Proceedings of the 2nd International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/55/20230962

Research on the Synthesis of Hong Kong NFT Index Using Principal Component Analysis and Index Prediction Based on LSTM-Modified ARMA-GARCH Model

Weidong He[1,a,*], and Jiahe Yu[1,b]

[1]*Department of Economics, Minzu University of China, Beijing, 10081, China*
a. 2456362783@qq.com, b. 1172882578@qq.com
**corresponding author*

Abstract: With the advent of the Web3.0 era, virtual assets have gained prominence in individuals' asset portfolios, making Non-Fungible Tokens (NFTs) increasingly significant within the financial trading landscape. To address the issue of multicollinearity in regression analysis, this paper employs Principal Component Analysis (PCA) to perform dimensionality reduction on five correlated foundational sectors. Moreover, to enhance the accuracy and reliability of predictive outcomes, the study combines the Long Short-Term Memory (LSTM) model with the Autoregressive Moving Average-Generalized Autoregressive Conditional Heteroskedasticity (ARMA-GARCH) model. Through the application of these methods and practical implementation, the study forecasts the NFT index of the Hong Kong stock market for the next 30 days. This forecasting of return volatility contributes vital insights for investment decision-making. The research complements and offers application recommendations in financial innovation, deepening, and regulation. By devising novel products and tools to meet investor demands, providing risk management and investment opportunities, the model's predictive outcomes can be utilized in regulatory and risk management strategies within the national financial trading market. This study provides regulatory guidance, policy formulation insights, and envisions further refinements of the research methodology by integrating information shock effects.

Keywords: NFT, principal component analysis, LSTM model, ARMA-GARCH model

1. Introduction

The NFT market in Hong Kong is rapidly growing, attracting more investors and institutions. The government of the Hong Kong Special Administrative Region has been actively working on enhancing regulatory supervision for digital assets. The Hong Kong Securities and Futures Commission (SFC) is in the process of licensing digital asset exchanges and has issued guidelines for compliance. In the NFT market, cultural institutions like the Hong Kong Museum of Art are adopting NFT technology to release historically significant digital artifacts. NFT artists in Hong Kong are also using this platform to showcase their work. This study combines the LSTM and ARMA-GARCH models to analyze the Hong Kong stock market's high-correlation sectors, including finance, telecommunications, consumer goods, and technology. The data set covers minute-by-minute data from September 14, 2020, to May 19, 2023, including the NFT World Index data. The findings offer

DOI: 10.54254/2754-1169/55/20230962

valuable insights for investors interested in the Hong Kong NFT index and can guide policymakers. The research methodology can be a reference for other financial markets. ARMA-GARCH models are constructed for various sectors, and a three-loop LSTM model refines predictions from these models. Using the LSTM-modified ARMA-GARCH model, the study predicts NFT return volatility. It provides recommendations for financial innovation, deepening, and regulation. This includes the development of NFT yield volatility-based options products, volatility trading strategies, risk management tools, and integrating predictive outcomes into regulatory frameworks. Model-predicted outcomes can inform market regulation and intervention policies, ensuring market fairness and integrity. Further research and validation are essential for practical applications, with potential refinements incorporating information shock effects to better understand NFT market volatility.

2. Theoretical Foundation

The advent of Non-Fungible Token (NFT) assets has garnered substantial attention from scholars both domestically and internationally. An increasing number of financial researchers have directed their focus towards the realm of NFTs. Presently, literature concerning NFTs employs various statistical econometric models to analyze and forecast the indicators of NFTs themselves [1][2]. Additionally, exploration into the transmission of returns and volatility between NFTs, cryptocurrencies, and conventional assets has ensued [3][4][5], contributing to the comprehension of trends and potentials within the NFT market.

Principal Component Analysis (PCA) finds extensive application in the realm of physical chemistry and other scientific fields. Utilizing PCA, complex datasets and interrelationships among independent variables, which are arduous to explicate, can be reduced into a lesser number of abstract factors known as principal components [6][7][8]. In the domain of economics, PCA is commonly employed for research in macroeconomics [9][10]. In this study, PCA will be employed to condense causally related sectors of the NFT index, culminating in the synthesis of a singular US NFT index, followed by the derivation of the Hong Kong NFT index.

The AutoRegressive Moving Average-Generalized Autoregressive Conditional Heteroskedasticity (ARMA-GARCH) model is widely recognized as a classical forecasting methodology. Nevertheless, this model, alongside several other conventional statistical methods, falls short in capturing the nonlinear features inherent in time series data [11][12]. Consequently, Artificial Neural Networks (ANN) have emerged as a popular tool for modeling nonlinear relationships and predicting indicators. Subsequently, Recurrent Neural Networks (RNN), owing to their consideration of the temporal influence of past information, have gained prominence for time series prediction [13]. Within RNN, the Long Short-Term Memory (LSTM) model has exhibited strong performance in predictive analysis related to time series data and is considered a comprehensive version of RNN [14]. LSTM capitalizes on historical data to benefit from a high degree of consistency in time series analysis [15]. Furthermore, when compared to either RNN or time series models in isolation, a blend of RNN and GARCH models often demonstrates superior efficacy [16][17]. Therefore, to enhance predictive capabilities, this study integrates the LSTM model with the ARMA-GARCH model to forecast the values of five foundational sectors for the forthcoming thirty trading days.

3. Principal Component Analysis for Dimension Reduction

3.1. Research Purpose and Ideas

The four basic sectors (technology, communication, finance and consumption) with a causal relationship with the NFT index will reduce their dimensions, and extract the main components to fit the US stock NFT index, and then construct the NFT representative index of Hong Kong stocks. The specific operation steps are as follows: test whether it is suitable for the main component analysis,

DOI: 10.54254/2754-1169/55/20230962

extracting the main components, fitting the US stock NFT index, and fitting the Hong Kong stock NFT index.

3.2. KMO and the Bartlett-Tests

In order to avoid the problem of multicollinearity in the model, this paper adopts the principal component analysis method to adopt the dimension reduction treatment for the five basic plates, and fits it to become an index to measure the NFT index of the US stock market. First, KMO test and Bartlett spherical test were used to determine whether the data are suitable for principal component analysis and do dimensionality reduction treatment. According to Table 1, KMO value of 0.785> 0.600 and P-value of Bartlett spherical test of 0.000 <0.001 is significant, indicating that there is correlation between variables and is suitable for main component analysis.

Table 1: KMO and Bartlett tests.

KMO price		0.785
Bartlett Sphelicity test	Approximate chi square	4309.887
	df	10
	P	0.000***

3.3. Principal Components Were Extracted

Secondly, In the variance interpretation table, at the principal component 3, the total variance interpretation was below 1, and the cumulative contribution of variable interpretation reached 86.154%, which is already greater than 85%, indicating that the extraction is sufficient. In this way, the original five variables are converted into three new and mutually independent composite indicators.

