

doi:10.16055/j.issn.1672-058X.2022.0002.011

# 基于 Attention 机制的 LSTM 股价预测模型

林 昕, 朱小栋

(上海理工大学 管理学院, 上海 200093)

**摘要:**针对时间序列分析方法和神经网络对于股价预测具有一定局限性的问题,将基于 Attention 机制的 LSTM 模型应用于股价预测;以 2014-01-02—2020-09-22 日的上证工业指数、上证环保指数等相关数据为样本,在 LSTM 模型中引入 Attention 机制,使模型聚焦于重要的股价特征信息,预测股票第二日的最高价;实证研究发现,相较于 MLP, RNN 和 LSTM 模型,基于 Attention 机制的 LSTM 模型的 RMSE 值比基准模型平均降低了 3%—45% 左右,4 种模型对环保企业的预测精度均高于污染企业;随后将空气质量指数,温度和湿度纳入特征,提升反映污染企业和环保企业股价规律相关特征的数据质量,实证结果发现在加入新的特征以后,4 种模型在预测上证工业指数和上证环保指数波动趋势时, RMSE 值均下降了 1% 左右。

**关键词:**空气质量;注意力机制;长短期记忆网络;股价预测

**中图分类号:** TP181

**文献标志码:** A

**文章编号:** 1672-058X(2022)02-0075-08

## 0 引言

股价波动可以反映市场和经济的变化情况,但由于股价的波动性,预测股价一直是研究的热点和难点之一。对股价预测,最为常见的就是时间序列分析方法,比如吴玉霞等<sup>[1]</sup>建立 ARIMA 模型对华泰证券收盘价变动的趋势和规律进行了预测;Lin Z<sup>[2]</sup>利用 GRACH 模型对股票波动率进行建模;张贵生等<sup>[3]</sup>在 ARMA-GRACH 模型中融入股价滞后值的近似微分信息,提高了预测准确度。但是这些方法存在局限性,要求数据满足的约束条件高,对于长期预测的效果不是太好,并且对于非线性非平稳序列,一般都需要经过对数化或者差分等前期处理。

现今随着机器学习的发展,越来越多的人将神

经网络应用于股票研究中。崔文喆等<sup>[4]</sup>采用 GARCH 模型和 BP 神经网络(Back Propagation)模型,利用上海 A 股 30 支股票分别进行短期、中期和长期预测,实验发现 BP 神经网络模型的每日与一周总体预测效果均优于 GARCH 模型,且总体差异具有高度统计学意义。Larry O O 等<sup>[5]</sup>对 34 个包括新兴市场和发达国家的股票数据,比较 BP 和浅层神经网络结构,发现股票数据集越大越复杂,BP 的表现越优于浅层神经网络结构。蔡方中等<sup>[6]</sup>在 BP 神经网络中引入主成分分析法和果蝇算法,提高了上证指数股价的预测效果。

BP 神经网络虽然相较时间序列分析方法有比较优异的性能,但学者们普遍认为 BP 神经网络可以实现的功能是有限的。因为股价受到多种因素影

收稿日期:2020-12-01;修回日期:2021-01-14.

基金项目:国家自然科学基金项目资助(71771152).

作者简介:林昕(1996—),女,安徽池州人,硕士研究生,从事深度学习和电子商务研究.

通讯作者:朱小栋(1981—),男,安徽太湖人,副教授,博士,从事深度学习和电子商务研究. Email:Zhuxd@ussA.edu.cn.

响,而这些因素之间彼此又会互相作用,信息长度各种各样且时时刻刻都在变化,故 BP 模型在学习复杂多变的股价波动规律时表现匮乏。学者们主要从两方面改进这一问题。

一方面是改进模型,长短期记忆网络是一种时间递归神经网络,是在循环神经网络(RNN, Recurrent Neural Network)的基础上发展来的。RNN 近年来文本翻译和序列预测问题中具有广泛应用,但其在训练网络过深时,容易产生梯度消失和梯度爆炸现象。LSTM 在解决这一问题的基础上,对长期信息记忆更加有效,故将它应用在时间序列预测上,预测准确度将会得到提高。彭燕等<sup>[7]</sup>搭建不同层次的 LSTM 模型,将股价预测准确度提高了约 30%。冯宇旭等<sup>[8]</sup>将 LSTM 用于沪深 300 指数的预测,相比 SVR 和 Adaboost 准确度更高。

