UNIVERSITY OF INFORMATION TECHNOLOGY

FACULTY OF INFORMATION SYSTEMS

Time Series Forecasting using Linear Regression, Non-Linear Regression, Garch, Sarimax and DeepAR

- PHẠM PHƯỚC TỶ-20522136
- VÕ THÀNH TÚ -20522105
- NGUYỄN ĐÌNH ĐAN THY 20522006

Time Series Forecasting using Linear Regression,Non-Linear Regression, Garch, Sarimax and DeepAR

A case study of predicting StockMarket

1st Phạm Phước Tỷ
Undergraduate Student of University of Information Technology – VNUHCM
Ho Chi Minh City, VietNam
20522136@gm.uit.edu.vn

2nd Võ Thành Tú
Undergraduate Student of University of Information Technology - VNUHCM
Ho Chi Minh City, VietNam
20522105@gm.uit.edu.vn

3rd Nguyễn Đình Đan Thy
Undergraduate Student of University of Information Technology - VNUHCM
Ho Chi Minh City, VietNam
20522006@gm.uit.edu.vn

***Abstract*—Vietnam's stock market has developed both in terms of the number of listed shares and trading value, but individual investors' knowledge of the influence of factors on investors' investment decisions Private investment is still limited leading to the risk of loss in investment activities is very large. This study focuses on the figure of stockmarket prices from January 2015 to April 2020 according to the Vietstock Economic and aim to forecast and analyze daily Stock price through the establishment of Linear Regression,Non-Linear Regression, Garch, Sarimax and DeepAR.**

Index Terms – Garch model, data forecasting, DeepAr on Stockmarket, Time Series.

I. INTRODUCTION

Vietnam's stock market has developed both in terms of the number of listed shares and trading value, but individual investors' knowledge of the influence of factors on investors' investment decisions Private investment is still limited leading to the risk of loss in investment activities is very large. The stock price is an important indicator of a company, and there are many factors that affect its value Different events can have different effects on the mood and mood of the public, which in turn can affect the development of stock market prices. Due to the dependence on various factors, stock prices are not static but dynamic, noisy and non-linear time series data. Due to its great learning capability for solving the investors' decision-making considerations. Besides the group of factors belonging to investors, the group of factors belonging to the macro environment, factors belonging to the potential of investment securities have an impact on the investment decisions of individual investors.[1] The stock market plays an important role in the allocation of resources and directly as a source of financing and as a determinant of the value and creditworthiness of companies . However, a growing body of empirical evidence has raised some doubts about whether equity markets are efficient in the sense of appropriately reflecting relevant and avaiiable information.' The large swings in equity prices in several countries during the 1980s provided additional evidence that market valuations were more variable than the earnings prospects of firms. These events have encouraged reform proposals* to limit volatility, as excessive volatility or mispricing can have adverse real-world consequences and lead to misallocation of resources.

II. RELATED WORD

- Box–Jenkins[1] used Time series analysis for forecasting and control. White [2] used Neural Networks for stock market forecasting of IBM daily stock returns. Following this, a range of studies reported on the efficacy of different learning algorithms and forecasting methods using ANN. Henry [3] used ARIMA model, to predict the daily close and morning open price, But all these forecasting methods have problems when there is nonlinearity in the time series
- Engle (1982), [4] studied on ARCH and Bollerslev (1986), [5] on GARCH models, and revealed that, these models were designed to deal with the assumption of non-stationarity found in real life financial data. He further pointed out that these models have become widespread tools for dealing with time series heteroscedasticity. ARCH and GARCH models treat heteroscedasticity as a modeled variance.
- Application of SARIMAX Model to Forecast Daily Sales in Food Retail Industry Nari Sivanandam Arunraj (*) is a scientific associate at Research and Technology Campus for Supply Chain Management and Data Analytics in Grafenau, Deggendorf

Institute of Technology, Germany. He received his bachelor of technology in chemical engineering from University of Madras, India, master of engineering in industrial engineering from Anna University, India, and PhD in industrial engineering and Management of Indian Institute of Technology Kharagpur,India. His research interests include predictive analytics, simulation and artificial intelligence.

- Michael Fernandes is a research assistant at the Supply Chain Management and Data Analytics Research and Technology Campus at the Deggendorf Institute of Technology in Grafenau, Germany. He obtained his physics degree at the University of Muenster, Germany. His research interests are predictive analytics, simulation, and artificial intelligence. He was a research associate at the Max Planck Institute for Atmospheric Research in Katlenburg-Lindau, Germany. He has two years of professional experience in calibration technology.. He also has more than five years of experience in the field of software design and development.
- Bengio et al. (1994) investigate methods such as simulated annealing, multi-grid random search, time-weighted pseudo-Newton optimization, and discrete error propagation. Their \latch" and \2-sequence" problems are very similar to problem 3a with minimal time lag 100 (see Experiment 3). Bengio and Frasconi (1994) also propose an EM approach for propagating targets. With n so-called \state networks", at a given time, their system can be in one of only n di
erent states. See also beginning of Section 5. But to solve continuous problems such as the \adding problem" (Section 5.4), their system would require an unacceptable number of states (i.e., state networks).
- Schmidhuber's hierarchical chunker systems (1992, 1993) do have a capability to bridge arbitrary time lags, but only if there is local predictability across the subsequences causing the time lags (see also Mozer 1992). For instance, in his postdoctoral thesis (1993), Schmidhuber uses hierarchical recurrent nets to rapidly solve certain grammar learning tasks involving minimal time lags in excess of 1000 steps. The performance of chunker systems, however, deteriorates as the noise level increases and the input sequences become less compressible. LSTM does not from this problem.