Table 2: Total variance interpretation.

	characteristic root		
ingredient	characteristic root	Variance interpretation rate (%)	Cumulative variance interpretation rate (%)
1	2.681	53.618	53.618
2	1	20.002	73.62
3	0.627	12.534	86.154
4	0.443	8.854	95.009
5	0.25	4.991	100

3.4. Fit the US Stock Market NFT Index

Table 3: The component matrix table.

name	ingredient		
	Component 1	Component 2	Component 3
Financial services sector LN yield	0.332	0.004	-0.138
Technology sector LN yield	0.326	0.002	-0.167
Communication service sector LN yield	0.292	-0.003	-0.633
Consumer sector LN yield	0.266	-0.028	1.071
The World NFT Index LN yield	0.006	1	0.029

DOI: 10.54254/2754-1169/55/20230962

The calculation formula for the principal components F1, F2, and F3 is obtained from the component matrix table (Table 3) as follows:

$$F_1 = 0.332X_1 + 0.326X_2 + 0.292X_3 + 0.266X_4 + 0.006X_5 \quad (1)$$

$$F_2 = 0.004X_1 + 0.002X_2 - 0.003X_3 - 0.028X_4 + X_5 \quad (2)$$

$$F_3 = -0.138X_1 - 0.167X_2 - 0.633X_3 + 1.071X_4 + 0.029X_5 \quad (3)$$

Table 4: Factor weight table.

name	Variance interpretation rate (%)	Cumulative variance interpretation rate (%)	weight (%)
Principal Component 1	0.536	53.618	62.234
Principal Component 2	0.2	73.62	23.217
Principal Component 3	0.125	86.154	14.549

According to Table 4, the comprehensive score is calculated with the variance contribution rate of each factor as the weight:

$$\mathrm{F} = 0.622F_1 + 0.232F_2 + 0.146F_3 \quad (4)$$

3.5. Fit the Hong Kong Stock NFT Index

To predict the trend of the securities market index, it is often only necessary to grasp the relationship between the change rate of various influencing factors and the change of the fluctuation trend of the index. In the context of economic globalization, in order to maximize profits, the frequent cross-border flow of international capital has a profound impact on the stock market. Especially after the subprime mortgage crisis in 2008, the economic ties between countries became closer, leading to the closer connection of the global stock market and the trend of global integration of the stock market. With the rapid development of the digital economy, the correlation and influence of the large basic sectors on the stock index in the stock markets of various countries are roughly the same. Therefore, it can be considered that the basic sector with the impact and correlation to the US stock NFT index plays the same role for the Hong Kong stock NFT index in China's stock market. In conclusion, fitting THE Hong Kong NFT index also uses the following comprehensive score algorithm:

$$\mathrm{F} = 0.622F_1 + 0.232F_2 + 0.146F_3 \quad (5)$$

4. LSTM Modified the ARMA-GARCH Model

4.1. Research Purpose and Ideas

The prediction results of ARMA-GARCH model are corrected with the three-cycle LSTM model, and then the prediction results of each sector and the results of principal component analysis are used to predict the index of Hong Kong NFT index in the next 30 trading days. The specific operation process is as follows: data preprocessing, partitioning data sets, training ARMA-GARCH model, fitting ARMA-GARCH model, constructing LSTM model, LSTM model training, model correction and prediction.

DOI: 10.54254/2754-1169/55/20230962

4.2. Model Data Selection

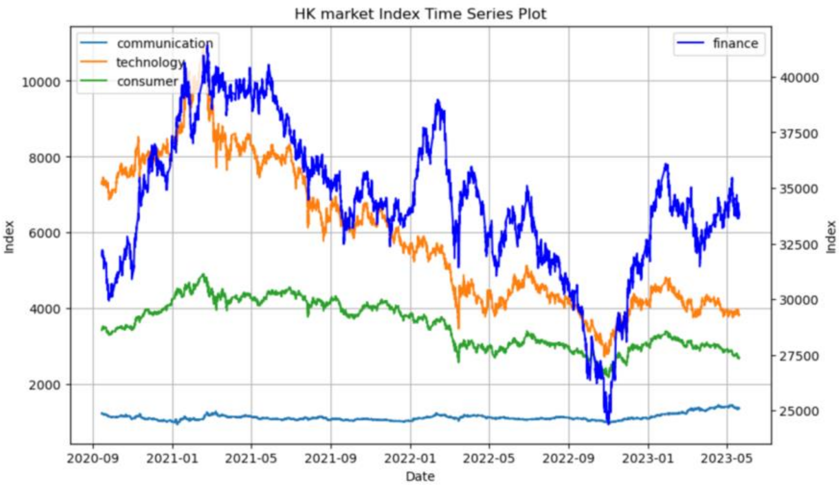

Figure 1: The four exponential timing plots.

To better predict the law between time series data and volatility, adopt high frequency trading data modeling prediction, select the hang seng financial, hang seng telecommunications, hang seng consumption and hang seng technology four index on September 14,2020-May 19,20,2023, every time-sharing data (data source: Choice database terminal), the index sequence diagram as shown above, the NFT world index also within the same range of each time-sharing data, subsequent modeling analysis to use the above data. The data were not normally distributed, considering the T distribution or the generalized difference distribution, and the T distribution was tested after the information criterion.

4.3. ARMA-GARCH Model Construction

4.3.1. Time-Series Stationarity Test

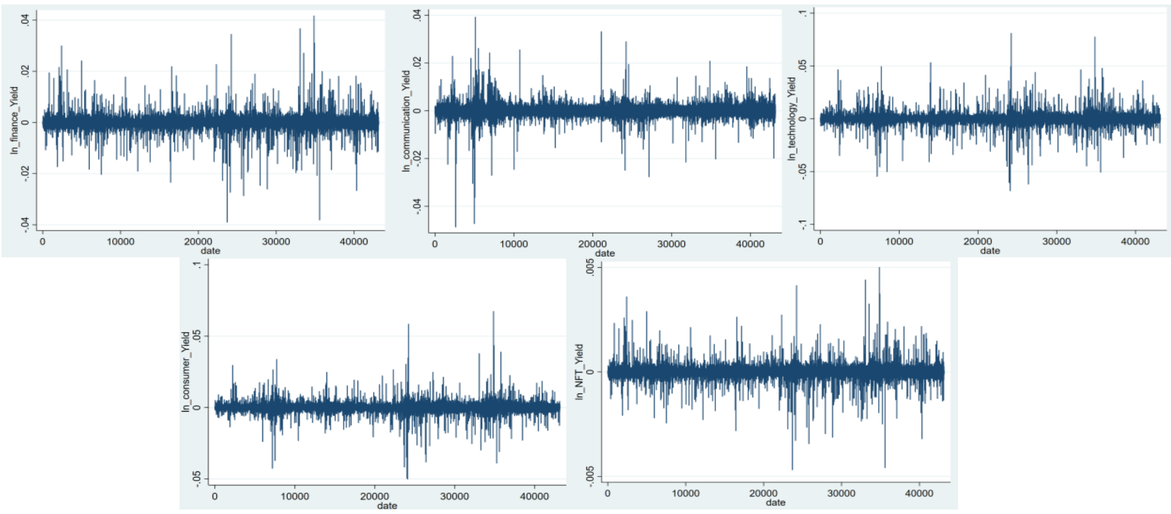

Figure 2: Time sequence chart of each index yield.

