

文章编号: 2095-2163(2022)05-0123-06

中图分类号: TP391

文献标志码: A

基于 LSTM 的股票价格预测分析

李桂城, 许丽, 张利

(贵州大学 大数据与信息工程学院, 贵阳 550025)

摘要: 本文主要对股票预测的方式进行改进,使预测结果更接近真实数据。以往的股票价格预测研究大多简单地将股票价格作为序列数据,通过模型进行训练预测,或者只是通过分析新闻文本、股民评论的情感倾向预测股票价格的涨跌,这都不能全面地对股票价格进行考量。本文通过参考影响股票实际价格的多种因素,对股票价格预测结果展开研究。本研究提出新的模型 M,该模型结合了时间序列预测以及文本情感分析两种方法,采用的均为 LSTM 的拓展模型,分别是 DA-RNN、BiLSTM-Attention。一方面,通过分析对比 XGBoost、LSTM 与 DA-RNN 的实验结果验证了 DA-RNN 在时序预测实验中的有效性;另一方面,采用 BiLSTM-Attention 模型作为调整股票预测值的主要方法,使预测结果更加具备可解释性。本研究通过改进的股票价格预测模型 M 来进行股票价格的预测。实验结果表明,在海康威视股票数据集中,当文本情感倾向的权重值设置为 0.01 时,MAE、RMSE 值均达到最小,即预测效果最好。

关键词: 股票价格预测; 时间序列; 文本情感分析; LSTM

Stock price forecast analysis based on LSTM

LI Guicheng, XU Li, ZHANG Li

(College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China)

[Abstract] This study mainly optimizes the stock forecasting method to make the stock price forecast closer to the real data. Most of the previous researches on stock price prediction simply take stock price as serial data and make training prediction through the model, or just predict the rise and fall of stock price by analyzing the emotional tendency of news text and shareholders' comments, which cannot comprehensively consider stock price. By referring to the factors that affect real stock prices, the paper investigates the results of stock price predictions. This study proposes a new model M, combines two methods of time series prediction and text sentiment analysis, and uses the extended model of LSTM, namely DA-RNN and BiLSTM-Attention. On the one hand, the experimental results of XGBoost, LSTM, and DA-RNN are analyzed and compared to verify the effectiveness of DA-RNN in the time series prediction experiment; on the other hand, BiLSTM-Attention is used as the main method to adjust the stock forecast value, and makes the results more interpretable. This study uses an improved stock price forecasting method M to predict stock prices. The results demonstrate that in the Hikvision stock data set, when the weight value of text sentimentality is set to 0.01, the MAE and RMSE values are both minimized. This means the prediction is the best.

[Key words] stock price forecast; time series; text sentiment analysis; LSTM

0 引言

随着人工智能技术的发展和金融市场的完善,利用大数据分析技术对股市进行预测将有着重要的研究价值。在股价预测方面,目前已有不少成熟的研究,比如,运用在时间序列里包含的信息对未来的走势做出预测,运用和股票有关的新闻信息对股价的涨跌做出预测等。其中,深度学习模拟神经网络的发展过程,使得股票预测的研究具有良好的可行性。

股票价格预测非常困难,因为影响股票价格的因素有很多^[1]。对于个股来说,股票价格数据的趋势、股票对应公司的运营状况、股民对该个股的评

价、国家的政策、突发的新闻事件等等,都会对股票价格产生影响。股票价格数据是一种时序数据,时序数据具有一定的趋势,股民常根据数据判断价格走势,此预测过程可通过深度学习方法实现^[1]。

在深度学习中,RNN 常会用来处理序列数据,如时间序列、文本数据等。但已有研究表明,RNN 可能导致梯度消失等问题,于是研究者引入了 LSTM(Long-term and Short-Term Memory Network),并在 RNN 的基础上添加了门控结构以及记忆单元,使得网络能够决定遗忘的信息以及往后传播的信息,因此也具有了解决梯度爆炸和梯度消失的能力。除了应用于图像分类、语音识别等领域,LSTM 在文

作者简介: 李桂城(1998-),男,硕士研究生,主要研究方向:时序预测、金融大数据;许丽(1998-),女,硕士研究生,主要研究方向:时序预测、金融大数据;张利(1987-),男,博士,特聘教授,主要研究方向:天文大数据、医疗大数据、金融大数据、图像处理。

