARA T, et al. Deep learning for stock prediction using numerical and textual information[EB/OL]. (2016-07-26) [2021-03-23]. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7550882>.
- [3] 彭燕,刘宇红,张荣芬. 基于LSTM的股票价格预测建模与分析[J]. 计算机工程与应用,2019,55(11):209.
- [4] 史建楠,邹俊忠,张见,等. 基于DMD-LSTM模型的股票价格时间序列预测研究[J]. 计算机应用研究,2020,37(3):662.
- [5] 赵红蕊,薛雷. 基于LSTM-CNN-CBAM模型的股票预测研究[J]. 计算机工程与应用,2021,57(3):203.
- [6] RAHMAN A S A, ABDUL-RAHMAN S, MUTALIB S. Mining textual terms for stock market prediction analysis using financial news[C]//International Conference on Soft Computing in Data Science. Singapore: Springer,2017.
- [7] ZHANG X, ZHANG Y J, WANG S Z, et al. Improving stock market prediction via heterogeneous information fusion [J]. Knowledge-Based Systems,2018,143:236.
- [8] 张明禄. 新闻事件驱动的市场预测研究[D]. 上海:上海交通大学,2013.
- [9] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]//Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE,2016:2818.
- [10] HENSMAN P, AIZAWA K. cGAN-based manga colorization using a single training image[EB/OL]. (2017-05-30) [2021-03-23]. <https://arxiv.org/pdf/1706.06918.pdf>.
- [11] 罗会兰,敖阳,袁璞. 一种生成对抗网络用于图像修复的方法[J]. 电子学报,2020,48(10):1891.
- [12] ZHOU X Y, PAN Z S, HU G Y, et al. Stock market prediction on high-frequency data using generative adversarial nets[J]. Mathematical Problems in Engineering,2018,special issue:1.
- [13] ZHANG K, ZHONG G Q, DONG J Y, et al. Stock market prediction based on generative adversarial network[J]. Procedia Computer Science,2019,147:400.
- [14] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation,1997,9(8):1735.
- [15] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Advances in Neural Information Processing Systems,2014,3:2672.
- [16] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of Wasserstein GANs[C]//Proceedings of the 31st Annual Conference on Neural Information Processing Systems. New York: Neural Information Processing Systems Foundation,2017:5769.
- [17] 李舟军,范宇,吴贤杰. 面向自然语言处理的预训练技术研究综述[J]. 计算机科学,2020,47(3):170.
- [18] 王子牛,姜猛,高建瓴,等. 基于BERT的中文命名实体识别方法[J]. 计算机科学,2019,46(增刊2):138.
- [19] CORREIA G M, NICULAE V, MARTINS A F T. Adaptively sparse transformers[EB/OL]. (2019-08-30) [2021-03-23]. <https://arxiv.org/pdf/1909.00015v1.pdf>.
- [20] ARACI D T. FinBERT: financial sentiment analysis with pre-trained language models[EB/OL]. (2019-06-25) [2021-03-23]. <https://arxiv.org/pdf/1908.10063.pdf>.
- [21] 侍冰雪,魏慧茹,朱韶东,等. 基于网络结构的股票相关性研究[J]. 浙江科技学院学报,2015,27(1):62.
- [22] 张杰,王苗. 基于傅里叶分析法的股票市场超高频数据相关性分析[J]. 数学的实践与认识,2010,40(7):63.