It can be seen from the timing chart of each index return rate (Figure 2) that there is variance aggregation effect and conditional heteroscedasticity. However, the p-value of the stationarity test of each index yield is <0.05, rejecting the null hypothesis, indicating that the time series is stable.

DOI: 10.54254/2754-1169/55/20230962

4.3.2. Judgment of ARMA Model Building

The time series of finance, communication, consumption and NFT index yield lag 12 order p-value value is less than 0.05, rejecting the null hypothesis, indicating that there is autocorrelation in each time series data, there are conditional mean laws that can be mined, and because the time series is stable, the ARMA model needs to be established.

The time-series of technology index of 12 is higher than 0.05, so the null hypothesis cannot be rejected, indicating that there is no autocorrelation in the time series data at 95% confidence level, and because the time series is stable, it is no need to establish ARMA model and directly consider whether there is arch effect.

4.3.3. Determine the ARMA Model Order and Parameters

Through finance, communication, consumption and NFT index yield AC and PAC chart can see no obvious cut trend, unable to judge, combined with ACF and PACF, through LL, AIC and BIC information criterion comparison, determine the financial, consumption and NFT index yield AR order 3, MA is order 2, determine model selection for ARMA (3,2) model, communication index yield AR order 3, MA is order 1, determine the model selection for ARMA (3,1) model, model parameters are as follows:

Table 5: The ARMA model parameters.

			Coef.
ln_finance_Yield	AR	L1	.4319108
		L2	-.9796393
		L3	-.4469863
	MA	L1	.9836452
		L2	-.0165851
ln_communication_Yield	AR	L1	.4421813
		L2	.0428584
		L3	.01693
	MA	L1	-.5121995
ln_consumer_Yield	AR	L1	-1.362824
		L2	-.6919797
		L3	.0311373
	MA	L1	1.392042
		L2	.7365798
ln_nft_yield	AR	L1	-.3192289
		L2	-.9608591
		L3	-.0980521
ln_nft_yield	MA	L1	.191668
		L2	.9752089

4.3.4. Self-Correlation Back Test

Financial, communication, consumption and NFT index yield model parameters except the constant items are significant, and through autocorrelation back, that the condition mean rule has been discovered, then consider the condition, the residual term autoregression test and LM test, found that financial, communication index yield the residual term autoregression coefficient of three significant,

DOI: 10.54254/2754-1169/55/20230962

consumption index yield residual term autoregression coefficient of two significant, considering the possible ARCH effect. It is found that the autoregression coefficient of the NFT index yield residue term was not significant. Considering that the ARCH effect may not exist, the model was established and the fitting effect of the model was considered.

The yield of technology index autoregresses the square of its own lag term, and finds that the first fourth order coefficient is significant. Considering the possible arch effect, the arch / gear model should be established to treat the conditional heteroscedastic phenomenon.

4.3.5. GARCH Model

Through the comparison of LL, AIC and BIC information criteria, the ARCH of finance, communication, technology and consumption index yield model is determined as order 1, GARCH is order 1, and the model is selected as GARCH (1,1) model. The model parameters are as follows:

Table 6: Model parameters of GARCH.

			Coef.
ln_finance_Yield	ARCH	L1	.2163816
	GARCH	L1	.6923909
ln_communication_Yield	ARCH	L1	.2051661
	GARCH	L1	.670413
ln_technology_Yield	ARCH	L1	.2882047
	GARCH	L1	.7453921
ln_consumer_Yield	ARCH	L1	.2993463
	GARCH	L1	.7313484

The model parameters were significant except the constant term, which were back-tested by LM with no asymmetric effect, indicating that the conditional variance law has been fully explored, the model establishment was completed, and the fitting effect of the model was considered.

4.3.6. For each Index Yield Model Equation

1) The ARMA (2,3) -GARCH (1,1) model equation for the financial index yield is as follows: ARMA part:

$$\ln finance\ yield_t = \&1.486018 \ln finance\ yield_{t-1} - 0.5186466 \ln finance\ yield_{t-2} - 1.539179\varepsilon_{t-1} + 0.5739443\varepsilon_{t-2} - 0.0039306\varepsilon_{t-3} \tag{6}$$

GARCH part:

$$\sigma_t^2 = 2.56e + 0.2163816\varepsilon_{t-1}^2 + 0.6923909\sigma_{t-1}^2 \tag{7}$$

2) The ARMA (3,1) -GARCH (1,1) model equation of the communication index yield is as follows:

ARMA part:

$$\begin{aligned} \ln communication\ yield_t = {} & 0.7419531 \ln communication\ yield_{t-1} \\ & +0.1542212 \ln communication\ yield_{t-2} \\ & +0.039845 \ln communication\ yield_{t-3} - 0.9451089\varepsilon_{t-1} \end{aligned} \tag{8}$$

DOI: 10.54254/2754-1169/55/20230962

GARCH part:

$$\sigma_t^2 = 3.07e + 0.2051661\varepsilon_{t-1}^2 + 0.670413\sigma_{t-1}^2 \tag{9}$$

3) The GARCH (1,1) model equation for the technology index yield is as follows:
GARCH part:

$$\sigma_t^2 = 3.92e + 0.7453921\sigma_{t-1}^2 \tag{10}$$

4) The ARMA (3,2) -GARCH (1,1) model equation of consumer index yield is as follows:
ARMA part:

$$\begin{aligned} \ln consumer\ yield_t = {} & 0.1828829 \ln consumer\ yield_{t-1} \\ & +0.8609489 \ln consumer\ yield_{t-2} \\ & -0.0587044 \ln consumer\ yield_{t-3} - 0.1091806\varepsilon_{t-1} \\ & -0.8707501\varepsilon_{t-2} \end{aligned} \tag{11}$$

ARCH part:

$$\sigma_t^2 = 0.2993463\varepsilon_{t-1}^2 + 0.7313484\sigma_{t-1}^2 \tag{12}$$

GARCH part:

$$\sigma_t^2 = 2.16e + 0.2993463\varepsilon_{t-1}^2 + 0.7313484\sigma_{t-1}^2 \tag{13}$$

