STOCK INDEX VOLATILITY PREDICTION

Introduction:

The realm of financial markets is a dynamic and complex ecosystem where investors and financial institutions navigate through a sea of data, attempting to make informed decisions in a volatile environment. The ability to understand and predict the behavior of financial indices is of paramount importance, as it enables market participants to assess risk, optimize portfolios, and make strategic investment choices. In this context, time series analysis and volatility modelling stand as essential tools in the financial analyst's toolkit, offering insights into the past, present, and potential future movements of these indices.

This report delves into a comprehensive study of four prominent global financial indices: the S&P 500, NIFTY 50, Hang Seng Index, and Nikkei. The goal of this analysis is to provide a deeper understanding of the dynamics of these indices by employing time series modelling techniques.

We combining two distinct models. First, the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models are employed to estimate and forecast volatility within each index. Volatility, a critical measure of risk, can greatly impact investment decisions and risk management strategies. By studying volatility, we aim to better understand the inherent risk profiles of these indices.

Second, we apply Vector Autoregressive (VAR) models on the percentage returns of these indices to analyze the interdependencies and causal relationships between them, and also on the volatility of these indices. VAR models, when applied to volatility, provide a deeper understanding of how shocks and fluctuations in volatility in one index can impact the others, shedding light on the dynamic relationships that govern the financial landscape.

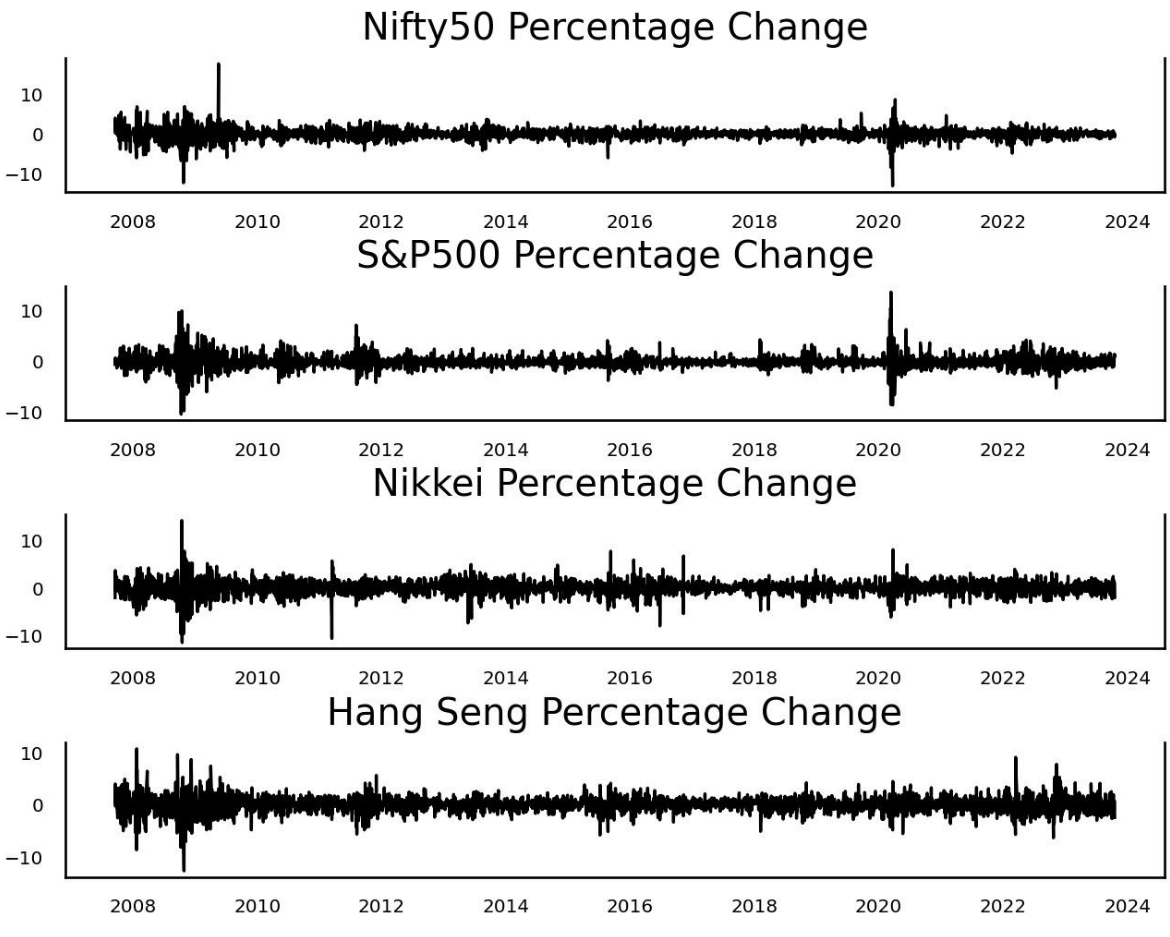
Data Collection and Preprocessing

In the data collection and preprocessing phase, we gathered historical financial data for four major indices (S&P 500, NIFTY 50, Nikkei, and Hang Seng) from yahoo finance in separate CSV files. To prepare the data for analysis, we first calculated the daily percentage change in closing prices for each index. This allowed us to create new columns in each dataset, for the Percentage Change. We then converted the date information into a standardized datetime format, which is crucial for effective time-based analysis and visualization.

Additionally, we applied column renaming to enhance clarity, ensuring consistency in column names across datasets. We simplified the column names to 'sp500_close', 'nsei_close', 'n225_close', and 'hsi_close' for S&P 500, NIFTY 50, Nikkei, and Hang Seng, respectively. To focus on the closing prices, we removed extraneous columns such as 'Open', 'High', and 'Low' from the dataframes. Lastly, to ensure data integrity, we eliminated rows containing missing values (NaN) from all datasets, providing clean and reliable data for subsequent analysis.

