

基于K线与均线的CAE和GRU模型的股价预测方法

黄蕊¹, 刘子捷², 崔骥³

¹南京信息工程大学数学与统计学院, 江苏 南京

²南京邮电大学物联网学院, 江苏 南京

³苏州博纳讯动软件有限公司, 江苏 苏州

收稿日期: 2022年12月28日; 录用日期: 2023年1月21日; 发布日期: 2023年1月31日

摘要

设计快速精确的股票预测方法以实现投资收益最大化是学界的研究热点。然而, 现有的股票预测方法忽略了不同K线和均线之间的关联对股价的影响, 从而导致预测效果不佳。为了解决该问题, 基于CAE和GRU模型提出了一种利用K线和均线特征的关联性预测股价的方法。与选取的现有股票预测方法相比, 本模型在上证指数数据集上具有更好的拟合和泛化性能, 能够有效地减小预测误差。

关键词

股票预测, K线和均线特征, CAE网络, GRU网络

Stock Price Forecasting Method Based on CAE and GRU Models of K-Line and Moving Average

Rui Huang¹, Zijie Liu², Ji Cui³

¹School of Mathematics and Statistics, Nanjing University of Information Science and Technology, Nanjing Jiangsu

²College of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing Jiangsu

³Suzhou Beyondcent Software Co., Ltd, Suzhou Jiangsu

Received: Dec. 28th, 2022; accepted: Jan. 21st, 2023; published: Jan. 31st, 2023

Abstract

Designing fast and accurate stock prediction methods to maximize investment returns is a hot re-

search topic in academia. However, the existing stock forecasting methods ignore the impact of the correlation between different K-lines and moving averages on stock prices, which leads to poor prediction results. To address this problem, based on CAE and GRU model, this paper proposes a method to predict stock price by using the correlation of K-line and moving average characteristics. Compared with the existing stock prediction methods selected, this model has better fitting and generalization performance on the Shanghai Stock index dataset, and can effectively reduce the prediction error.

Keywords

Stock Forecast, Characteristics of K-Line and EMA, CAE Network, GRU Network

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

不同于债券和基金, 股票投资是一种高风险、高收益的投资方式。由于股价受经济波动、行业趋势、市场情绪和意外事件等多方面因素的影响, 仅凭人为经验对股价进行预测准确度较低, 因此如何准确地预测股价一直备受研究人员与投资者的关注。不少研究者尝试采用基本面和资金面分析[1]、ARIMA [2]和 GARCH [3]等传统的方法来预测股票价格。然而传统股票预测方法存在滞后性强、无法捕捉非线性关系和预测精度低等问题。近年来, 随着机器学习技术的不断发展, 研究者们开始尝试将机器学习技术应用在股票预测领域。Lv J., 等人设计了一个 LightGBM 优化的 LSTM 来实现短期股价预测[4]; Mukherjee, S.等人利用卷积神经网络(CNN)模型预测股票价格[5]。

然而, 现有的基于机器学习的股价预测方法忽略了不同均线之间的关联对股价的影响。例如, 短期均线与一条或者多条中期均线交叉所形成的交叉点称为黄金十字, 在金融领域里称为“金叉”。当出现这种信号时, 则意味着股票价格即将上涨。这时, 投资者可以在交叉点刚刚形成的时候买入; 相反, 短期均线和一条或者多条中期均线交叉所形成的交叉点称为“死叉”。当出现这种信号时, 则意味着股票价格即将下跌。这时, 应该立即将股票卖出。此外, 均线与 K 线之间的位置关系也会对股市和个股价格产生影响。例如, 在上涨行情中, 5 日、10 日和 20 日均线自下而上依次顺序排列, 向右上方移动, 称为“多头排列”, 预示股价将大幅上涨; 在下跌行情中, 5 日、10 日和 20 日均线自上而下依次顺序排列, 向右下方移动, 称为“空头排列”, 预示股价将大幅下跌。针对现有的股价预测方法忽略不同 K 线和均线之间的关联对股价的影响而导致的预测精度不佳的问题, 本文基于卷积自动编码器(Convolutional Autoencoder, CAE)和门控循环神经网络(Gate Recurrent Unit, GRU)模型提出了一种利用 K 线和均线特征的关联性预测股价的方法。其中, K 线特征包括当日开盘价、收盘价、最高价和最低价。本方法通过烛台图的形式表征股票 K 线特征。CAE 模块从绘制的烛台图中提取短期、中期和长期深度 K 线特征。本方法选取 5 日、10 日和 20 日均线分别作为短期、中期和长期均线特征。GRU 模块将短期、中期和长期深度 K 线特征的均线数据作为输入以学习当前交易日的深度 K 线特征与均线特征。本方法将学习的当前交易日的深度 K 线特征与均线特征和选取的基本面特征送入全连接层, 得到预测的股价。

综上所述, 本文相关的贡献如下:

1) 本文对每个数据样本以 5 日、10 日和 20 日为单元进行划分, 绘制烛台图并利用 CAE 模块分别学

习短期、中期和长期深度 K 线特征。该方式保证 CAE 模块能够以不同的时间粒度提取准确的深度 K 线特征，为后续精确地股价预测奠定基础。

2) 本文将短期、中期和长期深度 K 线特征分别与短期、中期和长期均线数据堆叠作为 GRU 模块的输入以学习当前交易日的深度 K 线特征与均线特征。该方式使得 GRU 模块能够学习短期、中期和长期均线与 K 线的位置关系，有利于提高股价预测的精度。

3) 本文将 CAE 网络、GRU 网络和未经划分的 CAE + GRU 网络作为对比对象，分析了本方法与其在 MAE、MAPE 和 MSE 评价指标的性能。实验表明，本方法在上证指数数据集上具有更好的拟合和泛化性能，能够有效地减小预测误差。

