

# The Impact of The Development of Conversational AI on The Stock Indexes of Relevant Industries in The Chinese Stock Market

Yuxin Liu<sup>1, a</sup>

<sup>1</sup>School of Statistics, University of International Business and Economics, Beijing, 100029, China

<sup>a</sup>hhdedygg@163.com

---

**Abstract:** In recent years, the AI industry has been booming and has been favored by many investors. In 2023, generative AI has received a lot of attention due to the release of chat GPT, which has changed people's lives while also having a lot of impact on the stock market. In this paper, we will use the time series model and multivariate statistical analysis method to measure the impact on chat GPT on the stock price of related industries.

**Keywords:** ChatGPT, ARIMA(Auto-Regressive Moving Average Model), Multiple regression analysis, Event impact measurement.

---

## 1. Introduction

### 1.1. Subject Significance

#### 1.1.1. Stock index prediction is of great significance

The stock price trend directly affects the economic interests of investors, and also affects and reflects the macroeconomic policies of the country. Depending on the rise and fall of the index, the change trend of the stock price can be judged.

#### 1.1.2. AI, as a hot topic at present, has development potential and is worthy of in-depth study

After nearly 10 years of rapid development, AI technology has made great breakthroughs. With the increasing maturity of artificial intelligence theory and technology, the integration ability of artificial intelligence scene has been continuously improved. Therefore, commercial application has become the focus of AI technology enterprise layout in recent years. At present artificial intelligence technology has been in finance, medical care, security, education transportation, manufacturing, retail and other fields to achieve technology landing, and application scenarios are becoming more and more abundant. With the popularity of chat GPT various generative AI research companies around the world are also gradually making efforts, and the future market development is thriving, and its impact on economic development is not only reflected in the improvement of productivity, but also will have an impact on the stock market.

#### 1.1.3. Can be used as a typical case of the impact of one country's technological progress on the stock indexes of related markets in other countries, providing reference for future decisions

Due to the significant technological progress abroad, it will also have a significant impact on domestic investment. In the future, AI-related technological progress will also occur, but since AI is not a "long history" topic and there are few major breakthrough cases of the same scale studying the impact of conversational AI technology breakthroughs can help us make more accurate predictions about the stock market in the future when facing the improvement of important industries in the AI industry.

### 1.2. Overview of domestic and foreign research status

Generative AI, as a relatively novel topic, received public attention relatively late, and its large-scale development was concentrated in March and April in 2023. Due to the new topic, no relevant papers were published. At present, the existing relevant papers can be basically divided into three categories.

The first type of papers focus on evaluating whether the occurrence of sudden major event: has an impact on the stock market, For example, Chen Bo and Qian Huihui used ARIMA model to determine whether the novel coronavirus epidemic has an impact on the stock and bond market in research on the impact of the novel coronavirus pneumonia Epidemic on China's Stock and Bond Market.[1] Meanwhile, Zhu Min and Xu Zhongxiang used GARCH model in the impact of China-U.S Trade Frictions on China's Stock Market to study whether positive and negative events had an impact on abnormal returns and cumulative abnormal returns of different industries in China's stock market.[2] These articles focus on the impact of hot events on the stock market and answer the "YES OR NO" question without further research. At the same time, most articles focus on the impact of "black swan" events on the stock market.

The second type of paper tries to measure the impact of the event, The direct impact of events in one country on one country is observed, and the impact of events in one country on other countries is not looked at. For example, Andrew Ferguson, Peter Lam wrote 'Government policy uncertainty and stock prices: The case of Australia's uranium industry'.[3] The third category of papers focuses on the AI industry, conducts detailed research and forecasts of the future direction of the industry as a whole, focuses on the direct contribution of AI to productivity and social-economic development, and does not take into account the perspective of the stock market.

### 1.3. Innovation Points

First of all, what this paper does is not the traditional stock index forecast, but the time series correlation model is used to analyze the impact of events on stock indexes. More

importantly, the event we study is a technological advancement, not a widely studied global event, and at the same time, it is a research on open AI, which is a new event. Finally, this paper is to study the impact of overseas events on China's stock market, rather than the impact of domestic events on China.