但 LSTM 在学习输入序列的过程中,通常将其编码为固定长度的向量,当输入序列较长时,会限制模型的性能。Attention Mechanism(注意力机制)通过保留 LSTM 对输入序列的中间输出结果,然后训练一个模型对这些输入进行选择性的学习以解决这一问题。基于 Attention 机制的 LSTM 模型已经在多个领域被证实具有更优的性能,如 Youru L 等<sup>[9]</sup>在时间序列领域,将基于 Attention 机制的 LSTM 模型用于多个时间序列数据集做预测,实证结果证实模型比其他基准模型有更优的预测表现;胡荣磊等<sup>[10]</sup>在文本分类实验中,在 LSTM 和 CNN 的基础上加入了 Attention 机制,分类结果比未加入的有所提高;Dzmitry B<sup>[11]</sup>在机器翻译中引入注意力机制,以 LSTM 作为编码器和解码器,提升了机器翻译的准确度。因此本文通过将基于 Attention 机制的 LSTM 模型(AM-LSTM)用于预测股价的第二日最高价,并将其预测结果与 MLP(Multilayer Perceptron,多层感知机),RNN 和 LSTM 做出对比。

另一方面是选择更多与股价波动相关的特征或改进特征处理方式,以提升数据质量。在特征选择方面,近年来行为金融学的发展,学者们开始将投资者情绪加入股价数据来提升模型的预测性能。一些学者将人所处的环境,如光照、云层、温度等<sup>[12-13]</sup>纳

入影响因素进行考虑。其中部分国内外研究发现空气污染会通过引发投资者情绪进而影响股票市场的走势<sup>[14-16]</sup>,但总体上关于两者之间的研究并不多见。在改进特征处理方式方面,Cheong D 等<sup>[17]</sup>提出了一种基于 K-means 的投资组合优化方案,通过对投资者信息的聚类分析确定投资组合;陈佳等<sup>[18]</sup>在对纳斯达克和标普 500 指数预测完成特征选取后,分别应用系统聚类法进行特征分类、主成分分析法进行特征降维,提高了预测准确度和速度;李文星等<sup>[19]</sup>提出一种附有引力影响因子的半监督 K-means 核函数聚类算法,将其应用于多因子选股中,在选出较优的股票组合中具有优势。本文的目的在于预测上市污染企业和上市环保企业的股价波动趋势,故将与环境有关的指数加入预测模型,并利用 Attention 机制学习更重要的特征。

学者们将神经网络应用于股价波动趋势预测,取得了比时间序列方法更优的表现,但神经网络的固定编码向量长度会对其性能有所限制,故本文在神经网络中引入了注意力机制。现有文献探究了多种与股价波动相关的特征,拓宽了影响股价波动的研究思路,但分类研究不同类型股票的文献较少,本文对污染企业和环保企业的股价波动进行了分类研究。

在以往文献的基础上,主要做出了以下贡献:通过将基于 Attention 机制的 LSTM 模型用于股价预测,提升了预测性能;分别以上证工业指数作为上市污染企业的代表,上证环保指数作为上市环保企业的代表,分类研究不同类型股票价格预测的准确度;将空气质量指数、温度和湿度纳入模型特征,提高污染和环保企业股价的预测能力。

## 1 LSTM 模型设计

介绍了 LSTM 结构和 Attention Mechanism,并将 Attention 机制引入 LSTM 模型(AM-LSTM 模型),通过将基于 Attention 机制的 LSTM 模型用于股价预测,提升预测准确度。

### 1.1 LSTM

LSTM 是循环神经网络的一种变形,标准循环神经网络 RNN 相较于传统的 BP 神经网络,尽管可以利用时间序列的信息,但 RNN 在长距离传输下,容易发生梯度消失或梯度爆炸现象,导致反向传播过程中,神经网络无法训练权重,使得 RNN 难以学习长距离的信息。为解决这一问题, Hochreiter & Schmidhuber 于 1997 年提出了 LSTM<sup>[20]</sup>,通过增加了一个细胞单元和 3 种门机制,来控制信息流和避免长期依赖问题。