2. 相关工作

2.1. 传统的股票预测方法

如何准确预测股价一直以来成为国内外研究者关注的热点。研究者们先后采用基本面和资金面分析、ARIMA 和 GARCH 等传统的时间序列方法来预测股票的价格。Chen Y.J.等人考虑了公司财务状况、工业环境、宏观经济和金融新闻等因素，设计出一种基于基本面分析的股票市场预测方法[6]。该方法可用于计算财务指标的权重，根据股票特征评估和选择个股；方珂昊利用 ARIMA 模型对股票进行预测，将得到的预测值归为原始变量组，并加入相关变量，再结合偏最小二乘回归进行建模分析。通过实验，利用该组合方法能够得到更好的修正模型[7]；周丹文对恒生指数近 10 年的收盘价进行分析，建立 GARCH 模型，模拟了该指数波动性变化[8]。结果表明，在短期内，该模型拥有较好的预测效果。然而，传统的股票预测方法存在神经网络收敛速度慢、预测精度较低等问题。为了进一步提高预测精度，不少研究者利用机器学习技术进行股票预测。

2.2. 基于机器学习的股票预测方法

随着人工智能技术日渐成熟，越来越多的研究者采用机器学习技术来预测股价。Tabak B.M.等人利用 K 最近邻(KNN)和支持向量机(SVM)模型预测次日股价[9]。通过实验对比分析，SVM 模型具有更强的泛化能力；方昕等人首先利用粒子群算法提取股票数据的特征，然后用传统的随机森林算法和网格算法改进后的随机森林算法对沪深 300 和中证 500 指数进行预测[10]。通过实验对比分析，改进后的随机森林算法预测精度显著提升；Xie L.提出了一种基于 CAE 的无监督学习框架来预测金融时间序列领域的每日股价趋势。其中，CAE 主要通过编码和解码来提取高级语义特征[11]；Ni J.等人基于循环深度神经网络(RDNN)和 DCC-GARCH 模型开发了一种称为 DCDNN 的混合模型[12]。该混合模型能够准确预测股票之间的相关性；Hung C.C.等人提出了 DPP 框架，该模型以烛台图作为输入，接着通过 CAE 提取股票具有代表性的特征，最后通过 GRU 预测价格走势[13]；綦方中提出了一种基于 PCA 和 IFOA-BP 神经网络的股价预测模型[14]，与其它模型相比，该模型具有更快的收敛速度和更高的预测精度；许雪晨等人利用预训练模型 BERT 对股票新闻文本进行情感分析，并通过 LSTM 模型预测股价走势[15]。然而，现有的基于机器学习的股票预测方法忽略了不同 K 线和均线之间的关联对股价的影响。因此，希望通过考虑该因素使得预测精度得到进一步提升。

3. 方法

现有的基于机器学习的股票预测方法忽略了不同 K 线和均线之间的关联对股价的影响。为了解决该问题，本文提出了一种基于 CAE 和 GRU 模型的股价预测方法。该模型由 CAE、GRU 和全连接层组成。首先对每个数据样本以 5 日、10 日和 20 日为单元进行划分，绘制烛台图并利用 CAE 模块分别学习短期、

中期和长期深度 K 线特征;然后将 CAE 模块学习到的深度 K 线特征与均线特征堆叠作为 GRU 模块的输入以学习当前交易日的深度 K 线特征与均线特征;最后将选取的基本面特征和当前交易日的深度 K 线特征和均线特征送入全连接层,得到预测的股价。该方法的网络结构图如图 1 所示。

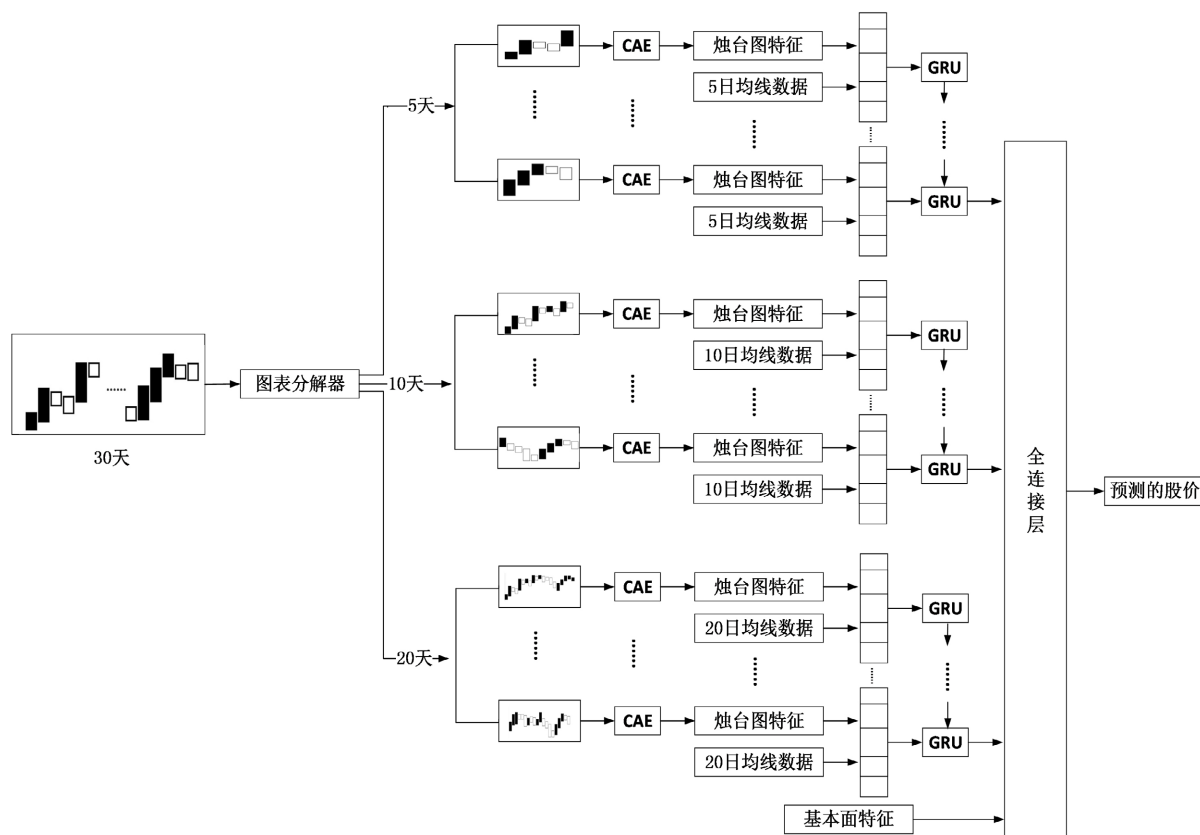


Figure 1. Network structure chart of stock price forecasting method based on CAE and GRU model