III. DATA AND METHODOLOGY

A. Data Sources

This study predicts the Price of Vinacapital Stock value, according to the Vietstock and Yahoo-Finance of the month of January 2022. These database was collected by Yahoo-finance of VietnamStock, Consisting of daily Stock price from January 1st 2017 to April 1st 2022.

Fig. 1. Daily VinaCapital's Stock price from 2017 to 2022

B. DeepAR

DeepAR is a supervised learning algorithm for time series forecasting that uses recurrent neural networks (RNN) to produce both point and probabilistic forecasts.[6] The DeepAR forecasting algorithm can provide better forecast accuracies compared to classical forecasting techniques such as Autoregressive Integrated Moving Average (ARIMA) or Exponential Smoothing (ES), both of which are implemented in many open-source and commercial software packages for forecasting. The DeepAR algorithm also supports other features and scenarios that make it particularly suitable for real-world applications.

DeepAR supports two data channels. The required train channel describes the training data set .The optional test channel describes the dataset that the algorithm uses to evaluate the accuracy of the model after training.

Method of DeepAR:

- Start—A string with the format YYYY-MM-DD HH:MM:SS. The initial timestamp cannot contain timezone information.
- Target - An array of floating point values or integers representing a time series. You can encode missing values as null literals or "NaN" strings in JSON or as nan float values in Parquet.
- dynamic_feat (optional) - An array of floats or integers representing the time series vector of the custom function (dynamic features). If you specify this field, all records must have the same number of internal tables (the same number of feature time series) In addition, each internal table must be as long as the associated target value plus forecast_length. Functions do not support missing values.

The DeepAR algorithm evaluates the trained model with different metrics. The algorithm calculates the root mean square error (RMSE) from the experimental data as follows.

$$RMSE = \sqrt{\frac{1}{nT}\sum_{i,t}\left((y_{i,t})^2 - y_{i,t}\right)\text{^}2}$$

Autoregressive Integrated Moving Average Model.ARIMA (p,d, and q), where pis the autoregressive term, qisthe number of moving average terms, and dis the number ofdifferences made when the time series becomes stationary. The prediction results can be adjusted by adjusting theaforementioned three parameters d,p, and q, so as to drawthe optimal model. The model calculation formula is as follows.

$$y_t = \theta_0 + \emptyset_1 y_{t-2} + \cdots + \emptyset_p y_{t-p} + \varepsilon \in_t - \theta_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

(1)

where y_t and ε_t are the actual value and random error of the time period t, respectively; Φi(i=1, 2,. . .,p) and θj(j=1,2,. . .,q) are the model parameters; p and q, the order of the model (p and q are integers), are also the model parameter mentioned earlier; the random error εt, whose mean value is0, is assumed to be independent and obey the same dis-tribution in the model. The variance of constant term is denoted as σ2. Equation (1) contains several important special cases of ARIMA series models.

If q=0, equation (1) can be simplified to an AR model p.

If p=0, the model can be simplified to an MA model of order q.

Among them, the model order (p,q) is a key link in ARIMA model construction, which determines the forecast accuracy of the model. The parameters of the operations AR and MA are defined by (p) and (q), respectively.

- Apply DeepAR on VinaCapital's Stock price
- First, we apply the data of VNC's Stock price

	Date	Open	High	Low	Close	Adj Close	Volume
0	2017-01-03	272.0	275.000000	272.000000	273.25	233.939148	170046
1	2017-01-04	272.0	275.000000	272.000000	272.75	233.511093	107823
2	2017-01-05	272.0	274.500000	270.250000	272.50	233.297043	391623
3	2017-01-06	272.5	273.750000	266.000000	273.00	233.725098	464329
4	2017-01-09	273.0	273.500000	270.000000	273.50	234.153183	276108
...	...	...	...	...	...	...	...
1323	2022-03-25	502.0	507.000000	502.000000	507.00	493.078217	265108
1324	2022-03-28	502.0	511.000000	502.000000	506.00	492.105652	115303
1325	2022-03-29	502.0	520.000000	500.000000	518.00	503.776154	149574
1326	2022-03-30	516.0	516.989990	507.666992	513.00	498.913422	65890
1327	2022-03-31	514.0	516.588989	506.000000	506.00	492.105652	68055

Fig 2. Data of VNC's Stock price

The Google training data has information from 1 Jan 2017 to 1 April 2022. There are 6 columns. The open column shows the price at which the stock opened when the market opened on a particular day. The Closing column refers to the price of an individual stock when the stock exchange closed the market for the day. The high column describes the highest price at which the stock traded during the period. The lowest column shows the lowest price for the period. Volume is the total amount of trading that took place over a certain period of time.

- Using LTSM for the prediction Stock Price

After the data are processed, in the hybrid model prediction experiment, [7] the ARIMA method is first employed in this article to process the S&P 500 index component stocks in the aspect of linear as the first step, and then the nonlinear part of the data residual value processed at the first step is used as the input data of the LSTM model. Finally, model establishment, data raining and testing is developed. The final prediction results of the correlation coefficient.

Predicting the results of the hybrid model alone is not enough to show that certain advantages of the model in the predictive performance of the research subjects, such as correlation coefficients [8] In order to make comparison between the proposed hybrid model proposed and other models for the accuracy of financial sequence forecasting, other commonly used fore-casting models are introduced as the reference group. Many studies have shown that the predictive power of the full sequential model is low. period In terms of forecasting economic sequences, three more commonly used forecasting models are also discussed, which are compared with the forecasting results of hybrid models..

Fig 3. Prediction of Data Stock Price of VNC

Comments: As you can see above, the model can predict the trend of the actual stock prices very closely. The accuracy of the model can be enhanced by training with more data and increasing the LSTM layers.

