

基于BO-XGBoost-PSO-TCN-BiGRU的黄金期货价格预测

谷爱玲¹, 莫俊铭¹, 黄绮晴²

¹广东工业大学数学与统计学院, 广东 广州

²五邑大学数学与计算科学学院, 广东 江门

收稿日期: 2024年11月10日; 录用日期: 2024年12月6日; 发布日期: 2024年12月12日

摘要

为了充分挖掘影响金融数据中时序性特征联系, 提高预测黄金期货价格的预测精度, 提出了一种基于粒子群优化(PSO)、时间卷积神经网络(TCN)和双向门控循环单元(BiGRU)模型相结合的优化双智能体深度学习模型的黄金期货价格预测方法。选择2018~2024年交易日成交价, 并添加六大类影响因素数据。通过贝叶斯优化XGBoost并结合SHAP算法进行可解释性特征筛选, 再建立PSO-TCN-BiGRU混合模型进行预测, 并通过与CNN、BiGRU、TCN、BiGRU-TCN、CNN-BiGRU、TCN-BiGRU模型对比分析, 结果表明所提预测模型的RMSE值、MSE值、MAE值都低于其他模型, R方值最高, 能更准确地预测黄金期货价格走势。

关键词

黄金期货价格, BO-XGBoost, PSO-TCN-BiGRU, 可解释性分析, 混合模型

The Prediction of Gold Futures Prices Based on BO-XGBoost-PSO-TCN-BiGRU

Ailing Gu¹, Junming Mo¹, Qiqing Huang²

¹School of Mathematics and Statistics, Guangdong University of Technology, Guangzhou Guangdong

²School of Mathematics and Computational Science, Wuyi University, Jiangmen Guangdong

Received: Nov. 10th, 2024; accepted: Dec. 6th, 2024; published: Dec. 12th, 2024

Abstract

The proposed optimized dual-agent deep learning model, based on particle swarm optimization (PSO), time convolutional neural network (TCN), and bidirectional gated recurrent unit (BiGRU)

文章引用: 谷爱玲, 莫俊铭, 黄绮晴. 基于 BO-XGBoost-PSO-TCN-BiGRU 的黄金期货价格预测[J]. 统计学与应用, 2024, 13(6): 2249-2260. DOI: 10.12677/sa.2024.136218

model, aims to fully explore the temporal characteristic relationships affecting financial data and enhance the prediction accuracy of gold futures price. The trading prices of 2018~2024 trading days were selected, along with six categories of influencing factor data. Bayesian optimization XGBoost combined with SHAP algorithm was used to screen interpretability features, followed by the establishment of a PSO-TCN-BiGRU mixed model for prediction. Comparison and analysis with other models such as CNN, BiGRU, TCN, BiGRU-TCN, CNN-BiGRU, TCN-BiGRU models revealed that the proposed forecasting model exhibited lower RMSE value, MSE value and MAE value than other models while achieving the highest R-square value. This indicates its ability to predict the trend of gold futures price more accurately.

Keywords

Gold Futures Prices, BO-XGBoost, PSO-TCN-BiGRU, Interpretability Analysis, Hybrid Model

Copyright © 2024 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

黄金在全球经济中拥有着无可替代的作用，因为它不仅兼具了商品属性与货币属性，而且具有良好的投资以及价值储藏功能，同时其衍生品黄金期货更是投资者关注的焦点之一，随着世界经济增长放缓，地缘政治不确定性加剧的情况下，投资者越来越看重避险资产，而黄金是众多避险资产中最受青睐的，能降低投资组合的风险度。黄金期货是以黄金为交易对象的金融衍生品，允许投资者在未来某个约定时间买卖一定数量的黄金。黄金期货市场是买卖这种远期交货合同的市场，是全球最重要的期货市场之一。鉴于黄金期货市场在金融市场中有着十分重要的位置，如何准确预测黄金期货价格走势成为一个重要研究课题，这对投资者来说是非常重要的。同时，还可以为相关部门，精确监控黄金市场的行情，保证我国的经济平稳发展。

以往学者们运用传统的统计学模型进行期货价格预测，这些传统的预测模型，包括 ARIMA 模型(移动平均自回归模型)、GARCH 模型(自回归条件异方差模型)、指数平滑模型等，在黄金价格预测研究上取得了不错的进展。费婧文[1]构造 ARIMA 模型有效预测了我国黄金期货价格的短期走势。在这些模型基础上，丁磊等人[2]选用 ARIMA-GARCH 模型对黄金价格的动态演化过程进行了实证分析，并发现该组合模型能较好地拟合黄金价格的运动路径。但是传统预测方法往往基于线性假设和固定参数模型，难以有效处理非线性、高维和动态变化的金融时间序列数据。此外，这些方法往往忽略了数据中的隐藏特征和潜在关系，导致预测精度有限。

近年来，深度学习技术的迅速发展为金融时间序列预测提供了新的思路和方法。深度学习模型通过模拟人脑神经网络的运作方式，能够自动提取数据中的复杂特征和规律，并学习数据的内在表示。在时间序列预测领域，多层前馈网络(BP)、循环神经网络(RNN)、长短期记忆网络(LSTM)和门控循环单元(GRU)等模型已被广泛应用于金融时间序列预测中，并取得了一定的成果。曾濂等人[3]提出了一种基于 BP 神经网络的改进模型实现了高精度仿真黄金期货价格。刘璐等人[4]建立多层 LSTM 以及双向 LSTM 模型预测未来黄金期货价格及其发展变化的规律和趋势。

