

| | | |
|---|--|--|
|  <p>ISSN NO. 2320-5407</p> | <p>Journal Homepage: - www.journalijar.com</p> <h2 style="text-align: center;">INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)</h2> <p style="text-align: center;">Article DOI: 10.21474/IJAR01/3327 DOI URL: http://dx.doi.org/10.21474/IJAR01/3327</p> |  <p>INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR) ISSN 2320-5407 Journal homepage: http://www.journalijar.com Journal DOI: 10.21474/IJAR01</p> |
|---|--|--|

RESEARCH ARTICLE

ONE YEAR STOCK PRICE PREDICTION AND ITS VALIDITY USING LEAST SQUARE METHOD IN MATLAB.

Asep Juarna.

Departemen of Informatics, Gunadarma University, Jakarta-Indonesia.

Manuscript Info

Manuscript History

Received: 28 December 2016
Final Accepted: 15 January 2017
Published: February 2017

Key words:-

LQ45, blue chip stock, time series data, two LQ45's criteria, interpolation, least square method, polynomial, extrapolation, relative RMSE.

Abstract

Given time series data of 43 LQ45's monthly stocks price during the period January 2013 to December 2013; this research interpolates the data into best polynomials as many the stocks, with each RMSE (root mean square error) representing degree of investment risk. The polynomial also is used to extrapolate (i.e. to predict) the next months stock price during the period January 2015 to January 2016; finally the prediction prices is compared with the 13 monthly actual prices and their RMSE (root mean square error) are computed. The computations are run in MATLAB (Matrix Laboratory) programming language implementing the curve fitting procedure based on least square method. The important results of computations are following: (1) degree of polynomials are in the range [28, 97] of maximum 100, with average 73.7209, (2) relative RMSE of interpolation (RMSE-i) are in the range [0.2248%, 2.6804%] with average 0.6719%, (3) relative RMSE of extrapolation (RMSE-e) are in the range [2.1838%, 56.9015%] with average 11.5455%. The average of relative RMSE-e is small enough; it means that the risk for 13 months investment during the period January 2015 to January 2016 is small enough. The value of relative RMSE-e also explains why a stock is removed from or still in LQ45 index, and this is valid for 40 of 43 stocks.

Copy Right, IJAR, 2017,. All rights reserved.

Introduction:-

Trend of stock price changing, i.e. the time series of the price, is interesting information for investor in investing his or her money in the stock; the others information that should be considered is performance of the associated company which is shown by its financial report, cash flow, and equity changing (Brigham, 2004; Xiao, 2014; Jianfeng, 2014). The expected (or predicted) stock price at certain future time, which is a continuum of the trend, can be derived from the time series using some prediction technique such as interpolation (curve fitting) one; this predicted price has some uncertainty that represents the risk of investment in the stock.

In this paper, the time series data is monthly stock price one of LQ45 index, released by IDX (Indonesian Stock Exchange). IDX has about twenties indexes such as LQ45 Index, Kompas100 Index, Jakarta Islamic Index, and InfoBank15. Some of the indexes consist of blue chip stocks, i.e. the stocks with stable return, issued by well-known and well-established companies having small liability and high liquidity (Investopedia, 2016). The monthly time series data can be interpolated to get the best polynomial, i.e. the one of certain degree with minimal relative RMSE

(root mean square error); this relative RMSE represents the risk of investment in the stock. The polynomial actually is an extrapolator to get the prediction price of the stock, together with its relative RMSE, at certain future months.

The objective of the research is *to investigate validity of one year stock price prediction*, especially of 43 blue chip stocks which are member of LQ45 index defined in period of February 2015 - July 2015¹⁾, here validity means degree of accordance of the prediction values and the actual ones, also to investigate why a stock is removed from or still in LQ45 index. The monthly data are the prices during the period January 2013 to December 2014. This historical data is interpolated to get the best polynomial for each stock; the chosen polynomial is the best one, i.e. the polynomial with certain degree over interval [1,100] having minimal relative interpolation-RMSE (RMSE-i). Each chosen polynomial is then used to extrapolate the stock price during the period January 2015 to January 2016; the resulted stock price is then compared with its actual data and the relative extrapolation-RMSE (RMSE-e) is determined; *the relative RMSE-e shows the validity of prediction* which in turn gives guidance for investor to keep track of his or her investment in the stock.