C. Linear Reregssion

-Linear regression is a supervised learning algorithm in which the prediction output is continuous and has a constant rate.

-It is used to predict values in a continuous range instead of classifying them into categories or groups.

-There are 2 main types:

+ Simple regression

+ Multivariate regression

-The goal of the linear regression model:

+ Find a model to describe an association between X and Y (X could be age, weight and Y could be BMD).

+ Adjust noise factors.

+ Prognosis: give value X how can predict Y value (Prognosis is predictive knowledge about some event or event or estimate its possible development in the future).[9]

Example: Given 2 points in 2D space

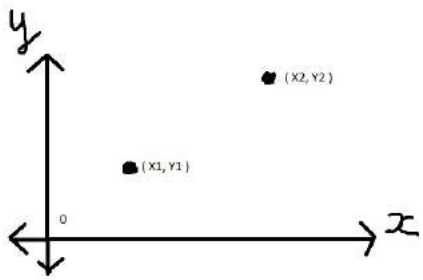

Fig 4. Example diagram of Linear Regression

Simple linear regression model is a regression model that estimates the relationship between one independent variable and one dependent variable using a straight line.

Y - Response variable, dependent variable,…

+ Y is a continuous variable

X - Predictor variable, independent variable,…

+X is a continuous or discontinuous variable

-Equation:

$$Y = f(X, B) + \in$$

α: intercept

β: slope/gradient

ϵ: random error

-There are 4 assumptions about the model

+ The relationship between X and Y is linear in terms of parameters

+X has no random error

+The values of Y are independent of each other

+ Random error: follows a normal distribution, has a mean of 0 and a constant variance [10]

ε~N(0,σ^2)

- Apply the Linear Regression for Data of VNC Stock Price

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                 Volume   R-squared:                       0.435
Model:                            OLS   Adj. R-squared:                  0.430
Method:                 Least Squares   F-statistic:                     95.84
Date:                Fri, 09 Dec 2022   Prob (F-statistic):           1.36e-31
Time:                        11:02:21   Log-Likelihood:                -4370.7
No. Observations:                 252   AIC:                             8747.
Df Residuals:                     249   BIC:                             8758.
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const       1.331e+07   4.87e+06      2.732      0.007    3.71e+06    2.29e+07
High        2.352e+06    1.7e+05     13.820      0.000    2.02e+06    2.69e+06
Low        -2.349e+06   1.72e+05    -13.665      0.000   -2.69e+06   -2.01e+06
==============================================================================
Omnibus:                      122.010   Durbin-Watson:                   1.104
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              714.808
Skew:                           1.871   Prob(JB):                    6.04e-156
Kurtosis:                      10.354   Cond. No.                     3.67e+03
==============================================================================
```

Fig 6. Result of Data after analyzed

Comments: Model 1 with all features fitted with R^2 value 0.435. This indicates open, high, low, volume and adj close are essential for predicting closing value accurately.This indicates prediction of close value is not affected with adj close. This study reveals with open, high, low and volume itself enough for finding approximate prediction of close value.

D. Non-Linear Regression

Nonlinear Regression is a form of regression analysis in which the data fits to a model [5] and that was expressed as a a graph represented by a line or curve. The Nonlinear Regression relates two variables (X and Y) in a nonlinear (curved) pattern.[11]

Equation:

$$Y = f(X, B) + \in$$

Where:

X is a vector of P predictors.

β is a vector of k parameters.

ϵ is the known regression function.

The term "nonlinear" refers to the parameters of the model, not the independent variables. There are unlimited possibilities to describe the deterministic part of the model. What a great flexibility provides a good ground on which to make statistical inferences.

The main difference of linear and nonlinear regression models lies in the calculating the leaset squares. Basically, the purpose of the nonlinear models is to minimize of the sum of the squares as least as possible using iterative numeric procedures.

The least squares estimator of θ, denoted by ˆθ, is the point in the parameter space such that f(ˆθ) is closest to yin the sample space among all feasible f(θ) in the solution locus. The least squares estimate is derived from minimizing the residual sum of squares.[12]

$$S(\theta) = \sum_{i=1}^{n} \left\{ y_i - f\left(x_i; \theta\right) \right\}^2, \theta \in R^P$$

- Apply the Non-Linear Regression for Data of VNC Stock Price

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                 Volume   R-squared:                       0.529
Model:                            OLS   Adj. R-squared:                  0.519
Method:                 Least Squares   F-statistic:                     55.23
Date:                Fri, 09 Dec 2022   Prob (F-statistic):           2.53e-38
Time:                        11:08:33   Log-Likelihood:                -4347.8
No. Observations:                 252   AIC:                             8708.
Df Residuals:                     246   BIC:                             8729.
Df Model:                           5
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const       7.488e+07   9.95e+06      7.527      0.000    5.53e+07    9.45e+07
Open       -1.248e+05   2.61e+05     -0.478      0.633   -6.39e+05    3.89e+05
High        1.817e+06   3.12e+05      5.827      0.000     1.2e+06    2.43e+06
Low        -1.767e+06    2.6e+05     -6.787      0.000   -2.28e+06   -1.25e+06
Close        1.02e+07    1.5e+06      6.814      0.000    7.25e+06    1.32e+07
Adj Close   -1.04e+07   1.49e+06     -6.978      0.000   -1.33e+07   -7.46e+06
==============================================================================
Omnibus:                      137.642   Durbin-Watson:                   1.076
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1052.932
Skew:                           2.058   Prob(JB):                    2.28e-229
Kurtosis:                      12.129   Cond. No.                     1.31e+04
==============================================================================
```

Fig 7. Result of Data after analyzed

Comments: To get accurate results from a non-linear regression model, you should ensure that the functions specified accurately describe the relationship between the independent and dependent variables. The starting value is always the required value. Non-standard initial values can lead to a model that does not converge or a solution that is only locally optimal but not completely optimal.

