Computer Fraud and Security
ISSN (online): 1873-7056

A Comparative Study on Gold Future Forecasting Based on Time Series Large Models and Classic Models

Xuecheng Wang[1], Ruixin Hu[2,*], Xiaoyan Chen[1]
[1]*School of Economics and Management, North China University of Technology, Beijing, China*
[2]*School of law and Economics, Zhongnan University of Economics and Law, Wuhan, China*
**Corresponding Author.*

Abstract:
In the realm of economic research, the application of large models has become increasingly prevalent. Given the substantial reliance on time series data in finance-related studies, it is essential to evaluate the suitability of large time series models by comparing them with traditional models. This comparative approach is crucial for validating the efficacy and relevance of these models in the context of financial data analysis. In this paper, the closing price of COMEX gold futures from January 1, 2008 to March 8, 2024 is selected for empirical research. The price of COMEX gold futures is predicted vertically and compared with the results of the forecast data after further introducing the dollar index, crude oil prices, and the stock market as exogenous variables. At the same time, the traditional time series model and the machine learning model are compared in terms of model quality and efficiency. It was found that the Time-GPT model performed better in prediction accuracy after introducing exogenous variables. In terms of horizontal comparison, the quality results of the model are generally better than those of the traditional results. The Time-GPT model is strong in capturing long-term dependencies, but the large number of parameters makes the model run inefficiently.

Keywords: time series analysis, gold price prediction, large model technology, Transformer, Time-GPT model

INTRODUCTION

Research Background

In recent years, large model techniques such as GPT and BERT have made remarkable achievements in the fields such as natural language processing. These models learn rich language and knowledge representation by pre-training on large-scale datasets. The success of this technology has led researchers to apply it to financial time series forecasting, including gold price forecasting. [1]

In the current global economic environment, uncertainties such as geopolitical tensions, economic cycle fluctuations, monetary policy instability and frequent public health emergencies have significantly increased the volatility of financial markets. In this context, gold as a long-honored safe haven asset, increasingly important. Investors and market participants are increasingly relying on gold to avoid risk and protect asset value, making the volatility of gold prices widely watched. Therefore, accurate forecasting of the gold price is of great significance for investors to develop effective asset allocation strategies, financial institutions to manage the risk, and policy makers to assess the macroeconomic environment.

Traditionally, gold price prediction mainly relies on time series analysis methods, such as ARMA, SARIMA and other models, which predict future trends by analyzing historical price data. With the development of technology and the improvement of data analysis ability, machine learning and deep learning technologies have gradually been introduced into the gold price prediction [2]. These advanced methods are able to handle larger, higher dimensional datasets and identify more complex patterns and relationships in the data, providing new tools and perspectives for gold price prediction. However, with the increase of data and the complexity of the model, how to select appropriate models and parameters, and how to improve the accuracy and stability of prediction, has become a new challenge for researchers to face. The ability of large models to understand and process complex data offers new possibilities for gold price predictions. However, applying these models directly to time series data requires solving a series of challenges such as data representation transformation, model training strategy adjustment, and interpretation of results. Moreover, the large parameter scale of large models and the high demand for computational resources also limit its wide generalization in practical applications.

This paper discusses the application and significance of the large model technology in the gold price prediction from the theoretical and practical levels, aiming to provide a new theoretical perspective and practical tool for the financial market prediction, to promote the development of financial technology and improve the decision-making quality of the financial market have certain reference.

By exploring and analyzing the application of large model technology in gold price prediction, the paper integrates the latest artificial intelligence technology into the traditional financial model, which provides a new perspective for understanding the dynamics of the financial market. This not only involves finance and economics, but also combines the latest achievements in deep learning and natural language processing in computer science. This interdisciplinary research method helps to promote knowledge exchange and technology integration between different fields, and provide new ideas and tools for solving complex financial problems.

By applying the large model to gold price prediction, this paper explores new methods in data preprocessing, model training strategy and results interpretation, which provides theoretical innovations for improving the accuracy and interpretability of financial time series prediction. And accurate gold price forecast is crucial for investors and financial institutions to develop effective investment strategies and risk management measures. By introducing large model technology, this paper aims to improve the prediction accuracy, thus helping market participants to make more scientific and rational decisions.

Literature review

In recent years, with the increasing uncertainty of economic environment, currency depreciation and inflation, gold as a safe haven asset, many investors will increase their investment in gold during these periods to protect their wealth from depreciation. Therefore, an accurate prediction of gold prices has always been important for market participants.