Research Methodology:-

LQ45 Index:-

Many blue chip stocks are member of famous indexes in IDX like LQ45, however not all member of LQ45 are blue chip stocks. A stock is member of LQ45 at a period if it fulfills the following “two LQ45’s criteria” (Fakhrudin, 2006):

for the first 30 stocks:-

- listed in IDX at least for 3 months
- the biggest 30 stocks in regular market trading

For the next 15 stocks:-

- part of 25 active stocks with the greatest number of trading days in regular market,
- then, part of 20 stocks with highest transaction value and frequency,
- finally, part of 15 stocks with biggest market capitalization

Membership each stock in LQ45 is evaluated every February and August. In August 2015, for example, two stocks with (in this research) codes “05” and “06” (ANTM, Aneka Tambang (Persero) Tbk, and CTRA, Ciputra Development Tbk, respectively) are removed from LQ45 and replaced by SRIL (Sri Rejeki Isman Tbk) and WTON (Wijaya Karya Beton Tbk), while in February 2016 stocks “17” (EXCL, XL Axiata Tbk), “23” (ITMG, Indo Tambangraya Megah Tbk), and WTON (Wijaya Karya Beton Tbk), are replaced by ANTM (Aneka Tambang (Persero) Tbk), HSMP (PT HM Sampoerna Tbk), dan MYRK (PT Hanson International Tbk) masuk (Febri, 2016; Kabar, 2016). Note that WTON is not one of 43 stocks under consideration in this research.

Least Square Method:-

Given m pairs data (x_i, y_i) ; least square method constructs polynomial of degree n , i.e. $p_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$, i.e. to get the coefficients $a_0, a_1, a_2, \dots, a_n$, with following “least square procedure”:

- compute $\sum_i x_i, \sum_i x_i^2, \sum_i x_i^3, \dots, \sum_i x_i^{2n}$
- compute $\sum_i y_i, \sum_i x_i y_i, \sum_i x_i^2 y_i, \sum_i x_i^3 y_i, \dots, \sum_i x_i^n y_i$
- construct the following linear system:

$$\begin{pmatrix} m & \sum_i x_i & \sum_i x_i^2 & \dots & \sum_i x_i^n \\ \sum_i x_i & \sum_i x_i^2 & \sum_i x_i^3 & \dots & \sum_i x_i^{n+1} \\ \sum_i x_i^2 & \sum_i x_i^3 & \sum_i x_i^4 & \dots & \sum_i x_i^{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_i x_i^n & \sum_i x_i^{n+1} & \sum_i x_i^{n+2} & \dots & \sum_i x_i^{2n} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \sum_i y_i \\ \sum_i x_i y_i \\ \sum_i x_i^2 y_i \\ \vdots \\ \sum_i x_i^n y_i \end{pmatrix} \Leftrightarrow MA = B \quad (1)$$

¹⁾ Two of 45 stocks, i.e. SILO (Siloam International Hospital Tbk) and SSMS (Sawit Sumbermas Sarana Tbk), do not included in this research because of incompleteness data.

4. solve the linear system for $A = [a_0, a_1, a_2, \dots, a_n]^T$, where X^T is transpose matrix of X .