E. GARCH

GARCH (Generalized AutoRegressive Conditional Heteroskedasticity): is a statistical model used in analyzing time-series data where the variance error is believed to be serially autocorrelated. GARCH models assume that the variance of the error term follows an autoregressive moving average process

- What is volatility
  In finance, volatility is a statistical measure of the dispersion of asset returns over time

It is often computed as the standard deviation or variance of price returns.

In general, the higher the volatility, the riskier a financial asset

We can compute the volatility as the standard deviation of price returns following three easy steps:
Step 1: Calculate returns as percetage of price changes

$$return = \frac{P1 - P0}{P0}$$

Step 2: Calculate the sample mean return of a chosen n-period

$$mean = \frac{\sum_{i=1}^{n} return_i}{n}$$

Step 3: Calculate the sample standard deviation

$$volatility = \sqrt{\frac{\sum_{i=1}^{n} (return_i - mean)^2}{n-1}} = \sqrt{variance}$$

- Compute volatility in Python

Use pandas pct_change() method: (To compute percentage as returns, apply"pct_change()" method from the "pandas"

return_data = price_data.pct_change()

Use pandas std() method: to the return data to compute the standard deviation

volatility = return_data.std()

Assume we measure volatility as the standard deviation of returns, then monthly volatility can be obtained.
Convert to monthly volatility from daily: (assume 21 trading days in a month)

$\sigma_{monthly} = \sqrt{21} * \sigma_d$ by multiplying daily volatility by the square root of 21, which is the average number or trading days in a month.
Conver to annual volatility from daily: (assume 252 trading days in a year)

The challenge of volatility modeling:

- **Heteroskedasticity:**

In ancient Greek:"different" (hetero) + "dispersion" (skedasis)

A time series demonstrates varying volatility systematically over time

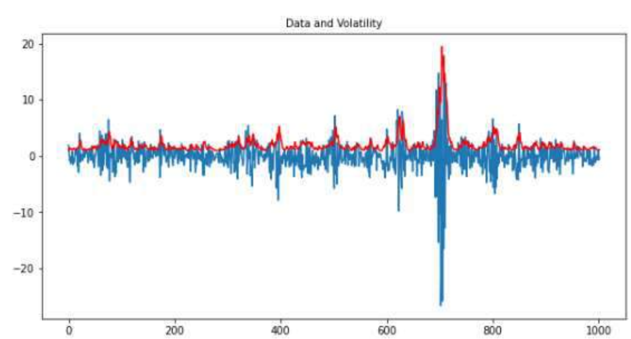

Fig 9. Volatility Prediction of Data VN StockPriceVNC Stock Price

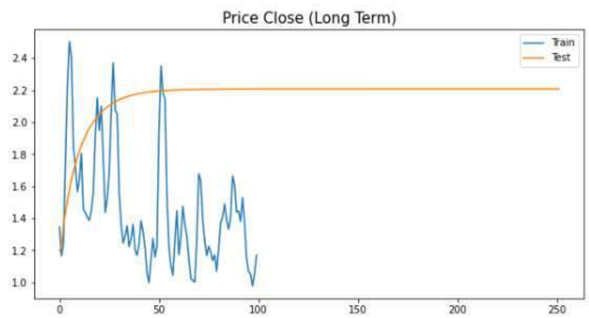

Fig 10. Long Term Volatility Prediction of Data VNC Stock Price

Volatility Prediction of Data VNC Stock Price

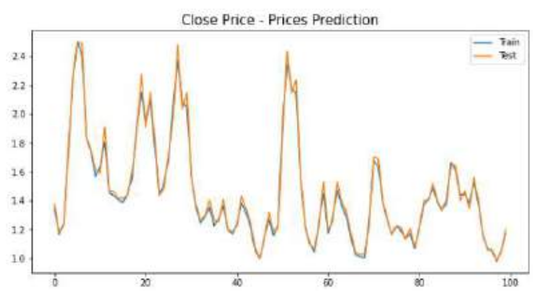

Fig 11. Volatility Prediction – Rolling Forecast of Data VNC Stock Price act

Comments: As you can see above, the model can predict the trend of the actual stock prices very closely. The Garch have showed the true volatility prediction of dataset, we can conclude that the algorithm has worked very well.

F. SARIMAX

(S)ARIMA(X): Seasonal Autoregressive Integrated Moving Average with Exogenous Factors

Autoregressive: A model that uses dependencies between observations and observations with multiple lags.

Integrated: Stationary time series using differencing raw observations (e.g., subtracting one observation from observations at a previous time step).

Moving average: A model that uses the dependence between observations and the residuals of a moving average model applied to lagged observations.

It includes the ARIMA model and its variants SARIMA and **SARIMAX**: Statistical Models for Forecasting.ARIMA stands for Autoregressive Integrated Moving Average.

It is a combination of two models: AR (autoregressive) model using lagged values of the time series for forecasting and MA (moving average) model using lagged values Predicted residual values.In other words, the model uses the dependencies between data values and historical error values to optimize forecasts.

ARIMA uses three parameters - ARIMA(p,d,q)

p: The number of lagged observations in the model; also called lag order. d: How often the original observations differ; also known as the degree of difference.

q: The size of the moving average window; also known as the moving average order.

