

基于Bagging集成算法的COMEX黄金期货价格预测

李 辰

河南工业大学土木建筑学院, 河南 郑州

收稿日期: 2025年2月26日; 录用日期: 2025年3月19日; 发布日期: 2025年3月28日

摘 要

针对全球金融市场复杂性问题, 构建准确的黄金价格预测模型显得尤为重要。文章选取多个影响黄金价格的主要因素, 考虑宏观经济、金融市场、政治风险等多个维度, 运用斯皮尔曼分析法筛选指标并建立指标体系框架。采用灰狼优化算法(Grey Wolf Optimizer, GWO)优化极端梯度提升模型(eXtreme Gradient Boosting, XGBoost), 构建GWO-XGBoost模型, 并将其作为基础学习器, 建立集成模型Bagging-GWO-XGBoost。实验结果显示, 优化算法提升了模型在COMEX黄金期货价格预测中的表现; 当加入集成算法后, 模型的RMSE、MAE和MAPE均有所改善。因此, 该模型能够为黄金价格预测提供更加精准的预测数据, 被视为一种实用高效的金融工具, 帮助金融市场进行预测与决策。

关键词

黄金期货, 价格预测, 机器学习, 集成学习

Prediction of COMEX Gold Futures Prices Based on the Bagging Integrated Algorithm

Chen Li

College of Civil Engineering and Architecture, Henan University of Technology, Zhengzhou Henan

Received: Feb. 26th, 2025; accepted: Mar. 19th, 2025; published: Mar. 28th, 2025

Abstract

Given the complexities of global financial markets, constructing an accurate model for gold price prediction is of significant importance. This study selects multiple key factors influencing gold

prices, considering various dimensions such as macroeconomics, financial markets, and political risks. Spearman's correlation analysis is employed to filter indicators and establish an indicator system framework. The Grey Wolf Optimizer (GWO) is utilized to optimize the eXtreme Gradient Boosting model (XGBoost), resulting in the development of a GWO-XGBoost model. This model is then used as a base learner to construct an ensemble model, Bagging-GWO-XGBoost. Experimental results demonstrate that the optimization algorithm enhances the model's performance in predicting COMEX gold prices. Furthermore, the integration of the ensemble algorithm leads to improvements in the model's RMSE, MAE, and MAPE metrics. Therefore, this model can provide more accurate predictive data for gold price forecasting and is considered a practical and efficient financial tool, aiding in market predictions and decision-making processes.

Keywords

Gold Futures, Price Prediction, Machine Learning, Integrated Learning

Copyright © 2025 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

黄金作为一种兼具货币属性与商品属性的贵金属，在人类经济发展史上始终扮演着重要角色。黄金不仅是商品交换的一般等价物，也是全球经济体系中对抗通胀的重要工具。作为避险资产，它能够帮助抵御全球经济的不确定性和系统性风险。此外，黄金在珠宝设计和航天领域的应用也彰显独特价值。近年来，随着地缘政治风险上升、货币政策不确定性增加以及全球经济结构性矛盾的凸显，黄金的避险属性愈发受到市场关注。在此背景下，各国央行持续调整外汇储备结构，增持黄金以增强金融体系的稳定性。中国黄金交易分为现货和期货交易市场，黄金期货因其杠杆和风险管理功能成为重要投资工具。

当前用于预测黄金价格的研究模型主要可分为三类：传统计量经济学模型、机器学习模型和组合模型。例如，徐静怡和孔梦奇[1]通过使用自回归积分滑动平均(Autoregressive Integrated Moving Average, ARIMA)模型对 COMEX 黄金历史价格数据进行时间序列分析，最终得出未来短期会小幅上涨的结论。但是，时间序列模型难以捕捉外部影响因素的信息，因此，又有学者使用深度学习方法进行预测，如尹晨曦[2]使用经粒子群优化算法(Particles Swarm Optimization, PSO)优化的 BP 神经网络模型对上海黄金交易所的黄金期货收盘价进行多变量预测，经过对比证明所提出的 PSO-BP 模型具有较强的预测能力。随着人工智能的发展，一些学者提出了将分解算法结合机器学习模型加以应用，如，梁龙跃和黄盈[3]使用自适应噪声的完全集合经验模态分解(Complementary Ensemble Empirical Mode Decomposition with Adaptive Noise, CEEMDAN)方法和长短期记忆(Long Short-Term Memory, LSTM)网络结合的方式对 COMEX 黄金期货价格进行预测，经过实验发现，对数据先做分解再做预测可以有效提升模型的效果。但是，这一方法的建模步骤比较复杂且耗时，导致预测时间较长。

针对当前研究的不足，本文提出了一种基于 Bagging 集成算法的预测模型。该模型以 GWO-XGBoost 为基模型，其优势在于结合了 GWO 算法对超参数的高效优化能力与 XGBoost 的强大预测能力，从而显著提升模型的性能和泛化能力，且该模型在处理复杂和非线性数据时表现优异，在具有噪声的数据集中具有良好的鲁棒性。基于此，本文将其与 Bagging 集成算法相结合，以进一步提升模型的预测精度和稳定性，为黄金价格预测提供更加高效和准确的解决方案。

