The Research on the Prediction of VIX Index Based on the Machine Learning Method

Ke Chen[1*]

[1] School of Finance, Shandong Technology and Business University, Shandong 264005 China

*2021212222@sdtbu.edu.cn

Abstract. As a barometer of quantitative investment, the volatility and future trend of VIX index can provide a reference for investors. This paper uses the data of CBOE from 1990 to 2022 to analyze its data characteristics. It finds the VIX index can be divided into two blocks, a flat zone and a rising zone. In the stable area, the VIX index satisfies the characteristics of mean reorientation, and the fluctuation range is basically near the historical average. In the climbing area, the unstable market environment leads to a surge in people's risk aversion, which leads to a significant increase in the VIX index. Besides, this paper further uses the machine learning method to predict the future VIX index. It suggests that the VIX fluctuated up and down a bit, but did not spike, indicating that the market is still in a normal state. Through the prediction of the trend of VIX index, the public can provide a forward-looking forecast, for the future need to avoid risk, to respond to risk in advance to provide a feasible forecast scheme.

Keywords: VIX index, Prediction, Machine Learning

1 Introduction

The VIX index has attracted much attention since it was introduced by the Chicago Board Options Exchange in 1990 as a barometer of stock market volatility and has been studied by many scholars since its introduction. Whaley (2009) brought up the definition of VIX. VIX was seen as a forward-looking record of the normal return unpredictability of the US S&P 500 Index over the course of the following 30 days. The VIX is inferred from the costs of S&P 500 file choices, prevalently involved by the market for of guaranteeing the worth of stock portfolios. Elevated degrees of the VIX reflect financial backer nervousness in regards to an expected drop in the securities exchange, similarly as high Hood insurance payments reflect mortgage holder uneasiness about conceivable harsh weather conditions [1]. Simonato et al. (2017) inspected in this how a model verbalized throughout a period changing non-Gaussian dispersion with restrictive skewness and kurtosis can add to the general clarification of the VIX elements [2]. Magafas et al. (2017) applied non-direct strategies to break down and anticipate the everyday VIX record which is perhaps of the main stock list on the planet. The point of the investigation is to quantitatively demonstrate it the relating time series is a deterministic turbulent one. Besides, it can be accomplished at least one days ahead forecast.

V. Escudero et al. (Eds.): BADS 2022, AHCS 11, pp. 1448–1456, 2023.
https://doi.org/10.2991/978-94-6463-102-9_152

The exploration utilizes Grassberger and Procaccia's technique in the time series examination to assess the connection and least implanting aspects of the comparing abnormal attractor in Non-Linear Properties of the VIX Index [3]. Bao et al. (2012) pointed out that through model comparison, it is found that LR model and LRJ model are superior to SR model and GBM model, and it is necessary to add stochastic volatility in LR model and LRJ model [4]. Pan et al. (2019) fostered a new summed up autoregressive restrictive heteroskedasticity (GARCH) model that records for the data overflow between two business sectors. This model is utilized to identify the convenience of the CBOE instability record (VIX) for working on the exhibition of unpredictability guaging and choice estimating. They track down the huge capacity of VIX to foresee stock unpredictability both in-example and out-of-test. VIX data additionally serves to extraordinarily decrease the choice estimating mistake [5]. Kyrtsou et al. (2019) in the expect to investigate the complicated connections between S&P500, VIX and volume they presented a Granger causality test utilizing the nonlinear measurement of Asymmetric Partial Transfer Entropy (APTE). Through a reenactment work out, it emerges that the APTE offers exact data on the idea of the network. The exact discoveries in this paper concretize the data stream that joins volume, S&P500 and VIX, and combine the influence impact and the unbalanced stock return-volume relationship into a brought together structure of investigation [6].