为了更加充分识别复杂时间序列中的信息，更全面地描述和解释数据的特征，一些学者开始尝试将不同的深度学习模型进行组合，提出了深度学习中的组合预测方法，这些混合模型结合了两种甚至多种

单一模型的优势, 通过将不同的模型进行组合, 混合模型可以充分利用各种模型的优点, 从而提高了模型的整体性能。通常具有更强的泛化能力, 可以学习到更多的特征和规律, 从而提高了对未知数据的预测准确性。杨光[5]通过建立 CNN-LSTM-ATTENTION 模型对黄金期货价格预测, 发现 CNN-LSTM-ATTENTION 模型对比 CNN-LSTM 可以提高价格预测的准确程度以及对价格预测的准确程度也有提高。

然而, 黄金期货价格预测是一个复杂且具有挑战性的任务。黄金期货的走势受到多种因素的影响, 这些因素包括供求关系、美元汇率、各国的货币政策、石油供求关系、国际政治局势的变化以及市场情绪等。因此需要对众多指标进行筛选, 尹晨曦[6]选取影响黄金价格的相关因素后, 利用 Lasso 对所选因素进行特征提取, 再引入 PSO-BP 方法有效实现黄金期货价格的短期预测。在黄金期货价格影响因素相关性研究中本文引入贝叶斯优化 XGBoost, 通过贝叶斯算法找到 XGBoost 的最佳超参数, 并进行特征筛选, 再使用 SHAP 方法[7]进行模型可解释性分析并可视化。筛选出相关性强的特征以便于后续建立预测模型。

本文提出了一种基于粒子群优化时间卷积网络(TCN)和双向门控循环单元(BiGRU)的黄金价格预测模型。TCN [8]和 BiGRU 分别具有强大的特征提取和时序建模能力, 能够捕捉数据中的复杂特征和长期依赖关系。而粒子群优化算法能够捕捉数据中的复杂特征和规律, 在多维空间中搜索最优解。该模型的预测精度得到了提升, 且评价指标最优。能够为黄金期货价格预测科学性探索提出参考。

2. 基础理论

2.1. 贝叶斯优化

贝叶斯优化[9]是由 Snoek 等提出的优化算法, 用来求解表达式未知函数的极值问题。在算法优化问题中用于寻找最优神经网络超参数。相关研究表明该优化算法优于网格搜索及随机搜索算法, 相比于这两个算法, 贝叶斯优化算法能更容易实现全局最优。

在超参数搜索空间 X 内, 贝叶斯优化器进行有限次试验, 不断迭代得出最佳超参数的组合。最佳超参数组合公式如下:

$$x^* = \arg \min f(x) \quad (1)$$

公式中 x^* 为优化后最佳超参数组合, $f(x)$ 为目标函数。

2.2. XGBoost 算法

XGBoost [10]算法是 GBDT [11]的一种高效实现, 是一种集成式机器学习算法。GBDT 回归模型的算法为:

输入: 训练集 $\text{data} = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, 基学习器数量以及损失函数。

(1) 初值估算一个常数 c , c 能最大限度地减小损失函数, 并创建仅有一个根结点的树;

(2) 对样本 $m = 1, 2, 3, \dots, n$ 求解损失函数的负梯度值以求出残差;

(3) 各叶结点的目标值的平均值 Y_{tk} 上对具有极低损失的最优拟合结果 O_{tk} 进行相加, 从而获得第 u 棵树模型该次损失的估计量 $p_u(x_m)$;

(4) 不断地进行迭代更新, 获得最后的强学习模型。

2.3. SHAP 方法

SHAP [12]是一种基于 Shapley value 的计算方法, 用于解释黑盒模型是如何衡量特征对最终结果值的影响。通过对各要素的作用大小进行分析, 对其进行重要度排序, 通过可视化的方式将其与各要素的关联展现出来, 从而加深对机器学习的认知。Shapley value 的计算方法来源于博弈论[13]方法, 在博弈论中,

该数值用来衡量每个参与者对最终结果做出的贡献。我们可以将 XGBoost 视为一个博弈的模型，数据特征看成是参与博弈的人。Shapley value 计算公式为：

$$\phi_i(v) = \sum_z \frac{|z|!(m-|z|-1)!}{m!} [v(z \cup \{i\}) - v(z)] \quad (2)$$

公式中 v 是一个合作博弈的收益函数，等于 XGBoost 模型的输出， m 为参与人的个数，等于数据特征个数， Z 是参与人的随机组合。

2.4. TCN 时间卷积网络

时间卷积网络是一种专门用于处理序列数据的卷积神经网络。与传统的卷积神经网络不同，TCN 采用了特殊的卷积结构和因果约束，使得它能够处理具有时间依赖性的数据。通过采用扩张卷积[14]和因果卷积[15]相结合的方式，TCN 能够在保持序列顺序的同时，有效地提取序列中的局部和全局特征。扩张卷积能够增加感受野的大小；而因果卷积则确保了网络在处理当前时刻的数据时，只能利用之前时刻的信息，从而避免了未来信息的泄露。通过堆叠多个卷积层，TCN 可以逐步增加网络的深度，从而捕获更高阶的特征。同时，由于其卷积操作的特性，TCN 在处理长序列时不会出现梯度消失或爆炸的问题，这使得它在处理长时间依赖关系的任务中表现出色。TCN 因果卷积结构如图 1。