如图 1 所示,贯穿于顶部的水平线称为细胞状态,信息可以沿着这条线不变的流动,以解决对长距离信息学习不足的问题。3 种不同的门分别为输入

门,遗忘门和输出门。遗忘门决定信息的丢弃,输入门决定信息的更新,输出门决定细胞状态的信息输出。通过门结构来保存重要的信息,遗忘不必要的信息,以提高对长时间序列的记忆。LSTM 各个门的计算如式(1)一式(5)所示:

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

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

$$o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o) \quad (3)$$

$$C_t = f_t \otimes C_{t-1} + i_t \otimes (W_c * [h_{t-1}, x_t] + b_g) \quad (4)$$

$$h_t = o_t \otimes \tanh(C_t) \quad (5)$$

其中  $\sigma$  是 sigmoid 激活函数,  $i_t, f_t, o_t, h_t$  分别为  $t$  时刻下各个门的状态,细胞状态和输出。

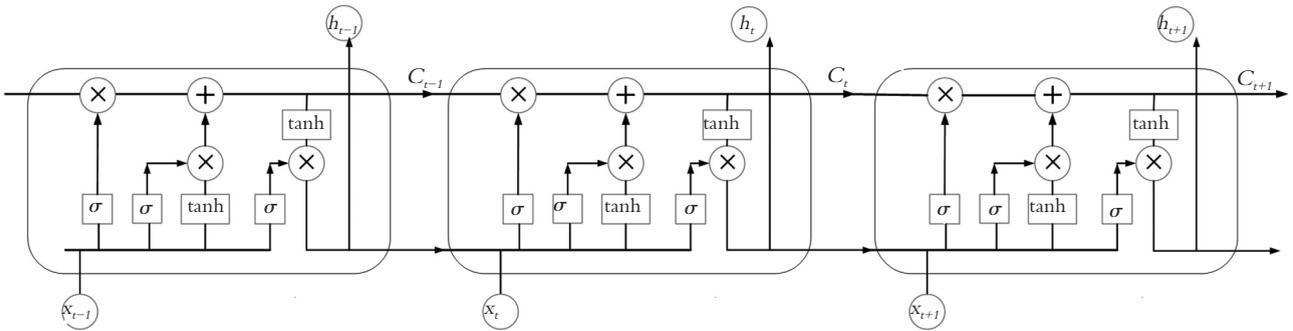


图 1 LSTM 结构图

Fig. 1 The structure of LSTM

### 1.2 Attention 机制

股票价格随着市场因素、环境因素等的变化而不断波动,不同时间点的因素变动对当前股价的影响程度也不同。距离当前时间点更近的信息,由于时间间隔短,对股价波动的影响越大;随着距离当前时间点的间隔越远,这种影响逐渐减弱。

Attention 机制可以很好地捕捉到股价变动的这一特征,它改变了传统的 encoder-decoder 结构,传统的解码器对每一个输入都赋予相同的权重,但现实中不同的输入重要性往往不同。Attention 通过在解码过程中,利用评分函数计算不同输入对预测值的影响程度,为其赋予不同的权重来解决这一问题。其本质通过计算当前的输入序列和输出向量的差异,差异越小就说明注意力应该在此处应该赋予更

大的权重。Attention 机制如图 2 所示。

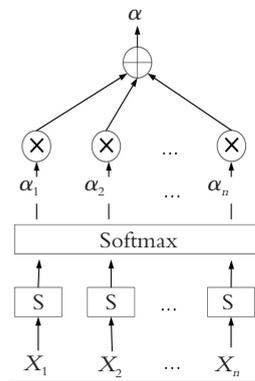


图 2 Attention 结构图

Fig. 2 The Structure of Attention

### 1.3 AM-LSTM 模型

通过输入历史时间序列数据  $(X_1, X_2, \dots, X_n)$  学习特征,对第  $n+k$  天的最高股价进行拟合。预测

第二日的最高股价,即设置  $k=1$ 。图 3 展示了 AM-LSTM 模型对股价预测的总体结构,共包括 4 个部分:数据预处理、LSTM 层、Attention 层和全连接层。

