

# 基于CNN-LSTM混合模型的股票价格预测

■ 邵景昊

(南京信息工程大学, 江苏 南京, 210044)

## 一、引言

股票价格未来趋势是投资者关注的焦点之一。近年来,随着计算机技术与人工智能的快速发展,深度学习算法在金融领域得到了广泛应用。本文提出了一种基于卷积神经网络(CNN)与长短期记忆神经网络(LSTM)的混合预测模型(CNN-LSTM),并选取贵州茅台这一股票进行短期股价预测。本文以循环神经网络(RNN)等模型作为对照组,通过实测值与预测值的差距评价模型效果。实验结果表明,本文构建的CNN-LSTM混合模型效果处于较优水准,其平均百分比误差(MAPE)仅为0.0342,显著低于RNN等神经网络模型。

目前,股票市场投资者们最关心的问题之一就是股票价格的未来趋势,这有助于他们做出正确的投资决策并缓释风险。然而,因为金融时间序列的非平稳、非线性和高波动性等特征,所以想要实现精准的股票市场价格预测是具有挑战性的。尽管如此,仍有不少学者在尝试研究股票价格预测的方法。传统时间序列分析常被应用于股票价格预测中,例如,刘红梅使用了ARIMA模型对股票价格进行预测。但是,这些分析方法是建立在一系列假设之上,包括数据的平稳性、自相关性以及时序数据的分布特征等。这些限制条件使得传统时间序列模型几乎不可能准确预测股票的未来价格。对此,随着计算机技术的发展,有不少学者为解决这一问题,使用了机器学习等方法。黄同愿与陈芳芳采用支持向量机(SVM)对股票未来的走势进行了预测。为了进一步改善预测效果,还有学者将智能算法与机器学习算法结合起来。孙秋韵等人使用遗传算法(GA)优化了支持向量机模型并得到了更好的预测结果。

近年来,深度学习与神经网络在各个行业得到了广泛应用。蔡红与陈荣耀应用PCA-BP神经网络对股票价格预测进行了研究。彭燕等人则应用长短期记忆神经网络(LSTM)对股票价格预测进行了研究,发现相较于BP神经网络和RNN循环神经网络,LSTM有着更好的表现。LSTM是可以充分反映输入时间序列数据中的长期历史过程的神经

网络,因此,LSTM在预测股票未来趋势时有较好的表现。然而,单个LSTM网络难以发现非连续数据间的潜在联系。

为了更有效地预测股票未来趋势,本文提出了一种基于卷积神经网络(CNN)与长短期记忆神经网络(LSTM)的混合模型(CNN-LSTM)。与单个LSTM网络相比,CNN-LSTM混合模型利用CNN特征提取的能力,更易于提取时序数据的隐藏特征。与此同时,因为LSTM相较于其他神经网络,如RNN、BP等,在处理时序数据时表现突出,所以混合模型CNN-LSTM将最大限度地提高预测精度。

## 二、理论基础

本文主要介绍了深度学习算法在股票价格预测上的应用。为进一步地了解本文提出的模型,本部分主要介绍CNN、LSTM和CNN-LSTM模型。

### (一) CNN

卷积神经网络(CNN)在图像处理、语音识别和语义分析等领域得到了广泛的运用。CNN是一种前馈式神经网络,具有多层架构,包括卷积层、池化层和全连接层等。其中,卷积层是卷积神经网络的核心。卷积核可以高效实现对输入数据的特征提取。此外,卷积层的权值在神经元之间共享,卷积层的正向传播公式如下:

$$Y = f(W \times X + b)$$

其中,Y为卷积运算后的输出值,W为权重矩阵,b为偏置向量,f为卷积层的激活函数,通常,激活函数采用修正线性单元(ReLu)函数。ReLu激活函数可以表示如下:

$$f(x) = \max(0, x)$$

池化层的作用是去除数据中的冗余特征,只保留最具代表性的特征,从而获得更概括的表达。因此,池化层能有效降低数据的维度,在减少模型的训练时间的同时防止模型过拟合。全连接层的作用是对CNN网络中前面几层提取到的特征进行非线性组合。

### (二) LSTM

长短期记忆神经网络(LSTM)是循环神经网络RNN的变种,在不同领域的时间序列数据研究

中表现突出。LSTM 拥有多个循环单元，包括输入门、输出门和遗忘门。由于其特殊的门状结构，LSTM 可以解决 RNN 网络中存在的梯度消失和梯度爆炸的问题，并更好地处理较长的时序数据。LSTM 的结构如图 1 所示。