图 1. 基于 CAE 和 GRU 模型的股价预测方法的网络结构图

3.1. 数据预处理

本文使用上证指数的开盘价、收盘价、最高价和最低价作为原始数据。数据集中的每个样本包含 30 个交易日内的开盘价、收盘价、最高价和最低价的数据。为了使得网络模型具备学习均线交叉以及均线与 K 线之间的位置关系对股价影响的能力,因此需要对原始数据进行预处理。

本方法通过绘制烛台图反映股票的 K 线特征。烛台图可由开盘价、收盘价、最高价和最低价表示,如图 2 所示。烛台图描述了给定时间段内股票价格的波动情况。其中,每根柱线代表一个交易日的信息。烛台的不同颜色代表不同的含义,如果开盘价高于收盘价,实体将以白色填充;否则,实体将以黑色填充。

本方法将每 4 列描述一个交易日的 K 线特征,它通过将开盘价、收盘价、最高价和最低价转换为二维数组的方式绘制给定时间段为 m 天的烛台图。因此,烛台图的列数为 $4m$ 。对于每一个交易日的 K 线特征,第一列代表开盘价;第二列代表高低价区间;第三列代表收盘价;第四列代表间隔列。烛台图的每行为给定时间段 m 天内最高价(记为 P_{\max})与最低价(记为 P_{\min})之间的均匀分割。假设某日的 K 线特征值为 P ,本方法按照公式(1)所示的规则构建二维数组。

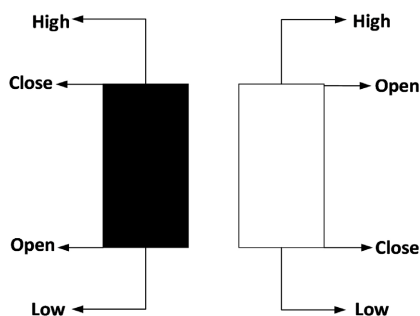


Figure 2. Candlestick diagram
图 2. 烛台图

$$a_{kl} = \begin{cases} 1, & \left(\frac{P_{\max} - P}{P_{\max} - P_{\min}} \cdot 4m - 1 < k \leq \frac{P_{\max} - P}{P_{\max} - P_{\min}} \right) \\ 0, & \left(k > \frac{P_{\max} - P}{P_{\max} - P_{\min}} \text{ or } k \leq \frac{P_{\max} - P}{P_{\max} - P_{\min}} \cdot 4m - 1 \right) \end{cases} \quad (1)$$

其中 k 指股票价格 ($0 \leq k < 4m$), l 指天数 ($0 \leq l < 4m$)。

本方法绘制的给定时间段为 5 天的烛台图示例如图 3 所示。

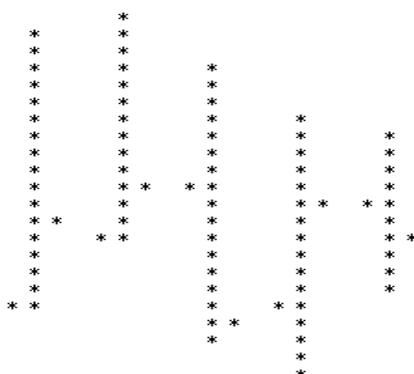


Figure 3. Example of a candlestick chart for a given time period of 5 days
图 3. 给定时间段为 5 天的烛台图示例

依据该方法, 本文以 5 日、10 日和 20 日为给定时间段分别绘制尺寸为 20×20 、 40×40 和 80×80 的烛台图来反映短期、中期和长期 K 线特征。

本文通过 5 日、10 日和 20 日均线分别表示短期、中期和长期均线特征。其中, 5 日均线指连续 5 个交易日的收盘价的平均值。公式如(2)所示:

$$MA5 = \frac{C_1 + C_2 + \dots + C_5}{5} \quad (2)$$

10 日均线指连续 10 个交易日的收盘价的平均值。公式如(3)所示:

$$MA10 = \frac{C_1 + C_2 + \dots + C_{10}}{10} \quad (3)$$

20 日均线指连续 20 个交易日的收盘价的平均值。公式如(4)所示:

$$MA20 = \frac{C_1 + C_2 + \dots + C_{20}}{20} \tag{4}$$

其中 c_i 指的是第 i 天的收盘价。

除此以外，本文额外选取 MACD、RSI、MOM 和成交量四种基本面特征，研究其对于股价的影响。

3.2. 模型构建

本方法将上述提及的烛台图、均线以及基本面特征作为模型的输入。该模型首先对每个数据样本以 5 日、10 日和 20 日为单元进行划分，绘制烛台图并利用 CAE 模块分别学习短期、中期和长期深度 K 线特征；然后将 CAE 模块学习到的深度 K 线特征与均线特征堆叠作为 GRU 模块的输入以学习当前交易日的深度 K 线特征与均线特征；最后将选取的基本面特征和当前交易日的深度 K 线特征和均线特征送入全连接层，得到预测的股价。