The gold price has long been the result of a combination of many factors, including both macroeconomic indicators, geopolitical and geopolitical events, market supply and demand conditions, and investor psychology. At the same time, the gold price, as a time series data, also has typical characteristics. Wu (2013) conducted a time series study on the data of gold price, analyzed the stability or shock of gold price, and expounded the characteristics of gold price fluctuation consistent with the fractal. The fractal has the characteristics of long memory, self-similarity and stability, which provides reference for the fluctuation of gold price. The gold price is sustainable by the value of Hearst index [3]. The traditional method of gold price prediction has been occupying an important position in this market, and the method is constantly improved according to the characteristic properties of gold. Xi (2014) used the time series correlation theory to establish the ARMA model of gold price, and found that it can dynamically describe the generation process of gold price data and better predict the gold price [4]. In view of the two characteristics of gold price, Wang et al. (2022) improved the traditional ARIMA model and constructed the SARIMA model of detrend and deseasonal factors. Prediction with the improved model and comparison with the ARIMA model show that the SARIMA model improves the prediction accuracy to some extent [5]. Chen et al. (2018) conducted an empirical study on the long memory of the gold price, used the Hurst index to confirm that there is indeed significant long memory in the gold price, and established the ARFIMA-GARCH model family, thus reflecting the fluctuation aggregation of the gold yield sequence. The empirical results prove that it does have long memory and heteroscedasticity, the prediction error is very small, and the forecast price fluctuation trend is basically the same [6]. He et al. (2023) proposed a deep learning ensemble-based financial time series forecasting model, which achieved superior performance in terms of forecasting accuracy and robustness compared with the benchmark individual models.[7]

With the constant change and complexity of financial markets, financial markets need more accurate and reliable prediction models to support investment and risk management decisions. Therefore, in a highly uncertain and volatile environment like the gold market, the emergence of machine learning methods has contributed greatly, providing new perspectives for time series analysis, especially showing great potential for handling large-scale, high-dimensional and complex data patterns. Ma et al. (2015) used the BP neural network model to study and predict the gold futures price from the perspective of the time series model, and found that the prediction results of the BP network model were good, and the prediction results of the ARIMA and the data were the best [8]. Liu et al.(2021) used the neural network method to conduct deep learning, and established multi-layer LSTM and two-way LSTM model prediction. Compared with the traditional model, the two-way LSTM model was better than all contrast models, and achieved good prediction effect [9]. Chen et al. (2023) choose the dollar index, nasdaq composite index and other international important index as the influence factors of international gold futures prices, based on the combination of EMD-LSTM model to forecast prices, and compared with the traditional machine learning, the results show that the model on the international gold futures price forecast, can more accurately predict the international gold futures prices [10].Mao et al. (2023) Deep learning methods are also gradually introduced into time series forecasting applications. The deep neural network uses multiple nonlinear layers to construct the feature representation of the previous time series to learn the internal change law of the time series [11]. Li et al. (2023) used the GluonTS time series forecasting framework to forecast Shanghai's export data and evaluate the effect. Experimental results show that the prediction effect of time series based on deep learning is significantly better than that of the traditional ARIMA model [12]. Based on the improved Transformer, Li et al. (2024) constructed a volatility prediction model TGC-FinTrans (TCN-BiGRU-CNN Finance Transformer). Experimental results show that the proposed model is superior to other baseline methods in predicting the volatility of financial data, and can predict volatility more accurately and capture the complex changes of the financial market, providing investors with a more accurate decision-making reference [13].

At the same time, the rise of large language models also marks an important milestone in the field of artificial intelligence. Large models have achieved remarkable results in natural language processing, computer vision and other fields, but have relatively few applications in traditional time series and spatiotemporal data analysis methods. Ming et al. (2023), through the latest progress of large model research on temporal and spatiotemporal data, emphasizing the importance of temporal data and spatiotemporal data analysis and its universality in real-world applications [14]. The application of large models in the field of time series prediction is divided into two kinds: the first one is to directly use the large model of NLP to do time series prediction. In such methods, time series prediction is made using NLP large models like GPT and Llama, focusing on how to transform the time series data into the input data suitable to the large model. In the financial field, such as Yu et al. (2023), the first exploration of timing prediction, using LLMs models, highlights the potential of these models in handling cross-sequence inference, integrating multimodal signals, and providing interpretable results [15]. By comparing with traditional models, LLMs can effectively integrate text news and numerical time series data, providing more in-depth analysis and understanding. The second is to train large models in the time series domain. In this method, large models such as GPT or Llama are trained jointly trained by establishing a large number of time series data sets, and used for downstream time series tasks. For example, Rasul et al. (2023) built a universal univariate probability time prediction model ag-Llam, which was trained on a large number of temporal data from the Monash Time Series library, and showed good zero sample prediction ability [16].Cao et al. (2023) presents an innovative multi-faceted model grounded in invariant learning principles, which significantly bolsters the comprehensive efficacy and adaptability of financial forecasting across both in-distribution and out-of-distribution data samples [17].