Procedure:

1.Plotting the percentage change for all the indices



2.Checking the stationarity of data using ADF Test

Percentage_Change_nifty_data

ADF Statistic: -20.614625996799372

p-value: 0.0

Percentage_Change_sp500_data ADF

Statistic: -13.10415298212619 p-

value: 1.6859705979257847e-24

Percentage_Change_nikkei_data ADF

Statistic: -13.10415298212619 p-

value: 1.6859705979257847e-24

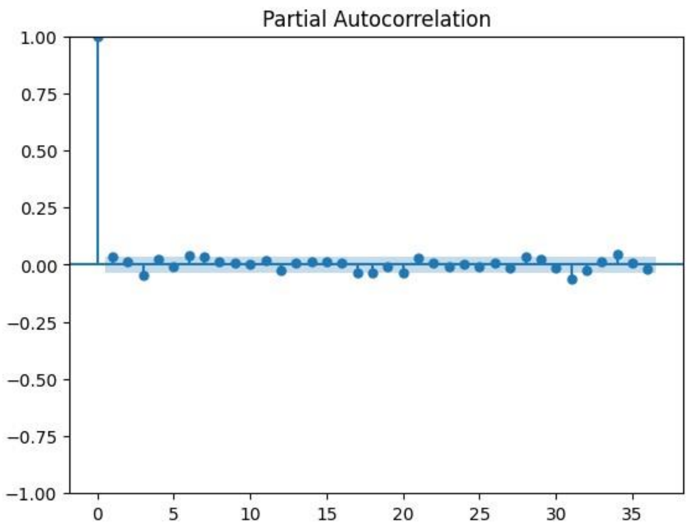
Percentage_Change_hsi_data ADF

Statistic: -13.10415298212619 p-

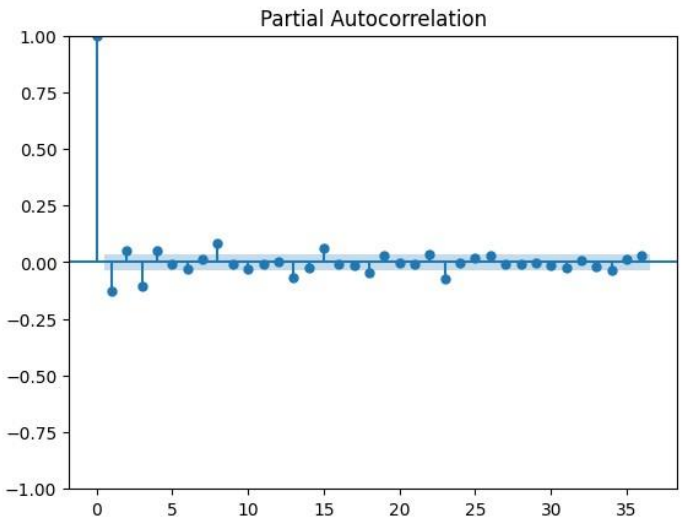
value: 1.6859705979257847e-24

As p-value<0.05, the data taken is stationary, and no further differencing is needed.

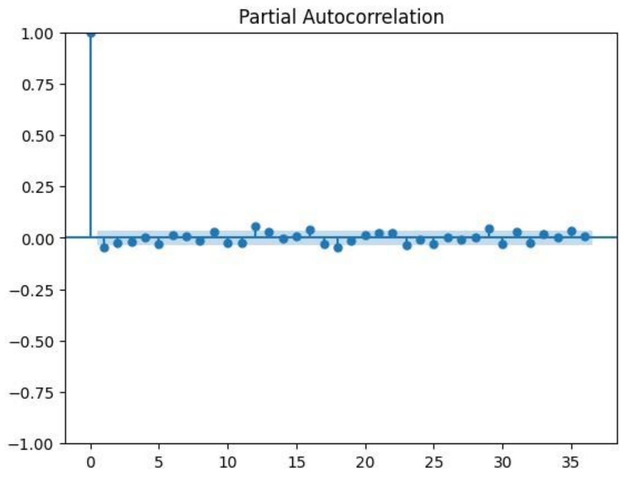
3. Plot the PACF graph for each index to get the orders p and q.

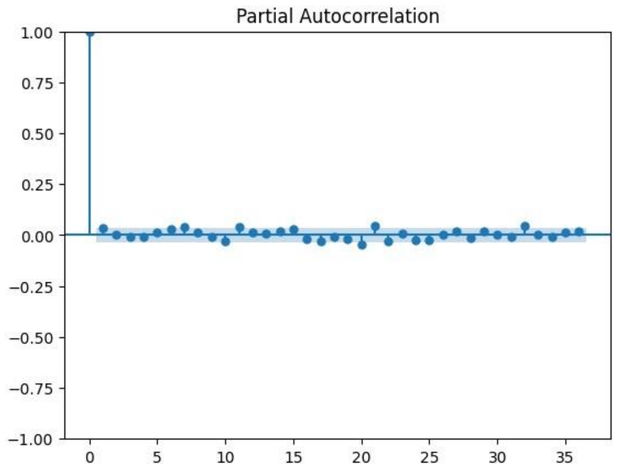


Partial Autocorrelation
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0
5
10
15
20
25
30
35



Partial Autocorrelation
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0
5
10
15
20
25
30
35