3.2.1. 卷积自动编码器

卷积自动编码器(CAE)是一种编码器 - 解码器架构的神经网络,被广泛用于图像和视频处理中的特征学习,但在股票预测应用中并不常见。本文采用的 CAE 具体结构及对应参数如图 4~6 所示。在编码过程中,本文使用多个二维卷积层和最大池化层从表征短期、中期和长期 K 线特征的烛台图中分别学习短期、中期和长期深度 K 线特征。在解码过程中,本文通过反池化操作重新构建烛台图,使得其与原烛台图尽可能相似。具体而言,本文采用的 CAE 模块的输入分别为 5 日、10 日和 20 日的烛台图,它们的维度大小分别为 $20 \times 20 \times 1$ 、 $40 \times 40 \times 1$ 和 $80 \times 80 \times 1$ 。本文采用的 CAE 模块的编码网络和解码网络由多个二维卷积层和最大池化层组成。其中,采用的二维卷积层和最大池化层的参数如图 4~6 所示。CAE 模块通过编码过程分别从 5 日、10 日和 20 日的烛台图中提取维度为 $5 \times 5 \times 64$ 的深度 K 线特征。

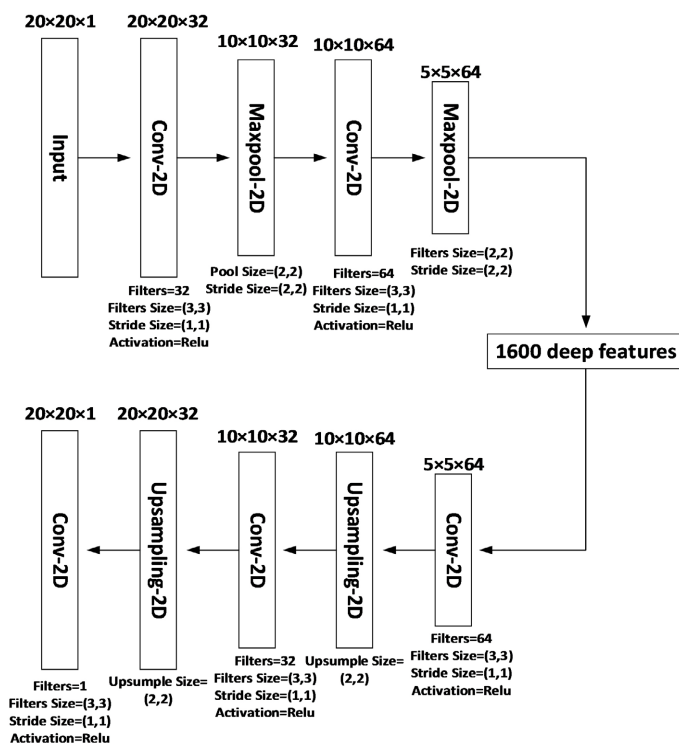


Figure 4. The specific structure and corresponding parameters of CAE with 5 days as the division unit
图 4. 以 5 日为划分单元的 CAE 具体结构及对应参数

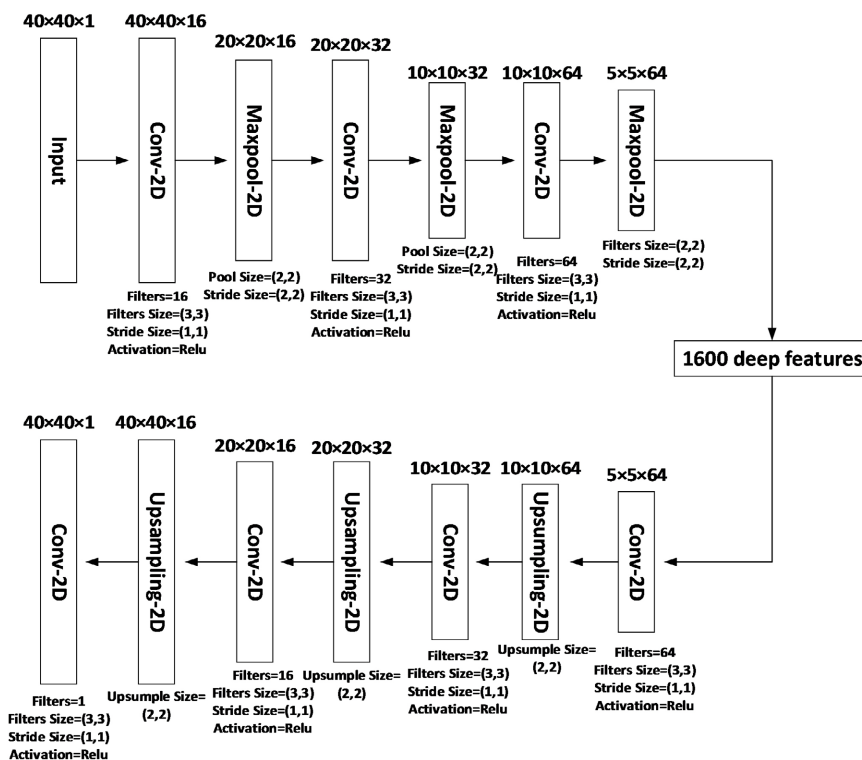


Figure 5. CAE specific structure and corresponding parameters with 10 days as division unit
 图 5. 以 10 日为划分单元的 CAE 具体结构及对应参数