## 2. Materials and Methods

### 2.1. Data Sources

The data set used in this article is from the wind database. In order to make the surveyed stocks fully reflect the impact of events on the development of generative AI, we use the CHATGPT index of China in the database for research and use the data at the close of the market. In terms of event measurement, the selection of news events is based on the official news release time rather than the local media release time, and the events are classified according to the article title.

### 2.2. Basic research methods

First of all, in order to eliminate random interference, this paper uses the singular spectrum analysis algorithm to separate the main trend component from the random fluctuation component in the original stock index time series, so as to avoid the interference of random fluctuation in the identification process of the main trend component. For the main trend component of the original stock index time series, the total variance explanation ratio is ensured to be above 95%.

In order to ensure that chat GPT related events really have an impact on the stock price of generative AI, this paper refers to the practice of Chen Bo, Qjan Huihui, and uses the historical normal closing price transaction data of the stock and bond market before the time node studied to establish ARIMA. Hope to establish a time series model to fit the inherent trend law of the stock index before the occurrence of a special event, so as to show that the event has a significant impact on the stock index after the occurrence. The data before the event is used as the basis for modeling, and the training set and test set are established to observe the prediction effect of the model on the stock index before the event to ensure the accuracy of the model prediction. Finally, the model is used to predict the stock index after the event and compared with the real value. Since the data used in the modeling comes from the data before the event, it should only predict the trend of the time series under the influence of the original trend and random fluctuations. In order to prove that the deviation of the stock index from the original trend is due to the influence of the event rather than random fluctuations, 95% confidence interval of the model prediction is set as a test. If the real data after the event does not fall into the 95% confidence interval predicted by the model, then it is reasonable to say that the event has a significant impact on the stock index that makes it deviate from the original change trend. Meanwhile, the Holt two-parameter exponential smoothing method was used for fitting and verification of the results.

Under the condition of ensuring that related events really

have an impact on the stock index the degree of impact is measured.

This paper uses the method of dummy variables and refers to Andrew Ferguson, Peter lam. Government policy uncertainty and stock prices: In The case of Australia's uranium industry, a regression model with dummy variables is used to measure the specific impact of specific events.

$$GPT_t = \beta_0 + \beta_1 MARKET_t + \beta_2 AI + \sum_{k=1}^K (\lambda_k^0 D_k^t + \lambda_k^1 D_k^{t-1} + \lambda_k^2 D_k^{t-2}) + \varepsilon_t \quad (1)$$

$$D_k^t = \begin{cases} 1, & \text{Specific event k occurs} \\ 0, & \text{Otherwise} \end{cases} \quad (2)$$

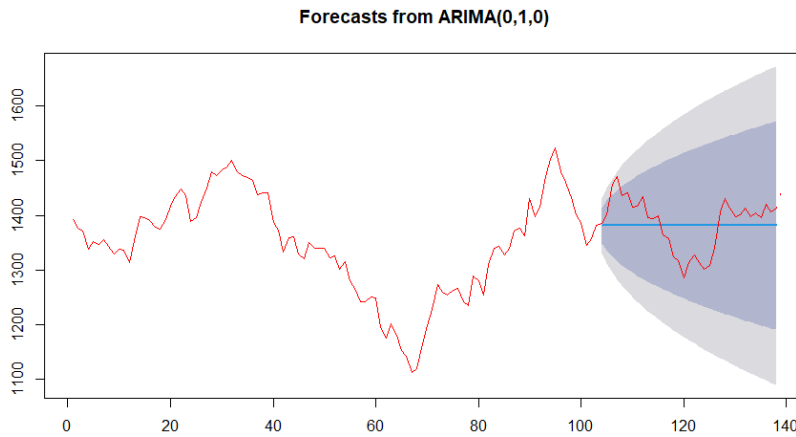
MARKET is a stock index that can measure the change of market situation and reflect the overall change of the market. It is proposed to use the CSI 800 stock index. AI is a variable that reflects the whole industry. It is proposed to use the AI stock index published by wind to ensure that MARKET and AI are orthogonal. In this model, the change of Chat GPT index can be divided into four parts: the change of the stock market as a whole, the change of the industry, the abnormal return brought by special events and the random disturbance. In order to measure the changes brought by abnormal returns, dummy variables were used, and an additional period of lag was added based on Andrew Ferguson, taking into account the reaction time to foreign news and the lag of news, and the two periods of lag of its impact were simultaneously considered in the model. After fitting the model, the variables that fail the test can be removed, and the events that have an impact on the stock index and their impact time can be obtained. The overall impact of the events can be obtained by summing up the impact coefficients of each period.