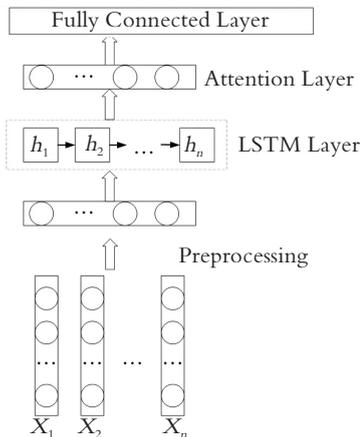


图 3 AM-LSTM 模型结构图

Fig. 3 The structure of AM-LSTM

### 1.3.1 数据预处理

预处理后的数据存储在 CSV 文件中,由于样本数据各自的量纲不同,为了提高实验的可信度,需要对数据进行归一化处理,归一化公式如式(6)所示:

$$X_{\text{new}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (6)$$

其中,  $X_{\text{new}}$  是归一化后的数据,  $X$  是原始数据,  $X_{\min}$ ,  $X_{\max}$  分别是原始数据的最小值和最大值。

### 1.3.2 LSTM 层

LSTM 在时间序列数据的分类和预测中都有很好的表现,故用于学习预处理后数据的深层特征。通过 LSTM 层的训练,可以更充分地学习时间序列数据的长期关系。LSTM 当前的隐藏层状态由当前输入  $X_t$  和  $(t-1)$  时刻的隐藏层状态  $H_{t-1}$  共同决定。输入第  $j$  时刻的序列信息  $X_j$ , LSTM 层的输出结果如式(7)所示:

$$H_j = \text{LSTM}(X_j, H_{j-1}) \quad (7)$$

### 1.3.3 Attention 层

Attention 层输入为上一层 LSTM 的输出,首先通过编码器(Encoder)进行编码,然后通过 Attention 对编码向量计算对应的权值  $(\alpha_1, \alpha_2, \dots, \alpha_N)$ ,用计算的权值对编码向量进行加权作为解码器(Decoder)的输入,最后经过解码器得到输出。通过不断学习优化相应的权值,突出输入特征中的关键信息,进一步深度挖掘变量之间的非线性特征。Attention 层相似度的计算如式(8)所示,分数越高说明关注度越高,注意力机制就会给其分配更大的权重:

$$u_j = \tanh(W_i H_j + b_i) \quad (8)$$

其中  $W_i$  是权重系数,  $b_i$  是偏置量,将  $u_j$  输入 softmax 函数中,进行归一化,如式(9)所示,得到归一化后的权重  $\alpha_j$ :

$$\alpha_j = \frac{\exp(u_j^T \mathbf{u}_w)}{\sum_{j=1}^T \exp(u_j^T \mathbf{u}_w)} \quad (9)$$

其中,  $\mathbf{u}_w$  是随机初始化的注意力权重矩阵,Attention 机制的  $S$  是不同权重与对应各个隐藏层状态的乘积的累加和,计算如式(10)所示:

$$S = \sum_{j=1}^T \alpha_j H_j \quad (10)$$

最终得到了各个变量被赋予了不同关注度的 Attention 层的输出,如式(11)所示:

$$F_j = H(S, H_j) \quad (11)$$

### 1.3.4 全连接层

全连接层的输入为 Attention 层的输出。全连接层对该层的输入进行计算,得到股价的预测结果。与真实的股价值相减,返回损失函数。通过损失函数的反向传播,不断地修正权重和预测结果,如图 3 所示。

## 2 数据描述和评估指标

### 2.1 股价波动数据

受文献[15]的启示,选取上证工业指数代表污染企业的发展,上证环保指数代表环保企业的发展。由于自 2014 年以来,雾霾问题的爆发以及人们对环境问题关注度的上升,故选取两种指数从 2014-01-02—2020-09-22 日的每日数据,共 1 642 笔数据,数据均来自于 Wind。涉及的基本特征包括开盘价、最高价、收盘价、最低价、成交量、成交额这 6 个变量,以此数据进行训练与测试,通过前一天的信息预测第二日的最高价。