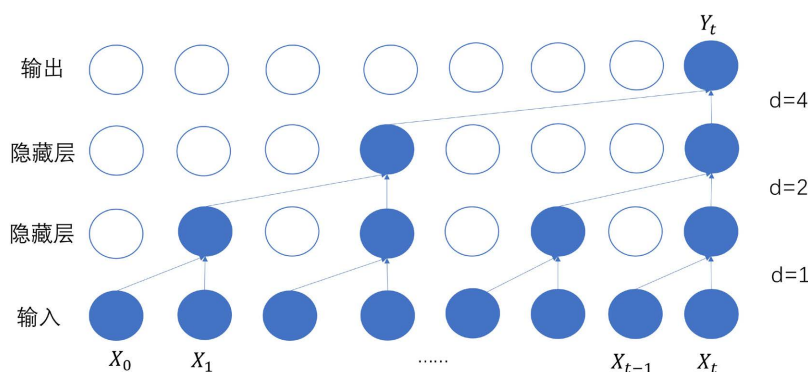


Figure 1. Causal convolution structure
图 1. 因果卷积图

假设输入序列为 X ，卷积核为 g 的一维时间序列 l 上的因果膨胀卷积运算 H 表达式为：

$$H(l) = (x * g)(l) = \sum_{g(i)} x_{l-d \times i} \quad (3)$$

式中： k 为卷积核的大小， $*$ 为卷积运算符， $l-d \times i$ 为膨胀卷积的输入向量元素， d 为扩张因子。TCN 可以通过增加 d 或增加 k 的数值来实现增加感受野的目的。

在黄金期货价格预测任务中，因果卷积和空洞卷积模块的叠加会使得神经网络层数加深，防止训练过程中出现梯度下降甚至消失的现象，引入残差块[16]连接，表达式为：

$$c = \sigma(x + g(x)) \quad (4)$$

式中： $g(x)$ 代表卷积层的输出， σ 是激活函数。

2.5. BiGRU 深度学习网络

GRU [17] 是循环神经网络的一种，能够解决 RNN 梯度爆炸或者梯度消失的问题。而 BiGRU [18] 是一个双向的 GRU 网络模型，通过对时间序列数据进行前向和后向两个方向的处理，捕获前后的信息，更

加全面地理解时间序列数据的前后关系,有利于提高时间序列数据特征的完整性和全局性。GRU 中的细胞结构如图 2。

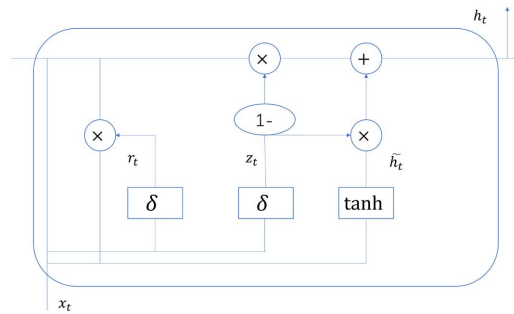


Figure 2. Diagram of GRU structure

图 2. GRU 结构图

在图 2 中,当前时刻的输入为 x_t , h_{t-1} 是前一个细胞状态。 W_r 、 W_z 分别为重置门和更新门的权重矩阵, σ 和 $\tanh$ 为激活函数,更新门[19]和重置门[19]的更新如下:

$$r_t = \sigma(W_r * [h_{t-1}, x_t]) \quad (5)$$

$$z_t = \sigma(W_z * [h_{t-1}, x_t]) \quad (6)$$

经过重置门和更新门后,GRU 单元的候选状态值 $\tilde{h}_t$ 和最终输出状态值 h_t 如下:

$$\tilde{h}_t = \tanh(W_{\tilde{h}} * [r * h_{t-1}, x_t]) \quad (7)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (8)$$

其中, $W_{\tilde{h}}$ 为候选状态值的权重矩阵。

GRU 单元对时间序列数据的每一时刻的状态计算,将输入序列中 useful 信息留下,将噪声去掉,最后的输出 y_t 公式如下:

$$y_t = \sigma(W_0 h_t) \quad (9)$$

其中 W_0 为输出值的权值矩阵。

BiGRU 能够在任意时刻,通过两个 GRU 的状态共同决定 BiGRU 的输出。BiGRU 的结构如图 3。

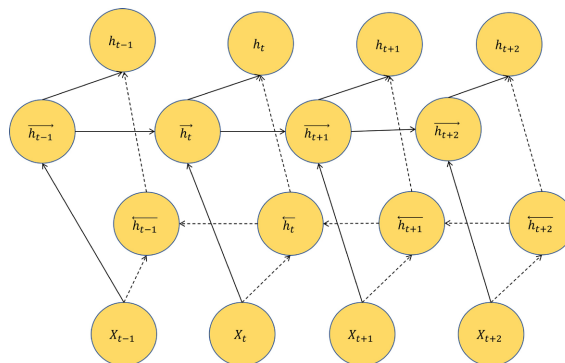


Figure 3. Diagram of BiGRU structure

图 3. BiGRU 结构图

如图 3, BiGRU 的 t 时刻隐藏层状态值通过前向隐藏层状态 $\overrightarrow{h_{t-1}}$ 和反向隐藏层状态 $\overleftarrow{h_{t-1}}$ 加权求和得出:

$$h_t = GRU(x_t, \overrightarrow{h_{t-1}}) \quad (10)$$

$$\overleftarrow{h_t} = GRU(x_t, \overleftarrow{h_{t-1}}) \quad (11)$$

$$h_t = W_t h_t + V_t \overleftarrow{h_t} + b_t \quad (12)$$

其中 W_t 和 V_t 分别是正向的权重参数和反向权重参数, b_t 是偏置项。

2.6. PSO

PSO [20] (粒子群算法) 是受对鸟群觅食行为研究结果得到启发, 提出的一种群智能算法技术, 目的就是找出最优解, 核心算法就是先在高维空间中随机初始化粒子群, 并定义优化目标的随机参数, 而这些参数决定了高维空间的维数。每个粒子在搜索空间中拥有自己的位置和速度。通过不断地调整和迭代就可以找到最优解。