5) The ARMA (3,2) model equation for the yield of the NFT index is as follows:
ARMA part:

$$\begin{aligned} \ln NFT\ yield_t = {} & -0.3192289 \ln NFT\ yield_{t-1} - 0.9608591 \ln NFT\ yield_{t-2} \\ & -0.0980521 \ln NFT\ yield_{t-3} + 0.191668\varepsilon_{t-1} + 0.9752089\varepsilon_{t-2} \end{aligned} \tag{14}$$

4.4. Model Fitting

According to the model of the index yield fitting, draw the timing diagram between the predicted value and the actual value, found the trend is basically the same, and select the last thirty days of data (1440 data) as a test set, calculate the root mean square error (RMSE), the average absolute error (MAE) and the average absolute percentage error (MAPE), the result output is as follows:

DOI: 10.54254/2754-1169/55/20230962

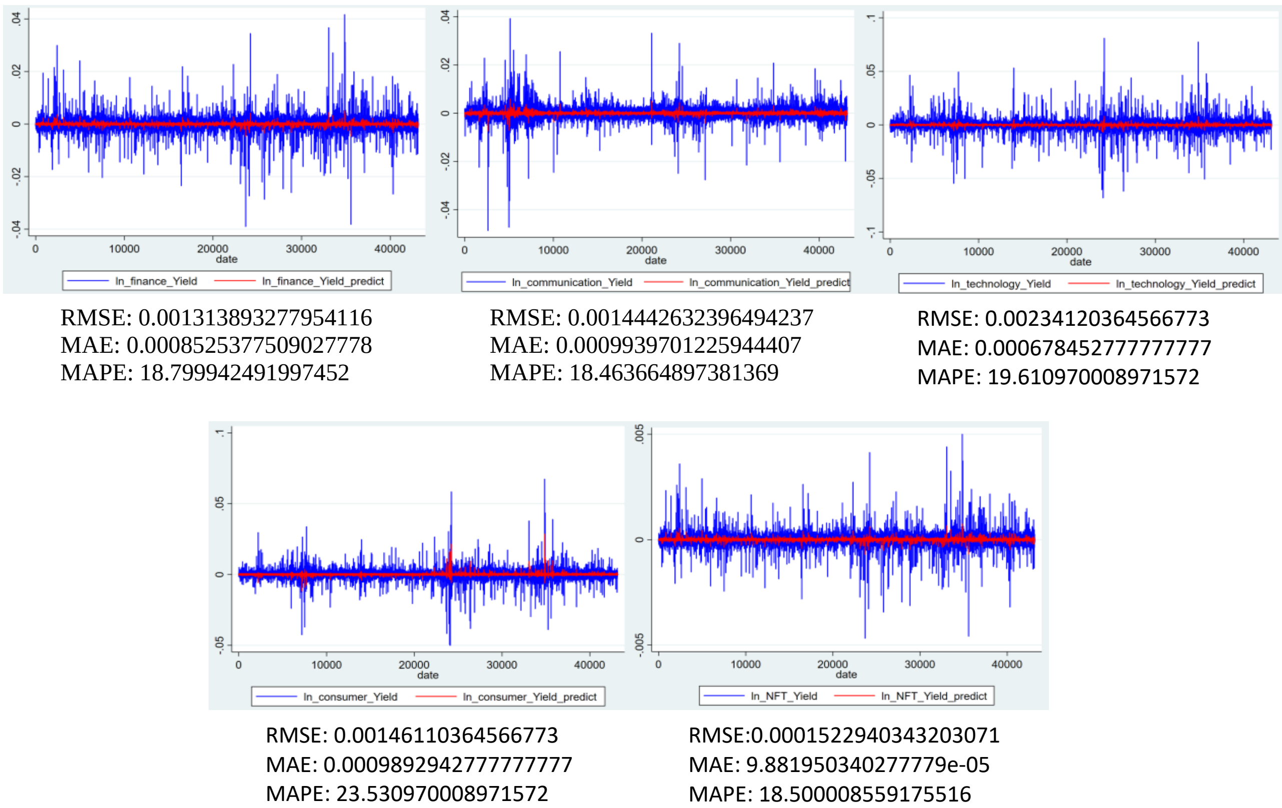

Figure 3: Time diagram of predicted value and actual value of each index.

The MAPE value was found to be 18.7%, Communication index yield test set MAPE value of 18.5%, Technology index yield test set MAPE value of 19.6%, Consumer index yield test set MAPE value of 23.5%, The NFT index yield test set MAPE value of 18.5%, Show that the fitting accuracy is not very high, Reneed to correct the prediction results of the model, Consider the prediction using the LSTM model for the exponential time series data of the data, And log-differential the predicted results, The time series of obtaining yields, Then, by combining the time series predicted by the ARMA-GARCH model and the LSTM model, Get the time series prediction results of the corrected yield.

4.5. Construct the LSTM Model

Using the exponential time series as the input data, an LSTM model was constructed using Python. The LSTM model can capture the dynamic features of the data by learning the dependencies between the sequences. Attempt to use LSTM models with different numbers of layers and compare their performance on validation data. The predictive power of the model can be compared using evaluation indicators such as root mean square error (RMSE) or mean absolute error (MAE). Looking at the performance of the model under different numbers of layers, finding a number of layers makes the model perform best on the validation set. The time series data of finance, communication, technology, consumption and NFT sector index is processed and the LSTM model is established. The fitting results are as follows:

DOI: 10.54254/2754-1169/55/20230962

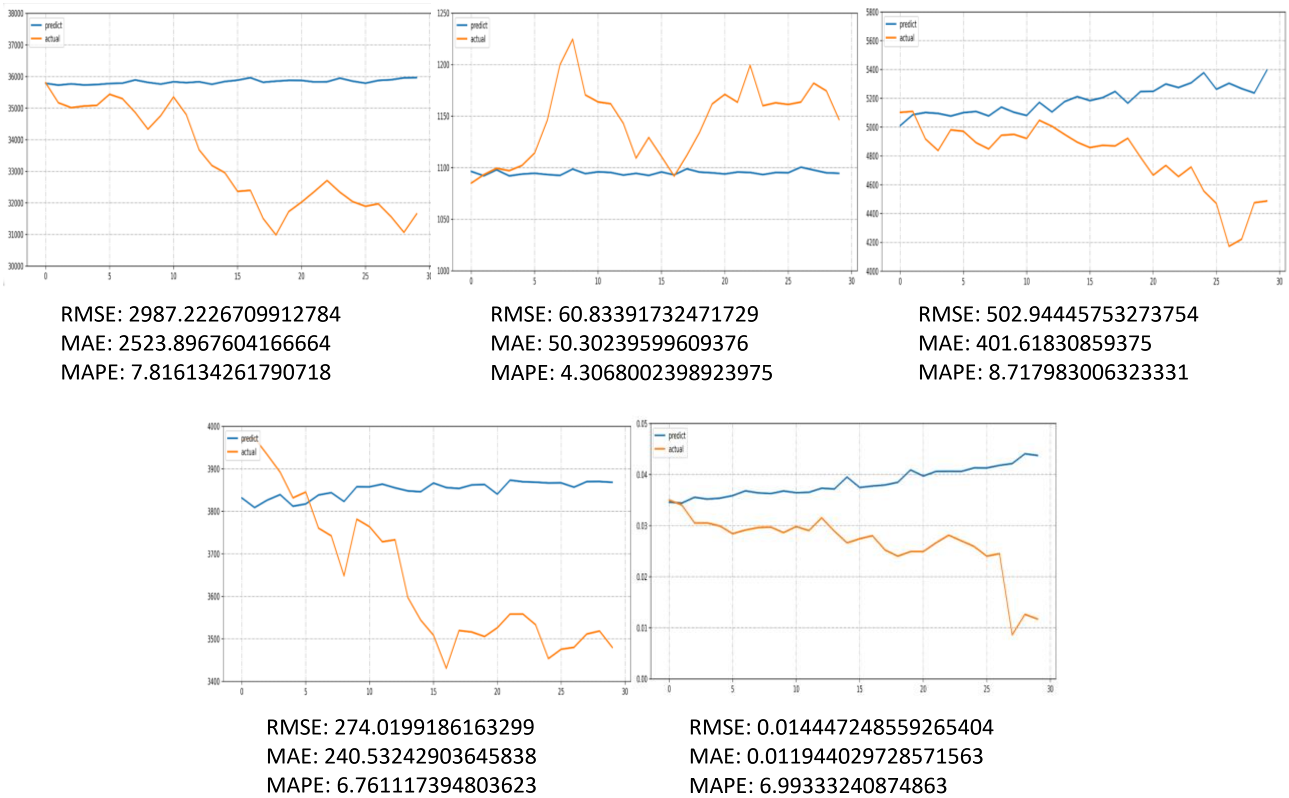

Figure 4: LSTM model fitting results for each exponential time series.

It is found that the MAPE value of financial index LSTM model is 7.81%, communication index time series LSTM model MAPE is 4.31%, technology index time series LSTM model is 8.72%, consumer index time series LSTM model MAPE is 6.76%, NFT index time series LSTM model MAPE is 6.99%, indicating that the index model has high fitting accuracy and can better describe the change trend of the index time series, which is suitable for time series prediction and the index prediction in the next 30 trading days.

Table 7: Comparison of fitting results before and after LSTM correction.

	The MAPE values of the ARMA-GARCH model	Model MAPE values from LSTM correction
ln_finance_Yield	18.7%	7.81%
ln_communication_Yield	18.5%	4.31%
ln_technology_Yield	19.6%	8.72%
ln_consumer_Yield	23.5%	6.76%
ln_NFT _Yield	18.5%	6.99%

4.6. Model Prediction Results

Each part of the LSTM model index time series prediction results for logarithmic difference get the yield prediction time series, and then with its ARMA-GARCH model yield prediction results for linear combination, get the five plate three days, reuse the coefficient of the principal component analysis method fitting Hong Kong NFT index. Besides,the data is drawn as a timing diagram, and the results are as follows:

DOI: 10.54254/2754-1169/55/20230962

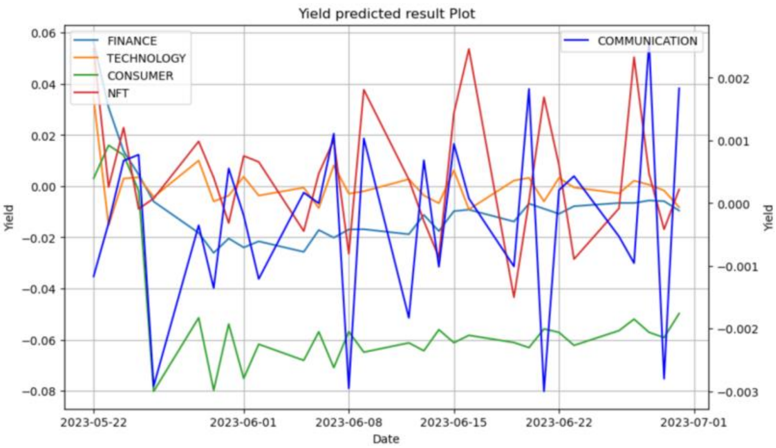

Figure 5: Time sequence chart of the forecast results of the five sectors in the next 30 trading days.

According to the above predicted return rate, the index prediction data of relevant sectors can be obtained through the inverse log difference [Pt = Pt-1 / (1-Exp (Yi))], and the coefficient in the principal component analysis method can be used to fit the NFT index of Hong Kong. The specific data are as follows:

Table 8: Index forecast data of each sector.

Date	FINANCE	COMMUNICATION	TECHNOLOGY	CONSUMER	NFT	The Hong Kong stock market NFT index
2023/5/22	35737.17982	1378.940418	5218.924413	2710.521647	0.012345	8603.167578
2023/5/23	34835.61564	1380.112668	4965.863429	2745.86599	0.011696	8400.306512
2023/5/24	34246.29321	1381.486397	5057.859067	2735.400147	0.011971	8303.211182
2023/5/25	33935.0032	1381.617795	5060.327728	2697.489299	0.011596	8233.410489
2023/5/26	33578.82562	1376.531013	5021.448769	2493.980457	0.011644	8095.129435
2023/5/29	33169.00204	1380.06463	5093.588885	2566.774009	0.011907	8054.544816
2023/5/30	32910.99276	1378.685247	5012.702524	2495.086665	0.01174	7969.030231
2023/5/31	33097.80346	1381.31683	5024.266618	2560.508576	0.011533	8026.94701
2023/6/1	32980.90604	1380.277578	5061.540758	2506.62013	0.01184	7994.6365
2023/6/2	33060.31942	1378.88428	5024.25252	2540.365339	0.011812	8013.357484
2023/6/5	32924.43145	1380.783525	5039.854989	2524.343647	0.011496	7985.81444
2023/6/6	33207.22269	1380.547875	4999.735438	2552.736286	0.011758	8040.533201
2023/6/7	33107.4089	1382.081527	5083.971871	2517.22615	0.011913	8025.844352
2023/6/8	33214.59734	1376.483158	5028.05402	2552.863645	0.011395	8046.659551
2023/6/9	33216.70046	1381.976142	5032.670655	2532.468162	0.01215	8041.934356
2023/6/12	33152.20726	1378.029899	5056.537733	2541.691436	0.011731	8036.683453
2023/6/13	33403.61499	1381.494983	5024.472903	2533.838377	0.011541	8075.863581
2023/6/14	33193.17516	1379.152621	5009.144853	2554.867569	0.01138	8040.133691
2023/6/15	33453.37025	1381.85749	5073.967549	2541.841349	0.012039	8096.589979
2023/6/16	33471.93924	1380.653745	4997.088524	2549.227712	0.012345	8088.540145
2023/6/19	33317.22259	1379.16032	5053.861677	2542.031291	0.011203	8067.316651
2023/6/20	33551.49468	1383.068082	5059.204996	2536.967211	0.011633	8110.896902
2023/6/21	33484.23595	1376.416714	5012.641885	2555.599428	0.012114	8095.25897
2023/6/22	33417.75466	1380.838671	5059.111584	2552.132074	0.011797	8090.417358
2023/6/23	33517.70777	1381.146439	5040.129361	2539.218352	0.011371	8101.697088

DOI: 10.54254/2754-1169/55/20230962

Table 8: (continued).