At the same time, big data is also facing great challenges, especially in the complexity of data in the financial field, the accuracy of prediction needs to be further explored. Xie et al. (2023) conducted extensive zero-sample learning ability analysis of ChatGPT in multimodal stock trend prediction. The results show that ChatGPT has limited success rate in predicting stock trend because it is not only lower than state-of-the-art methods, but also lower than traditional methods such as linear regression using price characteristics. In addition, more specialized training or fine-tuning of the model is needed [18]. Mishev et al. (2020) presents a comprehensive study on sentiment analysis in finance, evolving from traditional lexicon-based methods to state-of-the-art Transformer models. The study involves over a hundred experiments on publicly available datasets annotated by financial experts, showing that even with a relatively small dataset, significant results can be achieved [19]. The academic consensus acknowledges the complexity of training GANs and transformers, particularly those equipped with multiple attention layers [20]. Integrating these models presents a significant challenge due to their inherent training difficulties. Fu et al. (2022) utilize the convolutional networks with attention and the transformers. The attention-based GANs are tested on the S&P 500 index and option data, which not only reproduce the stylized facts, but also smooth the autocorrelation of returns [21]. In China, the application of the big model in the financial field is still in the initial stage of trial. Wu (2023) believes that the big model brings new opportunities for financial application scenarios, and will be the first application in the four levels of intelligent marketing, intelligent service, intelligent operation and intelligent risk control. It also emphasizes that the large model technology has not yet developed and mature, and special attention should be paid to the technical ethical issues such as data security in the application. On the premise of ensuring safety compliance, priority can be given to the relatively stable high-value scene, grasp the opportunities, and adhere to the integrity and innovation [22]. Zhang et al. (2024)consider that Large language models can not only process massive amounts of financial data, but also extract valuable information and insights from them, and automatically generate corresponding financial texts or financial decisions. From intelligent customer service to risk management to market analysis and personalized service, large language models play a key role. They help financial institutions increase efficiency, reduce risk, and improve the customer experience, while providing decision-makers with deeper market insights and decision support. As technology continues to advance, large language models will continue to play an important role in the financial sector as a key driver of financial innovation and digital transformation [23].

To sum up, an asset is a safe haven asset for market participants in an uncertain economic environment. This paper first analyzes the application of the large model in the financial time series, and uses the time series large model—Time-GPT to predict the gold futures price, and to compare with the traditional time series model, draws conclusions, and to further provide suggestions on the use of the model.

EMPIRICAL ANALYSIS OF GOLD FUTURES PRICE FORECAST

The Principle of Time-GPT Model

The Time-GPT model [24] developed is a generative pre-trained Transformer model [25] designed specifically for time-series prediction. The core advantage of this model compared with the previous mainstream recurrent neural network (RNN) and long and short memory network (LSTM) lies in its original attention mechanism. This mechanism can give the model the ability to identify key information in the text, allowing the model to pay more attention to important words, and this mechanism shows better parallel computing performance and scalability when dealing with long sequences. Therefore, the Transformer model quickly laid the foundation in the NLP field and achieved remarkable results in various text processing tasks. It combines self-attention mechanisms with specific local location encoding, built on a multiple encoder-decoder architecture, where each layer employs residual connectivity and hierarchical normalization techniques. The model is designed to process time series data with multiple frequency and attribute features, and can flexibly adapt to a variety of different input sizes and prediction windows. This also allows Time-GPT to make predictions based on historical data, making accurate predictions for new time series data without additional training. It has been trained on large data sets that bring together more than 100 billion data records in finance, meteorology, energy and the Internet. On data training, Time-GPT training is the largest collection of publicly available time series available, totaling over 100 billion data points. The training set contains a time series from a wide range of areas, including finance, economics, demographic, healthcare, weather, Internet of Things sensor data, energy, network traffic, sales, transportation and banking. Due to this different domain set, the training dataset contains a time series with a wide range of features. The overall logical structure of time series model is shown in Figure 1.

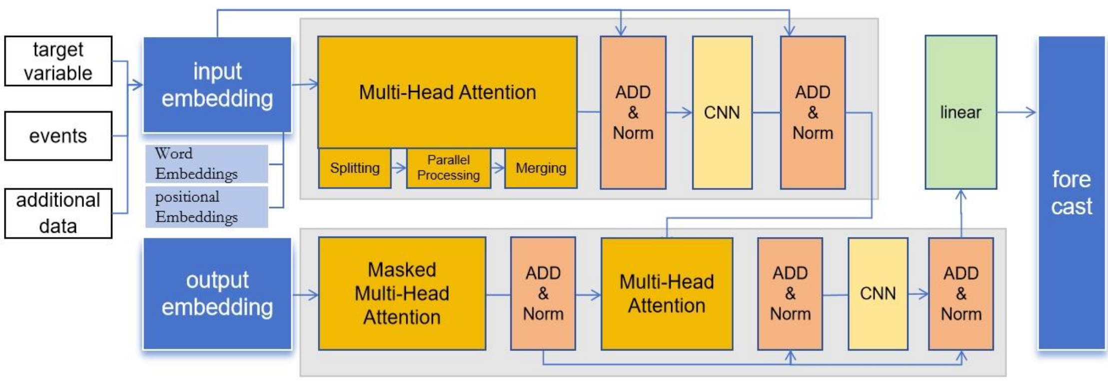

Figure 1. Structural diagram of the Time-GPT model

The Time-GPT model adopts the characteristics of transfer learning to understand it on the basis of previous research on time series prediction.