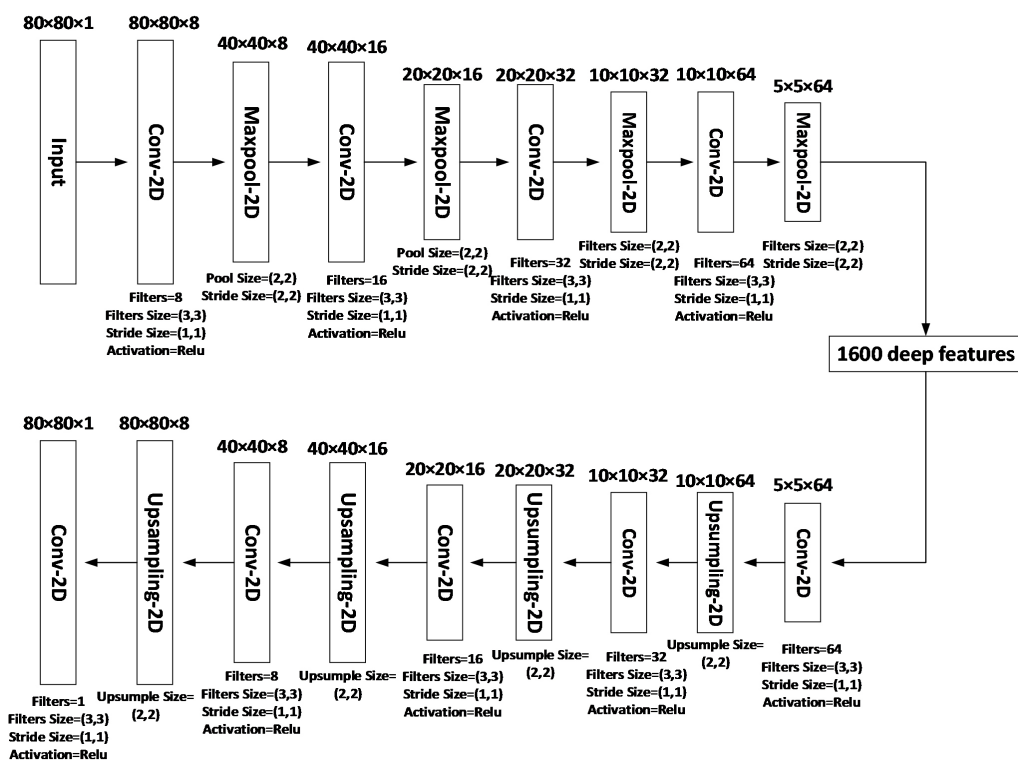


Figure 6. CAE specific structure and corresponding parameters with 20 days as division unit
 图 6. 以 20 日为划分单元的 CAE 具体结构及对应参数

3.2.2. 门控循环神经网络

相较于 LSTM, 门控循环神经网络(GRU)通过将输入门和遗忘门合并为更新门的方式进一步提高了运行效率。因此, 本文采用 GRU 模块根据 CAE 模块提取的深度 K 线特征以及数据预处理得到的均线数据, 学习当日的深度 K 线特征和均线特征以预测股价。具体而言, 本文将 CAE 模块根据烛台图提取的短期、中期和长期深度 K 线特征分别与数据预处理得到的短期、中期和长期均线数据进行堆叠以作为 GRU 模块的输入。如图 7~9 所示, 5 日、10 日和 20 日的 GRU 模块分别包含 26、21 和 11 个 GRU 单元。每一个 GRU 单元包含 128 个隐藏特征, 激活函数选用 tanh 激活函数。GRU 模块的输出为当前交易日的深度 K 线特征和均线特征。

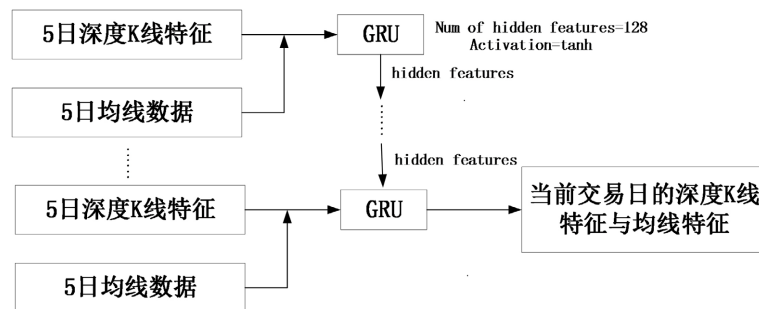


Figure 7. Specific structure and parameters of GRU module with 5 days as division unit

图 7. 以 5 日为划分单元的 GRU 模块具体结构及参数

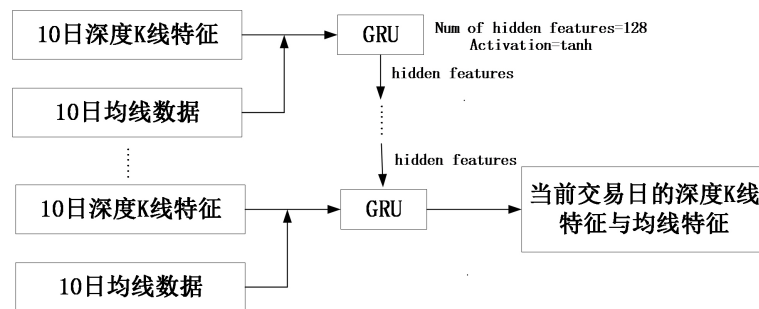


Figure 8. Specific structure and parameters of GRU module with 10 days as division unit

图 8. 以 10 日为划分单元的 GRU 模块具体结构及参数

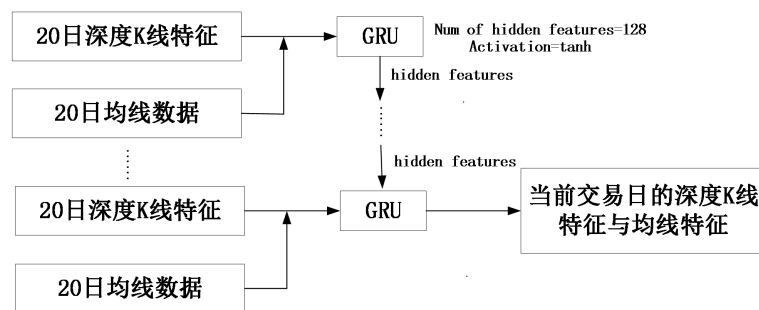


Figure 9. Specific structure and parameters of GRU module with 20 days as division unit