通讯作者: 张利 Email: lizhang.science@gmail.com

收稿日期: 2021-11-23

本情感分析、时序数据预测等方面也表现良好^[2]。

当下研究中,文献[1]将 RNN 模型用于预测股票在未来一天的涨跌情况,证明了将新闻序列与股票价格序列相结合能够带来更高的准确率。文献[3]把 LSTM 模式运用于对股市波动率的预测中,经过大量的对比实验,得出了 LSTM 模型的预测效果优于 18 种传统的经典预测模型的结论,与此同时,历史股票的增加会使得 LSTM 模型的预测结果更加稳定。文献[4]提供了一个基于 LSTM 技术的深度神经网络 DP-LSTM,其中包括了使用新闻文章作为隐藏信息并利用差异隐私机制,整合不同的新闻来源对标普 500 股票价格预测,大量实证研究表明提出的 DP-LSTM 的预测比平均 MPA 增加了 0.32%,对标普 500 指数的预测, MSE 改善达到了 65.79%。文献[5]提出了 CNN-LSTM 混合神经网络,通过结合多种可能影响股票价格的因素,对结果进行预测。另外,为提高结果的准确率,作者在文章中增加了注意力机制,将该模型以及几种经典的方法应用到 2 个真实的数据集上,通过横向以及纵向的比较验证了所提出模型的有效性。文献[6]运用 OpinionFinder 等工具对 Twitter 上的内容进行每一天的的大众情感分析,并将提取到的情感特征加入模型中用于预测股票价格的涨跌。

根据研究表明,很多学者在股票预测上要么单一地考虑影响股票价格的因素,要么只是根据历史值来预测未来值,这都不能全面地对股票价格进行考量。另外,大量研究发现个人情绪和感情因素会对人的决策产生影响,对此进行分析即可从中预测并获取股票的相关指标,基于此,很多学者利用文本情感分析来研究股票价格的变动^[7]。基于此,本文融合了多种影响股票实际价格的因素,提出的时序预测与文本情感分析结合方法具有一定重要的现实意义。

本文内容安排如下:首先简略地概述了深度学习以及文本情感分析在股票价格变动以及涨跌方面的发展及应用;然后,阐述了所应用的深度学习模型的方法及原理;接下来,将所提出的模型与传统方法应用于海康威视真实数据集上,以 RSME、MAE、准确率、召回率等作为评价指标,对各模型进行对比评价。最后,对本文内容进行总结概括。

1 算法原理及流程

股票价格预测指的是根据股票价格的历史数据以及与股票相关的市场信息,预测股票在接下来一段时间内的具体价格或者涨跌状况^[1]。以往的股

票价格预测研究大多简单地将股票价格作为序列数据,通过模型进行训练,或者只是通过分析新闻文本、股民评论的情感倾向进行预测。为了将更多因素添加到股票价格预测中,本文将这 2 方面结合起来考虑,首先采用时间序列预测模型预测股票价格,然后再通过提取新闻文本信息等,对先前预测结果进行调整,使预测值更趋近于真实值。

基于 LSTM 在文本、时序数据方面的优良表现,本文将使用基于 LSTM 的深度学习模型,应用于时序预测与文本情感分析两个方面,重点探索时间序列趋势预测以及新闻文本情感对股票价格的影响,使股票价格预测结果更加准确。本文的算法流程图如图 1 所示。

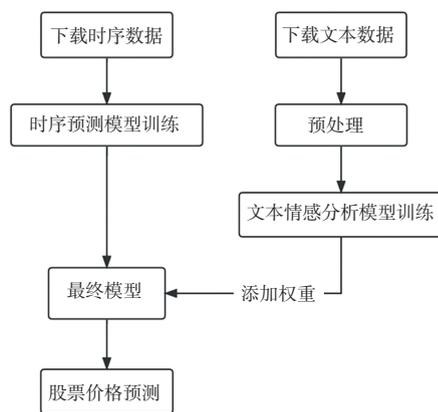


图 1 算法流程图

Fig. 1 Algorithm flow chart

1.1 时间序列预测

在时间序列预测中,本文引入了基于双阶段注意力机制的循环神经网络(DA-RNN)。DA-RNN 模型如图 2 所示。该模型包含 2 种注意力机制,第一种是在输入阶段中,使用注意力机制的编码器,第二种是在每个时间步上,采用时间注意力机制的解码器。输入阶段的注意力机制可以在每个时刻针对性地选择相关信息,而时间注意力机制可以在比较长的时间段内捕捉到该时间段内序列的关系^[8]。