$$\lambda_k = \lambda_k^0 + \lambda_k^1 + \lambda_k^2 \quad (3)$$

## 3. Results and Analysis

### 3.1. Time series model -- Determine whether the impact occurs

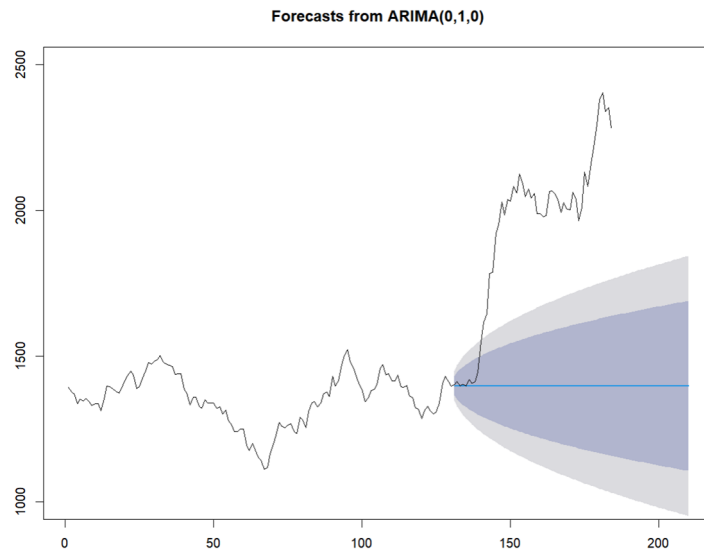
We first assume that the events from January to March 2023 have caused the stock index to deviate from its original trajectory, so we need to establish a time series model with the data before January 2023. After analyzing the original data through the singular spectrum analysis, we use the data from July 1, 2022 to December 1, 2022 as the training set, and the data from December 1, 2022 to January 1, 2023 as the test set, respectively using ARIMA and two-parameter exponential smoothing method to model and compare the effect of the model on the test set, Taking the ARIMA model as an example, Figure 1 shows the effect of the model built using the training set data on the test set, The fluctuations of the data are all within the 95% confidence interval, which conforms to our hypothesis that the stock index changes according to its inherent law before the event occurs.



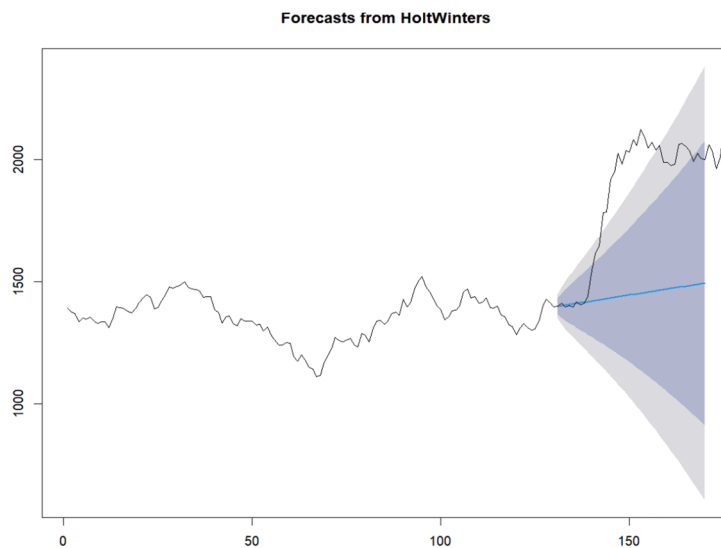
**Figure 1.** The effect of the model built using the training set data on the test set

On the test set, both models have good results, and the residuals pass the white noise test. In the test set, both models have high test accuracy, which are 0.985 and 0.984 respectively. ARIMA model in Figure 2 is slightly better than Holt smoothing method in Figure 3. When the model is applied to the data after the event (from January 1, 2023 to March 1, 2023), the result is shown in the figure below. It can be seen that the trend of the stock index after the event can no

longer be captured by the mode that does not contain the influence of the event and fits the data before the event: The real data broke out of the 95% confidence interval of the original model, confirming the previous assumption that the occurrence of the event changed the inherent trend of the relevant stock index so that it could no longer be predicted by the original model that effectively predicted it.