图 9. 以 20 日为划分单元的 GRU 模块具体结构及参数

3.2.3. 全连接层

本文通过在 GRU 模块后连接一个全连接层以学习当日的股票价格。具体而言,如图 10 所示,该全连接层的输入为 GRU 模块学习得到的当日深度 K 线特征和均线特征以及预处理得到的基本面特征;输出为表征当日股价的 1 个神经元。本文选取 Sigmoid 函数作为全连接层的激活函数,其公式如(5)所示:

$$S(x) = \frac{1}{1 + e^{-x}} \quad (5)$$

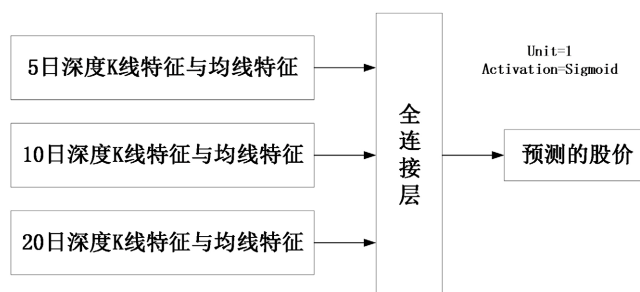


Figure 10. The specific structure and corresponding parameters of the full connection layer

图 10. 全连接层的具体结构及对应参数

3.2.4. 损失函数

本文的损失函数由 CAE 模块的损失函数和全连接部分的损失函数组成。为了使得 CAE 模块能够通过编码器从输入的烛台图中学习到准确的深度 K 线特征,需要保证通过 CAE 模块解码器得到的烛台图与输入的烛台图尽可能相似。因此,本文选取二元交叉熵作为 CAE 模块的损失函数,其公式如公式(6)所示:

$$H(p, q) = -(p * \log(q) + (1 - p) * \log(1 - q)) \quad (6)$$

其中 p 指真实分布, q 指预测分布。

为了使得本文采用的全连接层能够准确地预测股价,需要保证通过训练得到股价的预测值与真实值尽可能相似。因此,本文选取 MSE 作为全连接层的损失函数,其公式如公式(7)所示:

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (7)$$

其中, n 为样本数, Y_i 为样本真实值, \hat{Y}_i 为模型预测值。

4. 实验分析

4.1. 数据及模型参数

4.1.1. 数据来源

本文将近 30 年来(即从 1992 年 1 月 1 日至 2021 年 12 月 31 日)的中国上证指数作为数据集。其中,每个数据样本包含单个交易日的开盘价、收盘价、最高价和最低价。上证指数的交易数据从财经数据接口包 Tushare (<https://www.tushare.pro>)获取。

4.1.2. 数据预处理

为了消除不同的量纲和量纲单位在特征数据计算时产生的差异,在训练前,本文将原始数据样本中的特征数据进行归一化处理,数据归一化的公式如(8)所示:

$$\chi = (X - X_{\min}) / (X_{\max} - X_{\min}) \quad (8)$$

其中, X 为股票的价格, X_{\max} 为股票价格的最大值, X_{\min} 为股票价格的最小值。

另外, 由于财务报告的缺失、计算错误等原因, 很可能造成数据的丢失, 从而会影响数据分析的准确性。因此, 缺失值的预处理非常重要。本文采用填充法对缺失值进行处理, 如果当前交易日的某个特征数据为缺省值, 则用前一个非缺省的特征数据值进行替换。

经过预处理后, 数据根据比例划分为训练集、验证集和测试集。其中, 前 80% 的数据作为训练集用来训练模型, 之后 10% 的数据作为验证集用来调整模型的参数, 最后 10% 的数据作为测试集用来评估模型的预测效果。

4.1.3. 评价指标

在实验过程中, 为了评估模型的预测性能, 本文主要使用三种不同的评估指标: 平均绝对误差(MAE)、平均绝对百分比误差(MAPE)和均方误差(MSE)。三种指标的计算公式如公式(9)至公式(11)所示:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (9)$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|Y_i - \hat{Y}_i|}{Y_i} \quad (10)$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (11)$$

其中, n 为样本数, Y_i 为样本真实值, \hat{Y}_i 为模型预测值。

4.1.4. 实验环境与参数设置

本文的硬件实验环节为一台装备 4 个 Intel(R) Xeon(R) Silver 4114 CPU @ 2.20GHz (40 核 CPU), 4 块 16GB DDR4 memory (64 GB 内存), 4 块 WD Black 1TB 硬盘和 2 块 NVIDIA GTX 1660S GPU 的高性能服务器。

本文的软件实验环境为 Tensorflow-gpu 2.0.0 下的 keras 2.8.0 深度学习框架, 配置 NVIDIA CUDA 10.0, cuDNN7.6.4 深度学习库加速 GRU 计算, python 版本为 python3.7。在训练弱学习器时采用 RMSprop 作为优化器, 学习率设置为 0.001。模型采用 200 个 epochs, 每个 batch 的大小为 64。

4.2. 上证指数预测结果分析

本文的数据集由三部分组成。其中, 1992 年 1 月 1 日至 2015 年 12 月 31 日的股价数据作为训练集; 2016 年 1 月 1 日至 2018 年 12 月 31 日的股价数据作为验证集; 2019 年至 2021 年 12 月 31 日的股价数据作为测试集。本文将 CAE 网络、GRU 网络和未经划分的 CAE + GRU 网络作为对比网络模型, 从而说明本方法在预测上证指数股价的有效性。具体实验结果如表 1 所示。

Table 1. Standard test system result data

表 1. 标准试验系统结果数据

模型	MAE	MAPE	MSE
CAE	0.0154	3.2001	4.22E-04
GRU	0.0081	1.4056	1.30E-04
未经划分的 CAE + GRU	0.0067	1.2745	8.80E-05
本文方法	0.0058	1.2381	5.89E-05