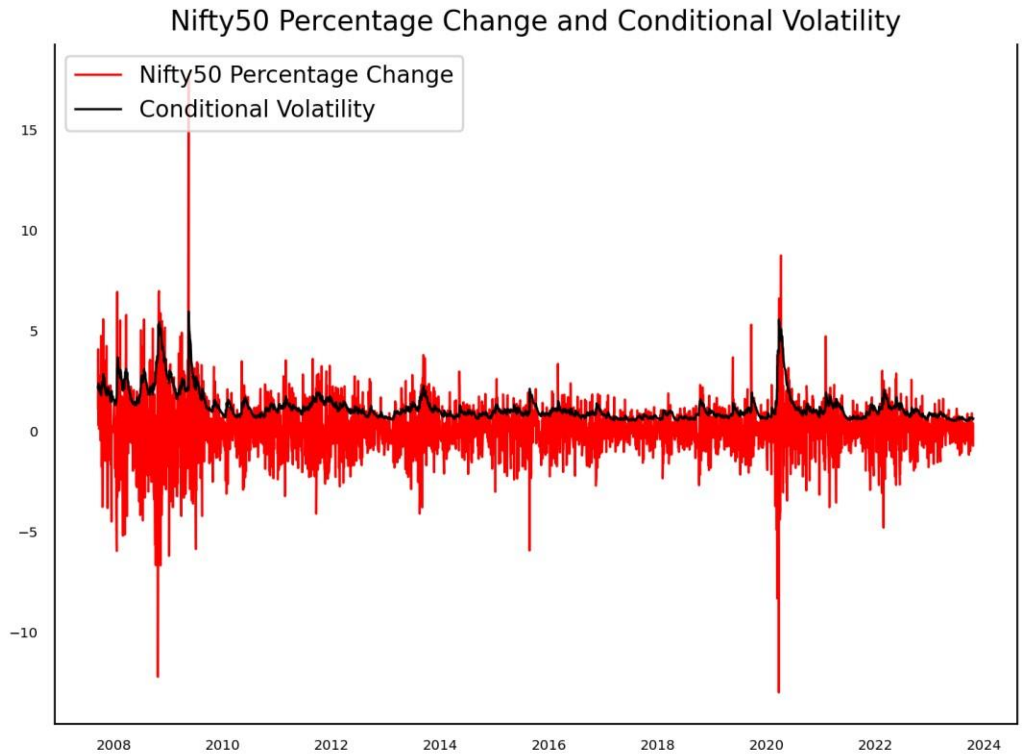


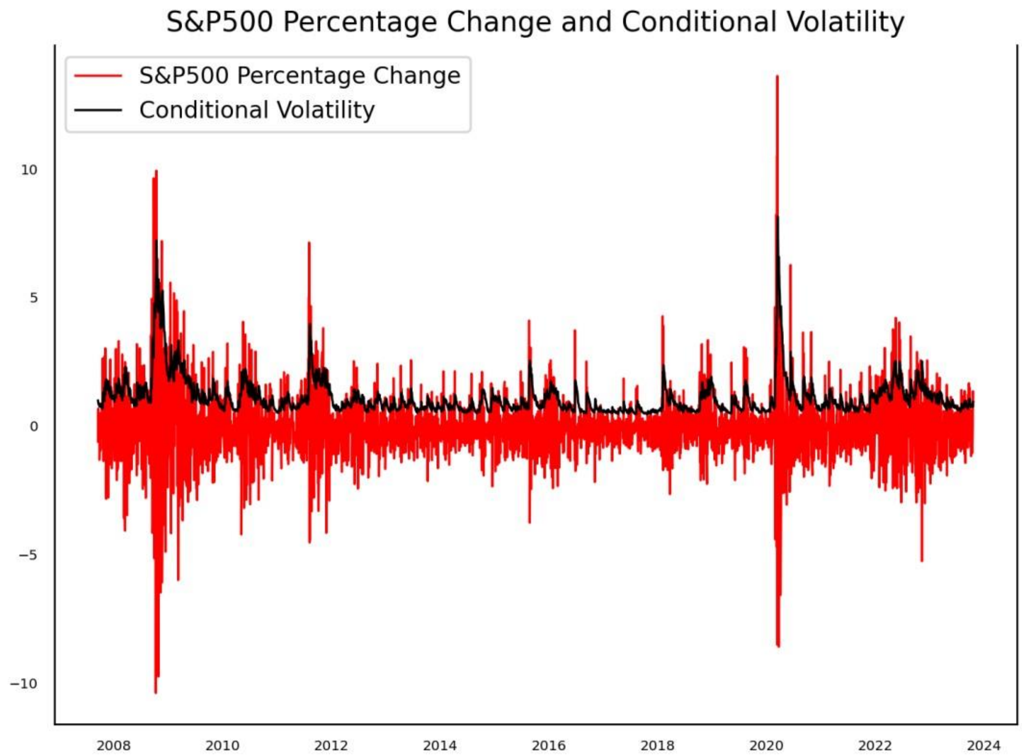
From PACF plot we got the order = (1,1)

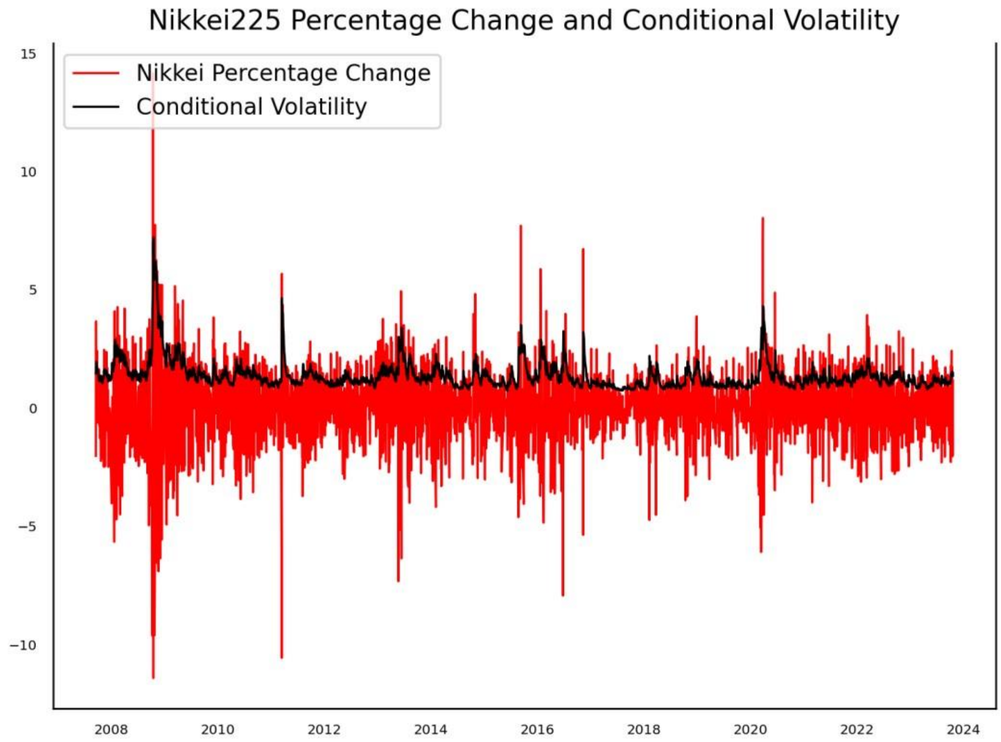
4. Apply the GARCH Model to get volatility of each index