### 2.2 空气质量数据

He 等<sup>[22]</sup>认为上海是中国的金融中心,拥有中国最大的股票交易量,其天气情况对中国股票市场的影响最显著;在 Levy 等<sup>[14]</sup>的论文中,其实验数据也证实了 He 等<sup>[22]</sup>所提出的观点。故选取上海市的空气质量数据建立预测模型。AQI 是世界范围通用的衡量空气质量的综合指标,囊括了 PM2.5, CO, SO<sub>2</sub> 等污染物的监测,通常被划分为 0~50, 51~100, 101~150, 151~200, 201~300 以及大于 300 六

个阶段,分别代表空气质量优、良、轻度污染、中度污染、重度污染和严重污染。故选取上海市的 AQI,温度和湿度数据作为空气质量的衡量指标。与股价数据相对应,AQI,温度和湿度数据也选取了 2014-01-01—2020-09-22 日的日度最大值,并剔除了周末及节假日等非交易日数据,以使空气质量数据周期与股价波动数据周期保持一致。相关数据来源于中国气象局。

### 2.3 评估指标

为保证对比实验的一致性,在实验过程中,所有的对比模型均将数据按 85%:15% 划分为训练集和测试集。1 640 个交易日的数据,共划分训练集 1 394 笔数据,测试集 246 笔数据。并采用均方根误差(Root Mean Square Error, RMSE)用作模型的评价指标, RMSE 值( $f_{\text{RMSE}}$ )越小,表明模型预测越准确。RMSE 计算公式如式(12)所示:

$$f_{\text{RMSE}} = \sqrt{\frac{\sum_{i=1}^n (y_{\text{pre}} - y_{\text{act}})^2}{n}} \quad (12)$$

其中,  $y_{\text{pre}}$  是预测值,  $y_{\text{act}}$  是实际值。

## 3 实验结果与分析

### 3.1 AM-LSTM 模型验证

为了验证 AM-LSTM 模型的性能,将 AM-LSTM 模型与 3 种基准模型进行了比较。参数设置如表 1 所示。

表 1 模型参数设置

Table 1 The setting of model parameters

模 型	主要参数	取 值
MLP	学习率	0.01
	优化函数	Adam
	激活函数	ReLU
	隐藏层神经元数量	5
	学习率	0.01
RNN	优化函数	Adam
	激活函数	ReLU、Linear
	隐藏层神经元数量	5
LSTM/ AM-LSTM	学习率	0.01
	优化函数	Adam
AM-LSTM	激活函数	ReLU、Linear
	隐藏层神经元数量	5
	时间步长	1

实验过程中,为了保证实验的准确和客观,将每种模型在同一训练集和测试集上均进行了 10 次实验,取 RMSE 的平均值作为模型最终结果。实验结果如表 2 所示。

表 2 不同模型结果对比

Table 2 The comparison of the results of different models

	上证工业 RMSE	上证环保 RMSE
MLP	24.07	21.10
RNN	17.89	15.03
LSTM	13.61	12.26
AM-LSTM	13.18	11.43

从表 2 可见,神经网络的表现均优于前馈神经网络,主要原因是其能够更好地捕捉时间序列特征。LSTM 和 AM-LSTM 的 RMSE 值低于 RNN,其原因是 LSTM 在学习远距离信息时表现更优。基于 AM-LSTM 模型的 RMSE 值在上证工业指数和上证环保指数的预测中均低于其余基准模型,其原因主要为该模型集成了 LSTM 记忆机制在处理时序模型方面以及 Attention 在赋予特征权重方面的优势。此外,4 种模型在测试集上对上证环保指数的预测误差均小于上证工业指数,说明模型对上证环保指数的特征具有更好的学习能力。