由表 1 可知, 本文所提出的模型在各评价指标中都取得了良好表现, 总体效果均优于其它对比模型。本文方法与 CAE 网络相比, 在预测股价时考虑 GRU 获取的股价时间序列数据的特征, 因此本文方法的预测效果较好; 本文方法与 GRU 网络相比, 在预测股价时考虑 CAE 从烛台图中提取的深度 K 线特征, 因此本文方法的预测效果较好; 本文方法与未经划分的 CAE + GRU 网络相比, 在预测股价时考虑短期、中期和长期均线数据与 K 线特征之间的关联, 因此本文方法的预测效果较好。

综上所述, 本文利用 CAE 模块提取的短期、中期和长期的深度 K 线特征以及 GRU 模块学习的短期、中期和长期均线数据与 K 线特征之间的关联进行股价预测。因此, 本文方法的预测效果很好, 更接近真实值, 预测误差更小。

图 11~14 分别为本文方法、CAE 网络、GRU 网络和未经划分的 CAE + GRU 网络在上证指数测试集上的股价预测结果。其中, 横坐标表示测试集数据样本的索引, 纵坐标表示经过归一化处理后的股票收盘价数值。由图可以看出, 本文所研究的模型在预测值与实际值的拟合度上表现较好, 其预测性能明显高于其他对比模型的预测性能。该模型不仅能获取股票未来趋势, 而且在股票的总体总势变化较大的拐点预测正确。因此, 通过参考股价收盘价的预测值进行股票交易, 能够有效提高股民的收益和降低投资风险。

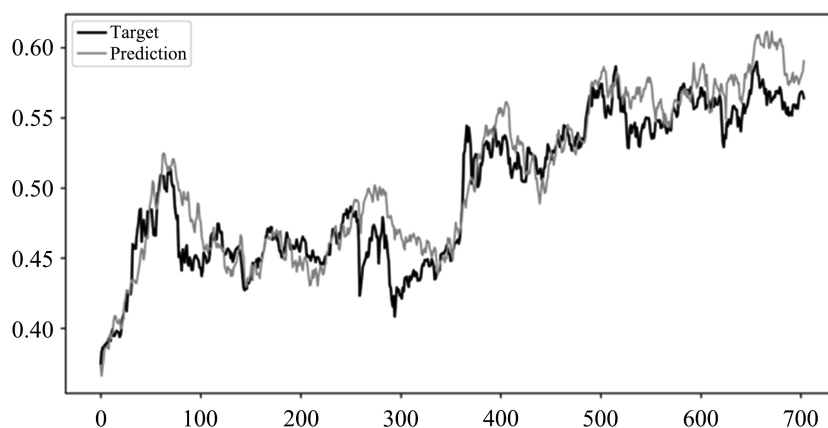


Figure 11. Effect chart of CAE network's stock price forecast on the Shanghai stock index test set

图 11. CAE 网络在上证指数测试集上的股价预测效果图

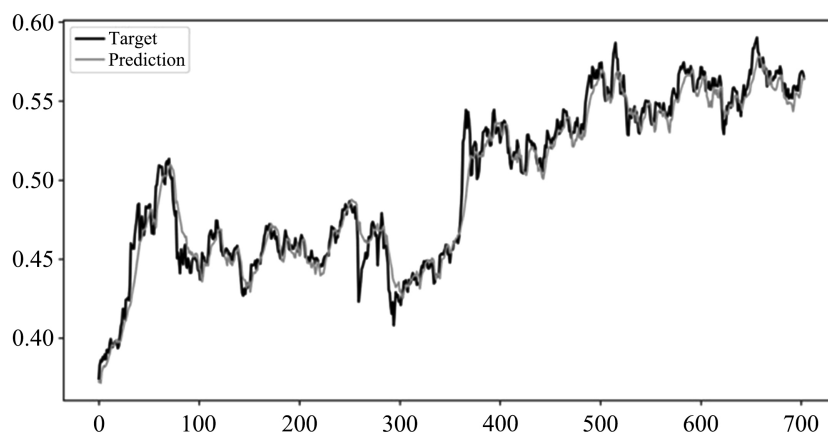


Figure 12. Effect chart of GRU network's stock price forecast on the Shanghai stock index test set

图 12. GRU 网络在上证指数测试集上的股价预测效果图

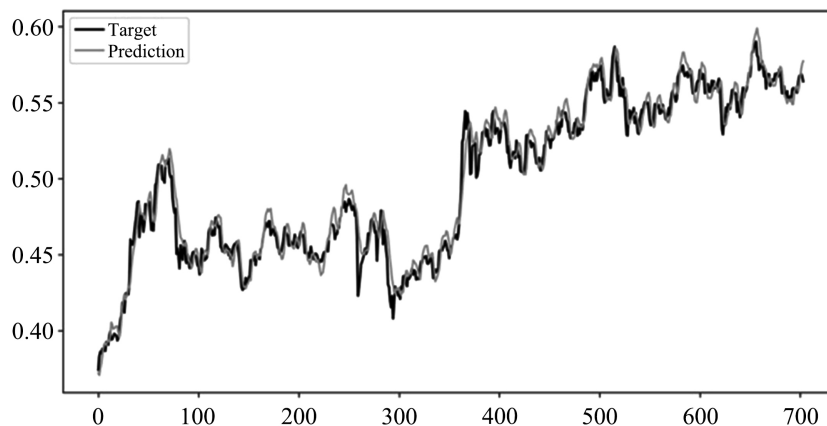


Figure 13. Effect chart of stock price forecast of undivided CAE + GRU network on shanghai stock index test set