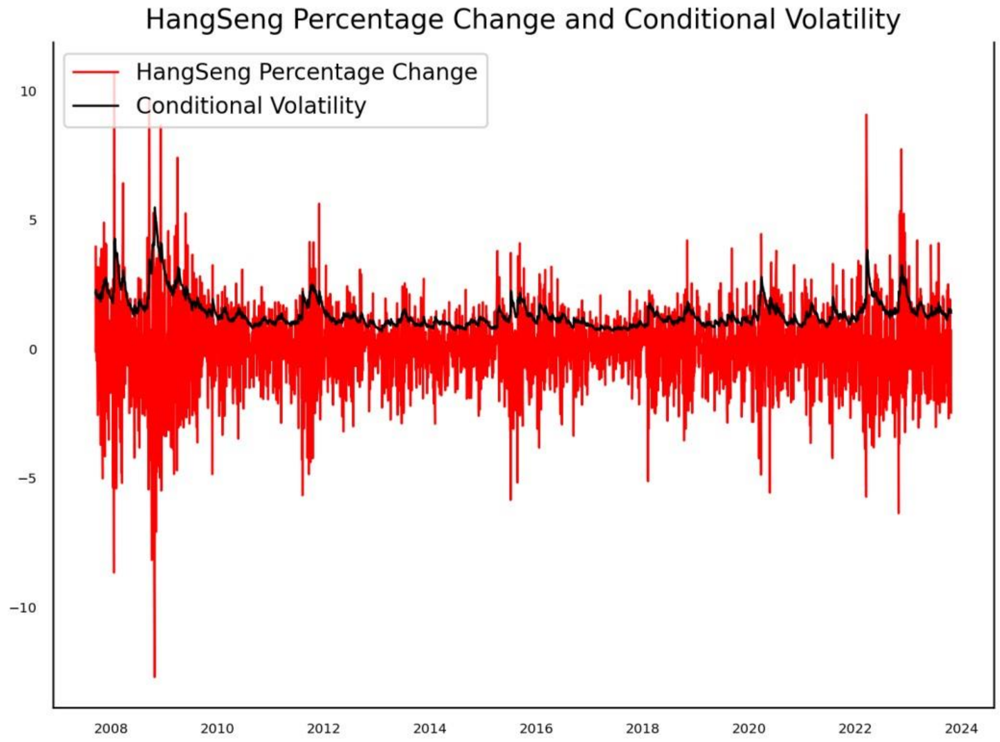
We utilize the GARCH Model to measure the conditional volatility of each index and then plot it against the percentage change calculated for each index.

GARCH(1,1) model was utilized, signifying that the model considered one lag of both returns and past volatility to estimate the current volatility. the GARCH modeling approach

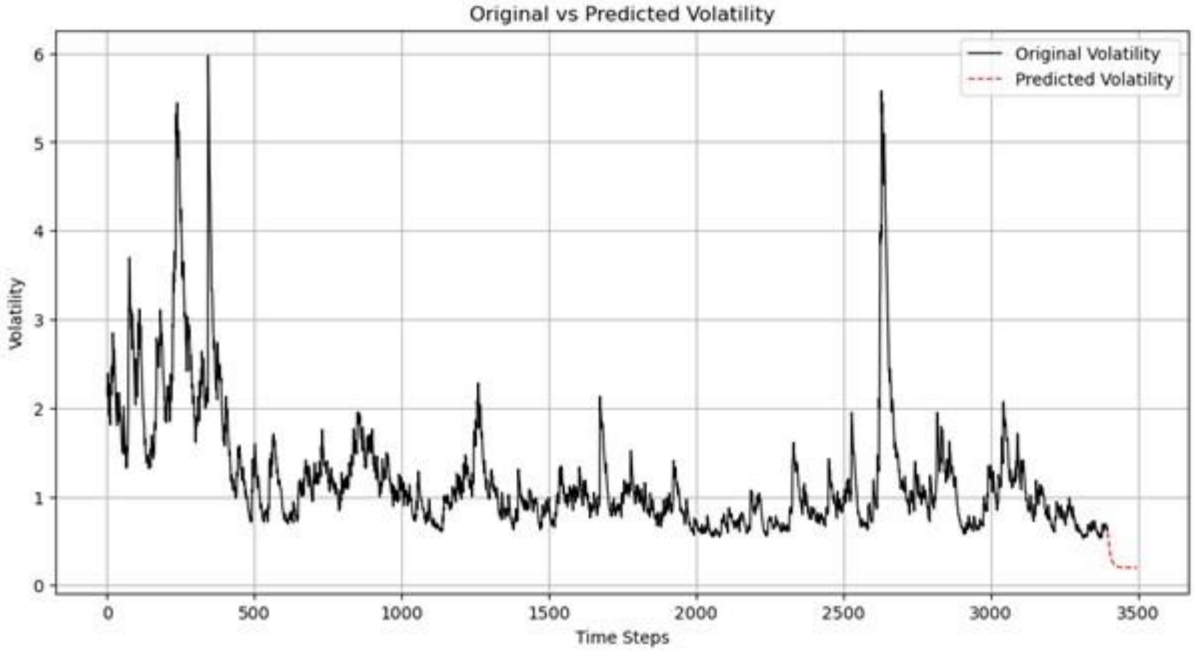

Nifty50 Percentage Change and Conditional Volatility
Nifty50 Percentage Change
Conditional Volatility
15
10
5
0
−5
−10
2008
2010
2012
2014
2016
2018
2020
2022
2024