2023/6/26	33559.8781	1379.818128	5028.63314	2553.912699	0.0116	8112.054708
2023/6/27	33559.36242	1379.231979	5053.571882	2565.400747	0.012306	8119.989247
2023/6/28	33592.84219	1384.097201	5045.719487	2552.169945	0.011758	8121.113566
2023/6/29	33583.60761	1376.693368	5034.119543	2546.96027	0.011504	8115.011281
2023/6/30	33460.51193	1383.080744	5001.313843	2571.03928	0.011685	8094.248031

5. Conclusions and Implications

5.1. Application Advice

5.1.1. Financial Innovation

Based on the research outcomes and methodology of predicting NFT return volatility using the LSTM-corrected ARMA-GARCH model, several viable applications can be explored, particularly within the realm of financial derivative product innovation:

1)NFT Options Products.

The design and introduction of options products based on NFT return volatility is conceivable. By applying the forecasted volatility levels and trends from the model to option pricing models, diverse types of NFT options can be tailored, including call options, put options, and combination strategies. These options can meet investors' demands for NFT market volatility, providing them with more flexible risk management and investment opportunities.

2)Volatility Trading Strategies.

Leveraging the model-predicted NFT return volatility, volatility trading strategies can be developed and executed. These strategies may encompass volatility arbitrage, volatility trading, option combination strategies, etc., to profit from fluctuations in NFT market volatility. Investors can employ the predictive outcomes of the model, coupled with appropriate trading strategies, for risk management and portfolio optimization.

3)Risk Management Tools.

The application of model-predicted NFT return volatility can be extended to the development of risk management tools. For instance, designing risk exposure indicators, risk measurement models, or dynamic risk management strategies based on the model's forecasted outcomes. These tools can assist investors and traders in comprehending and managing the volatility risk of the NFT market, thereby enhancing the accuracy and efficacy of investment decisions.

5.1.2. Financial Deepening

The research outcomes and methodology of predicting NFT return volatility using the LSTM-corrected ARMA-GARCH model offer avenues for exploring their viable applications within the context of financial deepening in the national financial trading markets. During the process of financial deepening, effective regulatory and risk management mechanisms are typically required to safeguard investor interests, maintain market stability, and foster market development. The application of the LSTM-corrected ARMA-GARCH model to predict NFT return volatility can be applied to the regulation and risk management of the NFT market in the following aspects:

1)Risk Assessment and Monitoring.

Utilizing the forecasted NFT return volatility from the model, regulators can monitor and assess the risk level of the NFT market. Regulatory authorities can periodically evaluate the volatility levels and trends in the market based on the model's predictions, thereby understanding the market's risk condition and implementing corresponding regulatory measures.

DOI: 10.54254/2754-1169/55/20230962

2)Risk Alerts and Interventions.

Building upon the forecasted NFT return volatility, regulatory agencies can establish risk alert mechanisms and appropriate intervention measures. When market volatility surpasses specific thresholds or exhibits abnormal fluctuations, regulatory bodies can promptly issue warnings and take suitable actions, such as restricting trading activities or intensifying regulatory reviews, to mitigate potential risks and maintain market stability.

3)Product Innovation and Standardization.

Relying on the model's predictive outcomes, the innovation and standardization of NFT derivative products can be promoted. Regulatory authorities can evaluate and oversee various NFT derivative products based on market volatility predictions, ensuring their alignment with market demand, investor protection, and market stability requirements.

4)Investor Education and Risk Management Guidance.

Leveraging the predictive results of the model, regulatory bodies can offer investor education and risk management guidance. Regulatory authorities can utilize the predictive outcomes to provide investors with information and recommendations concerning risks in the NFT market, aiding investors in better comprehending and managing market risks. This approach contributes to both investor protection and the stability of market participation.

The integration of the LSTM-corrected ARMA-GARCH model into the regulatory and risk management framework of the NFT market can enhance transparency, stability, and investor confidence, fostering a conducive environment for the healthy development of the financial sector. Nevertheless, the practical implementation of these applications should be conducted meticulously, aligning with the specific requirements and dynamics of the national financial trading markets.

5.1.3. Financial Regulation

Within the realm of financial regulation, ensuring market stability and safeguarding investor interests stands as a crucial mandate. The application of the LSTM-corrected ARMA-GARCH model to predict NFT return volatility can be harnessed for regulatory purposes and the maintenance of market stability in the NFT market in the following aspects:

1)Market Regulation and Intervention.

Building upon the forecasted NFT return volatility, regulatory policies and interventions can be designed. Regulatory bodies can implement appropriate market regulation measures when market volatility surpasses a specific threshold or experiences abnormal fluctuations. Measures such as restricting leverage trading and increasing margin requirements can be employed to prevent excessive market volatility and potential systemic risks.

2)Regulatory Guidance and Policy Formulation.

Utilizing the predictive outcomes of the model, regulatory guidance and policy formulation can be provided. Regulatory authorities can formulate corresponding regulatory policies and guidance based on the model's predictions, with the aim of safeguarding investor interests, promoting market fairness and transparency, and fostering the healthy development of the NFT market.

3)Market Behavior Monitoring and Manipulation Detection.

Drawing from the forecasted NFT return volatility, monitoring market behavior and detecting signs of manipulation can be enhanced. The model can assist regulatory bodies in identifying market manipulation and abnormal trading activities, thereby reinforcing market surveillance and investigations to maintain market fairness and integrity.

It is imperative to emphasize that the aforementioned are merely illustrative examples of potential viable applications. The specific applications will depend on market demands, risk management needs, financial derivative product innovations, as well as the requirements, regulatory framework, and market conditions set forth by national financial regulatory institutions. In practical

DOI: 10.54254/2754-1169/55/20230962

implementation, further research, testing, and validation are necessary to ensure the accuracy and applicability of the model's predictive results, in conjunction with the integration of other regulatory measures and tools.

5.2. Future Prospects of the Model

In analyzing NFT return volatility trends, we employed principal component analysis to synthesize returns from closely correlated sectors (finance, technology, consumer, and communication). The LSTM-corrected ARMA-GARCH model was used for forecasting, effectively capturing historical return volatility patterns and trends. However, model fit results showed that the MAPE values for predicted returns in all five sector indices exceeded 1%. This suggests that relying solely on historical return data doesn't fully encompass the market's information impact on NFT return volatility prediction. To address this, model modifications should account for information acquisition effects.