在 n 维空间中第 k 次迭代时粒子 m 的位置和速度的公式如下:

$$X_m(k) = [X_{m1}(k), X_{m2}(k), \dots, X_{mn}(k)] \quad (13)$$

$$V_m(k) = [V_{m1}(k), V_{m2}(k), \dots, V_{mn}(k)] \quad (14)$$

公式中的 $X_m(k)$ 和 $V_m(k)$ 分别为位置函数和速度函数。

粒子 mm 在第 $k+1$ 次迭代时的位置和速度的公式如下:

$$v_m(k+1) = qV_m(k) + c_1 r_1 [Lbest_m(k) - X_m(k)] + c_2 r_2 [Gbest(k) - X_m(k)] \quad (15)$$

$$X_m(k+1) = X_m(k) + V_m(k+1) \quad (16)$$

公式中的 $Lbest_m$ 为粒子 m 在本地解空间中获得的最佳解, $Gbest(k)$ 为算法中粒子 m 附近所有粒子获得的最佳解。 q 为惯性权重, c_1 和 c_2 是学习因子。在模型优化领域, PSO 常被用于自动调整模型的超参数 [21], 以找到使模型性能最优的参数组合。

3. PSO-TCN-BiGRU 优化混合神经网络模型

3.1. 混合神经网络基本结构

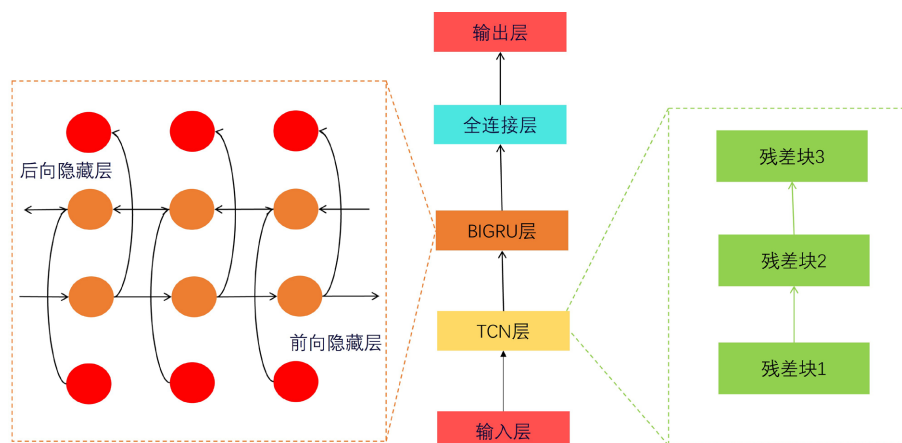


Figure 4. The structure of the combined forecasting model of TCN-BiGRU

图 4. TCN-BiGRU 预测网络模型结构

为了更好地融合 TCN 网络, BiGRU 网络的优点, 提出了 TCN-BiGRU 混合神经网络基本结构。首先使用 TCN 网络的一维因果卷积对时间序列数据进行提取, 扩张卷积进行时间序列特征提取, 再用拥有非线性[22]拟合能力的 BiGRU 网络进一步加深提取时间序列特征。图 4 为模型结构图。

首先, 通过堆叠 3 层 TCN 残差模块获得更大范围的输入时间序列数据的感受野以及对时间序列数据提取时间序列数据特征和数据降维[23], 避免出现梯度下降甚至消失的现象 BiGRU 获取 TCN 处理后的时间序列, 通过正反方向的双层 GRU 层连接起来, 一层按时间步从前往后对时间序列数据进行正向处理, 另一层从后往前对时间序列数据进行反向处理。正因为有这种结构 BiGRU 能更完整全面地探索时间序列的依赖关系, 获取上下文关联连接, 最后运输到全连接层实现从高维特征映射为预测结果。

3.2. 优化混合神经网络基本结构

为了更好地优化 TCN-BiGRU 网络模型, 提出了 PSO-TCN-BiGRU 混合神经网络基本结构。利用 PSO 算法凭借自身全局搜索能力强的优点找出 TCN-BiGRU 模型的最优超参数, 避免陷入局部最优问题, 使得 TCN-BiGRU 的优点发挥得更好, 帮助模型达到更好的预测性能, 增强模型的鲁棒性, 使得模型在处理黄金期货时间序列数据时更加准确, 图 5 为 PSO 优化模型算法流程图:

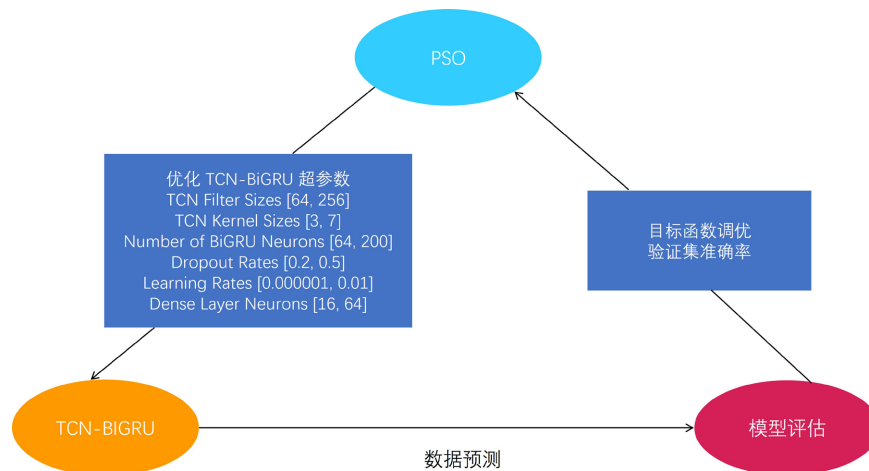


Figure 5. PSO optimization model algorithm flow chart
图 5. PSO 优化模型算法流程图

4. 实验结果与分析

4.1. 数据预处理与变量筛选

本实验中所有模型使用的数据选取以 2018 年 1 月 2 日至 2024 年 3 月 5 日六年每日的上海黄金期货交易所的黄金期货为基础, 并添加六大类数据, 包括经济政策不确定性、市场情绪描述、搜索引擎指数、股指期货变量、大宗期货变量以及宏观变量。经济政策不确定性数据有 GPR (地理政策风险), WLEMUIINDXD (股票市场相关经济不确定性指数), USEPUINDXD (美国经济政策不确定性指数)。市场情绪描述有 VIXCLS (标普 500 波动率指数), GVZCLS (黄金 ETF 波动率指数), OVXCLS (石油隐含波动指数)。搜索引擎指数有百度搜索指数。股指期货变量有 ACWI (MSCI 全球标准指数), CRB 现货综合指数, NASDAQ100 (纳斯达克 100 指数), SP500 (标准普尔 500 指数), 白银期货结算价, 道琼斯工业平均指数, 德国法兰克福 DAX 指数, 上证指数。大宗期货变量有 SPGSCI (标准普尔高盛综合商品指数), DHHNGSP (亨利港天然气现货价格), DCOILWTICO (库欣原油价格), DCOILBRETEU (欧洲原油价格),

黄金现货价格。宏观变量有 DEXUSEU (美元欧元汇率), DEXCHUS (人民币美元汇率)。考虑各变量间的交易日期存在差异, 本文以上海黄金期货交易所的黄金期货价格数据为基础进行筛选, 每个影响因素很可能受到滞后效应, 所以缺失的数值用上一个交易日的价格进行填补。筛选后的样本数据为 1490 条, 在实际工程问题中, 需要对数据进行预处理, 避免数据噪声对预测效果的影响, 本文采用 min-max 归一化 [24] 方法将数据缩放到 [0,1] 之间。

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{17}$$

其中: x 为处理前的数据, x' 为 min-max 归一化 [24] 后的数据。
因数据属性太多容易让预测模型产生数据冗余的现象, 因此需要进一步进行特征筛选操作。本文提出的特征筛选方法是可解释性 [25] 的贝叶斯优化的 SHAP-XGBoost 方法。贝叶斯优化 XGBoost 的调参的结果如表 1 所示:

Table 1. Bayesian optimization for XGBoost
表 1. 贝叶斯优化 XGBoost

参数名称	最优超参数
L2 正则项系数	0.00943
L1 正则项系数	0.10306
建立树时对特征采样的比例	0.6
每个决策树所用的子样本占总样本的比例	1.0
学习率	0.02
弱学习器个数	360
树的最大深度	13

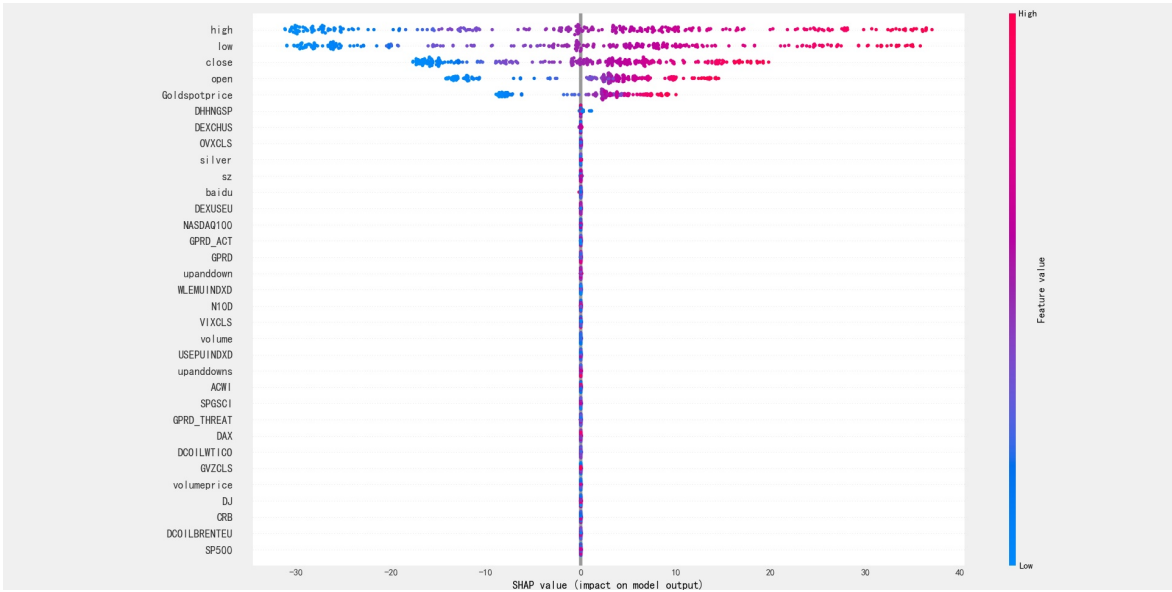


Figure 6. The global analysis graph of the importance of influencing factors
图 6. 影响因素重要性全局分析图