2. 算法介绍与模型构建

2.1. 灰狼优化算法

灰狼优化算法(Grey Wolf Optimizer, GWO)是由 Mirjalili 等[4]开发的一种基于群体智能的元启发式算法,其灵感来源于灰狼种群的社会结构。在灰狼的社会等级结构中,有 α 狼、 β 狼、 δ 狼和 ω 狼四个不同等级的灰狼,其社会地位从左到右依次递减,上级灰狼对下级灰狼具有绝对的支配权。四个等级的狼群分别代表了 GWO 优化过程中搜索到的四个解,其中 α 狼代表最优解, β 狼和 δ 狼代表两个次优解, ω 狼代表候选解。最出色的前三名分别为 α 狼、 β 狼和 δ 狼,它们将带领其他狼朝着目标前进。与此同时,其余的狼被归类为 ω 狼,它们会调整自己与 α 、 β 或 δ 狼的位置关系。该算法包括灰狼的社会等级、包围、狩猎和攻击猎物等步骤。

2.2. 极端梯度提升

极端梯度提升(eXtreme Gradient Boosting, XGBoost)是一种高效且强大的机器学习算法,通常用于解决分类和回归问题,由 Chen 和 Guestrin [5]开发。基于梯度提升决策树(Gradient Boosting Decision Tree, GBDT)的思想,通过在选择性生成中训练多个决策树模型来提高预测性能。与 GBDT 相比, XGBoost 通过二阶泰勒展开扩展了目标函数,从而保留了更多的客观信息并提高了其准确性。XGBoost 支持正则化技术,以避免过拟合问题,该算法还可以通过最小化损失函数的梯度来优化模型参数,从而提高训练效率和准确性。在精度、可解释性和鲁棒性等方面, XGBoost 是一种非常强大的机器学习算法,广泛应用于数据挖掘、金融风险管理等领域。由于该算法有很多超参数,如 `learning_rate`、`n_estimators`、`max_depth` 等,因此我们可以使用优化算法来调整 XGBoost 的参数以实现最佳预测。XGBoost 模型的核心在于通过构建新的决策树优化结构,将每棵树的预测分数相加作为预测模型的预测值,将预测结果的残差拟合并优化为预测分数。

2.3. Bagging 集成学习

Bagging [6]的核心思想是自助法采样和模型的集成。该方法通过对原始数据集进行自助法采样,生成多个不同的训练子集,并在每个子集上独立训练一个基学习器。由于每个基学习器是在不同的数据子集上训练的,因此它们具有多样性,能够有效降低整体模型的方差。在预测阶段, Bagging 通过平均所有基学习器的预测结果,获得最终的预测值。这种方法不仅提高了模型的稳定性,还能在一定程度上减少过拟合现象,使得 Bagging 在回归问题中展现出良好的性能。当数据集较大且基模型容易过拟合时,可以使用 Bagging 来降低过拟合风险并提高预测精度,故本文选取 Bagging 作为集成方法。

2.4. Bagging-GWO-XGBoost 模型构建

由于黄金价格序列数据具有较大的波动性和复杂性,引入 Bagging 算法可以降低过拟合的风险,因此,将 Bagging、GWO 和 XGBoost 模型进行组合,形成 Bagging-GWO-XGBoost 模型。图 1 为本次实验的流程图,具体建模步骤如下。

首先,数据预处理。将选定的 13 个外部影响因素和历史相关数据作为模型的输入特征,对时间序列数据进行差分以消除趋势,并对所有数据进行归一化以消除量纲差异。将处理后的数据按照 8:2 的比例划分为训练集和测试集。

其次,寻找最佳参数。使用灰狼优化算法为 XGBoost 模型寻找最佳参数值。首先初始化灰狼种群,并设定迭代次数和种群大小;评估每个灰狼个体的适应度,适应度函数基于回归性能评价指标 MSE 来评估;之后不断迭代更新灰狼位置,直到达到最大迭代次数。

最后，模型集成。采用 Bagging 算法对经过 GWO 优化后的 XGBoost 模型进行集成操作。使用自助法从训练集中抽取多个不同的数据子集，并在每个子集上使用 GWO 算法找到的最优参数来训练基学习器。最后，通过平均所有基学习器的预测结果来得到集成模型的最终预测值。