Information cascade theory posits that market participants' behavior is influenced by others and information transmission. When some participants adjust their decisions due to information impact, it triggers a chain reaction affecting price trends. Additionally, market micro-structure theory examines trading mechanisms and participant interactions. It suggests that the presence of high-frequency traders may amplify the impact of information shocks on prices, given their rapid trading and responsiveness to market fluctuations. As financial markets deepen and quantitative trading grows, understanding the impact of information shocks on index return volatility becomes increasingly crucial for model refinement.

Information shocks refer to sudden impacts on the market caused by unexpected events, news, or other information. They can significantly influence the price trends of a specific sector. Information shocks impact index return volatility through the following transmission mechanisms:

1)Market Reaction Mechanism.

Information shocks prompt immediate reactions from market participants, resulting in heightened trading activity. As participants adjust their investment decisions based on the content and interpretation of information, trading volume and prices might experience substantial fluctuations.

2)Information Dissemination Mechanism.

Information shocks rapidly spread through channels like media, news outlets, and social media. The speed and reach of information dissemination determine the degree of awareness and reaction time among market participants. Different information holds varying impacts on different sectors. When critical information reaches a broader array of participants, they might adjust their price expectations for that sector, thereby influencing price trends.

3)Information Interpretation Mechanism.

The impact of information shocks on market participants depends on their interpretation and comprehension of the information. Diverse participants might interpret the same information differently based on their perspectives and information processing capabilities. Disparities in information interpretation might heighten buying and selling pressures, thus impacting price fluctuations.

Consider using the LDA (Latent Dirichlet Allocation) model to capture the information propagation mechanism, and using the preprocessed text data to train the LDA model (LDA is an unsupervised machine learning algorithm to discover topics from text data, where topics can be viewed as different information shocks.), Then the trained LDA model is used for topic classification (such as a news article), and the LDA model gives the probability distribution of each document belonging to a different topic. Then the impact effect is evaluated, and according to the results of the theme classification, the impact effect of different information shocks on different plates is analyzed. You can calculate the distribution of each topic in different plates, and the relative weight of the subject in a specific plate. This allows the impact of different themes on different plates. At the same time, attention should be paid to the interpretation and verification. According to the results of the impact effect evaluation, the impact mechanism of different information on different plates can be further explained and verified with the actual situation. For example, if a theme has a high weight in the financial sector, it means that the theme has a greater impact on the financial sector.

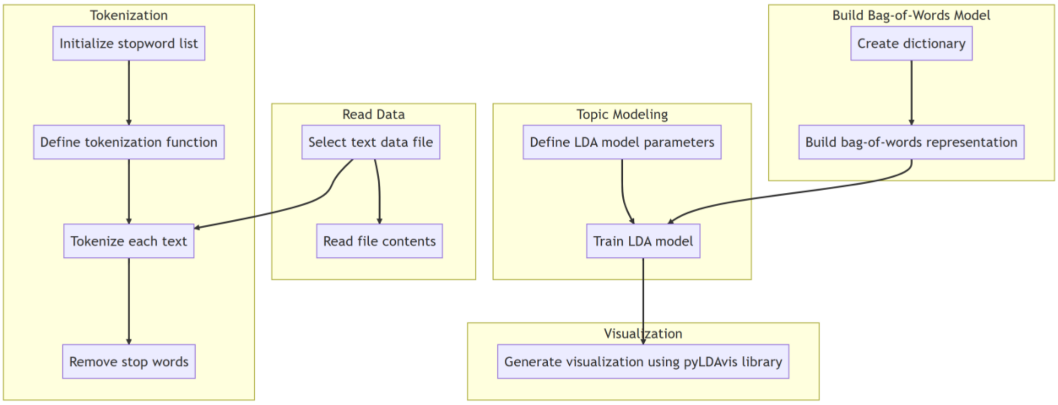

Figure 6: The LDA model construction logic.

Can consider the use of information emotional tendency analysis NLP (Natural Language Processing) model to reflect information interpretation mechanism, using pretreatment text data training emotion analysis model (emotion analysis is a text classification task, to judge the emotional tendency in the text, such as positive, negative or neutral emotion), for each text (such as a news report or a social media comments), using trained emotion analysis model for emotion classification. The model will give the emotional tendency of each text, and the corresponding emotional score or probability. Then, according to the emotion classification results, the impact effect of different information on market participants is analyzed. The distribution of emotional tendencies for different information types, and the relative weight of emotional tendencies among specific market participants can be calculated. This allows the extent to which different information affects market participants' expectations. At the same time, attention should be paid to interpretation and verification. According to the results of the impact effect evaluation, the impact mechanism of different information on the expectations of market participants can be further explained and verified with the actual situation. For example, if the distribution of emotional tendencies of a certain information type shows high negative emotions, it indicates that the information has a large negative impact on the expectations of market participants.

DOI: 10.54254/2754-1169/55/20230962

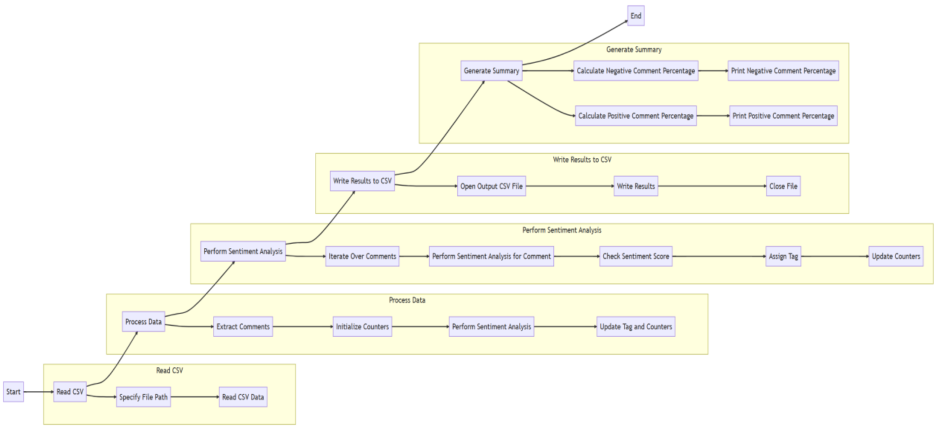

Figure 7: NLP Model Building Lics.

To sum up, considering the research method and conclusion of prediction of NFT yield volatility, can be used for financial product innovation, financial market deepening and financial regulatory response, so need to reflect the impact effect of market information, so consider the LSTM revised ARMA-GARCH model, on the basis of can introduce the analysis of NLP model and information impact theme classification LDA model, model correction.