The ARIMA model is extended to use exogenous inputs and produces an ARIMAX (autoregressive integrated moving average with exogenous input model) model. In this model, time series are modeled using other independent variables as well as the time series itself. The equation of the ARIMAX looks like.

$$\Delta P_t = c + \beta X + \phi_1 \Delta P_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t$$

The SARIMA model represents the Seasonal Autoregressive Integrated Moving Average. It contains additional parameters related to the seasonal part.

In fact, we can think of a SARIMA model as a combination of two ARIMA models: one dealing with the non-seasonal component and the other dealing with the seasonal component.

So the SARIMA(p,d,q)(P,D,Q,S) model has all the parameters mentioned above (non-seasonal parameters) and P,D,Q,S which are seasonal parameters,...not

- Seasonal order:

p: autoregressive order.

d: Differentiation order.

q: moving average order.

Seasonal order

P: Seasonal autoregressive order.

D: Order of seasonal differences.

Q: Seasonal moving average order S: Seasonal period length.ply the SARIMAX Algorithm for Data of VNC Stock Price

Based on Sarimax, we have charted open price, high price, low price, close price of the stock in recent years.

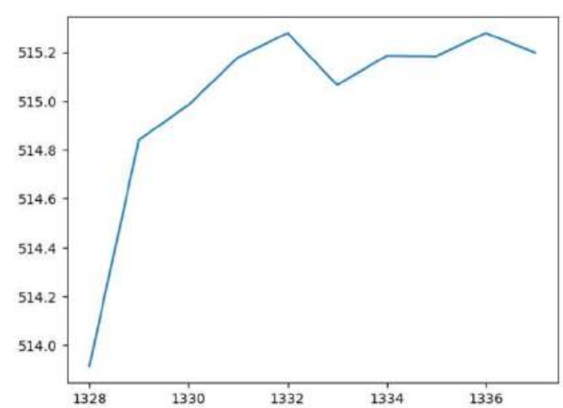

Fig 12. Open Price Chart Prediction of Data VNC Stock Price

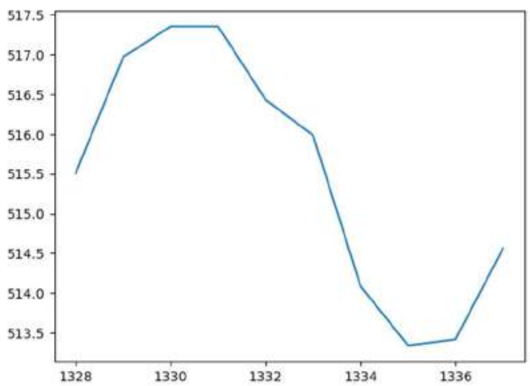

Fig 13. Open Price Chart Prediction of Data VNC Stock Price

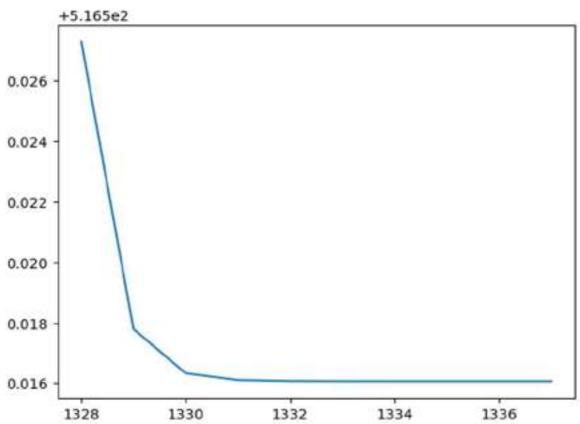

Fig 14. High Price Chart Prediction of Data VNC Stock Price

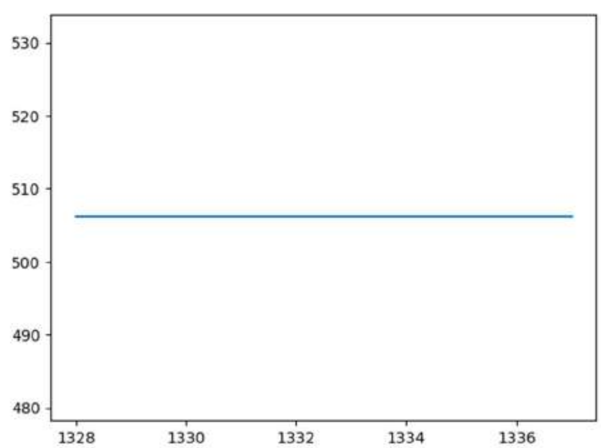

Fig 15. Low Price Chart Prediction of Data VNC Stock Price

Fig 16. Close Price Chart Prediction of Data VNC Price

Comments: As we can see above, the pattern has been shown to predict the outcome of the closing price, opening, high and low, the correlation between the charts clearly

Train/Test Model

MAPE

The average percentage error is known as MAPE (Mean Absolute Percentage Error). A statistical indicator of a forecasting system's accuracy is the mean absolute percentage error (MAPE). It can be calculated as the average absolute percentage inaccuracy for each time period less the true values divided by the true values. It expresses this accuracy as a percentage.

$$MAPE = \frac{100\%}{N} \sum_{T=1}^{N} \left| \frac{A_{t-F_t}}{A_t} \right|$$

RMSE

RMSE is a commonly used measure of the difference between the values (sample or population values) predicted by a model or an estimator and the observed values..

$$RMSE = \sqrt{(f-o)\hat{}2}$$

- **Result of RMSE AND MAPE**

Model	Train-test	RMSE	MAPE
SARIMAX	7,3	139.6506657	27.047614
	8,2	40.57683013	6.9945017
	9,1	46.7259246	8.405408
LR	7,3	45.77245708	9.8309474
	8,2	40.61952407	8.6326988
	9,1	41.20263642	8.7777659
NLR	7,3	0.299649135	33.317255
	8,2	0.117059369	10.216799
	9,1	0.081480421	7.6709892
DEEPAR	7,3	450.1158122	99.838725
	8,2	479.27049	99.82499
	9,1	502.6273704	99.827834

According to the result above:
Sarimax Method
We can see that the train/test (8-2) give the lowest MAPE.
Linear – Regression
We can see that the train/test (8-2) give the lowest MAPE.
NonLinear-Regression
We can see that the train/test (9-1) give the lowest MAPE.
DeepAR-Method
We can see that the train/test (9-1) give the lowest MAPE.