其中， $f_t$ 、 $i_t$ 、 $o_t$  分别代表遗忘门、输入门和输出门。 $x_t$  是当前 LSTM 单元的输入， $h_{t-1}$  是前一单元的输出， $C_{t-1}$  是前一单元的状态。

LSTM 的第一个门为遗忘门  $f_t$ ，遗忘门用来控制前一单元的单元状态  $C_{t-1}$  被遗忘的比例，其输入值包括前一单元的输出  $h_{t-1}$  和当前单元的输入值  $x_t$ 。计算公式为：

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

其中， $\sigma$  为激活函数 sigmoid。这个激活函数使得输出值的范围为 (0,1)。

输入门  $i_t$  的输入值包括上一单元的输出值  $h_{t-1}$  和当前时刻的输入值  $x_t$ ，经过计算得到输入门的输出值和候选单元状态，如下式所示：

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

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

其中， $C_t$  为候选单元状态，输入门  $i_t$  决定了候选单元状态所保留的比例，将其与前一时刻保留的信息相加，可得到更新后的单元状态信息，如下式所示：

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t$$

输出门  $o_t$  的输入值包括上一单元的输出值  $h_{t-1}$  和当前时刻的输入值  $x_t$ ，它能决定更新后的单元状态信息传递到下一单元的比例，从而得到当前单元的最终输出，如下式所示：

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

$$h_t = o_t \cdot \tanh(C_t)$$

### (三) CNN-LSTM 混合模型

在实际情况下，由于宏观因素的影响，股票价格时序数据呈现出非平稳、低信噪比的特点，单一模型难以准确提取其数据特征，预测精度会大大降低。因此，考虑到卷积神经网络提取局部重要信息特征的能力和 LSTM 在处理时序数据方面的优势，本文采用混合模型 CNN-LSTM 对股票价格进行预测。CNN-LSTM 的网络结构如图 2 所示。

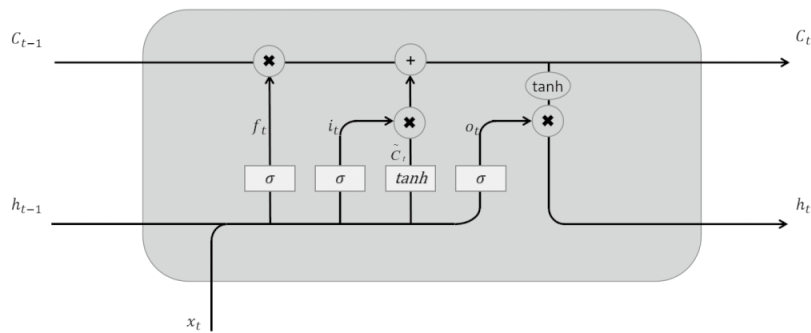


图 1 LSTM 结构图

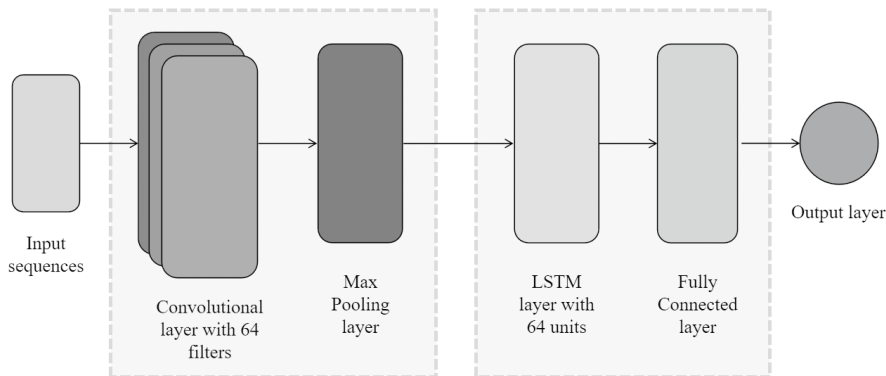


图 2 混合模型 CNN-LSTM 结构图

### 三、实证分析

为了证明 CNN-LSTM 混合模型预测的准确性，本部分主要介绍 CNN-LSTM 的预测结果以及与其他方法的对比。本次实验通过 python 3.6 完成，由深度学习工具 keras 搭建模型。

#### (一) 数据来源

本次实验主要选取了贵州茅台这一股票上市以来的日收盘价数据(数据来源为同花顺 APP)。数据集包括了从 2001 年 8 月 27 日至 2021 年 12 月 3 日共 4845 个交易日的数据。其中，前 80% 的数据集作为训练集，后 20% 的数据集作为测试集。其价格数据如图 3 所示。