5.3. Summary

The primary objective of this study was to analyze the volatility of Hong Kong Stock Exchange's (HKEX) NFT index returns. As the Hong Kong virtual asset exchange has yet to formally commence NFT trading, direct relevant data remains unavailable. Therefore, we adopted a strategy wherein associated high-impact basic sector index returns were used to synthesize an HKEX NFT index correlated with the global NFT sector index. To achieve this, four core sectors, namely financial services, technology, communication services, and consumer goods, were paired with the world NFT sector index to develop separate ARMA-GARCH models for return volatility prediction. However, these models exhibited suboptimal fitting. Subsequently, a three-loop LSTM model was constructed to forecast index prices for the five sectors. By logarithmically differencing these predictions, the LSTM and ARMA-GARCH forecasted return data were linearly combined, yielding a corrected set of forecasted returns. Leveraging principal component analysis, the forecasted return results for the HKEX NFT index were synthesized.

Based on our research into forecasting NFT return volatility using the LSTM-modified ARMA-GARCH model, we formulated practical suggestions and prospects for application. In terms of financial innovation, options products rooted in NFT return volatility, volatility trading strategies, and risk management tools could be designed and introduced to cater to investors' need for risk management and investment avenues within the NFT market. Concerning financial deepening, the model's predictions could be employed in national financial transaction markets for risk assessment, surveillance, risk alerting and intervention, product innovation and standardization, investor education, and risk management guidance. In the domain of financial regulation, the predictive outcomes could guide market control and intervention policies, offer regulatory guidance and policy references, and aid in market behavior monitoring to preserve fairness and integrity.

Nevertheless, it's imperative to acknowledge that the aforementioned applications are mere exemplifications. Their specific feasibility hinges on market demands, risk management requirements, financial derivative product innovations, and the specific circumstances of national financial

DOI: 10.54254/2754-1169/55/20230962

regulatory bodies, regulatory frameworks, and market dynamics. In practical application, further research, testing, and validation are prerequisites to ensure the accuracy and applicability of predictive results, harmonizing them with other regulatory tools and measures.

As we gaze into the future of this model, incorporating information acquisition influences for enhancement is advisable. Information cascade and market microstructure theories underscore the impact mechanisms of information shocks, encompassing market reaction, information dissemination, and information interpretation mechanisms. To this end, employing a topic-classifying LDA model for capturing information dissemination, alongside an NLP sentiment tendency model for assessing various information shocks' sectoral impacts, could offer a more comprehensive understanding and validation of NFT market volatility.

In summary, this study, rooted in the LSTM-modified ARMA-GARCH model for NFT return volatility prediction, presents feasible application suggestions for financial innovation, deepening, and regulation. Future efforts should refine the model by integrating other factors and methods to better address challenges within the NFT market, thereby fostering its sound development.

References

[1] *Dowling, M. (2021). Fertile LAND: Pricing non-fungible token. Finance Research Letters, Article 102096.*

[2] *Vidal-Tomás, D. (2022). The new crypto niche: NFTs, play-to-earn, and metaverse tokens. Finance Research Letters, Article 102742.*

[3] *Yizhi Wang, Volatility spillovers across NFTs news attention and financial markets, International Review of Financial Analysis, Volume 83, 2022, 102313, ISSN 1057-5219, https://doi.org/10.1016/j.irfa.2022.102313.*

[4] *Ko, H., Son, B., Lee, Y., Jang, H., & Lee, J. (2022). The economic value of NFT: Evidence from a portfolio analysis using mean-variance framework. Finance Research Letters, 47, Article 102784.*

[5] *Yousaf, I., & Yarovaya, L. (2022). Static and dynamic connectedness between NFTs, defi and other assets: Portfolio implication. Global Finance Journal, 53, Article 100719.*

[6] *Fu Xiaojie, Wang Zhen, Zhang Ge. Spatiotemporal evolution of Guangli River based on principal component analysis [J]. Water supply and drainage, 2022,58(S1): 812-814.DOI:10.13789/j.cnki.wwe1964.2021.06.18.0005.*

[7] *Zhao Keying, Mou Kai. Discussion on the evaluation method of mud shale reservoir based on grey correlation analysis and principal component analysis [J]. Geology and Exploration, 2023,59 (02): 443-450.*

[8] *Cui Xinying, Lv Zhiyuan, Zhang Mengmeng, Liu Yutao, Qin Bingwei, Zhao Qiaozhen, Li Xiaojie, Li Piwu. Establishment of the aroma substance evaluation model of medium and high temperature large koji based on principal component analysis method [J]. Journal of Food Safety and Quality Testing, 2023,14(07): 279-287.DOI:10.19812/j.cnki.jfsq11-5956/ts.2023.07.052.*

[9] *Miao Yingchun, Li Rui. Analysis of satisfaction of community residents based on principal component analysis —— Take Taiyuan, Shanxi Province as an example [J]. The Jinyang Academic Journal, 2023(01): 109-117.DOI:10.16392/j.cnki.14-1057/c.2023.01.008.*

[10] *Lu Yixuan. Comprehensive evaluation of employment quality level in China based on principal component analysis [J]. Journal of Hebei North University (Social Science Edition), 2023,39 (01): 39-44.*

[11] *Kumar, M., Thenmozhi, M., (2007). A comparison of different hybrid arima-neural network models for stock index return forecasting and trading strategy. In: Proceedings of 20th Australasian Banking and Finance Conference, Sydney, Australia.*

[12] *Abbasimehr, H., Shabani, M., Yousefi, M., (2020). An optimized model using LSTM network for demand forecasting.Comput.Ind.Eng.143, 106435.*

[13] *Tapia, S., & Kristjanpoller, W. (2022). Framework based on multiplicative error and residual analysis to forecast bitcoin intraday-volatility. Physica A: Statistical Mechanics and its Applications, 589(00), http://dx.doi.org/10.1016/j.physa.2021.126613.*

[14] *Zhuanxin Ding, Clive W.J.Granger, Robert F.Engle, A long memory property of stockmarket returns and a new model, Journal of Empirical Finance, Volume 1, Issue 1, 1993, Pages 83-106, ISSN 0927-5398, https://doi.org/10.1016/0927-5398(93)90006-D.*

[15] *Guo, X., Gao, Y., Li, Y., Zheng, D., Shan, D., (2021). Short-term household load forecasting based on long-and short-term time-series network. Energy Rep.7, 58–64.*

[16] *Kristjanpoller, R.W., & Hernández, P.E. (2017). Volatility of main metals forecasted by a hybrid ANN-GARCH model with regressors. Expert Systems with Applications, 84,290–300. http://dx.doi.org/10.1016/j.eswa.2017.05.024.*

DOI: 10.54254/2754-1169/55/20230962

[17] *Hu, Y., Ni, J., & Wen, L. (2020). A hybrid deep learning approach by integrating LSTMANN networks with GARCH model for copper price volatility prediction. Physica A: Statistical Mechanics and its Applications, 557(00), http://dx.doi.org/10.1016/j.physa.2020.124907.*