$f_\theta : \mathcal{X} \mapsto \mathcal{Y}$. Where Y is the feature space and X is the dependent variable space. $X = \{y_{[0:t]}, x_{[0:t+h]}\}$, $Y = \{y_{[t+1:t+h]}\}$. H is the predicted range, y is the target time series, and x is the exogenous covariate. The goal of the prediction task is to estimate the following conditional distributions:

$$\mathbb{P}\left(y_{[t+1:t+h]} \middle| y_{[0:t]}, x_{[0:t+h]}\right) = f_\theta\left(y_{[0:t]}, x_{[0:t+h]}\right) \quad (1)$$

$Ds = \{(X, y) | X \in \mathcal{X}, y \in \mathcal{Y}\}$.Transfer learning refers to pre-training a model on a (usually large) source dataset to improve its performance on new prediction tasks with the target dataset.

Data Sources and Descriptive Statistics

The data selected in this study are derived from the closing price of COMEX gold futures from January 2008 to March 2024, considered one of the benchmarks for global gold prices, offer almost all-weather trading hours, allowing investors around the world to trade at almost any time. This high degree of accessibility, combined with the transparency of the market, provides investors with an efficient and liquid trading environment. At the same time, the volatility of gold futures price is affected by many factors, including the DOLLAR index, stock market, crude oil price and so on. These factors also appear frequently in related studies, so this study selected these variables as exogenous variables to explore the prediction of gold futures price, among which the Dow Jones Index is selected as the representative of the changing trend of the stock market. In order to ensure the homogeneity and reliability of the data, the above data use the same frequency of daily data. Descriptive statistics for the above data were followed, and the results are shown in Table 1:

Table 1: Analysis of descriptive data

variable	COMEX price of gold futures	The dollar index	Crude Oil Price (WTI)	Dow Jones Index
sample size	4039	4039	4039	4039
min value	711.300	71.310	-37.63	6547.05
max value	2168.700	114.149	145.29	39131.53
lower quartile	1209.600	80.516	52.875	12652.51
upper quartile	1721.150	97.350	91.665	26957.83
average	1421.900	89.941	72.817	20475.63
median	1327.400	92.521	72.510	17918.15
standard error of the mean	5.076	0.153	0.364	136.415
variance	104082.8	94.673	533.977	75162782.085
standard deviation	322.618	9.730	23.107	8669.64
skewness	0.139	-0.045	0.138	0.406
kurtosis	-0.927	-1.139	-0.517	-1.122

Based on the descriptive statistical analysis of 4,039 samples, the COMEX gold futures closing price, DOLLAR index, crude oil price (WTI) and Dow Jones Index showed different statistical characteristics. The average closing price of COMEX gold futures was $1421.900, standard deviation of 322.618, skewness of 0.139, and kurtosis of-0.927, indicating that the price distribution is relatively symmetric and relatively flat. The mean of the dollar index is 89.941, standard deviation of 9.730, skewness of-0.045 and kurtosis of-1.139, showing features of a near normal distribution but are slightly biased to the left. The average value of crude oil price (WTI) is $72.817, standard deviation is 23.107, skewness is 0.138, and peak degree is-0.517, indicating that its price distribution is also close to normal distribution with a certain symmetry. The mean of 20475.63, standard deviation of 8669.64, skewness of 0.406 and kurtosis of-1.122, showing a right skewed distribution characteristic and low kurtosis. These descriptive statistics reveal the distributional properties of different economic indicators, and provide a basis for further analysis of their dynamic change trends.

Data Preprocessing

For large language models, especially when dealing with numerical, time series data. Although the Transformer model is somewhat robust in handling data at different scales and ranges, data normalization or standardization is still needed. First normalization helps the optimization algorithm to find the solution with the smallest error more smoothly and faster, which can significantly reduce the training time. Secondly, if the data scale is too large or too small, it may lead to the gradient disappearance or explosion, which will affect the efficiency of model learning. Through normalization, the dominant influence of some features on the model learning during the training process can be reduced, so that the model can learn the relative importance of features rather than the absolute numerical size, which helps to improve the adaptation ability and generalization of the model to new data.

This paper selects the common normalization method: Z-score standardization. Z score standardization is a common data preprocessing technique, and the main purpose of this method is to adjust the feature scale of the data with zero mean and unit variance, and thus improve the comparison and algorithm performance among different features. The specific calculation formula is as follows:

$$Z = \frac{(X-\mu)}{\sigma} \tag{2}$$

Where X is the original data point, μ is the sample mean, and σ is the sample standard deviation.