输入阶段的注意力机制实现即为每个时刻输入的影响因子赋予一定的权重,首先使用当前时刻的输入 $x_t \in R^n$ 以及前一个时刻编码器的隐藏状态 h_{t-1} ,更新当前时刻编码器的隐藏层状态 h_t ,更新公式如下:

$$h_t = F(h_{t-1}, x_t) \quad (1)$$

其中, F 是一个非线性激活函数,此处使用的是 LSTM,利用其特性来捕获长距离依赖关系。每个 LSTM 单元的更新公式总结如下:

$$f_t = \sigma(W_f[h_{t-1}; x_t] + b_f) \quad (2)$$

$$i_t = \sigma(W_i[h_{t-1}; x_t] + b_i) \quad (3)$$

$$o_t = \sigma(W_o[h_{t-1}; x_t] + b_o) \quad (4)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_c[h_{t-1}; x_t] + b_c) \quad (5)$$

$$h_t = o_t \odot \tanh(c_t) \quad (6)$$

其中, $[h_{t-1}; x_t] \in R^{m+n}$ 是前一个隐藏层状态 h_{t-1} 与当前输入 x_t 的拼接; $W_f, W_i, W_o, W_c \in R^{m \times (m+n)}$, 以及 $b_f, b_i, b_o, b_c \in R^m$ 等是需要学习的参数。为了自适应地选择相应的特征, 这里引入了注意力机制。即对于每个时间步的输入 x_t , 为相应的每个影响因子添加一定的权重, 即 α_t^k 。得到新的输入 \tilde{x}_t , 具体计算方式如下:

$$\tilde{x}_t = (\alpha_t^1 x_t^1, \alpha_t^2 x_t^2, \dots, \alpha_t^n x_t^n) \quad (7)$$

其中, α_t^k 可以由 e_t^k 归一化得到, 而 e_t^k 是根据上一个时刻的隐藏层状态 h_{t-1} 和单元状态 c_{t-1} 计算而来, 数学公式如下:

$$e_t^k = v_e^T \tanh(W_e[h_{t-1}; c_{t-1}] + U_e x_t^k) \quad (8)$$

使用 *softmax* 函数进行归一化, 即:

$$\alpha_t^k = \frac{\exp(e_t^k)}{\sum_{i=1}^n \exp(e_t^i)} \quad (9)$$

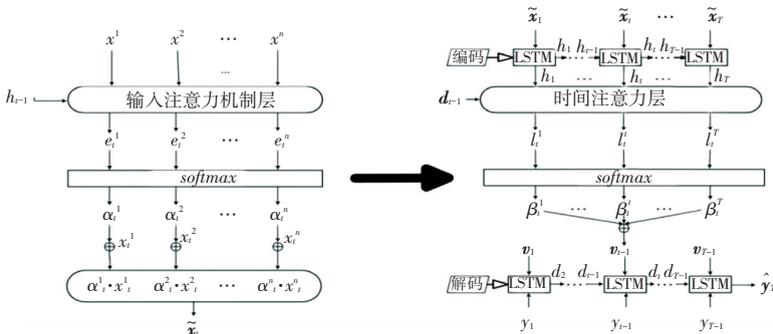


图 2 DA-RNN 模型

Fig. 2 DA-RNN model

1.2 文本情感分析

在文本情感分析中, 本文引入了基于注意力机制的双向 LSTM 模型, 双向 LSTM 包含前向和后向两个方向上的 LSTM, 前向 LSTM 主要获取上文的信息特征, 后向 LSTM 主要获取下文的信息特征, 因此该模型包含了前向与后向的所有信息, 并使用 Attention 注意力机制有选择性地聚焦比较重点的词语, 最终对句子的情感进行分类。既解决了多维度的信息爆炸以及传统循环神经网络的梯度消失等问题, 也有效地避免了传统模型对重点词语在上下文中提取信息的缺失^[9]。该模型总共包含输入层、嵌入层、LSTM 层、注意力层、输出层五个组成部分。