At the same time, many scholars use machine learning to predict the future prices of some commodities. Chowdhury et al. (2019) had utilized AI approach utilizing the product Rapidminer, where they had taken on various calculations like the choice tree, gathering learning technique and brain organization. It had been seen that, the expectation of close cost utilizing AI is basically the same as the one got utilizing BSOPM. AI approach stands apart to be a superior indicator over BSOPM, on the grounds that Black-Scholes-Merton condition incorporates hazard and profit boundary, which changes ceaselessly. They had additionally mathematically determined unpredictability. With the stock's cost went up because of overpricing, unpredictability increments at a gigantic rate and when instability turns out to be extremely high; generally, the market will be falling, which can be noticed and decided utilizing our changed BSOPM. The proposed changed BSOPM has additionally been made sense of in light of the relationship of Schrodinger condition (and intensity condition) of quantum physical science [7]. Adegboye & Kampouridis, (2021) proposed an original DC-based structure, which utilizations AI calculations to foresee when a pattern will switch. This permits merchants to be in a situation to make a move before this occurs and hence increment their productivity. They join our methodology with a clever DC-based exchanging system and play out an inside and out examination, by applying it to 10-min information from 20 unfamiliar trade markets more than a 10-month time frame. The all out number of tried datasets is 1,000, which permits us to contend that our outcomes can be summed up and are broadly appropriate. They contrast our outcomes with ten benchmarks (both DC and non-DC based, like specialized investigation and purchase and-hold). The findings in this paper show that the proposed approach can not only obtain a fundamentally higher benefit, but also diminished risk, and genuinely beat the other exchanging techniques various different execution measurements [8]. Barboza et al. (2017)tried AI models (support vector machines, sacking, helping, and arbitrary timberland) to foresee

chapter 11 one year before the occasion, and contrast their presentation and results from discriminant examination, calculated relapse, and brain organizations[9]. Arne Thorvald Gierløff Hollum et al. (2013) investigated the utilization of AI strategies in arranging monetary news to foresee stock cost developments [10].

Besides, VIX index is a barometer of quantitative investment, and predicting the future trend of VIX will be of great help to the analysis of market conditions and how to analyze investment strategies. Therefore, this paper hopes to predict the future trend of VIX index and provide feasible solutions through machine learning.

The main contributions of this paper include that theoretically enriching the research on the prediction of VIX index and practically being helpful for investors to prevent future risks in advance, hedge risks and reduce investors' losses.

The remaining of this study is structured as follows. Section 2 analyses the VIX Index's characteristic. Section 3 forecasts VIX Index. Section 4 is the cause analysis. Section 5 concludes.

2 Characteristic analysis of VIX index

2.1 Definition of VIX Index

The VIX index is also known as the fear index. Introduced by the Chicago Board Options Exchange in 1990, it is a benchmark measure of the future volatility of S&P 500 index options. The VIX index is a real-time measure of the market's expectations for volatility over the next 30 days. It can warn of impending crashes, and thus reflect investment psychology.

The technique for ascertaining the VIX Index is utilizing standard SPX choices and week by week SPX choices that are recorded for exchanging on CBOE Options. Standard SPX choices lapse on the third Friday of every month and week by week SPX choices terminate on any remaining Fridays. Just SPX choices with Friday terminations are utilized to ascertain the VIX Index. Just SPX choices with over 23 days and under 37 days to the Friday SPX termination are utilized to ascertain the VIX Index. This paper set the SPX choices to be weighted to yield a consistent development 30-day proportion of the normal unpredictability of the S&P 500 Index. The summed up equation utilized in the VIX Index computation is:

$$\sigma^2 = \frac{2}{T}\sum_i \frac{\Delta K_i}{K_i^2} e^{RT} Q(K_i) - \frac{1}{T}\left[\frac{F}{K_0} - 1\right]^2 \quad (1)$$

The VIX Index isn't extraordinary as a financial exchange instability list, however, is just quick to have been presented and subsequently has a first-mover advantage. The procedure for processing the record isn't interesting to the costs of S&P 500 file choices. It can be adopted to any list choice market. Without a doubt, the CBOE has proactively applied the philosophy to make instability record for the NASDAQ 100(VXN) and for the DJIA(VXD). The just significant imperative is that the fundamental list choice market has profound and dynamic exchanging across an expansive scope of activity costs. Of course, the VIX has likewise drawn in global imitators. The NYSE Euronext has

applied a similar philosophy to file choices recorded on the AEX (a list of 25 stocks exchanged Amsterdam), the BEL20 (a record of 20 Belgium stocks), the CAC40 (a list of 40 French stocks), and the FTSE 100 (a record of 100 stocks exchanged the U.K.).