With this transformation, the mean of each feature changes to 0 and the standard deviation to 1. This means that the distribution of data will be centered on zero, and each data point will be changed in standard deviation. Practical operation examples results in Table 2:

Table 2: Normalization treatment

Date	The COMEX gold price z-score standardization	Dow Jones index z-score normalization	crude oil price z-score standardization	The dollar index z-score standardization
1/2/2008	-1.751	-0.857	1.160	-1.430
1/3/2008	-1.722	-0.856	1.141	-1.445
1/4/2008	-1.732	-0.885	1.086	-1.442
1/7/2008	-1.743	-0.882	0.964	-1.415
1/8/2008	-1.686	-0.910	1.018	-1.423
1/9/2008	-1.682	-0.893	0.989	-1.384

Gold Price Forecast Based on Time-GPT

In this paper, the gold price prediction model is constructed based on the Time-GPT model respectively. The specific construction ideas are as follows:

1. For the closing price of gold futures on each forecast target date, the closing price of gold futures on the 15 trading days prior to that date is used as the input feature. Such a sliding window approach helps the model to capture the time series properties and recent trends of gold prices.

2. For the direct prediction results, the COMEX gold futures price is trained to obtain the final result, and the future actual value of the exogenous variables is added to obtain the final results.

According to the previous ideas, two gold futures price prediction models are constructed respectively. After completing the training, the two models were used to predict the test set, and the prediction results are presented in the following line graph together with the actual data. More details in Figure 2.

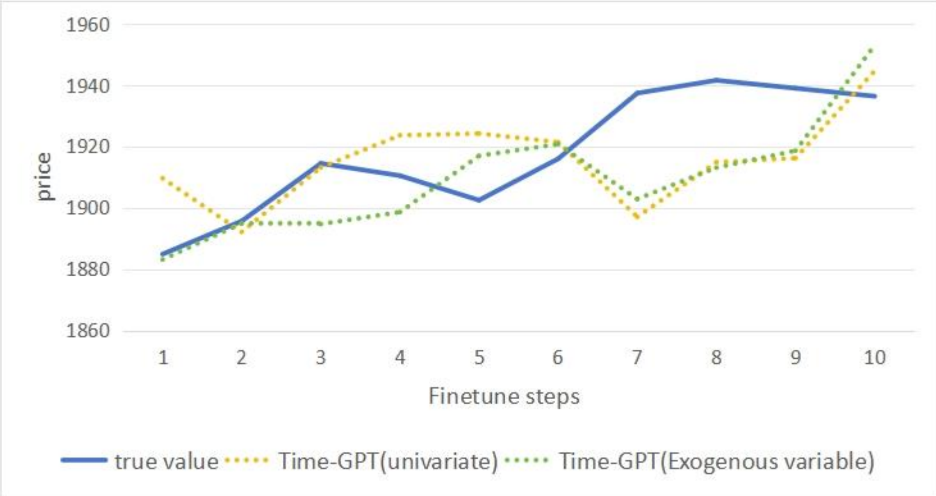

Figure 2. Results presentation

Table 3: Comparison of the 3Time-GPT results

	single argument		Introduction of exogenous variables	
actual result	settlement	average error	settlement	average error
1884.9	1909.6	0.79%	1883.2	0.72%
1895.7	1892.2		1894.9	
1914.5	1913.2		1894.7	
1910.5	1923.7		1898.6	
1902.5	1924.2		1917.0	
1916.0	1921.4		1920.8	
1937.4	1897.1		1902.9	
1941.6	1914.9		1913.1	
1939.0	1916.2		1918.7	
1936.4	1944.6		1953.0	

In For most dates, the model introducing exogenous variables predicted closer to the actual price. In particular, on February 23rd and on March 5th, this model predicted very close to the actual results, showing a lower mean error. There are also some exceptions. For example, on February 28 and March 1, the predicted value of this model was much lower than the actual value, which may be due to the influence of exogenous variables or the fact that the model failed to correctly explain some emergencies.

Table 3 shows that the models introduced exogenous variables provided more precise predictions in most cases, improved predictive power with the inclusion of more relevant economic indicators. This suggests that integrating multiple data sources may improve prediction accuracy in complex market environments. However, even improved models may not perform well in certain specific situations, underscoring the importance of continuously optimizing the models and strategies.

COMPARATIVE ANALYSIS

Comparison of Model Results

The Time-GPT model was selected to predict the COMEX gold futures price, and then we chose to compare it with classic models, such as ARIMA, machine learning, random forest and XGBOOST models. Results can be found in Figure 3.

The reason for choosing these specific models for comparison is that ARIMA financial time series analysis, mainly suitable for the prediction of linear trend and seasonal patterns, and machine learning has the characteristics of flexibility in data processing, its ability in processing non-linear data has been gradually applied to time series prediction. By comparing these models, we can better compare the properties of Time-GPT in processing complex and nonlinear time-series data, especially in the application of modern financial technologies.

The above results are obtained by constructing all kinds of models. Due to the large volatility of the prediction set, the introduction of fewer exogenous variables shows some limitations, and further optimization and adjustment are needed to improve the accuracy and stability of the prediction.