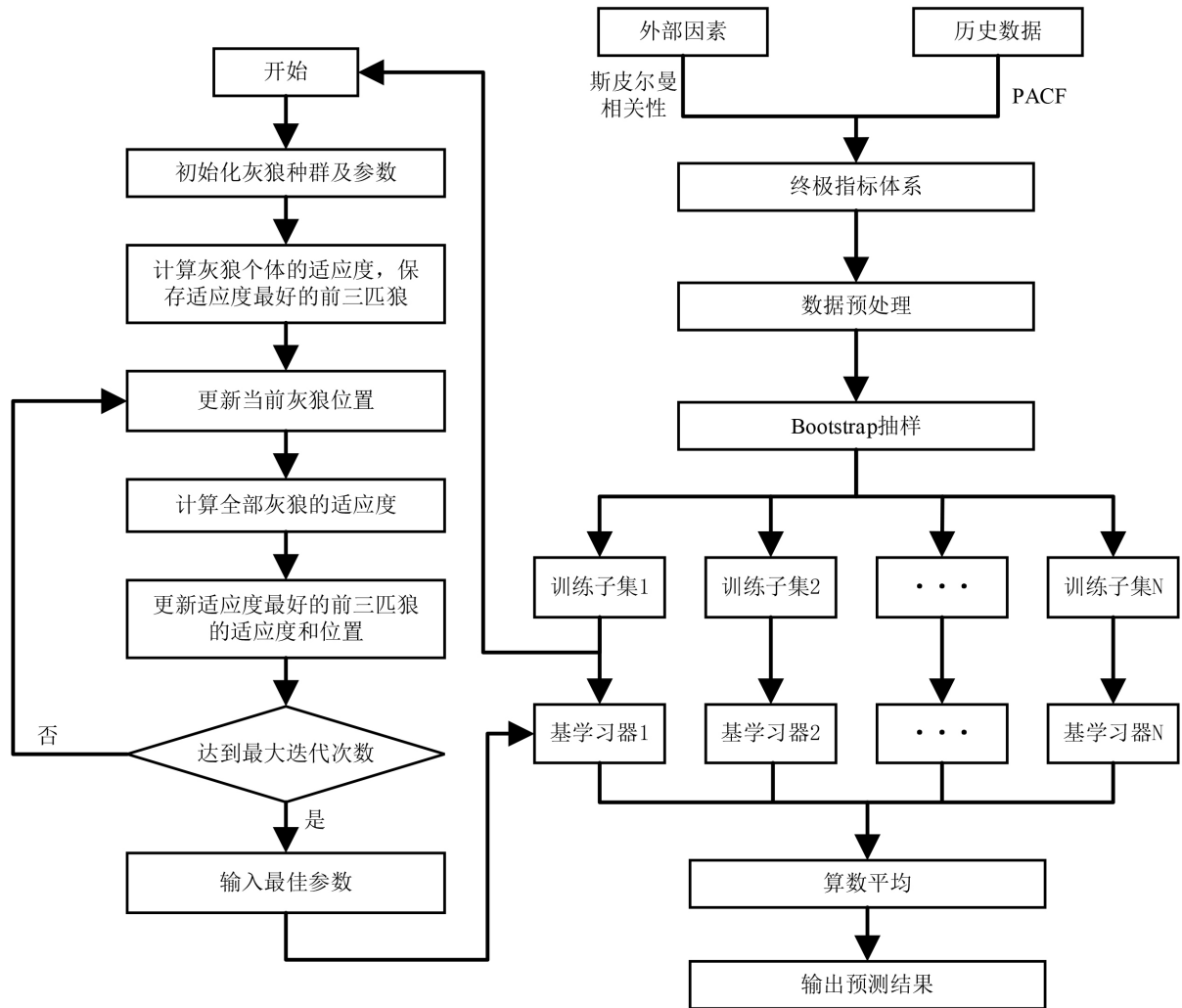


Figure 1. Bagging-GWO-XGBoost model prediction flowchart
图 1. Bagging-GWO-XGBoost 模型预测流程图

3. 指标体系建立

3.1. 数据来源

黄金期货市场种类繁多，其中纽约金、伦敦金和沪金作为主流黄金期货备受关注。由于纽约商品交易所(COMEX)具有高交易量和活跃度，成为全球重要黄金期货市场，其合约价格显著影响全球金价。鉴于其历史地位与与现货市场的配合，纽约金期货交易占主导地位，因此本研究选择 COMEX 黄金期货收盘价为研究对象。数据涵盖了 2017 年 12 月 1 日至 2024 年 3 月 20 日的每日交易数据，共计 1620 个样本，数据来源于 Choice 金融终端。在实验中，数据集的前 80%作为训练集，后 20%作为测试集。图 2 展示了所选取时间范围内的 COMEX 黄金期货收盘价走势。

由于黄金价格受到如白银价格、原油价格、美元等多方面因素影响，因此，有必要建立影响 COMEX 黄金期货价格的指标体系。通过对有关文献进行梳理，结合多方面影响，最终归纳出六类能够影响金价走势的外部因素作为初级指标体系，如表 1 所示。其中，X1~X12 来源于 Choice 金融终端软件，X13 来源于网站(<https://www.policyuncertainty.com/>)。



Figure 2. COMEX gold futures closing price trend
图 2. COMEX 黄金期货收盘价走势

Table 1. External factors affecting COMEX gold futures prices
表 1. 影响 COMEX 黄金期货价格的外部因素