此外,为了进一步探讨 AM-LSTM 相较于其他 3 种模型对于预测误差的降低程度,基于测试集的结果提出了相对误差,计算如式(13)所示,式(13)中  $i$  为第  $i$  种股价指数。表 3 为相对误差的结果。由表 3 可见,AM-LSTM 相较于 MLP,对两种指数预测准确度的提升幅度均超过了 45%;相较于 RNN,也均超过了 20%;对比 LSTM 也有小幅的提升,进一步说明本文提出的模型在预测多变量时间序列问题中更具有针对性。

$$\text{Relative RMSE}_i = \left| \frac{\text{RMSE}_{\text{AM-LSTM},i} - \text{RMSE}_{\text{Other-model},i}}{\text{RMSE}_{\text{Other-model},i}} \right| \quad (13)$$

表 3 模型相对误差

Table 3 The relative errors of different models

	上证工业/%	上证环保/%
AM-LSTM 相对 MLP	45.24	45.83
AM-LSTM 相对 RNN	26.33	23.95
AM-LSTM 相对 LSTM	3.16	6.77

### 3.2 样本纳入空气质量数据验证

图 4 和图 5 为 4 种模型加入 AQI、温度和湿度特征前后,上证环保指数和上证工业指数测试集的 RMSE 值对比结果。可以看出,在纳入 AQI、温度和湿度特征后,4 种模型的 RMSE 值均有不同程度的降低。

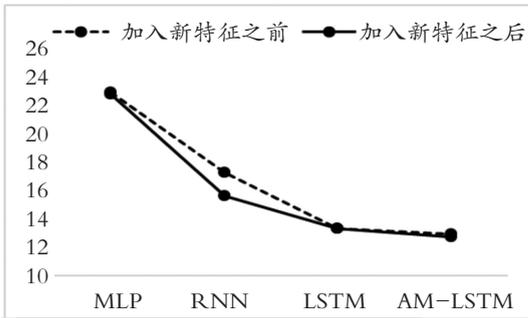


图 4 上证工业指数对比图

Fig. 4 The comparison of Shanghai Stock Industry Index

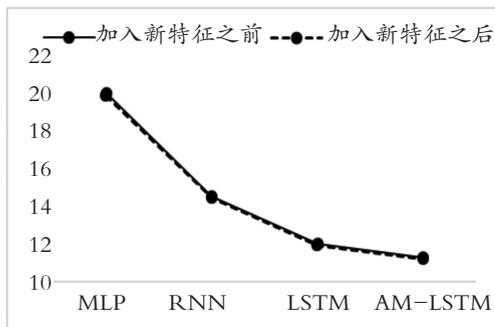


图 5 上证环保指数对比图

Fig. 5 The comparison of Shanghai Environmental Protection Index

在上证工业指数中,LSTM 和 AM-LSTM 的 RMSE 值约下降了 0.7%,RNN 的 RMSE 值下降幅度略高于前两者,为 1%。这可能是由于 LSTM 和 AM-LSTM 在未纳入该特征前,相对于 RNN 已经学习更多的信息,表现更优,所以在加入特征后, RMSE 值的下降幅度小于 RNN。

在上证环保指数中,4 个模型的 RMSE 值表现出了与上证工业指数相似的变化趋势。LSTM 和 AM-LSTM 的 RMSE 值的下降幅度接近 1%,RNN 的 RMSE 值下降幅度高于前两者,略高于 1%。

实验结果表明:将空气质量有关数据纳入模型特征,对提升两种股价的预测准确度有一定作用。这一结果也是符合投资者心理预期的,上证环保指数代表了环保企业的表现,空气质量的优劣一定程度上会向投资者释放这些企业经营状况的信号;而

工业行业的生产均会产生被纳入空气质量指标的有害气体;即这些指数对空气质量的变化较敏感,因此空气质量的变化也会传递能被模型学习的特征。

## 4 结 论

以上证工业指数代表污染企业,上证环保指数代表环保企业,通过将基于 Attention 机制的 LSTM 用于预测股价的第二日最高价。并将 AQI、温度和湿度作为新的特征纳入模型,得出结论:

(1) 通过对比 MLP、RNN、LSTM 和 AM-LSTM 模型,AM-LSTM 的预测准确度最高,表明其在处理多维非线性问题时,具有更好的预测性能。且说明不同变量对股价预测的重要性不同,未来可以进一步探讨哪些变量具有更高价值的信息。