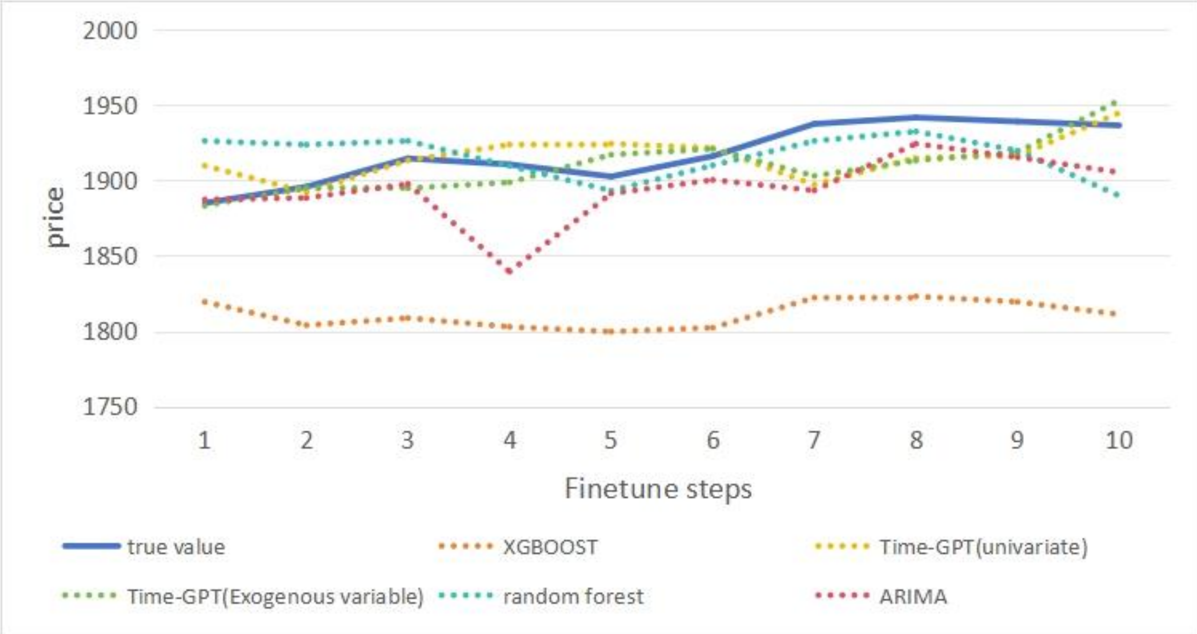

Figure 3. The prediction results of each model

In order to more scientifically evaluate the performance of the Time-GPT model proposed in this paper in the prediction of gold futures price, and to take into account the characteristics and limitations of different evaluation criteria in the application of actual financial prediction, it is difficult to measure the effect of the model comprehensively and comprehensively by relying on a single evaluation index. Therefore, this study will use multiple performance indicators, including prediction accuracy, model stability, and computational efficiency, to comprehensively evaluate the performance of the selected models. Such an evaluation method can not only provide a more comprehensive comparison of effects, but also better guide the model selection and application in practice.

Comparison of Model Quality

In the study of gold futures price prediction, the comparison of model quality is crucial to ensure that the selected prediction tools are not only accurate, but also show a high degree of reliability and stability in various market conditions.

First, in order to comprehensively evaluate the accuracy of the proposed model in predicting the price of gold futures, we will use three statistical indicators: mean square error (MSE), root mean square error (RMSE) and mean absolute error (MAE). These indicators can provide the error size of the model prediction results, and thus measure the accuracy of the prediction. The following are the specific calculation formulas for these indicators:

mean squared error (MSE):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \tag{3}$$

$Y_i\hat{Y}_i$Where: is the actual value, is the predicted value, and n is the total number of samples. MSE measures the average of the sum of squares of the difference between the predicted values and the actual values, which can very well reflect the overall magnitude of the prediction error.

Root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2} \quad (4)$$

The RMSE is the square root of the MSE, which keeps the units of error consistent with the units of the original data and is therefore easier to interpret.

mean absolute error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right| \quad (5)$$

MAE is the average of the absolute value of the difference between the observed value and the predicted value, which provides the average level of prediction error and is another common method to measure the accuracy of prediction.

Through the calculation and comparison of these indicators, we can get the performance difference of different models in the gold futures price prediction, which can provide a scientific basis for further model optimization and selection.

Table 4: Model comparison results

model	MSE	RMSE	MAE
Time-GPT univariate	422.870	20.563	16.798
Time-GPT introduces exogenous variable variables	345.840	18.596	15.331
ARIMA	1922.706	43.849	26.476
random forest	656.482	25.622	19.063
XGBOOST	311.99	17.663	11.82

Through comparison, from the whole point we found that after the introduction of exogenous variables model results better than univariate model, at the same time the effect of Time-GPT than traditional model for the better, at the same time relative to the machine learning model, large model in the spatial and time structure data overall performance, feature extraction ability is stronger. Results can be found in Table 4.

Comparison of Model Efficiency

In evaluating the efficiency of the model, the focus is on the training time and prediction speed of the model, which is crucial for feasibility and practicality in practical application scenarios.