由于贝叶斯优化后的 XGBoost 模型的黑盒特性不利于实际生产应用, 因此本文引入 SHAP 方法对该

模型提供全局模型解释，局部模型解释以及利用可视化的方式返回每个特征变量对数据的贡献。图 6 为影响因素重要性全局分析图。

由影响因素重要性全局分析图可知前五个对黄金期货成交价最关键的特征分别是最高价“high”，最低价“low”，收盘价“close”，黄金现货价格“Goldspotprice”，开盘价“open”。

在建立模型之前，需要建立滑动时间窗口[26]构建样本，窗宽为 30，预测未来一天的黄金期货结算价。建立了样本数据后，按照 8:2 的比例对所有数据划分为训练集和测试集以便于训练网络模型。

4.2. 混合神经网络效果指标

在预测过程中，测试样本分别输入到所建立的七个神经网络中，将 7 个模型中误差最小的预测值作为该测试样本对应的黄金期货价格预测值。预测效果评价指标用四个指标均方误差(MSE) [27]、均方根误差(RMSE) [28]、平均绝对误差(MAE) [29]、决定系数(R²) [30]来表示。

MSE 的计算公式：

$$MSE = \frac{1}{m} \sum_{i=1}^m (x^i - y^i)^2 \tag{18}$$

RMSE 的计算公式：

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (x^i - y^i)^2} \tag{19}$$

MAE 的计算公式：

$$MAE = \frac{1}{m} \sum_{i=1}^m |x^i - y^i| \tag{20}$$

R² 的计算公式：

$$R^2 = 1 - \frac{\sum_{i=1}^m (x^i - y^i)^2}{\sum_{i=1}^m (x^i - \bar{y})^2} \tag{21}$$

式子中的 m 为观测值和真实值的个数；xⁱ 是观测值；yⁱ 是真实值； $\bar{y}$ 是均值。

4.3. 实验结果与分析

Table 2. Comparison of network predictive performance
表 2. 网络预测性能对比

算法	MAE	MSE	RMSE	R ²
CNN	8.4998	102.1839	10.1086	0.7619
BiGRU	8.3537	84.7488	9.2059	0.8025
CNN-BiGRU	5.4785	38.3592	6.1935	0.9106
TCN	5.1406	38.1166	6.1739	0.9112
BiGRU-TCN	4.5070	32.8927	5.7352	0.9233
TCN-BiGRU	4.3661	27.6335	5.2570	0.9356
PSO-TCN-BiGRU	3.9151	21.7743	4.6663	0.9493

将 PSO 优化算法的迭代次数设置为 200，种群大小为 100。PSO 算法的 c₁ 设置为 1.5，c₂ 设置为 1.5，

q 设置为 0.9。通过 PSO 算法得到 TCN 第一层滤波器尺寸大小为 128, 内核尺寸为 5, 第二层滤波器尺寸大小为 256, 内核尺寸为 5, 第三层滤波器尺寸大小为 256, 内核尺寸为 5, BiGRU 神经元数量为 100, Dropout 设置为 0.3, 学习率为 0.001, Dense 为 58 和 16, 迭代次数为 200。

为了验证 PSO-TCN-BiGRU 黄金期货价格预测算法的高效性, 构建 CNN、BiGRU、TCN、BiGRU-TCN、CNN-BiGRU、TCN-BiGRU 模型进行对比验证, 每个算法进行 200 次迭代实验, 计算预测效果评价指标, 具体结果如表 2 所示。

对比表 2 结果可以知: PSO-TCN-BiGRU 模型的 MAE、MSE、RMSE 比 CNN 模型降低了 4.5847、80.4096、5.4423, 而 R^2 比 CNN 模型提高了 18.74%。PSO-TCN-BiGRU 模型的 MAE、MSE、RMSE 比 BiGRU 模型降低了 4.4386、62.9745、4.5396, 而 R^2 比 BiGRU 模型提高了 14.68%。PSO-TCN-BiGRU 模型的 MAE、MSE、RMSE 比 CNN-BiGRU 模型降低了 1.5634、16.5849、1.5272, 而 R^2 比 CNN-BiGRU 模型提高了 3.87%。PSO-TCN-BiGRU 模型的 MAE、MSE、RMSE 比 TCN 模型降低了 1.2255、16.3423、1.5076, 而 R^2 比 TCN 模型提高了 3.81%。PSO-TCN-BiGRU 模型的 MAE、MSE、RMSE 比 BiGRU-TCN 模型降低了 0.5919、11.1184、1.0689, 而 R^2 比 BiGRU-TCN 模型提高了 2.6%。PSO-TCN-BiGRU 模型的 MAE、MSE、RMSE 比 TCN-BiGRU 模型降低了 0.451、5.8592、0.5907, 而 R^2 比 TCN-BiGRU 模型提高了 1.37%。PSO-TCN-BiGRU 模型的各项误差值都是最小的, R^2 是最高的, 与单一的神经网络相比, 本文提出的混合模型的准确性得到了提高。结果表明, 所采用的 PSO-TCN-BiGRU 混合模型可以提高模型的准确性, 并具有更好的预测性能。