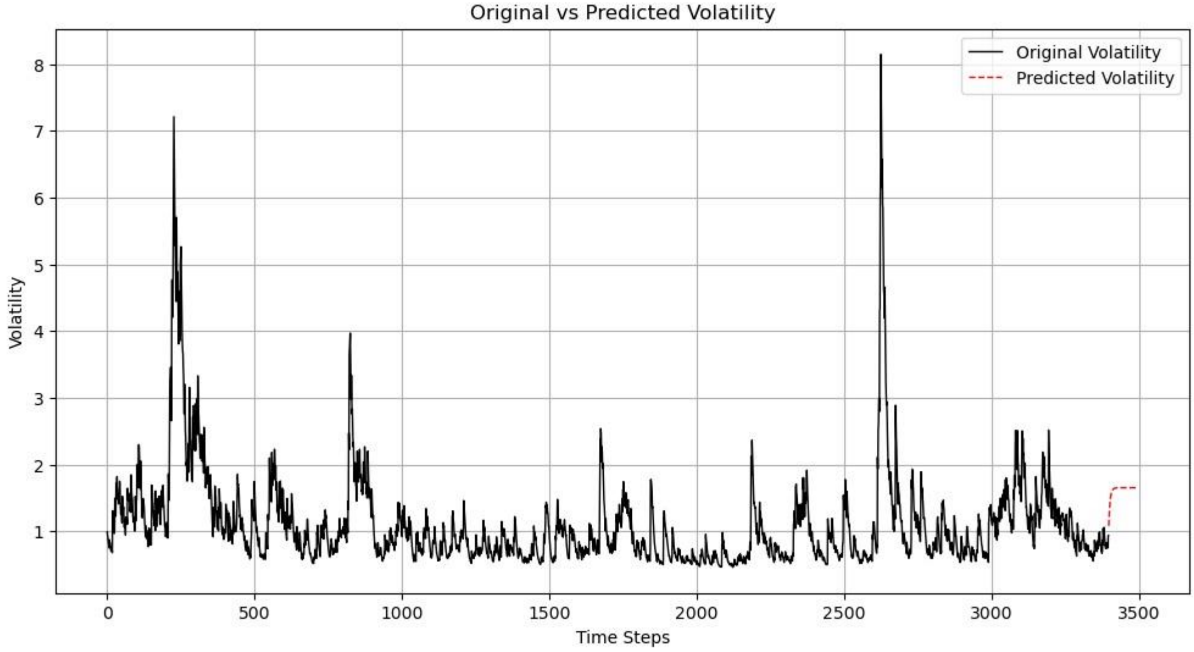

S&P500 Percentage Change and Conditional Volatility
S&P500 Percentage Change
Conditional Volatility
10
5
0
−5
−10
2008
2010
2012
2014
2016
2018
2020
2022
2024

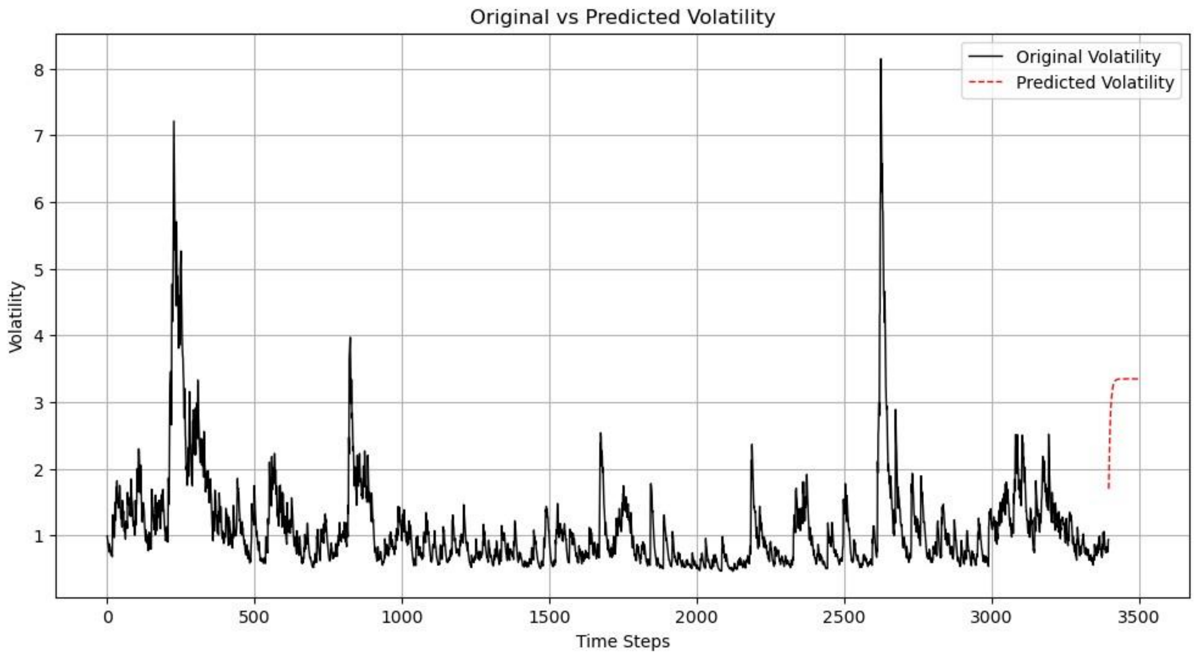

Nikkei225 Percentage Change and Conditional Volatility
Nikkei Percentage Change
Conditional Volatility
15
10
5
0
−5
−10
2008
2010
2012
2014
2016
2018
2020
2022
2024


HangSeng Percentage Change and Conditional Volatility
HangSeng Percentage Change
Conditional Volatility
10
5
0
−5
−10
2008
2010
2012
2014
2016
2018
2020
2022
2024