Table 5: Comparison results of model efficiency and effects

model	training time	Forecast time
Time-GPT	49.62s	2.37s
Time-GPT introduces exogenous variable variables	15.85s	4.71s
ARIMA	0.34s	3.28s
random forest	0.56s	4.31s
XGBOOST	4.46s	3.29s

Training time: For deep learning-based Time-GPT and machine learning models, a long training time is usually required due to their complex network structure and a large number of parameters. These models are usually performed on a GPU or any other high-performance computing device to speed up the training process. In contrast, the ARIMA model is usually fast for training on the CPU due to its relatively simple mathematical structure. This type of model does not need to handle a large number of parameter adjustments and thus generally outperforms deep learning models in training efficiency.

Prediction speed: In terms of prediction speed, although Time-GPT and machine models are slow during the training phase, these models are able to quickly generate predictions once trained, especially after they have been optimized and deployed to the appropriate hardware environment. ARIMA models also show high efficiency when generating prediction results, especially when the data volume is not huge, and these models can generate prediction results in almost real-time.

From Table 5, we can observe that although Time-GPT and machine learning may have advantages in prediction accuracy, they may not be as cost-efficient as the ARIMA model in model training and resource consumption. Simpler models might outperform transformers on smaller datasets [26]. This difference in efficiency should be considered when selecting models to ensure that the selected model meets the needs of a particular application scenario.

CONCLUSION

Amid economic cycle fluctuations and unstable monetary policies, gold, as a traditional safe-haven asset, has become increasingly important. Investors and market participants rely on gold to mitigate risks and protect asset value, directly driving gold price volatility. Traditional gold price prediction methods such as ARMA and SARIMA are primarily based on time series analysis, while modern technology has introduced machine learning and deep learning techniques into gold price prediction.

This paper first reviews the application of large models in financial time series analysis, detailing the characteristics of the Transformer model, and focusing on the construction principles and applications of the Time-GPT model in financial markets. This analysis clarifies the effectiveness and potential of large models in handling complex financial data.

The study uses the Time-GPT model to predict gold futures prices, selecting the closing prices of COMEX gold futures from January 1, 2008, to March 8, 2024, as the target variable, and referencing the Dollar Index as an exogenous variable. After normalizing the data, a sliding window method is used to select the closing prices from the 15 days prior to each target prediction date as features, enabling dynamic gold price predictions. The study found that the Time-GPT model, with exogenous variables, showed better accuracy in univariate predictions.

Finally, the paper conducts a detailed comparative analysis of the Time-GPT model and other commonly used models in predicting gold futures prices, evaluating the quality and efficiency of different models. Although the Time-GPT model may require more resources and longer training times than traditional models, it demonstrates high efficiency in the prediction phase, making it particularly suitable for financial markets requiring rapid responses. Time-GPT model performed well in smaller learning rate, which is consistent with Brown et al. [27].

The main innovations of this study are reflected in the following aspects: Model Selection and Application. This study applies the Time-GPT model as a method to financial time series data, especially gold price prediction, demonstrating its advantages in processing complex financial data. It effectively expands the application scope of the traditional financial model, and also provides a new perspective for the future financial market analysis; Integration of deep learning techniques Through integrating and combing traditional models and the latest deep learning techniques, this research optimizes the accuracy and efficiency of prediction models. Especially in the feature processing and data analysis that introduce time series, the Time-GPT model is able to capture the subtle patterns in gold price changes by learning large amounts of historical data; The applied study of exogenous variables introduced the dollar index as an exogenous variable to investigate its influence on the gold price. This approach improves the comprehensiveness and accuracy of the predictions and provides a new perspective for understanding the relationship between macroeconomic factors and financial market dynamics.

Despite these innovations, the study has some limitations: the "black box" nature of the Time-GPT model may pose challenges in practical applications, especially in fields requiring high interpretability, such as finance. Additionally, external factors like macroeconomic policies and market sentiment impact the accuracy of gold futures predictions. Future work should focus on improving model generalization, reducing computational resource demands, and enhancing model interpretability.

ACKNOWLEDGMENT

This paper was supported by the Yuxiu Innovation Project of NCUT (Project No.2024NCUTYXCX211).

REFRENCES

[1] Konstantinos Benidis, Syama Sundar Rangapuram, Valentin Flunkert, Yuyang Wang, Danielle Maddix, Caner Turkmen, Jan Gasthaus, Michael Bohlke-Schneider, David Salinas, Lorenzo Stella, François-Xavier

Aubet, Laurent Callot, and Tim Januschowski. Deep learning for time series forecasting: Tutorial and literature survey. ACM Comput. Surv., apr 2022:1-33.

[2] Sezer, Omer Berat, U. Gudelek, and M. Ozbayoglu. Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. Applied Soft Computing 90. May 2020 (2020): 106181.

[3] Wu Yanqiu. Research on the fractal characteristics of gold price based on re-standard range analysis method. Journal of Tonghua Normal University, 2013, 34(08): 16-18.

[4] Xi Jing. Gold price prediction based on the time-series model. Modern Computer (Professional Edition), 2014, (28): 9-14.