为预测输出 \hat{y}_T , 第二阶段使用了基于时间注意力机制的解码器, 该注意力机制类似于 encoder-decoder 结构中的机制, 这里的注意力层将前一个时间点隐藏层的状态 d_{t-1} 和单元状态 c'_{t-1} 作为输入, 计算得到该层的输出 l'_t , 并进行归一化得到 β_t^i 。这里涉及的数学公式可写为:

$$l'_t = v_d^T \tanh(W_d[d_{t-1}; s'_{t-1}] + U_d h_t) \quad (10)$$

$$\beta_t^i = \frac{\exp(l'_t)}{\sum_{j=1}^T \exp(l'_j)} \quad (11)$$

接下来, 将所有编码器隐藏层状态根据权重计算得到文本向量 $v_t = \sum_{i=1}^T \beta_t^i h_i$, 这里 v_t 在不同时刻是不同的。此后, 将目标序列结合得到 $\tilde{y}_{t-1} = \tilde{w}^T [y_{t-1}; v_{t-1} + \tilde{b}]$ 。利用 \tilde{y}_{t-1} 更新解码器中的隐藏层状态 $d_t = f_2(d_{t-1}, \tilde{y}_{t-1})$, 此处的 f_2 为 LSTM。最终, 得到预测结果 \tilde{y}_T , 即:

$$\tilde{y}_T = F(y_1, \dots, y_{T-1}, x_1, \dots, x_T) \quad (12)$$

1.2.1 文本向量化

使用深度学习模型进行文本情感分析首先是将文本向量化, 将大量的文本向量嵌入来表示文本数据, 以此作为模型的输入。给定一个包含若干词组成的序列, 将每个词转换成一个实数向量。对于序列中的每个词, 寻找得到词的嵌入式矩阵 $W^{word} \in R^{d^w \times |V|}$, 该矩阵是一个需要学习的参数, 其中 V 是一个固定大小的词汇, d^w 是词嵌入的矩阵。通过使用矩阵向量乘积可以将一个词 x_i 转换成相应的词嵌入 e_i 。公式如下:

$$e_i = W^{word} v^i \quad (13)$$

其中, v^i 是 $|V|$ 大小中的一个向量, 该值只在

索引 e_i 处为 1, 其他位置均为 0。所以句子可以表示为向量 $emb_s = \{e_1, e_2, \dots, e_r\}$ 。

1.2.2 BiLSTM 模型

BiLSTM 模型表示双向 LSTM, 该模型通过前向和后向对序列进行分析处理, 更加有效地提取文本数据前后文的信息特征, 最终将双向 LSTM 的输出加到一起。研究推得的数学公式如下:

$$h_i = [\vec{h}_i \oplus \overleftarrow{h}_i] \quad (14)$$

1.2.3 注意力机制及情感分析

基于注意力机制的网络在很多任务中表现不错, 如知识问答、机器翻译等。假设 H 是由 LSTM 产生的向量集合, 即 $H = \{h_1, h_2, \dots, h_T\}$, 其中 T 是句子的长度。 V 是一个已经训练的向量, R 是句子的表示, 由输出向量加权和得到。因此可得到如下计算公式:

$$M = \tan(H) \quad (15)$$

$$\alpha = \text{softmax}(V^T M) \quad (16)$$

$$R = H\alpha^T \quad (17)$$

本次研究获取句子级别的特征用来进行分类, 对应的数学公式则可写作如下形式:

$$S^* = \tanh(R) \quad (18)$$

至此, 采用 softmax 分类器计算每个类别的概率, 并以概率最大的作为预测的结果。BiLSTM-Attention 模型如图 3 所示。对于一个句子 S , 将 S^* 作为输入, 最终预测的结果标签为 \hat{y} 。则计算过程为:

$$\hat{p}(y | S) = \text{softmax}(V^{(S)} S^* + b^{(S)}) \quad (19)$$

$$\hat{y} = \text{argmax}_y \hat{p}(y | S) \quad (20)$$

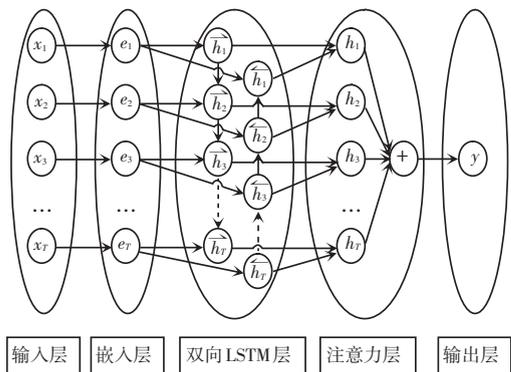


图 3 BiLSTM-Attention 模型

Fig. 3 BiLSTM-Attention model

2 实验过程

为验证本文提出的模型 M 的有效性, 首先获取

实验所需的相关数据并进行数据预处理, 通过设置相应的参数, 进行时序预测、文本情感分析等实验, 实验结果使用 $RMSE$ 等多种评价指标进行检验。

2.1 实验数据集

2.1.1 时序数据

时序数据主要通过 Python 第三方库 BaoStock 获取。BaoStock 是开源的证券数据服务平台, 其返回的数据格式为 DataFrame 类型。也可以将数据保存到本地文件后再进行分析。通过设置相关参数可得到相应时间内的所有股票交易数据, 如股票代码、查询的日期范围等。

本文采用的时序数据参数包含股票的开盘价、收盘价、最高价、最低价等。本文将第二天的收盘价作为预测值 y , 因为损失和收益一般是当天的收盘价计算得到的。实验的数据集为海康威视对应的股票信息。数据集包含了从 2010 年 8 月 30 日至 2021 年 8 月 19 日、总共 2 668 天的数据。时序数据样例见表 1。

表 1 时序数据样例

Tab. 1 Time series data samples

日期	开盘价	收盘价	最高价	最低价	预测值
2021-08-17	56.70	56.72	54.00	54.91	54.99
2021-08-18	54.51	55.50	53.32	54.99	54.91
2021-08-19	54.80	55.95	54.15	54.91	53.57

2.1.2 文本数据

论文采用的新闻文本数据来自于新浪财经网站, 该网站包含股票各方面的信息, 如股票实时涨跌、公司咨询、政策新闻等等。通过爬虫获取(描述新浪数据集具体信息)新闻标题、日期、涉及股票等。数据预处理过程包含清除中性标题数据、对标题文本进行人工标注, 标签包含 0 和 1。其中, 0 表示消极、1 表示积极。该数据集共包含 6 000 条, 这里, 表示积极的文本共 3 373 条, 表示消极的文本共 2 627 条。

2.2 评价指标

在时序预测实验以及最终结果实验中, 选用均方根误差 ($RMSE$)、平均绝对误差 (MAE) 作为评价标准。最终预测结果中选择的评价指标与时序预测实验所选的一致。公式如下:

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \quad (21)$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (22)$$

在文本情感分析实验中, 选择召回率 R 、准确率

P 以及综合准确率和召回率两者的 F_1 作为评判结果好坏的准则。数学定义的公式可分别表示为:

$$R = \frac{TP}{TP + FN} \quad (23)$$

$$P = \frac{TP}{TP + FP} \quad (24)$$

$$F_1 = \frac{2 * P * R}{P + R} \quad (25)$$

其中, TP 、 TN 、 FP 、 FN 含义参见表 2。

表 2 混淆矩阵

Tab. 2 Confusion matrix

	积极 (Positive)	消极 (Negative)
正确 (True)	TP	TN
错误 (False)	FP	FN

2.3 参数设置

在股票历史数据时序预测实验以及最终结果实验中, 参数设置见表 3。在文本情感分析实验中, 参数设置见表 4。

表 3 时序预测及最终结果实验参数

Tab. 3 Time series forecast and final result experimental parameters

参数	参数描述	时序预测实验及最终结果实验参数值
r	学习率	0.001
$batch_size$	数据批处理量	128
$Epoch$	周期数	200
train:test	训练集:测试集	3 : 1

表 4 文本情感分析实验参数

Tab. 4 Text sentiment analysis experimental parameters

参数	参数描述	文本情感分析实验参数值
r	学习率	0.001
$batch_size$	数据批处理量	50
$Epoch$	周期数	20
train:test	训练集:测试集	14 : 1