图 13. 未经划分的 CAE + GRU 网络在上证指数测试集上的股价预测效果图

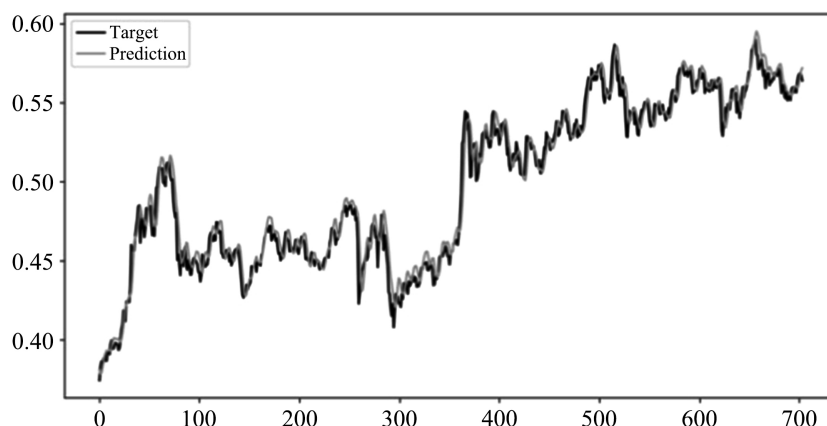


Figure 14. Effect chart of stock price forecast based on the test set of Shanghai stock exchange index

图 14. 本文方法在上证指数测试集上的股价预测效果图

5. 结语

本文基于 CAE 和 GRU 模型提出了一种利用 K 线和均线特征的关联性预测股价的方法。

本方法首先将训练样本中的 K 线数据按照 5 日、10 日和 20 日的划分单元进行划分，并根据划分后的 K 线数据绘制烛台图；其次，本文采用的 CAE 模块根据绘制的烛台图数据提取短期、中期和长期的深度 K 线特征；然后将提取到的深度 K 线特征分别与短期、中期和长期均线数据堆叠作为 GRU 模块的输入以学习当前交易日的深度 K 线特征和均线特征；最后本文将选取的基本面特征和当前交易日的深度 K 线特征和均线特征送入全连接层，得到预测的股价。由实验结果可知，本文所提出的方法相较于 CAE 网络、GRU 网络和未经划分的 CAE + GRU 网络具有更高的拟合度和预测精度。但是总体的预测性能还需要进一步的提升，后期将在股市预测模型中仍需考虑网络舆情、经济水平和国家政治等因素来更好地反映股票价格的涨跌情况，进一步为投资者提供合理有效的参考。

基金项目

国家自然科学基金面上项目(62076137)。

参考文献

- [1] 顾锋娟, 金德环. 基本面、资金面对中国股市波动影响分析[J]. 上海金融, 2012(10): 85-90+46+118.
- [2] 张亚婕. 基于 ARIMA 模型对股票和指数预测结果的简单比较分析[J]. 市场研究, 2019(11): 23-26. <https://doi.org/10.13999/j.cnki.scyj.2019.11.008>
- [3] 耿娟, 刘怡超. 基于 GARCH 模型的股票收益率分析及预测[J]. 河北企业, 2019(10): 61-62.
- [4] Lv, J., Wang, C., Gao, W., et al. (2021) An Economic Forecasting Method Based on the Light GBM-Optimized LSTM and Time-Series Model. *Computational Intelligence and Neuroscience*, **2021**, Article ID: 8128879. <https://doi.org/10.1155/2021/8128879>
- [5] Mukherjee, S., Sadhukhan, B., Sarkar, N., et al. (2021) Stock Market Prediction Using Deep Learning Algorithms. *CAAI Transactions on Intelligence Technology*, **2021**, 1-13. <https://doi.org/10.1155/2021/8128879>
- [6] Chen, Y.J., Chen, Y.M. and Lu, C.L. (2017) Enhancement of Stock Market Forecasting Using an Improved Fundamental Analysis-Based Approach. *Soft Computing*, **21**, 3735-3757. <https://doi.org/10.1155/2021/8128879>
- [7] 方珂昊, 赵凌. 基于偏最小二乘方法的 ARIMA 模型在股票指数预测中的应用[J]. 四川文理学院学报, 2018, 28(5): 7-11..
- [8] 周丹文. MATLAB 中基于 GARCH 模型对股票指数的拟合与预测——以恒生指数为例[J]. 对外经贸, 2020(4): 77-80.
- [9] Tabak, B.M., Silva, T.C., Zhao, L., et al. (2020) Applications of Machine Learning Methods in Complex Economics and Financial Networks. *Complexity*, **2020**, Article ID 4247587. <https://doi.org/10.1155/2021/8128879>
- [10] 方昕, 陈玲玲, 曹海燕. 基于改进的随机森林算法股票收益率预测研究[J]. 杭州电子科技大学学报(自然科学版), 2020, 40(1): 35-40. <https://doi.org/10.13954/j.cnki.hdu.2020.01.007>
- [11] Xie, L. and Yu, S. (2021) Unsupervised Feature Extraction with Convolutional Autoencoder with Application to Daily Stock Market Prediction. *Concurrency and Computation: Practice and Experience*, **33**, e6282. <https://doi.org/10.1002/cpe.6282>
- [12] Ni, J. and Xu, Y. (2021) Forecasting the Dynamic Correlation of Stock Indices Based on Deep Learning Method. *Computational Economics*, 1-21. <https://doi.org/10.1007/s10614-021-10198-3>
- [13] Hung, C.C. and Chen, Y.J. (2021) DPP: Deep Predictor for Price Movement from Candlestick Charts. *PLOS ONE*, **16**, e0252404. <https://doi.org/10.1371/journal.pone.0252404>
- [14] 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+156.
- [15] 许雪晨, 田侃. 一种基于金融文本情感分析的股票指数预测新方法[J]. 数量经济技术经济研究, 2021, 38(12): 124-145. <https://doi.org/10.13653/j.cnki.jqte.2021.12.009>