[5] Wang Peizhi, Mei Ling, Gao Wei. Gold Price Forecast based on the SARIMA model. Electronic Technology and Software Engineering, 2022, (23): 233-237.

[6] Chen Peng, Li Xingye. Gold price prediction analysis based on the ARFIMA-GARCH model family. E-commerce, 2018, (05): 42-44.

[7] He, Kaijian, Qian Yang, Lei Ji, Jingcheng Pan, and Yingchao Zou. 2023. Financial Time Series Forecasting with the Deep Learning Ensemble Model. Mathematics 11, no. 4: 1054.

[8] MA Baozhong, CHEN Chuanming. An Empirical Analysis of BP Neural Network Model in Gold Futures Price Prediction. Enterprise Herald, 2015, (11): 20+19.

[9] Liu Lu, Lou Lei, Liu Xianjun, Shi Sanzhi. A Comparative Study on COMEX Gold Futures Price Prediction Based on LSTM. Journal of Changchun University of Science and Technology (Natural Science Edition), 2021, 44(02): 129-135.

[10] Yang Chen, Chen Guizi. International Gold Futures Price Forecast based on EMD-LSTM. Journal of South-Central University for Nationalities (Natural Science Edition), 2023, 42(06): 857-864.

[11] Mao Yuanhong, Sun Chenchen, Xu Luyu, Liu Xi, Chai Bo, He Pengchao. Survey of Time Series Forecasting Methods Based on Deep Learning. Microelectronics & Computer, 2023, 40(04): 8-17.

[12] Li Wen, Deng Sheng, Duan Yan, Du Shouguo. Time Series Forecasting and Deep Learning: A Literature Review and Application Examples. Journal of Computer Applications and Software, 2020, 37(10): 64-70+84.

[13] Li Wenying, Pan Qiao, Yan Xiping. Financial Market Volatility Prediction Model Based on Deep Learning. Intelligent Computers and Applications, 2024, 14(07): 79-84.

[14] Ming Jin, Qingsong Wen, Yuxuan Liang, Chaoli Zhang, Siqiao Xue, Xue Wang, James Zhang, Yi Wang, Haifeng Chen, Xiaoli Li, Shirui Pan, Vincent S. Tseng, Yu Zheng, Lei Chen, and HuiXiong. Large Models for Time Series and Spatio-Temporal Data: A Survey and Outlook. CoRR abs/2310.10196 (2023).

[15] Yu, Xinli, Zheng Chen, Yuan Ling, Shujing Dong, Zongying Liu and Yanbin Lu. Temporal Data Meets LLM - Explainable Financial Time Series Forecasting. ArXiv abs/2306.11025 (2023).

[16] Kashif Rasul, Arjun Ashok, Andrew Robert Williams, Arian Khorasani, George Adamopoulos, Rishika Bhagwatkar, Marin Bilos, Hena Ghonia, Nadhir Vincent Hassen, Anderson Schneider, Sahil Garg, Alexandre Drouin, Nicolas Chapados, Yuriy Nevmyvaka, Irina Rish:Lag-Llama: Towards Foundation Models for Time Series Forecasting. CoRR abs/2310.08278 (2023)

[17] Defu Cao, Yixiang Zheng, Parisa Hassanzadeh, Simran Lamba, Xiaomo Liu, and Yan Liu. 2023. Large Scale Financial Time Series Forecasting with Multi-faceted Model. In Proceedings of the Fourth ACM International Conference on AI in Finance (ICAIF '23). Association for Computing Machinery, New York, NY, USA, 472–480.

[18] Xie, Qianqian and Han, Weiguang and Lai, Yanzhao and Peng, Min and Huang, Jimin.The Wall Street Neophyte: A Zero-Shot Analysis of ChatGPT Over MultiModal Stock Movement Prediction Challenges.arXiv preprint arXiv:2304.05351(2023)

[19] Mishev, Kostadin, et al. "Evaluation of Sentiment Analysis in Finance: From Lexicons to Transformers." IEEE Access PP.99(2020):131662-131682.

[20] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 6000–6010.

[21] Fu, Weilong, Ali Hirsa and Jörg Osterrieder. Simulating financial time series using attention. ArXiv abs/2207.00493 (2022).

[22] Wu Xiaoru. Cognitive big model new technology helps financial innovation. New Finance, 2023, (10): 26-31.
[23] Zhang Xiaoyan, WU Huihang. Application of Large Language Models in the Financial Field. Tsinghua Financial Review, 2024, (05): 22-26.
[24] Garza A, Mergenthaler-Canseco M. TimeGPT-1. arXiv preprint arXiv:2310.03589 (2023)
[25] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. Advances in neural information processing systems, 30, 2017.
[26] Ailing Zeng, Muxi Chen, Lei Zhang, and Qiang Xu. Are transformers effective for time series forecasting? In Proceedings of the AAAI conference on artificial intelligence, volume 37, pages 11121–11128, 2023.
[27] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. Advances in neural information processing systems, 33: 1877–1901, 2020.