(2) 将 AQI、温度和湿度纳入模型特征,模型的 RMSE 值进一步降低,表明空气质量的变化能够传达出上市污染企业和上市环保企业经营的信号。其中对环保指数预测准确度高于工业指数。

(3) 虽然将 Attention 机制引入 LSTM,模型的预测性能进一步提升,但仅仅增加了 AQI、温度和湿度进行实验,在股价受到多种因素影响的情况下,准确度提高的并不显著,未来可以考虑尽可能地将其其他多种因素纳入模型中,并结合聚类、分类等多种特征处理的方法,进行更深入的探讨。

### 参考文献 (References):

- [1] 吴玉霞,温欣. 基于 ARIMA 模型的短期股票价格预测 [J]. 统计与决策,2016(23):83—86.  
WU Yu-xia, WEN Xin. Short-term stock price prediction based on ARIMA model [J]. Statistics and Decision, 2016(23):83—86.
- [2] LIN Z. Modelling and forecasting the stock market volatility of SSE composite index using GRACH models [J]. Future Generation Computer Systems, 2018(79): 960—972.
- [3] 张贵生,张信东. 基于微分信息的 ARMAD-GARCH 股价预测模型 [J]. 系统工程理论与实践,2016,36(5): 1136—1145.  
ZHANG Gui-sheng, ZHANG Xin-dong. A differential-information based ARMAD-GRACH stock price

- forecasting model [J]. *Systems Engineering Theory & Practice*, 2016, 36(5): 1136—1145.
- [4] 崔文喆,李宝毅,于德胜. 基于 GARCH 模型和 BP 神经网络模型的股票价格预测实证分析[J]. *天津师范大学学报(自然科学版)*, 2019, 39(5): 30—34.
- CUI Wen-zhe, LI Bao-yi, YU De-sheng. Empirical analysis of stock price forecast based on GARCH model and BP neural network model [J]. *Journal of Tianjin Normal University (Natural Science Edition)*, 2019, 39(5): 30—34.
- [5] LARRY O O, MING C S, TIE J M, et al. Comparing the effectiveness of deep feedforward neural networks and shallow architectures for predicting stock price indices [J]. *Expert Systems with Applications*, 2019, 139(17): 449—465.
- [6] 綦方中,林少倩,俞婷婷. 基于 PCA 和 IFOA-BP 神经网络的股价预测模型[J]. *计算机应用与软件*, 2020, 37(1): 116—121.
- QI Fang-zhong, LIN Shao-qian, YU Ting-ting. A prediction model of stock price based on PCA and IFOA-BP neural network [J]. *Computer Applications and Software*, 2020, 37(1): 116—121.
- [7] 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.
- PENG Yan, LIU Yu-hong, ZHANG Rong-fen. Modeling and analysis of stock price forecast based on LSTM [J]. *Computer Engineering and Application*, 2019, 55(11): 449—465.
- [8] 冯宇旭,李裕梅. 基于 LSTM 神经网络的沪深 300 指数预测模型研究[J]. *数学的实践与认识*, 2019, 49(7): 308—315.
- FENG Yu-xu, LI Yu-mei. A Research on the CSI 300 index prediction model based on LSTM neural network [J]. *Mathematics in Practice and Theory*, 2019, 49(7): 308—315.
- [9] YOURU L, ZHU Z F, et al. EA-LSTM: evolutionary attention-based LSTM for time series prediction [J]. *Knowledge-Based Systems*, 2019 (181): 104785. 1—104785. 8.
- [10] 胡荣磊,芮璐,齐筱,等. 基于循环神经网络和注意力模型的文本情感分析[J]. *计算机应用研究*, 2019, 36(11): 3282—3285.
- HU Rong-lei, RUI Lu, QI Xiao, et al. Text sentiment analysis based on recurrent neural networks and attention model [J]. *Application Research of Computers*, 2019, 36(11): 3282—3285.
- [11] BAHDANAU D. Neural machine translation by jointly learning to align and translate [J]. *Computer Science*, 2014.
- [12] SAUNDERS E M J. Stock prices and the Wall Street weather [J]. *American Economic Review*, 1993, 83(3): 1337—1345.
- [13] APERGIS N, GUPTA R. Can (Unusual) weather conditions in New York predict south african stock returns [J]. *Research in International Business & Finance*, 2017(41).
- [14] LEVY T, YAG I L J. Air pollution and stock returns in the US [J]. *Journal of Economic Psychology*, 2011, 32(3): 374—383.
- [15] 郭永济,张谊浩. 空气质量会影响股票市场[J]. *金融研究*, 2016(2): 71—85.
- GUO Yong-ji, ZHANG Yi-hao. Can air quality affect the stock market [J]. *Financial Research*, 2016(2): 71—85.
- [16] 邓晓,张晗. 空气质量对股票市场影响的实证分析[J]. *统计与决策*, 2019, 35(1): 163—165.
- DENG Xiao, ZHANG Han. Empirical analysis of effect of air quality on stock market [J]. *Statistics and Decision*, 2019, 35(1): 163—165.
- [17] CHEONG D, KIM Y M, BYUN H W, et al. Using genetic algorithm to support clustering-based portfolio optimization by investor information [J]. *Applied Soft Computing*, 2017(61): 593—602.
- [18] 陈佳,刘冬雪,武大硕. 基于特征选取与 LSTM 模型的股指预测方法研究[J]. *计算机工程与应用*, 2019, 55(6): 108—112.
- CHEN Jia, LIU Dong-xue, WU Da-shuo. Stock index forecasting method based on feature selection and LSTM model [J]. *Computer Engineering and Applications*, 2019, 55(6): 108—112.
- [19] 李文星,李俊琪. 基于多因子选股的半监督核聚类算法改进研究[J]. *统计与信息论坛*, 2018, 33(3):