**Figure 2.** The ARIMA model is applied to the data after the event



**Figure 3.** The Holt smoothing model is applied to the data after the event

### 3.2. Multiple regression model with dummy variables -- event influence measure

In order to explain the abnormal fluctuations of relevant stock indexes and their relationship with related events, we collected relevant news from the internet that occurred from January 1, 2023 to March 1, 2023. The news can be broadly divided into four categories: actual updates, technical

cooperation with other companies, investment and capital flows, and academic comments.

First of all, all events are set as dummy variables according to the above model, and all events are sorted into four dummy variables according to classification for regression. Among them, the data of three stock indexes are logarithmically processed, and the results are shown in Table 1.

**Table 1.** Results of the first regression analysis

Ingpt	Coef.	Std.Err	t	P> t	95% Conf. Interval	
lnai	1.828653	0.051984	35.18	0	1.720217	1.93709
lnmarket	-1.22451	0.331804	-3.69	0.001	1.720217	-0.53238
up	0.00922	0.007681	1.2	0.244	-0.0068	0.025241
in	-0.03297	0.0104	-3.17	0.005	-0.05466	-0.01128
aca	-0.00648	0.006269	-1.03	0.313	-0.01956	0.006595
pro	0.022346	0.009854	2.27	0.035	0.001791	0.042902
constant	0.276908	2.300613	0.12	0.905	-4.52209	5.075902

Up is the actual update, in is the investment and capital flow, aca is the evaluation of the academic community, and pro is the technical cooperation with other companies.

It can be seen that for the two events of academic evaluation and investment, the P-value is greater than 0.05, and the T-test does not reject the null hypothesis, indicating

that the two events have no significant impact on the stock index and are regarded as having no significant impact on the change of the stock index.

After removing the two non-significant events and the intercept term, the results obtained are shown in Table 2:

**Table 2.** Regression results after removing non-insignificant variables

Ingpt	Coef.	Std.Err	t	P> t	95% Conf. Interval	
lnai	1.789372	0.047278	37.85	0	1.691569	1.887175
lnmarket	-1.13782	0.057911	-19.65	0	-1.25762	-1.01802
in	-0.03445	0.009597	-3.59	0.002	-0.0543	-0.0146
pro	0.019488	0.008208	2.37	0.026	0.002509	0.036468

Where  $R^2$  is greater than 0.99 and p value of F test is less than 0.01, indicating that the regression equation is significant. Test the variance inflation factor VIF, whose mean value is 1.8, without collinearity problem. The white heteroscedasticity test is performed on it with a p-value of 0.6837, and the null hypothesis of homoscedasticity is not rejected.

It can be seen from the results that the open investment and capital inflow have a negative impact on the growth of the stock index, the technical cooperation with other companies has a positive impact on the growth of the stock index, while the academic evaluation and the actual technical update have no significant impact on the growth of the stock index.

## 4. Conclusion

It can be seen from the test results of the time series model that the mode that does not take into account the impact of the event has a good performance in predicting the stock index before the event occurs, but fails to predict the event after the event occurs, indicating that the occurrence of the event does have an impact on the stock index, so the impact of various events on the stock index can be further studied. According to the results of multiple regression analysis, it can be seen that among the four major categories of events, public investment

and capital inflow and technical cooperation with other companies have significant impact on the stock index. For public investment and capital inflow, it is speculated that such events may cause people who want to invest in the related stocks to switch to the stocks of the industry that has injected capital into the industry, thus reducing the demand and decreasing the stock price. For technical cooperation with other companies, it is speculated that such incidents increase the investor's confidence that the industry has the potential for growth and broad application scenarios, so the demand for purchases increases and the stock price rises.

## References

- [1] Chen Bo & Qian Huihui. (2021). Study on the impact of COVID-19 on China's stock and bond market. *Industrial Technology and Economics*, (11), 53-60.
- [2] Zhu Min & Xu Zhongxiang. (2021). The impact of Sino-US trade Friction on China's stock market. *Research in International Finance*, (04), 3-12.
- [3] Andrew Ferguson & Peter Lam. (2016). Government policy uncertainty and stock prices: The case of Australia's uranium industry. *Energy Economics*, (60), 97-111.