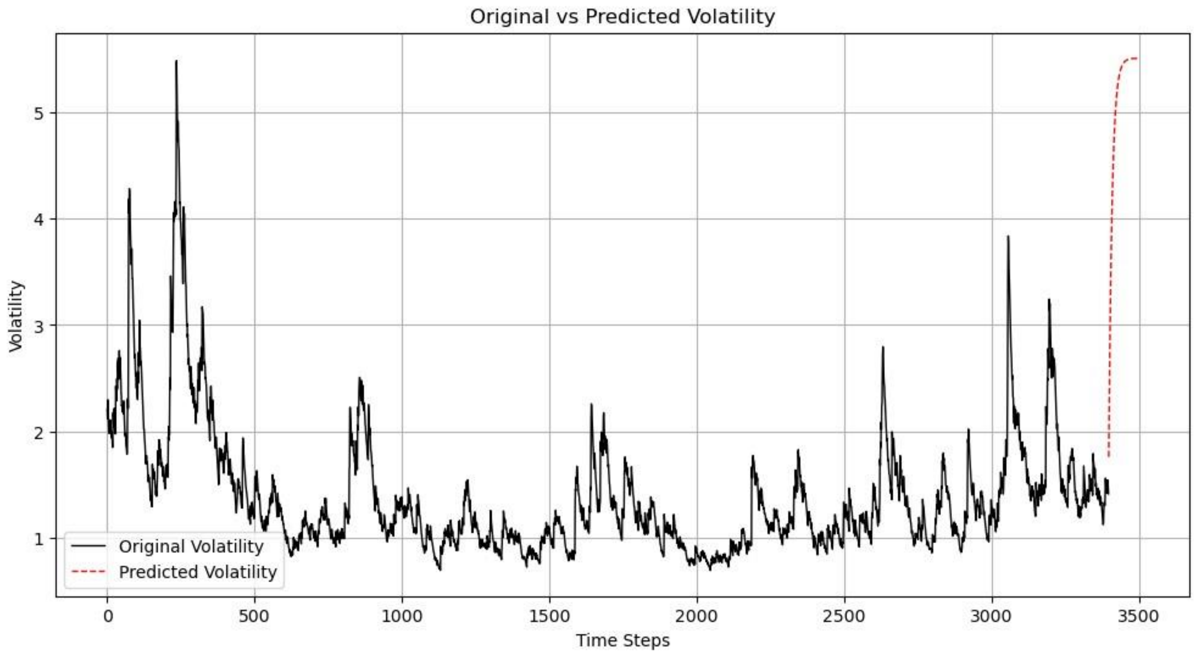
5. Predict the trend line based on the conditional volatility for each index







Original vs Predicted Volatility
Original Volatility
Predicted Volatility
Volatility
Time Steps
0
500
1000
1500
2000
2500
3000
3500
1
2
3
4
5
6
7
8



Original vs Predicted Volatility
Original Volatility
Predicted Volatility
Volatility
Time Steps
0
500
1000
1500
2000
2500
3000
3500
1
2
3
4
5

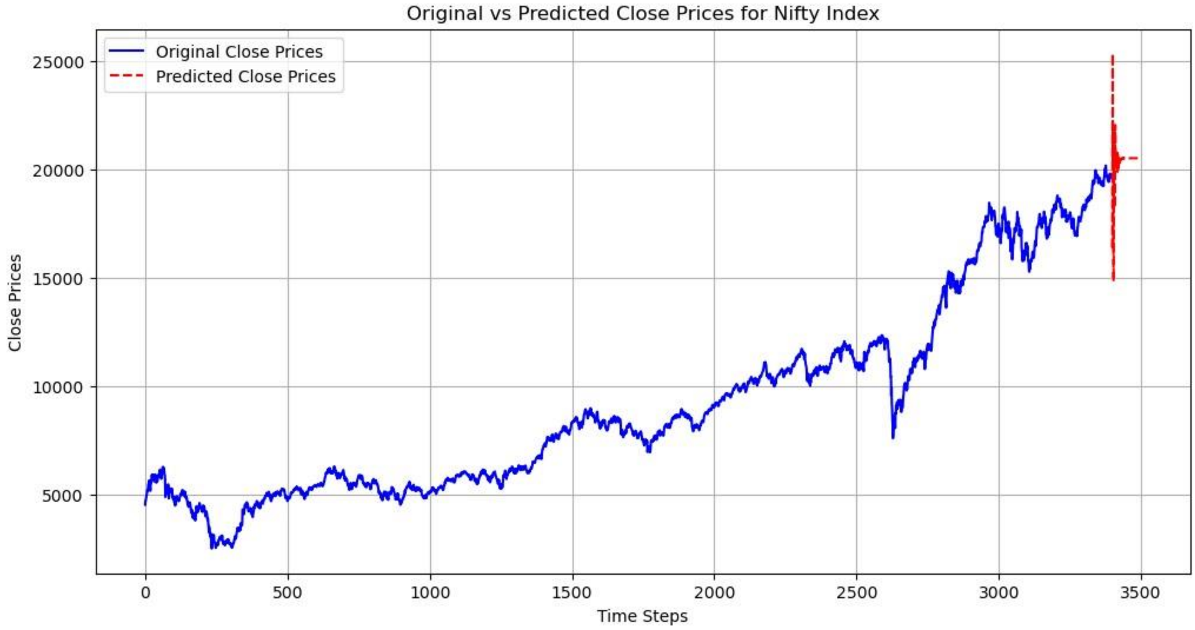
6. Apply different VAR Models and select the best model with least MSE