Conclusion: According to the result table, we can conlude that the **Sarimax method** have a train/test (9-1) give the lowest MAPE. Because of that we decide to use this algorithm method to predict the price of Stock Values for next 30days.

DeepAR

Train/Test (70% - 30%)

Fig 17. DeepAR-LSTM Tranning Test Model

Train/Test (80% - 20%)

Fig 18. DeepAR-LSTM Tranning Test Model

Train/Test (90% - 10%)

Fig 19. DeepAR-LSTM Tranning Test Model

Linear-Regression

Train/Test (70% - 30%)

Fig 20. Linear-Regression Tranning Test Model

Train/Test (80% - 20%)

Fig 21. Linear-Regression Tranning Test Model

Train/Test (90% - 10%)

Fig 22. Linear-Regression Tranning Test Model

Non-Linear Regression

Train/Test (70% - 30%)

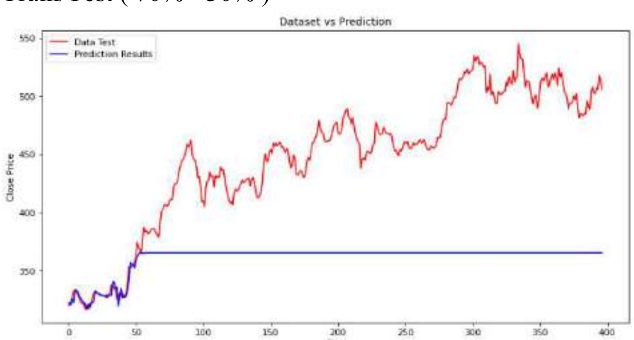

Fig 23. NonLinear-Regression Tranning Test Model

Train/Test (80% - 20%)

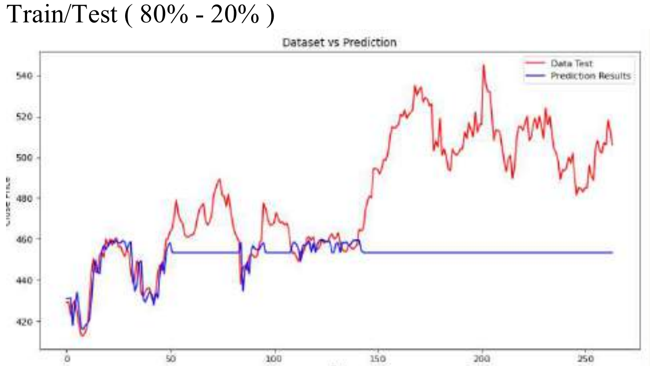

Fig 24. NonLinear-Regression Tranning Test Model

Train/Test (90% - 20%)

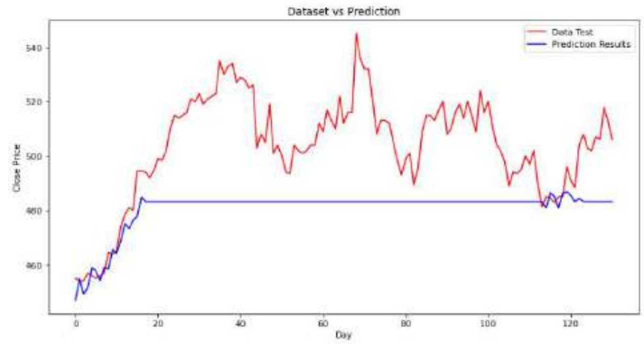

Fig 25. NonLinear-Regression Tranning Test Model

Sarimax

Train/Test (70% - 30%)

SARIMAX Results

Dep. Variable:	Close	No. Observations:	930
Model:	ARIMA(0, 1, 0)	Log Likelihood	-2497.556
Date:	Tue, 03 Jan 2023	AIC	4997.113
Time:	20:31:21	BIC	5001.947
Sample:	0	HQIC	4998.956
	- 930		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
sigma2	12.6669	0.246	51.399	0.000	12.184	13.150

Ljung-Box (L1) (Q):	0.86	Jarque-Bera (JB):	3440.98
Prob(Q):	0.35	Prob(JB):	0.00
Heteroskedasticity (H):	2.21	Skew:	-0.37
Prob(H) (two-sided):	0.00	Kurtosis:	12.40

Fig 26. Sarimax Tranning Test Model

Train/Test (80% - 20%)

SARIMAX Results

Dep. Variable:	Close	No. Observations:	1062
Model:	ARIMA(0, 1, 0)	Log Likelihood	-2916.176
Date:	Tue, 03 Jan 2023	AIC	5834.352
Time:	20:29:30	BIC	5839.319
Sample:	0	HQIC	5836.234
	- 1062		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
sigma2	14.2842	0.289	49.474	0.000	13.718	14.850

Ljung-Box (L1) (Q):	4.16	Jarque-Bera (JB):	2339.56
Prob(Q):	0.04	Prob(JB):	0.00
Heteroskedasticity (H):	2.75	Skew:	-0.14
Prob(H) (two-sided):	0.00	Kurtosis:	10.27

Fig 27. Sarimax Tranning Test Model

Train/Test (90% - 10%)