2.2 Characteristic of VIX from 2000-2020

As a forward-looking pointer, the VIX Index likewise will in general be mean returning; over the long run it will for the most part get or move once again to its verifiable normal. Unpredictability can't move higher in ceaselessness. It additionally can't move to nothing, however when you crunch the authentic data, you can see that the VIX Index spiked pointedly in 2008 and 2020.

Fig. 1. VIX index from 2000 to 2020.Retrieved from https://www.cboe.com

The VIX Index corresponding options implied volatility, so it can response the risk aversion of investors, when the stock market fell sharply, it will cause investors to buy put options to hedge risk, so at this time, implied volatility will also rise. According to the figure 1, the VIX index can be divided into two periods, one is from 2000 to pre-2008 and 2010 to early 2020, the VIX index volatility is small, and has been reversion to the historical mean, this is the VIX index plateau. The other is in 2008 and 2020, because the outbreak of the 2008 global financial crisis and the outbreak of COVID-19, leading to the peak of risk aversion, the VIX index also soared, this is the VIX index spike period. And in 2008 the outbreak of the financial crisis in 2020 and the outbreak of the COVID-19 lead to the stock market is not stable. As a result, investors' risk aversion continues to increase, reaching two peaks in historical data. At the same time, by comparing the VIX value of 2008 and the VIX value of 2020, it can be found that COVID-19 has a greater impact on the public in 2020.

3 Forecast of VIX index

3.1 Characteristic of VIX from 2000-2020

For the prediction of VIX index, machine learning method is adopted in this research. The main steps are data mining, processing, establishment of prediction model, model fitting and visualization of prediction data.

3.2 Data processing

The data used in this research study are historical data from the Chicago Board Options Exchange from January 2, 1990 to February 14, 2022.

Data processing includes three steps: data extraction integration, data analysis and data interpretation. Firstly, the historical data of VIX index is processed. There is no missing data in the provided data, but after careful study, it is found that there is no label field in the stock data. Therefore, the problem of label data needs to be solved before the linear regression model is created. Consider the amount of data to determine the time range. There are a total of 8090 rows of data. It is assumed that the number of predicted days is determined by dividing the amount of data by 100, and then the corresponding label value is constructed. Then the time translation operation is carried out with PANDAS. The weighted closing price field was extracted and converted into DataFrame data type. Shift (-5) transformation was further applied and assigned to the new field. The new field was subtracted from the original field to obtain the data with a difference of 5 days. And apply it to the entire VIX Index dataset.

3.3 Simulation

After the construction of the label field data is completed, the correlation between the feature field and the label field is checked, and the thermal map is used to display the output as shown in figure 2.

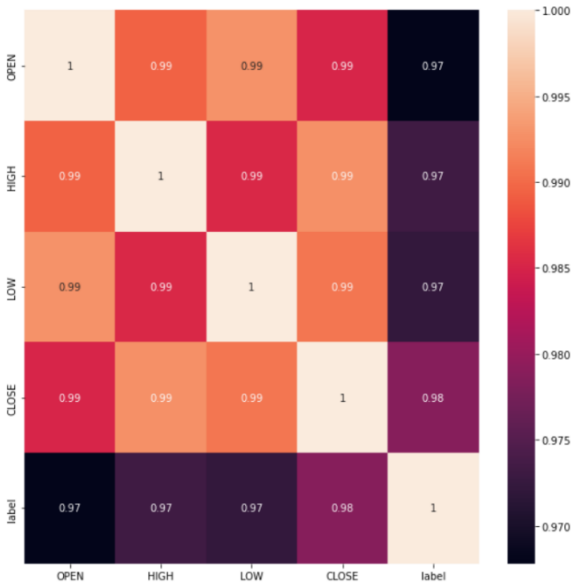

Fig. 2. Correlation thermal map. The picture is self-drawn

According to the results of correlation analysis, the field data is extracted again, and the data set is constructed. Two new fields are created, namely the change difference between high and low prices and the rise and fall, and the missing values are processed. The feature data X and label data Y are divided, and then this paper divides the training set and test set by using the ratio of 9:1. After the model is created and fitted, the prediction is made.