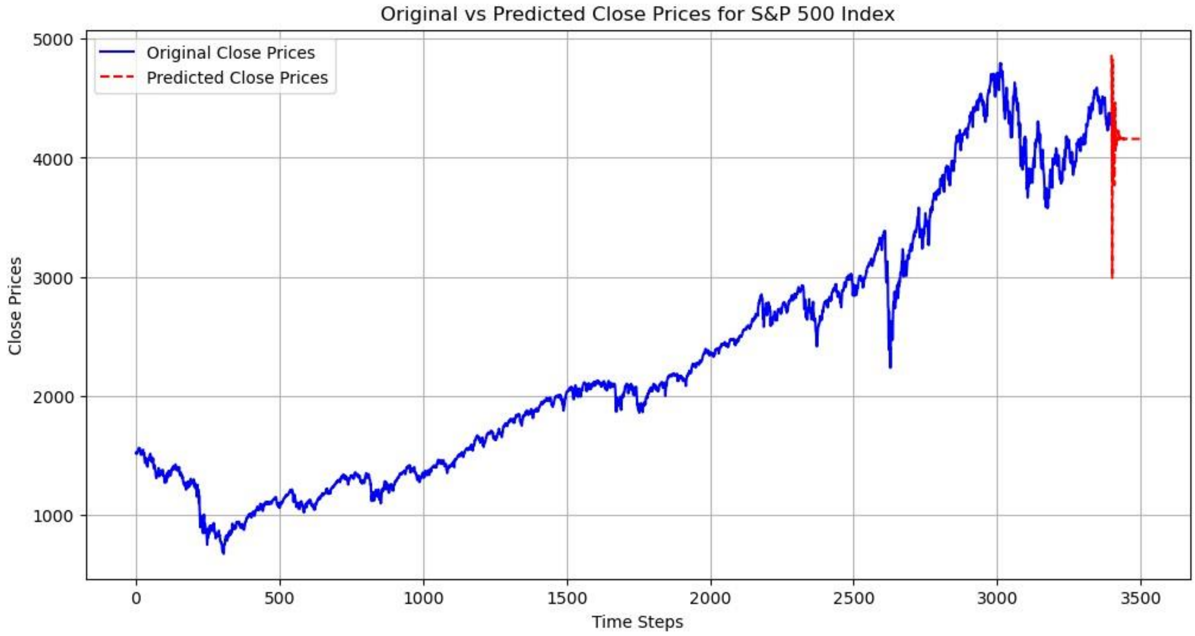
	p	mse
0	14	6.708690
1	13	6.724936
2	11	6.755462
3	10	6.780706
4	9	6.791478
5	8	6.806506
6	7	6.830063
7	6	6.842670
8	5	6.864401
9	4	6.878492
10	3	6.899045
11	2	6.957217
12	1	7.570246
13	12	8.140538

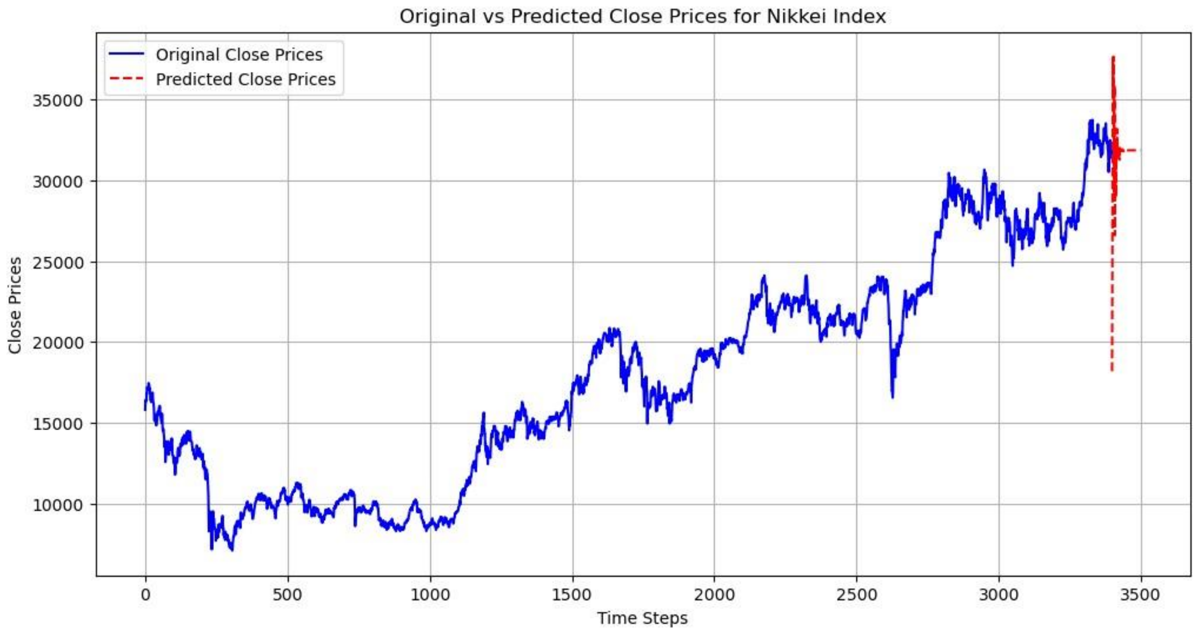
7. Run the best model for time series and fit the var model for time series data

```
                                                                                  Statespace Model Results
===========================================================================================================================================================
=============================
Dep. Variable:     ['Percentage_Change_nifty_data', 'Percentage_Change_sp500_data', 'Percentage_Change_nikkei_data', 'Percentage_Change_hsi_data']   No. Obser
vations:                3397
Model:                                                                                                                                         VAR(14)   Log Likel
ihood              -22006.185
                                                                                                                                           + intercept   AIC
44488.370
Date:                                                                                                                                 Sun, 22 Oct 2023   BIC
45947.464
Time:                                                                                                                                         16:22:37   HQIC
45009.895
Sample:                                                                                                                                              0
                                                                                                                                                - 3397
Covariance Type:                                                                                                                                   opg
===================================================================================================
Ljung-Box (L1) (Q):     0.01, 0.00, 0.00, 0.00   Jarque-Bera (JB):   28014.36, 12610.22, 9441.27, 2014.39
Prob(Q):                0.91, 0.98, 0.98, 0.95   Prob(JB):                     0.00, 0.00, 0.00, 0.00
Heteroskedasticity (H): 0.42, 0.69, 0.51, 1.05   Skew:                        0.64, 0.88, -0.16, 0.08
Prob(H) (two-sided):    0.00, 0.00, 0.00, 0.40   Kurtosis:                  17.01, 12.27, 11.16, 6.77
                         Results for equation Percentage_Change_nifty_data
```

8.Using the VAR Model we predict the trend lines for different indices






Original vs Predicted Close Prices for Nikkei Index
Original Close Prices
Predicted Close Prices
Close Prices
Time Steps
35000
30000
25000
20000
15000
10000
0
500
1000
1500
2000
2500
3000
3500


Original vs Predicted Close Prices for HangSeng Index
Original Close Prices
Predicted Close Prices
Close Prices
Time Steps
30000
25000
20000
15000
10000
0
500
1000
1500
2000
2500
3000
3500