为了更直观地表现出各种模型的差异, 本文显示 PSO-TCN-BiGRU 预测算法和其他预测算法的真实值和预测值的局部拟合效果。模型局部拟合效果图如图 7。

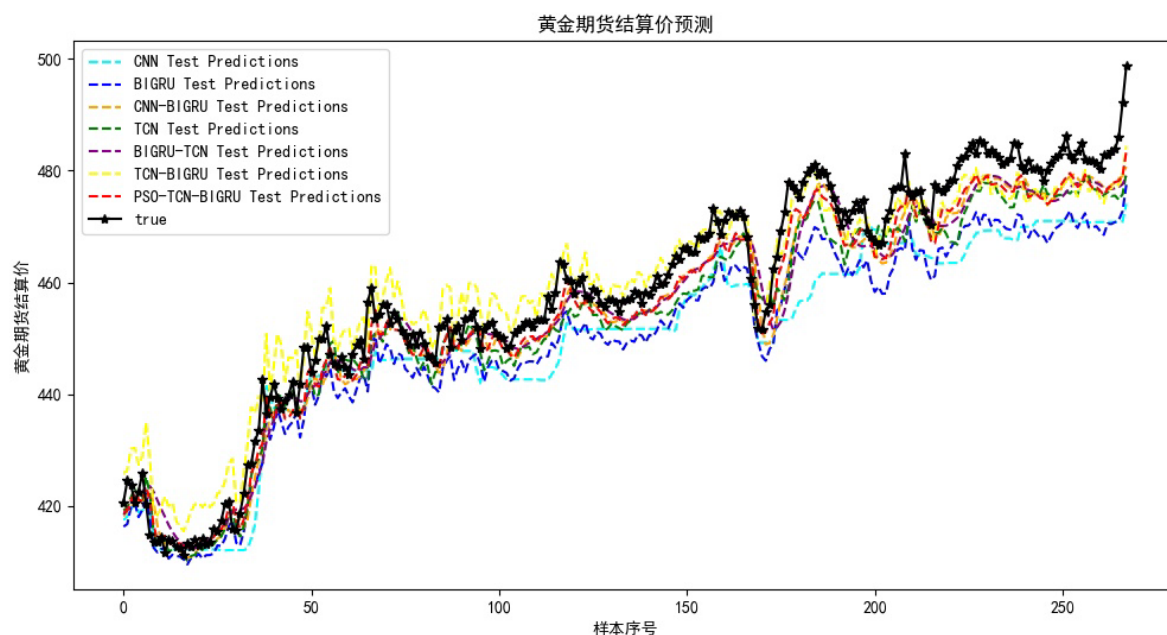


Figure 7. Effect of local fitting of the model

图 7. 模型局部拟合效果图

从图 7 可以看出这几种模型都可以预测出黄金期货结算价的大概走势, 但是 CNN、BiGRU、TCN、BiGRU-TCN、CNN-BiGRU、TCN-BiGRU 模型的预测结果没有本文提出的 PSO-TCN-BiGRU 模型拟合得更接近真实值。本文提出的 PSO-TCN-BiGRU 模型预测效果非常平稳, 与原数据变化保持一致, 拥有高

效率的跟随价格波动能力, 模型误差最小, 拟合优度更高。二者结合进一步说明了 PSO-TCN-BiGRU 模型比 CNN、BiGRU、TCN、BiGRU-TCN、CNN-BiGRU、TCN-BiGRU 模型的预测效果更佳, 精度更高。

5. 结论

本文寻找了六大类与黄金期货价格相关的特征, 为了提取关键变量引用最新的可解释机器学习模型贝叶斯优化 XGBoost-SHAP 方法提取前 5 个特征向量, 并将其导入 PSO-TCN-BiGRU 的深度学习神经网络模型, 对黄金期货价格进行预测, 通过设置对比分析实验可知: 本文提出的 PSO-TCN-BiGRU 预测模型相对于 CNN、BiGRU、TCN、BiGRU-TCN、CNN-BiGRU、TCN-BiGRU 模型具有更好的预测精度, 具有更好的拟合性, R^2 为 0.9493, MAE 为 3.9151, MSE 为 21.7743, RMSE 为 4.6663。从而验证了所提出的方法在黄金期货价格预测的有效性, 可为深度学习模型在金融预测领域探索发展提供参考决策。