SARIMAX Results

Dep. Variable:	Close	No. Observations:	1195
Model:	ARIMA(0, 1, 0)	Log Likelihood	-3335.019
Date:	Tue, 03 Jan 2023	AIC	6672.038
Time:	20:32:25	BIC	6677.124
Sample:	0	HQIC	6673.954
	- 1195		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
sigma2	15.6179	0.313	49.936	0.000	15.005	16.231

Ljung-Box (L1) (Q):	7.46	Jarque-Bera (JB):	2035.34
Prob(Q):	0.01	Prob(JB):	0.00
Heteroskedasticity (H):	2.90	Skew:	-0.11
Prob(H) (two-sided):	0.00	Kurtosis:	9.39

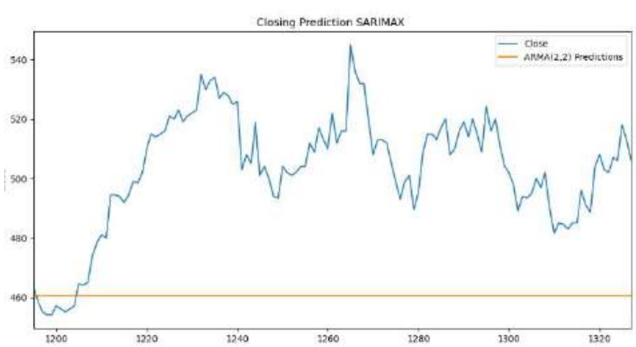

Fig 28. Sarimax Tranning Test Model

Prediction Next 30 Days

Fig 29. Real Chart Price of Data VNC Stock Price On Next 30 Days Later

DeepAR

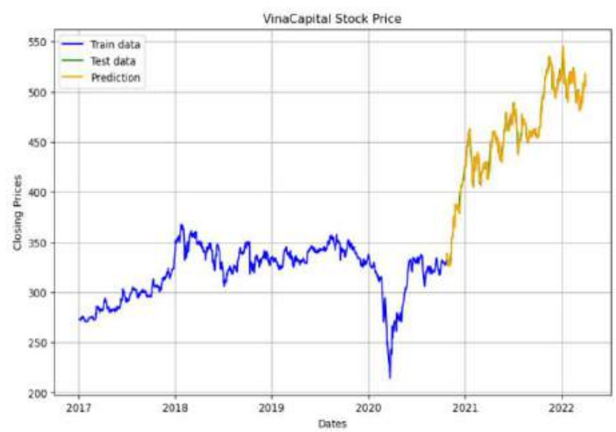

Fig 30. DeepAR Prediction of VNC's Stock Price Next 30 Days Later

Comments: The LSTM-DeepAR model having shown VNC's stock price prediction level after 30 days, we can see that the predicted value has a low difference, which means that the difference between the two graphs is low.

LSTM – Univariate

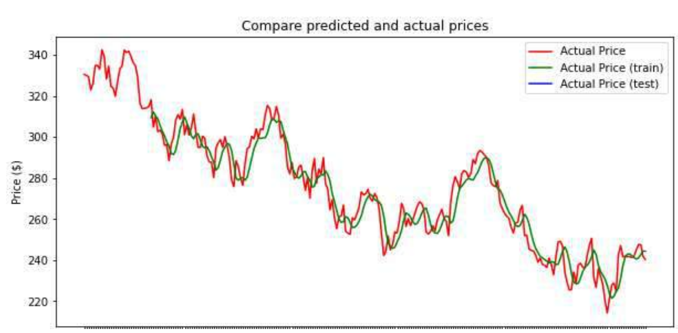

Fig 31. LSTM-Univariate Prediction of VNC's Stock Price Next 30 Days Later

Comments: The LSTM-Univariate model have shown VNC's stock price prediction next 30 days, we can see that the stock's price has a decreasing trend, in some sections there is still an increase, but in general the downtrend is the main, compared to the actual value of the stock, we can see that the difference is quite high.

Sarimax

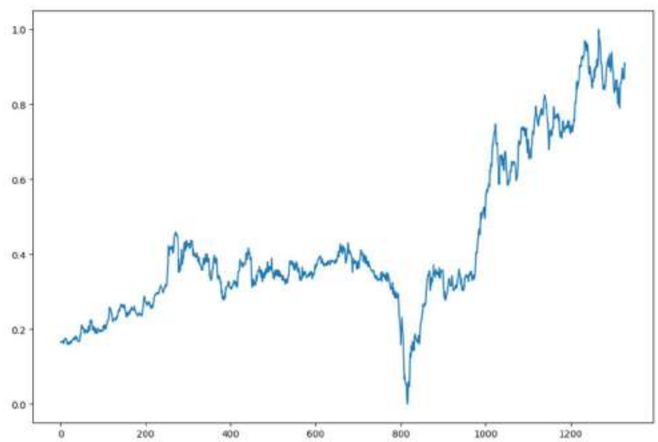

Fig 32. Sarimax Prediction of VNC's Stock Price Next 30 Days Later

Comments: The Sarimax prediction model showing that VNC's Stock Price after 30 days, we can in the early days, the model tends to decrease, but in the second half, it tends to increase again, the ratio of the difference with the real value is average.

REFERENCES

[1] Alexeev, V., & Tapon, F. (2011). Testing weak form efficiency on the Toronto Stock Exchange. Journal of Empirical Finance, 18(4), 661-691

[2] Alphonse, P., & Nguyen, H. (2013). Momentum Effect: Evidence from the Vietnamese Stock Market. Macrothink Institute. Asian Journal of Finance & Accounting, 5(2), 183-202. https://doi.org/10.5296/ajfa.v5i2.4310

[3] Henry M. K. Mok, "Causality of interest rate, exchange rate and stock prices at stock market open and close in Hong Kong," Asia Pacific Journal of Management, Vol. 10 (2), pp. 123– 143,1993.