2.4 实验

2.4.1 时序预测实验

本文时序预测实验数据集为海康威视股票信息, 数据集时间跨度包含了从 2010 年 8 月 30 日至 2021 年 8 月 19 日、总共 2 668 天的数据。在本实验中, 2 001 条数据作为训练集, 667 条数据作为测试集。使用 XGBoost、LSTM 进行对比实验, 海康威视数据集实验结果见表 5。DA-RNN 在海康威视测试集上的预测值与真实值曲线如图 6 所示。

表 5 海康威视数据集实验结果

Tab. 5 Hikvision datasets experimental results

模型	平均绝对误差 (MAE)	均方根误差 (RMSE)
XGBoost	3.020	3.642
LSTM	2.470	2.967
DA-RNN	1.584	1.981

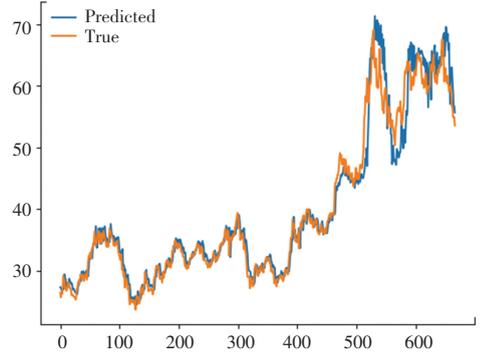


图 4 DA-RNN 在海康威视测试集上的预测值与真实值

Fig. 4 Predicted value and real value of DA-RNN on Hikvision test datasets

通过比较 XGBoost、LSTM 以及本文采用的 DA-RNN 模型的实验结果, 可以得知, DA-RNN 通过双阶段注意力机制, 既考虑输入阶段的信息特征, 也考虑长时间距离的信息, 相较于前 2 种模型表现更好, 在 MAE 、 $RMSE$ 上分别从 3.020、3.642 下降到 1.584、1.981。证明了 DA-RNN 模型在股价时序数据上的有效性。

2.4.2 文本情感分析实验

本文文本情感分析实验包含 6 000 条已经标注的新闻标题, 其中, 5 000 条标题文本数据作为训练集, 600 条作为验证集, 400 条作为测试集。为证明方法的有效性, 将本文中的 BiLSTM-Attention 模型与 SVM、LSTM 方法在相同实验环境下做对比实验, 结果见表 6。

表 6 文本情感分析结果

Tab. 6 Text sentiment analysis experimental results

模型	召回率/%	准确率/%	F_1 值/%
SVM	85.00	86.00	85.00
LSTM	80.00	81.00	80.00
BiLSTM-Attention	88.75	88.76	88.75

通过比较 SVM、LSTM、BiLSTM-Attention 算法的实验结果可知, 本文采用的 BiLSTM-Attention 通过双向的 LSTM, 同时考虑了上下文信息, 并且加入了注意力机制, 相较于单向的 LSTM 以及传统机器学习算法支持向量机 SVM 表现更优, 在准确率评价指标上分别提高了 7.76% 和 2.76%。BiLSTM-

Attention 方法在标题文本数据集上的 3 个指标中表现均是最好的,表明该模型有着较高的可用性。

2.5 模型 M 的实验结果

首先获取最终结果预测阶段相关数据集,本文选择 2021 年 8 月 20 日至 2021 年 9 月 22 日海康威视的时序数据集以及文本数据集。通过以上文本情感分析模型 BiLSTM-Attention 预测对应标题文本的情感倾向,对其添加权重,与 DA-RNN 模型预测的对应天数的值进行叠加,得到最终结果,见表 7。

表 7 最终实验结果

Tab. 7 Final experimental results

模型	权重(参数)	MAE	RMSE
DA-RNN	0	1.584	1.981
M	0.001	1.583	1.980
M	0.005	1.575	1.972
M	0.010	1.374	1.700
M	0.020	1.712	2.038
M	0.030	2.297	2.868

通过比较表 7 中的实验结果,可以得出,在本实验中,当权重值为 0.01 时,MAE、RMSE 值达到最小。本实验表明新闻文本数据对股票价格数值存在一定影响,当通过对其附加一定的权重时,可使股票价格预测值更接近真实值,说明本模型的可用性。