基金项目

国家自然科学基金(71971070, 71903036)。

参考文献

- [1] 费婧文. 基于 ARIMA 模型对我国黄金期货价格分析与预测[J]. 当代经济, 2017(9): 148-150.
- [2] 丁磊, 郭万山. 基于 ARIMA-GARCH 族混合模型的黄金价格预测研究[J]. 许昌学院学报, 2019, 38(6): 124-129.
- [3] 曾濂, 马丹岷, 刘宗鑫. 基于 BP 神经网络改进的黄金价格预测[J]. 计算机仿真, 2010, 27(9): 200-203.
- [4] 刘璐, 娄磊, 刘先俊, 等. 基于 LSTM 的 COMEX 黄金期货价格预测对比研究[J]. 长春理工大学学报(自然科学版), 2021, 44(2): 129-135.
- [5] 杨光. 基于 CNN-LSTM 的黄金期货价格预测及风险度量研究[D]: [硕士学位论文]. 上海: 上海财经大学, 2023.
- [6] 尹晨曦. 基于 Lasso-PSO-BP 方法的中国黄金期货价格短期预测[J]. 洛阳理工学院学报(自然科学版), 2022, 32(4): 81-87.
- [7] 李进, 魏艳龙, 薛红新. 基于随机森林模型与 SHAP 算法的流感样病例影响因素分析研究[J]. 信息技术与信息化, 2024(2): 3-6.
- [8] 宋继勋, 周春雷, 沈子奇, 等. 基于 TCN-GRU 的 Handle 标识解析系统负载均衡算法[J]. 福建师范大学学报(自然科学版), 2024, 40(2): 64-73.
- [9] 贾颖, 赵峰, 李博, 等. 贝叶斯优化的 XGBoost 信用风险评估模型[J]. 计算机工程与应用, 2023, 59(20): 283-294.
- [10] Yang, D., Wang, L., Yuan, P., An, Q., Su, B., Yu, M., et al. (2023) Cocystal Virtual Screening Based on the XGBoost Machine Learning Model. *Chinese Chemical Letters*, 34, Article ID: 107964. <https://doi.org/10.1016/j.cclet.2022.107964>
- [11] 王向章, 牟瑞芳, 王赫, 等. 基于 GBDT-GS 方法的民机重着陆风险预测[J/OL]. 北京航空航天大学学报: 1-12. <https://doi.org/10.13700/j.bh.1001-5965.2023.0443>, 2024-12-08.
- [12] 李培坤, 陈旭梅, 鲁文博, 等. 基于 XGBOOST-SHAP 的地铁建成环境与站点出行距离的非线性关系研究[J]. 铁道科学与工程学报, 2024, 21(4): 1624-1633.
- [13] 林志军, 李敏, 贺珊, 等. 基于博弈论-贝叶斯网络的煤矿瓦斯爆炸风险评估[J]. 煤炭学报, 2024, 49(8): 3484-3497.
- [14] 孙丹铭, 陈长征, 孙业彭. 结合 CNN 与 TCN 神经网络的滚动轴承寿命预测[J]. 机械设计与制造, 2024(8): 160-165.
- [15] 张铮, 高森, 张泽扬. 基于特征构造预处理与 TCN-BiGRU 的池塘溶解氧预测模型[J]. 上海海洋大学学报, 2023, 32(5): 943-952.
- [16] 姜广君, 段张伟, 穆东明, 等. 基于时间卷积网络的机床齿轮箱轴承剩余寿命预测[J]. 机床与液压, 2024, 52(12): 224-230.
- [17] 王益, 杨剑波, 李锐, 等. 基于时空序列的 CNN-GRU 气温预测模型[J]. 计算机应用, 2023, 43(z2): 54-59.
- [18] 吴子伯, 崔云霞, 曹炜琦, 等. 基于 CEEMD-BiGRU 模型的徐州市大气污染物浓度预测[J]. 环境工程, 2022, 40(9): 9-18.

-
- [19] 熊亨, 戚湧, 张伟斌. 基于 DCGRU-RF 模型的路网短时交通流预测[J]. 计算机科学, 2020, 47(5): 84-89.
- [20] 朱茂桃, 刘欢, 吴余胤, 等. 混合策略改进的粒子群算法[J]. 重庆理工大学学报(自然科学), 2024, 38(1): 110-121.
- [21] 闫涵, 仲崇权, 吴玉虎, 等. 一种基于高斯过程与粒子群算法的 CNN 超参数自动搜索混合模型优化算法[J]. 信息与电子工程前沿(英文版), 2023, 24(11): 1557-1573.
- [22] 殷克东, 张凯, 杨文栋. 基于概率累加的离散 GM(1, 1)模型及其在海洋天然气产量预测中的应用[J]. 系统工程理论与实践, 2024, 44(8): 2733-2748.
- [23] 杨政, 吴浩成, 张靖, 等. 降维数据预测组合能提高股票超额收益的预测性能吗? [J]. 计量经济学报, 2023, 3(3): 828-847.
- [24] 刘铭, 单玉莹. 基于 EMD-LSTM 模型的股指收盘价预测[J]. 重庆理工大学学报(自然科学), 2021, 35(12): 269-276.
- [25] 温廷新, 白云鹤. 融合 RF-GA-XGBoost 和 SHAP 的虚假新闻群体互动质量可解释模型[J/OL]. 数据分析与知识发现: 1-18. <http://kns.cnki.net/kcms/detail/10.1478.G2.20240117.1108.018.html>, 2024-12-08.
- [26] 郭杨莉, 马锋. 基于马尔科夫和混频数据模型的黄金期货市场波动率预测研究[J]. 中国管理科学, 2024, 32(1): 13-22.
- [27] 陈云峰, 于雪, 刘吉成, 等. 基于 TSO-LS-SVM 模型的电煤库存风险评价研究[J/OL]. 中国管理科学: 1-8. <https://doi.org/10.16381/j.cnki.issn1003-207x.2023.1065>, 2024-12-08.
- [28] 何少芳, 沈陆明, 谢红霞. 生成式对抗网络的土壤有机质高光谱估测模型[J]. 光谱学与光谱分析, 2021, 41(6): 1905-1911.
- [29] 刘思佳, 黄海. 基于人工智能的多样性通用数据周期推荐算法[J]. 系统科学学报, 2023, 31(1): 61-65.
- [30] 王可睿, 邵必林. 基于 Time2Vec-LSTM-TCN-Attention 的天然气负荷组合预测[J/OL]. 南京信息工程大学学报: 1-13. <https://doi.org/10.13878/j.cnki.jnuist.20240104002>, 2024-12-08.