类别	外部影响因素	简称	序号	文献依据
金属	期货收盘价：COMEX 银	Silver	X1	[7] [8]
能源	期货收盘价：WTI 原油	Oil	X2	[9]
宏观经济	美国 CPI	CPI	X3	[10]-[12]
	美国 PPI	PPI	X4	[12] [13]
	美国联邦基金利率	FFR	X5	[14]-[16]
	美国政府财政赤字	FRI	X6	[17]
金融市场	标准普尔 500 波动率指数	VIX	X7	[18] [19]
	美元指数	USDI	X8	[20]
	美元兑人民币汇率	USD_CNY	X9	[2] [21]
	美国十年期国债收益率	TY10	X10	[20] [22]
景气调查	密歇根大学消费者信心指数：终值	CCI	X11	[18] [23]
	Sentix 投资信心指数	SICI	X12	[18]
风险	地缘政治风险指数	GPRD	X13	[16] [18]

3.2. 斯皮尔曼相关性分析

为进一步探究所选取初级指标体系的有效性,采用斯皮尔曼相关性分析法进行验证。使用 SPSS27 软件进行分析,结果如表 2 所示。可以发现,本实验选取的 9 个变量的 p 值均小于 0.05,即与 COMEX 黄金期货收盘价均具有显著相关性。其中, X1、X2、X3、X4、X5、X6、X7、X8、X9、X10、X13 和金价显示为正相关, X11 和 X12 和金价显示为负相关。

Table 2. Spearman correlation analysis
表 2. 斯皮尔曼相关性分析

变量	斯皮尔曼相关性系数	P 值
Silver (X1)	0.830**	0.000
Oil (X2)	0.366**	0.000
CPI (X3)	0.674**	0.000
PPI (X4)	0.664**	0.000
FFR (X5)	0.112**	0.000
FRI (X6)	0.145**	0.000
VIX (X7)	0.210**	0.000
USDI (X8)	0.317**	0.000
USD_CNY (X9)	0.201**	0.000
TY10 (X10)	0.141**	0.000
CCI (X11)	-0.713**	0.000
SICI (X12)	-0.382**	0.000
GPRD (X13)	0.199**	0.000

注: **为 $p < 0.05$ 。

3.3. 建立终极指标体系

由于黄金价格具有非平稳性,因此对数据进行平稳性检验是有必要的,否则可能导致伪回归问题[24]。采用 SPSSPRO 软件对数据进行 ADF 单位根检验分析,结果如表 3 所示。可以发现, COMEX 黄金期货收盘价原始数据是不平稳的,而在一阶差分后, ADF 统计量值为-10.919, 小于 1%水平的临界值, p 值为 0.000, 可以认为显著的拒绝原假设, 这表明一阶差分后的数据具有平稳性。此外, 本研究还加入了历史滞后数据作为模型的输入变量, 图 3 展示了 COMEX 黄金期货收盘价数据在一阶差分后的偏自相关性 (PACF)结果。设 Y_i 为输出特征, 则 $\{Y_{i-7}, Y_{i-9}, Y_{i-13}, Y_{i-15}\}$ 可以作为数据集的历史输入变量。因此, 本文所提出的预测模型的最终输入变量为 $\{X1, X2, X3, X4, X5, X6, X7, X8, X9, X10, X11, X12, X13, Y_{i-7}, Y_{i-9}, Y_{i-13}, Y_{i-15}\}$ 。

Table 3. ADF test for COMEX gold futures closing price data
表 3. COMEX 黄金期货收盘价数据的 ADF 检验

差分阶数	t	P	临界值		
			1%	5%	10%
0	-0.648	0.860	-3.434	-2.863	-2.568
1	-10.919	0.000***	-3.434	-2.863	-2.568
2	-13.481	0.000***	-3.434	-2.863	-2.568

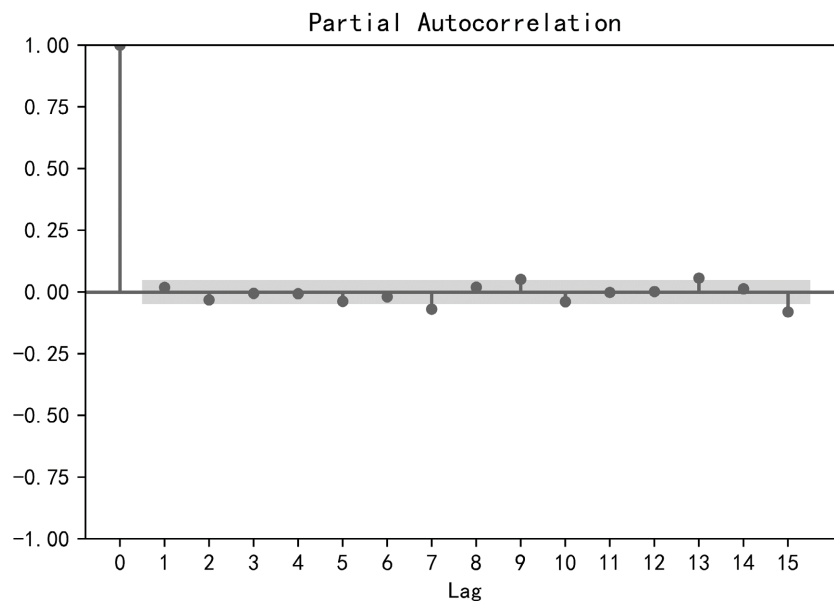


Figure 3. Results of partial autocorrelation function
图 3. 偏自相关函数结果

3.4. 归一化

由于各个指标数据之间存在不同量值和单位,为了使数据具有可比性,有必要将数据做归一化处理。详见公式(1):