3 结束语

本文提出的模型 M,通过叠加影响股票价格预测的多种因素,如历史数据、新闻文本情感倾向等,采用对比实验等方式,得到最终实验结果。相比原始 DA-RNN 模型,在评价指标 RMSE、MAE 上分别下降了 13.3%、14.2%,证明了所提出模型 M 的有效性。

(上接第 122 页)

4 结束语

沙特 SWCC 输水管道项目 SCADA 系统的应用,实现了输水管线的自动化建设。该系统的应用,使得管线及沿线场站的工艺生产数据可以实时采集并发送给工作人员,对现场的电气仪表设备进行实时监控与控制,通过网络系统可以对远程场站进行分析、控制与管理。自动化是现代工业的重要标志,SWCC 项目的高度自动化技术应用,能够提高生产效率,降低人力成本,方便企业管理决策,在确保安全生产中起到了极其重要的作用^[9]。

参考文献

[1] 杨海光. SCADA 控制系统在城市供水调度系统的应用[J]. 微

然而,本文算法主要考虑了新闻文本情感倾向作为时序数据的影响因素,后期工作会考虑将其他因素融入研究,如股民对个股信息的情感倾向、同一行业类别股票的涨跌趋势等等。另外,本文只从 4 个维度的特征数据进行时间序列预测,后续工作可以添加其它维度的数据,如交易量等,通过更丰富的时序数据进行预测,进一步提高准确率。

参考文献

- [1] 张泽亚,黄丽明,陈翀,等. 基于新闻特征抽取和循环神经网络的股票预测方法[J]. 文献与数据学报,2020,2(01):45-56.
- [2] 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):209-212.
- [3] 陈卫华. 基于深度学习的上证综指波动率预测效果比较研究[J]. 统计与信息论坛,2018,33(05):99-106.
- [4] LI Xinyi, LI Yinchuan, YANG Hongyang, et al. DP-LSTM: Differential privacy - inspired LSTM for stock prediction using financial news[J]. arXiv preprint arXiv:1912.10806, 2019.
- [5] YANG Can, ZHAI Junjie, TAO Guihua. Deep learning for price movement prediction using convolutional neural network and long short-term memory[J]. Mathematical Problems in Engineering, 2020, 2020(6):1-13.
- [6] BOLLEN J, MAO Huina, ZENG Xiaojun. Twitter mood predicts the stock market[J]. Journal of Computational Science, 2011, 2(1):1-8.
- [7] SCHUMAKER R P, CHEN H. Textual analysis of stock market prediction using breaking financial news: The AZFin text system[J]. ACM Transactions on Information Systems (TOIS), 2009, 27(2):1-19.
- [8] QIN Y, SONG D, CHEN H, et al. A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction[J]. 2017.
- [9] 黄奇景,付伟. 基于 Attention-BiLSTM 的情感分类模型[J]. 电子技术与软件工程,2020(03):14-15.

计算机信息,2010,26(19):66-67.

- [2] 李峰,唐胜安. 分公司站场数据采集、传输与通信标准化[R]. 成都:中国石油西南油气田公司,2009.
- [3] 李旭. 浅谈 SCADA 系统在城市燃气门站输配中的应用[J]. 城市燃气,2012(10):20-22.
- [4] SCADA/RTU. System functional design specification[Z]. Iraq: TSC,2021.
- [5] 陈海霞,任庆海,冯俊杰. SCADA 系统在污水泵站中的应用[J]. 电子测试,2018(01):93-95.
- [6] 雷宁桥,周亮亮,苏健,等. 基于 SCADA 系统的管线监测应用与探索[J]. 化工设计通讯,2017,43(11):36.
- [7] 邓彬,张宗包,郝蛟,等. 电网运行 SCADA 数据质量管控平台的研发与应用[J]. 电信科学,2017,33(S1):156-161.
- [8] 宋春花. 总线控制系统在输油管线 SCADA 系统中的应用[J]. 石油化工自动化,2017,53(05):39-42.
- [9] 曹永乐. 基于西气东输管道 SCADA 系统的应用分析[J]. 自动化应用,2017(08):64-66.