[4]] R. F. Engle and J. A. Patton. What is a Good Volatility Model? Quantitative Finance, 1:237–245, 2001.

[5] T. Bollerslev. Generalized Autoregressive Conditional Heteroscedasticity. Journal of Econometrics, 31:307–327, 1986.

[6] S. Siami-Namini and A. S. Namin, "Forecasting Economicsand Financial Time Series: ARIMA vs. LSTM," Papers, 2018,https://arxiv.org/abs/1803.06386.

[7] X. Liang, Z. Ge, L. Sun, M. He, and H. Chen, "LSTM withwavelet Transform based data preprocessing for stock priceprediction," Mathematical Problems in Engineering, vol. 2019,Article ID 1340174, 8 pages, 2019.

[8] S. Borovkova and I. Tsiamas, "An ensemble of LSTM neuralnetworks for high-frequency stock market classification,"SSRN Electronic Journal, vol. 01, 2018

[9] Farrar, D. E., & Glauber, R. R. (1967). Multicollinearity in regression analysis: the problem revisited. The Review of Economic and Statistics, 92-107

[10] McCallum, B. T. (1970). Artificial orthogonalization in regression analysis. The Review of Economics and Statistics, 110-113.

[11] Bates, D. M. and Watts, D. G. (1981). A relativeoffset orthogonality convergence criterion fornonlinear least squares. Technometrics,23,179-183.

[12] Cook, R. D. and Tsai, C. L. (1985). Residuals innonlinear regression. Biometrika,72, 23-29.

Another Reference Sources:

Buguk, C., & Brorsen, B. (2003). Testing weak-form market efficiency: Evidence from the Istanbul stock exchange. International Review of Financial Analysis, 12, 579-590. https://doi.org/10.1016/S1057-5219(03)00065-6

Chaudhuri, K., & Wu, Y. (2003). Random Walk Versus Breaking Trend in Stock Prices: Evidence from Emerging Markets. Journal of Banking and Finance, 27(4), 575-592.

Y. Romahi and Q. Shen, "Dynamic financial forecasting with automatically induced fuzzy associations," In Proceedings of the 9th international conference on fuzzy systems., pp. 493– 498,2000.

Y. Romahi and Q. Shen, "Dynamic financial forecasting with automatically induced fuzzy associations," In Proceedings of the 9th international conference on fuzzy systems., pp. 493– 498,2000.

Olha Zubarieva, Ivan Deriuge, Vadym Holosko, et al. Space Balancing in Online Handwriting Recognition Postprocessing using Deep Bidirectional LSTM [C]// 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR). Niagara Falls, NY, USA: IEEE,2018: 576-581

Zbornik Radova Ekonomskog Fakulteta U Rijeci: ČAsopis Za Ekonomsku Teoriju I Praksu, 27 (1) (2009), pp. 171-196

Volodymyr Turchenko, Patrizia Beraldi, Francesco De Simone, et al. Short-term stock price prediction using MLP in moving simulation mode [C] //Proceedings of the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems. Prague, Czech Republic: IEEE, 2011, 2: 666-671

Liu Jun, Wang Gang, Duan Lingyu, et al. Skeleton-Based Human Action Recognition With Global Context-Aware Attention LSTM Networks [J]. IEEE Transactions on Image Processing , 2017, 27(4): 1586-1599

Yasufumi Moriya; Gareth J. F. Jones. LSTM Language Model Adaptation with Images and Titles for Multimedia Automatic Speech Recognition [C]// 2018 IEEE Spoken Language Technology Workshop (SLT). Athens, Greece, Greece: IEEE, 2018: 219-226

Yi Jiangyan, Wen Zhengqi, Tao Jianhua, et al. CTC Regularized Model Adaptation for Improving LSTM RNN Based Multi-Accent Mandarin Speech Recognition [J]. Journal of Signal Processing Systems, 2018, 90(7): 985-997

Aistis Raudys. Optimal negative weight moving average for stock price series smoothing [C] //CIFEr. London, UK: IEEE, 2014: 239-246

Avraam Tsantekidis, Nikolaos Passalis, Anastasios Tefas, et al. Forecasting Stock Prices from the Limit Order Book Using Convolutional Neural Networks [C] // 2017 IEEE 19th Conference on Business Informatics (CBI). Thessaloniki, Greece: IEEE, 2017, 1: 7-12

Watrous, R. L. and Kuhn, G. M. (1992). Induction of nite-state languages using second-order recurrent networks. Neural Computation, 4:406{414.

Williams, R. J. (1989). Complexity of exact gradient computation algorithms for recurrent neural networks. Technical Report Technical Report NU-CCS-89-27, Boston: Northeastern University, College of Computer Science.

Hochreiter, S. and Schmidhuber, J. (1997). LSTM can solve hard long time lag problems. In Advances in Neural Information Processing Systems 9. MIT Press, Cambridge MA. Presented at NIPS 96.

Hochreiter, S. and Schmidhuber, J. (1995). Long short-term memory. Technical Report FKI-207- 95, Fakult at f ur Informatik, Technische Universit at Munchen

. Bengio, Y., Simard, P., and Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. IEEE Transactions on Neural Networks, 5(2)

Almeida, L. B. (1987). A learning rule for asynchronous perceptrons with feedback in a combinatorial environment. In IEEE 1st International Conference on Neural Networks, San Diego, volume 2, pages 609

Baldi, P. and Pineda, F. (1991). Contrastive learning and neural oscillator. Neural Computation, 3:526