9. Using Granger Causality test to check bidirectional causality for the indices

```
Percentage_Change_sp500_data causes Percentage_Change_nifty_data?

------------------

Granger Causality
number of lags (no zero) 1
ssr based F test:         F=375.2806, p=0.0000  , df_denom=3393, df_num=1
ssr based chi2 test:   chi2=375.6124, p=0.0000  , df=1
likelihood ratio test: chi2=356.2551, p=0.0000  , df=1
parameter F test:         F=375.2806, p=0.0000  , df_denom=3393, df_num=1

Percentage_Change_nifty_data causes Percentage_Change_sp500_data?

------------------

Granger Causality
number of lags (no zero) 1
ssr based F test:         F=0.7115  , p=0.3990  , df_denom=3393, df_num=1
ssr based chi2 test:   chi2=0.7121  , p=0.3987  , df=1
likelihood ratio test: chi2=0.7120  , p=0.3988  , df=1
parameter F test:         F=0.7115  , p=0.3990  , df_denom=3393, df_num=1
```

Change in S&P500 ❼ Change in Nifty50

```
Percentage_Change_sp500_data causes Percentage_Change_nikkei_data?

------------------

Granger Causality
number of lags (no zero) 1
ssr based F test:         F=185.5671, p=0.0000  , df_denom=3393, df_num=1
ssr based chi2 test:   chi2=185.7312, p=0.0000  , df=1
likelihood ratio test: chi2=180.8302, p=0.0000  , df=1
parameter F test:         F=185.5671, p=0.0000  , df_denom=3393, df_num=1

Percentage_Change_nikkei_data causes Percentage_Change_sp500_data?

------------------

Granger Causality
number of lags (no zero) 1
ssr based F test:         F=5.3132  , p=0.0212  , df_denom=3393, df_num=1
ssr based chi2 test:   chi2=5.3179  , p=0.0211  , df=1
likelihood ratio test: chi2=5.3138  , p=0.0212  , df=1
parameter F test:         F=5.3132  , p=0.0212  , df_denom=3393, df_num=1
```

Change in S&P500 ❼← Change in Nikkei 225

```
Percentage_Change_nikkei_data causes Percentage_Change_hsi_data?

------------------

Granger Causality
number of lags (no zero) 1
ssr based F test:         F=0.6618  , p=0.4160  , df_denom=3393, df_num=1
ssr based chi2 test:   chi2=0.6624  , p=0.4157  , df=1
likelihood ratio test: chi2=0.6624  , p=0.4157  , df=1
parameter F test:         F=0.6618  , p=0.4160  , df_denom=3393, df_num=1

Percentage_Change_hsi_data causes Percentage_Change_nikkei_data?

------------------

Granger Causality
number of lags (no zero) 1
ssr based F test:         F=16.1957 , p=0.0001  , df_denom=3393, df_num=1
ssr based chi2 test:   chi2=16.2101 , p=0.0001  , df=1
likelihood ratio test: chi2=16.1715 , p=0.0001  , df=1
parameter F test:         F=16.1957 , p=0.0001  , df_denom=3393, df_num=1
```

Change in hsi ❼ Change in Nikkei 225

```
Percentage_Change_nikkei_data causes Percentage_Change_nifty_data?

------------------

Granger Causality
number of lags (no zero) 1
ssr based F test:         F=4.3517  , p=0.0370  , df_denom=3393, df_num=1
ssr based chi2 test:   chi2=4.3556  , p=0.0369  , df=1
likelihood ratio test: chi2=4.3528  , p=0.0369  , df=1
parameter F test:         F=4.3517  , p=0.0370  , df_denom=3393, df_num=1

Percentage_Change_nifty_data causes Percentage_Change_nikkei_data?

------------------

Granger Causality
number of lags (no zero) 1
ssr based F test:         F=54.7495 , p=0.0000  , df_denom=3393, df_num=1
ssr based chi2 test:   chi2=54.7979 , p=0.0000  , df=1
likelihood ratio test: chi2=54.3605 , p=0.0000  , df=1
parameter F test:         F=54.7495 , p=0.0000  , df_denom=3393, df_num=1
```

Change in NIFTY 50 ❼← Change in Nikkei 225

```
Percentage_Change_hsi_data causes Percentage_Change_nifty_data?

------------------

Granger Causality
number of lags (no zero) 1
ssr based F test:         F=0.0953  , p=0.7575  , df_denom=3392, df_num=1
ssr based chi2 test:   chi2=0.0954  , p=0.7574  , df=1
likelihood ratio test: chi2=0.0954  , p=0.7574  , df=1
parameter F test:         F=0.0953  , p=0.7575  , df_denom=3392, df_num=1

Percentage_Change_nifty_data causes Percentage_Change_hsi_data?

------------------

Granger Causality
number of lags (no zero) 1
ssr based F test:         F=0.8328  , p=0.3615  , df_denom=3392, df_num=1
ssr based chi2 test:   chi2=0.8335  , p=0.3613  , df=1
likelihood ratio test: chi2=0.8334  , p=0.3613  , df=1
parameter F test:         F=0.8328  , p=0.3615  , df_denom=3392, df_num=1
```

Change in NIFTY 50 ❼/← Change in HIS

```
Percentage_Change_sp500_data causes Percentage_Change_nifty_data?

------------------

Granger Causality
number of lags (no zero) 1
ssr based F test:         F=375.2806, p=0.0000  , df_denom=3393, df_num=1
ssr based chi2 test:   chi2=375.6124, p=0.0000  , df=1
likelihood ratio test: chi2=356.2551, p=0.0000  , df=1
parameter F test:         F=375.2806, p=0.0000  , df_denom=3393, df_num=1

Percentage_Change_nifty_data causes Percentage_Change_sp500_data?

------------------

Granger Causality
number of lags (no zero) 1
ssr based F test:         F=0.7115  , p=0.3990  , df_denom=3393, df_num=1
ssr based chi2 test:   chi2=0.7121  , p=0.3987  , df=1
likelihood ratio test: chi2=0.7120  , p=0.3988  , df=1
parameter F test:         F=0.7115  , p=0.3990  , df_denom=3393, df_num=1
```

Change in S&P500 ❼ Change in NIFTY 50