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

其中, x^* 是归一化值, x 是原始数据, $x_{\min}$ 是数据集中的最小值, $x_{\max}$ 是数据集中的最大值。归一化后的数据在[0, 1]范围内。

3.5. 评价指标

为了直观地评估每个模型的预测能力,采用四种评估指标以全面反应模型的性能,包括均方根误差(RMSE)、平均绝对误差(MAE)、平均绝对百分比误差(MAPE)和决定系数(R^2)。当 MAE、RMSE 和 MAPE 值较低且 R^2 值接近于 1 时,模型就更准确。各指标的计算原理详见公式(2)~(5):

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

$$MAE = \frac{1}{n} \sum_{i=1}^n |(\hat{y}_i - y_i)| \quad (3)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (4)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\bar{y}_i - y_i)^2} \quad (5)$$

其中, n 为样本数, y_i 表示 COMEX 黄金期货收盘价的实际值, $\hat{y}_i$ 表示 COMEX 黄金期货收盘价的预测

值， $\bar{y}_i$ 表示 COMEX 黄金期货收盘价的平均值。

4. 实证分析

4.1. 实验环境与参数设置

本次仿真实验是在 Anaconda2.4.0 中 Python3.10 环境中进行，使用的是具有 Intel Core i5-13500 H CPU、Intel (R) Iris (R) Graphics GPU 和 16 G RAM 的 Windows 11 系统笔记本电脑。为验证所提出的“集成 - 优化”模型是最佳模型类型，本文使用多组模型进行对照试验，从单一机器学习模型，到加入优化算法的模型，再到加入集成技术，逐步验证当前研究模型的优越性。根据之前的研究[25]，本实验的优化算法参数设置为 30 次迭代，种群大小为 20。表 4 显示了各算法的参数设置，其中 WOA-XGBoost 和 GWO-XGBoost 显示了各自优化后的参数数值。

Table 4. Parameter settings
表 4. 参数设置

模型名称	参数值
SVR	默认值
XGBoost	默认值
WOA-XGBoost	n_estimators = 27, colsample_bytree = 0.428434, learning_rate = 0.101271, max_depth = 3.00281, subsample = 0.897977
GWO-XGBoost	n_estimators = 20, colsample_bytree = 0.410239, learning_rate = 0.106116, max_depth = 3.6425, subsample = 0.835404
Bagging-GWO-XGBoost	random_state = 80, n_estimators = 13 (Bagging); n_estimators = 20, colsample_bytree = 0.762746, learning_rate = 0.102715, max_depth = 3.65665, subsample = 0.88951

4.2. 实验结果

本研究使用机器学习和集成相结合的方法，考虑外部影响因素和历史数据作为模型的输入特征，并引入四个对比模型作为基准，体现所提出模型的有效性。实验结果如图 4 和表 5 所示，结果表明，与其他基准模型相比，Bagging-GWO-XGBoost 模型对 COMEX 黄金期货收盘价的预测表现出优异的预测性能。通过比较实证结果，可以得到以下结论。

首先，对于单一机器学习模型，我们设置了 SVR 与 XGBoost 进行对比。XGBoost 在此次实验中表现出更好的性能，表明此模型在处理黄金价格预测问题时能够更有效地捕获数据中的非线性关系。XGBoost 作为一种基于决策树的集成学习方法，它通过构建多棵决策树并进行集成，从而提高了模型的预测能力和鲁棒性。在黄金价格预测中，存在许多影响价格变动的因素，包括全球经济状况、货币政策、地缘政治事件等，这些因素之间的相互作用和影响往往是复杂且非线性的，故 XGBoost 模型在处理这种复杂性方面可能更具优势。

其次，优化算法在提升模型预测能力方面具有较强重要性。与没有加入优化算法的模型(模型 2)相比，加入优化算法(模型 3 和模型 4)可以显著提升模型的预测性能，评价指标均得到了大幅提升，说明优化算法有助于识别出模型的最佳参数以提高预测效果。此外，根据本次实验，GWO 比 WOA 在寻找最佳参数方面具有更强的能力。

最后，集成算法能够在原有模型的基础上再次提升预测效果。与模型 4 相比，模型 5 的 R^2 仍有进一

步的提升，且 RMSE、MAE、MAPE 分别提升了 15.56%、7.81%、7.83%。由于 Bagging 是同质集成模型，通过组合多个 GWO-XGBoost 模型进行集成，可以显著降低单一机器学习模型可能存在的过拟合现象，并且通过平均的方式综合多个模型的预测结果，使得整体模型的预测性能得到提升。