#### (二) 模型评价

本次实验选取了均方根误差 (RMSE)、平均百分比误差 (MAPE) 和平均绝对误差 (MAE) 这三个评价指标来评价模型的预测精度。这三个指标越小，说明预测误差越小。具体公式如下：

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

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right|$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|$$

公式中，n 为预测值个数， $\hat{y}_i$  为预测值， $y_i$  为真实值。

#### (三) 参数设定

因为股票价格波动较大，所以原时间序列数据之间存在较大的差距。为了避免极端值对预测结果的影响，在数据输入 CNN-LSTM 模型之前需要

对原时间序列数据进行归一化处理。本文采用了极差变化，具体公式为：

$$x_{norm} = \frac{x_i - x_{min}}{x_{max} - x_{min}}$$

其中， $x_{norm}$  为归一化后的数据， $x_i$  为原始数据， $x_{max}$  为原时间序列的最大值， $x_{min}$  为原时间序列的最小值。

在对数据进行归一化之后，将数据输入到 CNN-LSTM 网络中进行训练。其中，CNN 层卷积核大小设置为 3，卷积核数设置为 64，为了避免梯度消失，CNN 层的激活函数设置为 Relu。LSTM 层隐藏层设置为 1 层，隐藏神经元数设置为 64。LSTM 层激活函数设置为 tanh。本次实验采用滑动窗口法，时间步长为 5，即使用连续五个交易日的价格预测下一交易日的价格。具体参数设定如表 1 所示。

表1 参数设定

参数	值
卷积核大小	3
卷积核数	64
卷积层激活函数	Relu
LSTM 层隐藏单元数	64
LSTM 层激活函数	tanh
优化器	Adam
学习率	0.00001
迭代次数	250