30—36.

LI Wen-xing, LI Jun-qi. Improvement of semi-supervised kernel clustering algorithm based on multi-factor stock selection[J]. Statistics & Information Forum, 2018, 33(3):30—36.

- [20] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 122(9):1735—1780.

- [21] 宋志圣. 空气污染对中国股市的影响分析[J]. 金融发展研究, 2018(9):80—85.

SONG Zhi-sheng. Analysis of the impact of air pollution on China's stock market [J]. Financial Development Research, 2018(9):80—85.

- [22] HE X B, LIU Y. The public environmental awareness and the air pollution effect in Chinese stock market[J]. Journal of Cleaner Production, 2018(185):446—454.

## Attention-mechanism-based LSTM Model for Stock Price Predicting

**LIN Xin, ZHU Xiao-dong**

(Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

**Abstract:** Time series analysis methods and neural networks have certain limitations for stock price prediction. Therefore, Attention mechanism was introduced into the LSTM model to predict the stock price. The related data of Shanghai Stock Industry Index and Shanghai Environmental Protection Index from January 2, 2014 to September 22, 2020 was collected as samples, and based on these samples, the LSTM combined with Attention mechanism was used to focus on important stock price characteristic information and predict the highest stock price of the second day. The empirical results found that compared with MLP, RNN and LSTM models, the RMSE value of the LSTM model combined with the Attention mechanism was lower than the benchmark model by about 3%-45% on average. The four models all had higher prediction accuracy for environmental protection companies than polluting companies. Subsequently, the air quality index, temperature and humidity were included into the characteristics to improve the quality of the data that reflects the characteristics of polluting companies and environmental protection companies. The empirical results found that after adding new features, the RMSE values of the four models all dropped by about 1% when predicting the fluctuating trend of the Shanghai Industrial Index and the Shanghai Environmental Protection Index.

**Key words:** air quality; attention mechanism; long short-term memory; stock price prediction

责任编辑:田 静

引用本文/Cite this paper:

林昕,朱小栋. 基于 Attention 机制的 LSTM 股价预测模型[J]. 重庆工商大学学报(自然科学版), 2022, 39(2):75—82.

LIN Xin, ZHU Xiao-dong. Attention-mechanism-based LSTM model for stock price predicting [J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2022, 39(2):75—82.