According to the forecast results, the VIX will fluctuate between 15% and 31% over the next 81 days, and there is a 96.1% accuracy rate to put the price in the range of 15% to 31%.

Table 1. Forecast data sheet. Tables are self-drawn tables

Date	VIX	Date	VIX	Date	VIX	Date	VIX
2022-02-15	15.8216	2022-03-16	17.1274	2022-04-14	21.4899	2022-05-13	23.5596
2022-02-16	15.6047	2022-03-17	17.5460	2022-04-15	23.1442	2022-05-16	25.0061
2022-02-17	15.1288	2022-03-18	17.9386	2022-04-18	21.1068	2022-05-17	28.2199
2022-02-18	15.4971	2022-03-21	19.0068	2022-04-19	18.7986	2022-05-18	29.8957
2022-02-21	15.3529	2022-03-22	19.4434	2022-04-20	18.0259	2022-05-19	31.0733
2022-02-22	16.0010	2022-03-23	18.7027	2022-04-21	17.7978	2022-05-20	31.1240
2022-02-23	16.9295	2022-03-24	27.8803	2022-04-22	17.6562	2022-05-23	30.3322
2022-02-24	16.6129	2022-03-25	22.9826	2022-04-25	17.0399	2022-05-24	27.8989
2022-02-25	16.3725	2022-03-28	26.7066	2022-04-26	17.3467	2022-05-25	25.1350
2022-02-28	16.5318	2022-03-29	29.8301	2022-04-27	17.3107	2022-05-26	22.1740

2022-03-01	16.1435	2022-03-30	27.9533	2022-04-28	16.7277	2022-05-27	21.9012
2022-03-02	15.2010	2022-03-31	30.0320	2022-04-29	16.9659	2022-05-30	24.0717
2022-03-03	15.5008	2022-04-01	27.2934	2022-05-02	19.4191	2022-05-31	23.1858
2022-03-04	16.4156	2022-04-04	22.0503	2022-05-03	19.6420	2022-06-01	22.8563
2022-03-07	17.2274	2022-04-05	20.0651	2022-05-04	18.8630	2022-06-02	21.5566
2022-03-08	17.8190	2022-04-06	21.3911	2022-05-05	19.5392	2022-06-03	20.0684
2022-03-09	18.6341	2022-04-07	18.8683	2022-05-06	18.5220	2022-06-06	23.4143
2022-03-10	17.7175	2022-04-08	20.1973	2022-05-09	17.70145	2022-06-07	26.8499
2022-03-11	16.3987	2022-04-11	21.6399	2022-05-10	20.00061		
2022-03-14	16.6214	2022-04-12	19.4572	2022-05-11	19.3123		
2022-03-15	16.4495	2022-04-13	20.3436	2022-05-12	22.5687		

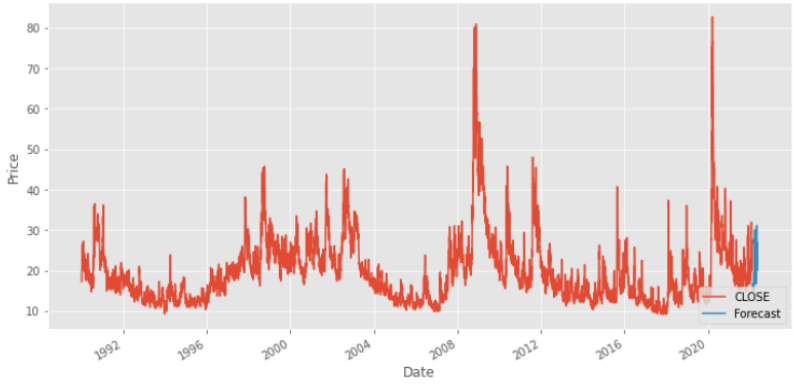

Fig. 3. Prediction data visualization. The picture is self-drawn

After visualizing the forecast data, it is found that the future price of VIX futures does not rise or fall too much, and still satisfies the mean reversion model, moving closer to the historical average. Indicates that the market is still in normal condition.