### 四、结果分析

为了验证本文提出的混合模型在预测股票价格数据方面的准确性，本部分使用 RNN、CNN-

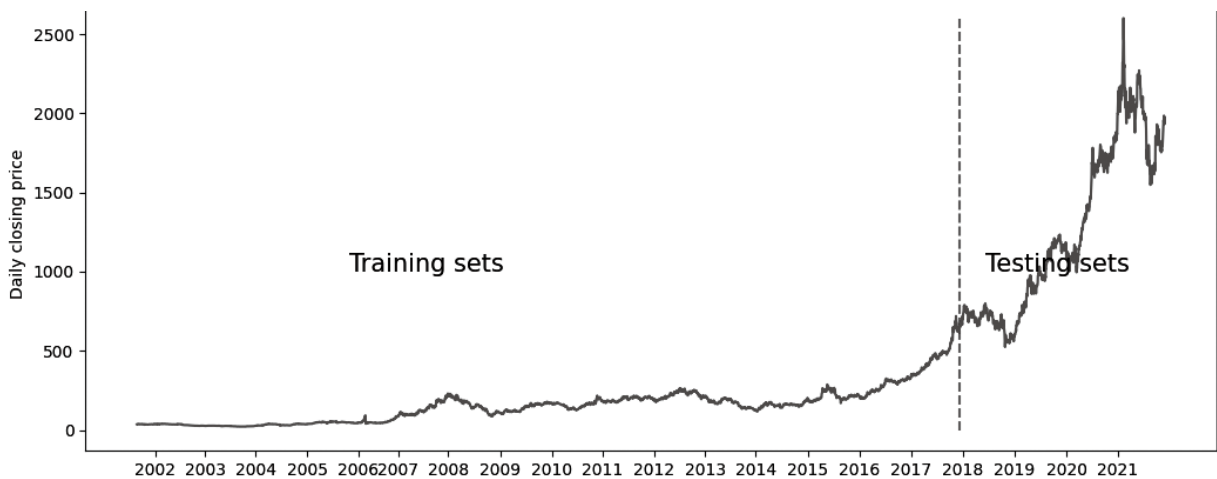


图 3 股票的日价格序列图

RNN、LSTM 与本文提出的 CNN-LSTM 模型进行对比。所有模型均采用相同的运行环境和数据集。

(一) 实验结果对比

为直观地了解模型的预测结果，图 4—7 展示了各种模型的预测结果与实测值的比较。

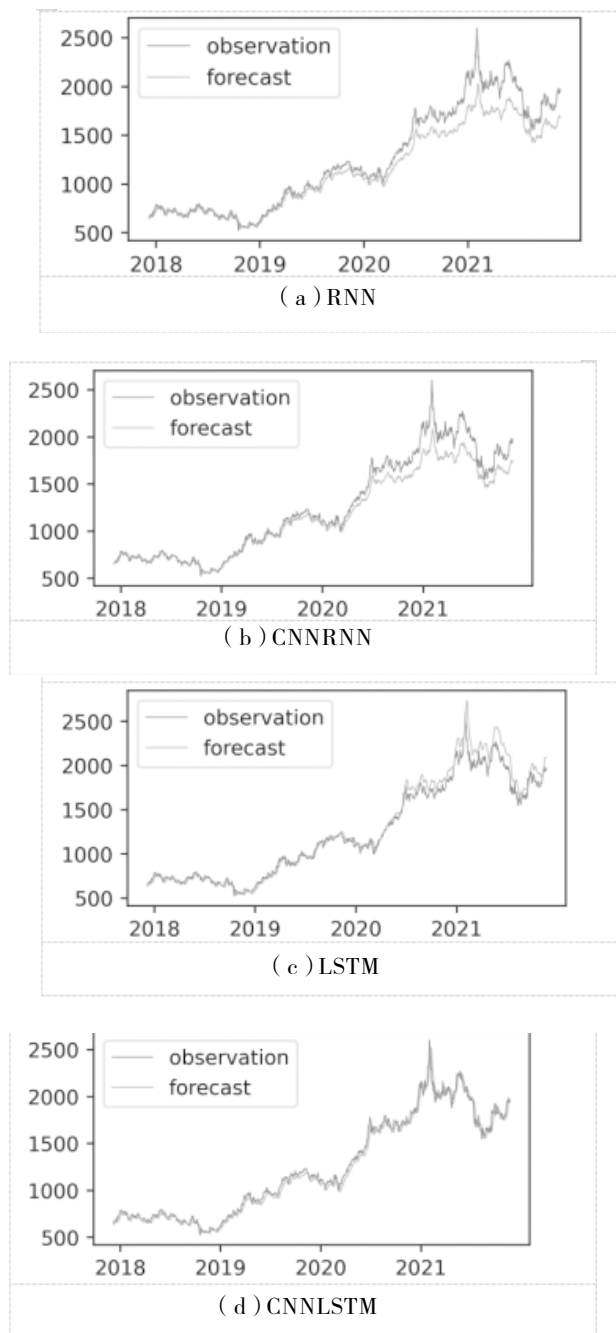


图 4 各种模型预测值与实测值比较

由图 4 可知，本文提出的 CNN-LSTM 模型在预测股票价格上的表现优于单个 LSTM 模型，并且优于 CNN-RNN 混合模型。其中，在疫情期间，股票价格波动较大，RNN 与 CNN-RNN 模型的预测精度显著低于本文所提出的模型。为了更进一步地了解各个模型的预测精度，表 3 展示了不同模型的评价指标的对比。其中，RMSE、MAPE、MAE 的数值越小，模型预测精度就越高。

表2 不同模型评价指标的对比

模型	RMSE	MAPE	MAE
RNN	159.1	0.0718	112.7
CNN-RNN	128.9	0.0548	88.3
LSTM	93.9	0.0429	61.8
CNNL-STM	54.5	0.0342	41.4

由表 2 可知，LSTM 模型的三个评价指标均低于 RNN，因此，在处理本次实验所选取的股票价格数据时，LSTM 更具有优势。与此同时，CNN-LSTM 混合模型的表现最佳。相较于单个模型 LSTM，CNN-LSTM 模型的 RMSE 降低了 42%，MAPE 降低了 20%，MAE 降低了 33%；相较于 CNN-RNN 模型，CNN-LSTM 模型的 RMSE 降低了 57%，MAPE 降低了 38%，MAE 降低了 53%。这些结果表明了 CNN-LSTM 模型在预测股票价格上的优越性，它可以更加精确地预测出股票价格数据，给市场中的投资者提供参考。

五、结语

为解决股票价格未来趋势预测的难题，本文建立了 CNN-LSTM 混合神经网络模型，对贵州茅台等股票进行了预测。通过实验结果可知，CNN-LSTM 混合神经网络模型在股票价格的短期预测上有着较好的效果，并且优于 RNN、CNN-RNN 和 LSTM 神经网络。因此，本文提出的股价预测模型对市场中的投资者有一定参考价值。

【作者简介】邵景昊(2001—)，男，江苏苏州人，本科在读，南京信息工程大学，金融工程专业。