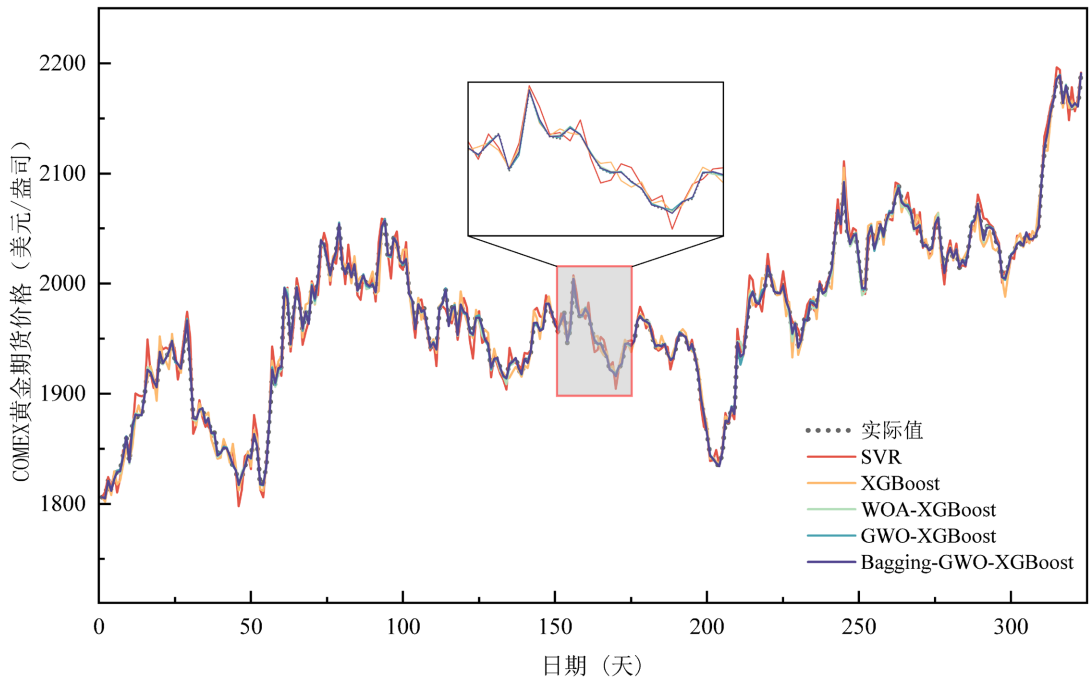


Figure 4. Comparison of predicted and actual values of the models
图 4. 各模型预测值与实际值的比较

Table 5. Predicted results of evaluation indicators
表 5. 评价指标预测结果

序号	名称	RMSE	MAE	MAPE	R ²
1	SVR	8.335946	6.361137	0.323745	0.988038
2	XGBoost	7.917456	5.892560	0.299844	0.989209
3	WOA-XGBoost	2.025161	1.115733	0.056828	0.999294
4	GWO-XGBoost	1.249284	0.754294	0.038360	0.999731
5	Bagging-GWO-XGBoost	1.083062	0.708447	0.036033	0.999798

4.3. 特征重要性分析

沙普利加法解释(SHapley Additive exPlanations, SHAP)是一种在机器学习模型中用于解释预测模型的工具，可以计算变量的重要性，解释输入变量对输出的影响，为模型的预测结果提供直观的解释，有助于理解模型的特征依赖性。蜂群图作为 SHAP 图像的一种表现形式，凭借其直观性和全面性，成为了全局解释中的经典图像。根据图 5 可以发现，占主导的因素是相关历史时间序列，除此之外，比较重要的外部变量还有消费者信心指数、美国十年期国债收益率和白银价格。

消费者信心指数是衡量经济预期的关键指标，显著影响黄金价格。当该指数上升时，经济增长预期增强，投资者就会减少避险资产需求，因此就可能引发黄金价格下跌。反之，当指数下跌，即经济增长

预期减弱，投资者避险情绪上升，黄金价格就会相应上涨。因此，消费者信心指数对于黄金价格的预测影响较大，我们应对该指数的发展趋势多加关注，以便于更准确评估黄金市场的走势。