4 Cause analysis

As can be seen from the forecast results, although the VIX index fluctuates up and down, the fluctuation range is not particularly large, indicating that the current market conditions are good, people's risk aversion is not particularly high, and the market is still in a normal state of operation.

The VIX has certain forward-looking, it can reflect the investor's risk aversion, future prediction can be found that the VIX is still maintain a relatively stable state in the future, that future market conditions will remain relatively good condition, by machine learning to predict the VIX, can grasp the future market conditions, and make corresponding countermeasures. VIX, can grasp the future market conditions, and make corresponding countermeasures.

5 Conclusion

As a barometer of quantitative investment, VIX index has attracted much attention since its launch. This paper analyses the historical data of VIX index. Moreover, the future trend of VIX index is modeled, fitted and predicted by the method of machine learning. There are three conclusions in this study. Firstly, by analyzing the characteristics of VIX index, it is found that, the VIX index can be divided into two blocks, a flat zone and a rising zone. In the stable area, the VIX index satisfies the characteristics of mean reorientation, and the fluctuation range is basically near the historical average. In the climbing area, the unstable market environment leads to a surge in people's risk aversion, which leads to a significant increase in the VIX index. Secondly, by predicting of the VIX index, it is found that the VIX index has little fluctuation in the future and still shows a state of mean reversion, indicating that the market is still in a normal state, Thirdly, after building the model and fitting the model, it is found that the prediction accuracy of the prediction model is 96.14%.This paper analyzes and predicts the future trend of VIX index through historical data through the method of machine learning, it helps to provide feasible suggestions for investors to hedge risks in the future. It is helpful for other scholars to better predict and analyze the future situation of the market.

References

1. Whaley, R. E. . (2009). Understanding the vix. Journal of Portfolio Management, 35(3), 98-105.
2. Simonato, Jean-Guy, Lalancette, & Simon. (2017). The role of the conditional skewness and kurtosis in vix index valuation. European Financial Management the Journal of the European Financial Management Association.
3. Magafas, L. , Hanias, M. , Tavlatou, A. , & Kostantaki, P. . (2017). Non-linear properties of the vix index. International Journal of Productivity Management and Assessment Technologies, 5(2), 16-24
4. Bao Q., Chen S.and Li S..(2012). Pricing and correction of Vix options. Financial Theory and Practice (4), 4.
5. Pan, Z. , Wang, Y. , Liu, L. , & Wang, Q. . (2019). Improving volatility prediction and option valuation using VIX information: a volatility spillover garch model. Journal of Futures Markets, 39(6), 744-776.
6. Kyrtsou, C. , Kugiumtzis, D. , & Papana, A. . Further insights on the relationship between sp500, vix and volume: a new asymmetric causality test1.
7. Chowdhury, R. , Mahdy, M. , Alam, T. N. , Dastegir, G. , & Quaderi, A. . (2019). Predicting the Stock Price of Frontier Markets Using Modified Black-Scholes Option Pricing Model and Machine Learning.
8. Adegboye, A. , & Kampouridis, M. . (2021). Machine learning classification and regression models for predicting directional changes trend reversal in fx markets. Expert Systems with Applications, 114645.
9. Barboza, F. , Kimura, H. , & Altman, E. . (2017). Machine learning models and bankruptcy prediction. Expert Systems with Applications, 83.

10. Arne Thorvald Gierløff Hollum, Mosch, B. P. , & Zoltán Szlávik. (2013). Economic sentiment: text-based prediction of stock price movements with machine learning and wordnet. Springer Berlin Heidelberg.