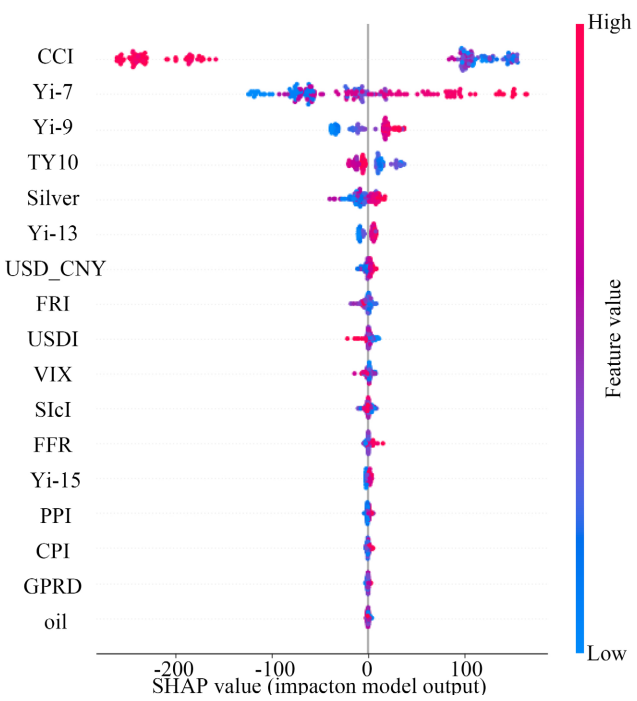


Figure 5. Feature importance analysis
图 5. 特征重要性分析

美国十年期国债收益率对 COMEX 黄金价格的影响主要反应在市场情绪和经济政策方面。当国债收益率上升，可以认为无风险的收益提升，即市场对美国经济前景的乐观预期增高，避险情绪减弱，投资者更倾向于将资金投入固定收益资产，而不是贵金属等避险资产，因此会导致金价下跌。

白银与黄金作为贵金属，在金融市场中表现出显著的价格相关性，二者价格通常同向变动。在金融市场动荡时，投资者往往会将贵金属视为风险对冲的重要工具。在白银与黄金的比较中，由于白银的价格相对较低，对于投资者而言，往往具备更高的投资性价比，可成为更具吸引力的选择。白银市场的价格波动能够间接影响黄金市场的供求平衡与价格变动，表明黄金与白银之间存在显著的正向价格影响关系。这一关系不仅体现了贵金属市场内部的紧密联系，也反映了投资者在不确定市场环境下的投资偏好与策略选择。

4.4. 附加实验

为了进一步验证模型的普适性和泛化能力，本研究进行了一次补充实验，选取两个不同时间范围的数据样本。首先，对 COMEX 黄金收盘价数据进行 ADF 检验，发现在一阶差分后的数据呈现平稳性，基于此，对一阶差分后的数据进行偏自相关性检验，结果显示，时间范围 1 的偏自相关函数结果为历史数据的第 14 天；时间范围 2 的历史数据为第 4 天和 13 天。之后，通过斯皮尔曼相关性检验分析剔除相关性不显著的指标。综上所述，本文所提出的预测模型的最终输入变量分别为{X1, X2, X3, X4, X5, X6, X7, X8, X9, X10, X12, Yi-14}和{X1, X2, X3, X4, X5, X6, X7, X8, X9, X10, X11, X12, Yi-4, Yi-13}。具体预测结果如表 6 所示。可以看出，在新的时间范围内，模型的预测结果和原始实验的结论相符，随着优化算法和集成算法的增加，模型的预测效果得到显著提高。因此，可以认为本研究所提出的“集成 - 优化”模

型对于预测黄金价格走势有重要意义。

Table 6. Predicted results of supplementary experiments
表 6. 补充实验的预测结果

序号	数据集	名称	RMSE	MAE	MAPE	R ²
1	2017 年 8 月 29 日至 2019 年 12 月 27 日	XGBoost	4.818809	3.839638	0.258251	0.976725
		WOA-XGBoost	1.961723	1.525340	0.102202	0.996143
		GWO-XGBoost	1.050806	0.827475	0.055564	0.998893
		Bagging-GWO-XGBoost	0.993206	0.792751	0.053300	0.999011
2	2022 年 2 月 28 日至 2024 年 7 月 31 日	XGBoost	8.672165	6.741818	0.291710	0.995441
		WOA-XGBoost	2.456421	1.769838	0.076497	0.999634
		GWO-XGBoost	1.681314	1.202422	0.052000	0.999829
		Bagging-GWO-XGBoost	1.528968	1.146389	0.049687	0.999858

5. 结论

本文基于外部影响因素和历史滞后时间数据，提出了一个“集成 - 优化”模型用于 COMEX 黄金收盘价预测。和其他模型相比，所提出的模型具有较好的预测能力，能够为金融市场提供一个准确的预测模型。在实证分析的基础上，本文的主要结论可归纳如下：

1) XGBoost 是一个比较适合用于黄金价格预测领域的模型，该模型基于 Boosting 集成学习，具有较强的泛化能力。在该模型的基础上加入智能优化算法，有助于基础模型找到最佳参数，从而大幅度提升模型的预测效果。

2) Bagging 算法通过集成多个基学习器的预测结果，有效降低了模型的方差，提高了预测的稳定性和准确性。该算法通过随机抽样生成不同训练子集，并在其上独立训练基学习器，从而降低了过拟合风险。金融时间序列数据具有复杂结构和噪声，在处理问题时，Bagging 算法展现出显著的优越性，能够捕捉全局趋势并减少预测偏差。

3) 外部影响因素对于黄金价格的预测至关重要。黄金价格的波动性较大，仅考虑历史价格数据不足以准确预测，为研究对象建立指标体系能够更全面、系统地考虑金价波动的原因，能够提升预测模型的准确性和实用性。通过深入分析全球经济形势、货币政策、地缘政治事件等外部因素，结合技术指标和市场情绪，构建一个综合的预测框架，可以更加精确地捕捉黄金价格的动态变化。

4) SHAP 能够增强模型的可解释性。由于黄金价格受多重因素影响，SHAP 方法可以从特征贡献程度解释各指标对输出变量的影响。根据本研究的实验结果，对 COMEX 黄金收盘价影响较大的因素有历史价格、消费者信心指数、美国十年期国债收益率以及白银价格等。

在当前金融市场波动不断的背景下，本文提出的集成预测模型能够为投资者提供一个强大的金融分析工具。该模型综合考虑了宏观经济因素、金融市场指数、地缘政治风险等多种影响因素，并运用“集成 - 优化”学习算法，为金融风险管理提供坚实的科学依据。

参考文献

[1] 徐静怡, 孔梦奇. 基于 ARIMA 模型对纽约 COMEX 黄金期货价格的研究[J]. 中国商论, 2022(18): 123-125.
[2] 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.

- [3] 梁龙跃, 黄盈. 黄金期货价格短期预测方法研究——基于 CEEMDAN 与 LSTM 模型的 COMEX 黄金期货价格数据分析[J]. 价格理论与实践, 2023(9): 164-168.
- [4] Mirjalili, S., Mirjalili, S.M. and Lewis, A. (2014) Grey Wolf Optimizer. *Advances in Engineering Software*, **69**, 46-61. <https://doi.org/10.1016/j.advengsoft.2013.12.007>
- [5] Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [6] Breiman, L. (1996) Bagging Predictors. *Machine Learning*, **24**, 123-140. <https://doi.org/10.1007/bf00058655>
- [7] Bouri, E. and Jalkh, N. (2019) Conditional Quantiles and Tail Dependence in the Volatilities of Gold and Silver. *International Economics*, **157**, 117-133. <https://doi.org/10.1016/j.inteco.2018.10.001>
- [8] Koullis, A. and Kyriakopoulos, C. (2023) On Volatility Transmission between Gold and Silver Markets: Evidence from a Long-Term Historical Period. *Computation*, **11**, Article No. 25. <https://doi.org/10.3390/computation11020025>
- [9] Mo, B., Nie, H. and Jiang, Y. (2018) Dynamic Linkages among the Gold Market, US Dollar and Crude Oil Market. *Physica A: Statistical Mechanics and its Applications*, **491**, 984-994. <https://doi.org/10.1016/j.physa.2017.09.091>
- [10] 曹万解. 国内黄金市场发展现状及黄金期货投资机会研究[J]. 商展经济, 2023(24): 97-100.
- [11] Qian, Y., Ralescu, D.A. and Zhang, B. (2019) The Analysis of Factors Affecting Global Gold Price. *Resources Policy*, **64**, Article ID: 101478. <https://doi.org/10.1016/j.resourpol.2019.101478>
- [12] 张永康, 晏乐惠, 张萌, 等. 我国焦煤期货市场研究[J]. 中国市场, 2024(15): 26-29.
- [13] 周子朝. 国际原油与黄金价格对我国通货膨胀的影响[J]. 北方经贸, 2022(10): 41-44.
- [14] 万博, 张华硕. 后金融危机时代国际黄金价格影响因素实证分析[J]. 科技和产业, 2014, 14(8): 168-172.
- [15] 吴子弘. 中美贸易摩擦对国际黄金价格变动影响研究[D]: [硕士学位论文]. 昆明: 云南财经大学, 2023.
- [16] 郑明贵, 曹天琦, 曾健林. 国际黄金期货价格影响因素研究[J]. 黄金科学技术, 2021, 29(4): 510-524.
- [17] 刘瑶. 财税政策与人民币国际化的互动路径研究[J]. 国际金融, 2023(6): 19-27.
- [18] 刘晓星, 汤淳, 杨广义. 金融安全风险演化: 理论与实证[J]. 兰州财经大学学报, 2023, 39(6): 18-30.
- [19] 温彬, 李鑫. 金价长期走势的六个典型化事实[J]. 银行家, 2024(7): 92-100.
- [20] 张若钦, 王刚. 美元指数和美国国债收益率对国际黄金价格的非线性影响——基于 STR 模型的研究[J]. 价格月刊, 2013(6): 37-40.
- [21] 杨雪倩. 基于 CNN-BiGRU 模型的黄金期货价格预测[D]: [硕士学位论文]. 开封: 河南大学, 2023.
- [22] 翁赫. 中国黄金期现货价格关联及基差影响因素分析[J]. 黄金, 2023, 44(8): 5-9.
- [23] 马文淑. 基于组合模型的人民币汇率预测实证研究[D]: [硕士学位论文]. 南京: 南京审计大学, 2022.
- [24] 吴虹晓. 中国黄金期货和现货价格关系的实证研究[J]. 中国货币市场, 2024(2): 71-75.
- [25] Duan, Y., Zhang, J. and Wang, X. (2023) Henry Hub Monthly Natural Gas Price Forecasting Using CEEMDAN-Bagging-HHO-SVR. *Frontiers in Energy Research*, **11**, Article ID: 1323073. <https://doi.org/10.3389/fenrg.2023.1323073>