In MATLAB environment the 3-rd and 4-th steps are executed in two following instructions:

$$1. \text{ assign } M = \begin{pmatrix} m & \sum_i x_i & \sum_i x_i^2 & \cdots & \sum_i x_i^n \\ \sum_i x_i & \sum_i x_i^2 & \sum_i x_i^3 & \cdots & \sum_i x_i^{n+1} \\ \sum_i x_i^2 & \sum_i x_i^3 & \sum_i x_i^4 & \cdots & \sum_i x_i^{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_i x_i^n & \sum_i x_i^{n+1} & \sum_i x_i^{n+2} & \cdots & \sum_i x_i^{2n} \end{pmatrix}, B = \begin{pmatrix} \sum_i y_i \\ \sum_i x_i y_i \\ \sum_i x_i^2 y_i \\ \vdots \\ \sum_i x_i^n y_i \end{pmatrix}$$

$$2. A = M \setminus B$$

In this research the input is matrix M of size 24×43 where the cell of i -th row and j -th column holds the prices of i -th month of j -th stock; the matrix is shown by Table 1.

Table 1:- Format of the monthly price data of 43 stocks for 2 years

| Month | Price of Stock-1 | Price of Stock-2 | | Price of Stock-43 |
|---------------|------------------|------------------|----------|-------------------|
| $x_1 = 1$ | $y_{1,1}$ | $y_{1,2}$ | | $y_{1,43}$ |
| $x_2 = 2$ | $y_{2,1}$ | $y_{2,2}$ | | $y_{2,43}$ |
| \vdots | \vdots | \vdots | \ddots | \vdots |
| $x_{24} = 24$ | $y_{24,1}$ | $y_{24,2}$ | | $y_{24,43}$ |

Computation Algorithm:-

Let "dt1314.mat" be MATLAB file that holds matrix M of size 24×43 , i.e. the prices of 43 stocks during the period January 2013 to December 2014, and "dt156.mat" MATLAB file that holds matrix $M1$ of size 13×43 , i.e. the prices of 43 stocks during the period January 2015 to January 2016; this algorithm construct 43 best polynomials, one for each stock, and so on, an finally compute each relative-RMSE-i and relative-RMSE-e:

Compute relative-RMSE-i and get the degree deg of the best polynomial:-

for stock $j = 1$ to 43 do

- $x_j = [1:24]$, $y_j = \text{dt1314}(1:24,j)$
- for degree of polynomial $n = 1$ to 100 do
 - construct matrix S and vector B of equation (1) in Section 2.2
 - solve the linear system $SA = B$ for $A = [a_0, a_1, \dots, a_n]$
 - for $i = 1$ to 24, compute interpolation price $y_{j,\text{int}} = a_0 + a_1x + a_2x^2 + \dots + a_n x^n$
 - compute relative RMSE-i(n) between y_j and $y_{j,\text{int}}$
- find the minimum relative RMSE-i(n) and get the degree deg of that best polynomial

Compute relative-RMSE-e for the best polynomial of degree deg:-

for stock $j = 1$ to 43 do

- $x_j = [1:13]$, $y_j = \text{dt156}(1:13,j)$
- for $i = 1$ to 13, compute extrapolation price $y_{j,\text{ext}} = a_0 + a_1x + a_2x^2 + \dots + a_n x^d$
- compute relative RMSE-e between y_j and $y_{j,\text{ext}}$

Result and Discussion:-

Main outputs of computation of every stock are: degree deg of the best polynomial, the relative RMSR-i, the relative RMSR-e, the curves of interpolation and extrapolation. Following are the statistic of the output:

- The interval of degree of the best polynomials is $[28, 97]$ with average 73.7209.
- The interval of the relative RMSE-i is $[0.2248\%, 2.6804\%]$ with average 0.6719%.
- The interval of the relative RMSE-e is $[2.1838\%, 56.9015\%]$ with average 11.5455%.

The average of relative RMSE-e of all 43 stocks is small enough; it indicates that *investment in the stocks on average is save enough*. The complete statistic is shown in the Table 2.

Table 2 also tells more information like follow. As is stated in the previous section, two stock namely “05” and “16” are removed from LQ45 on August 2015 while “17” and “23” are removed on February 2016; Table 2 shows that the relative RMSE-e of the four stocks are, respectively, 29.7395%, 7.3793%, 9.4617%, and 48.2903%. The relative RMSE-e of stocks “01” and “23” give a reason why the two stocks are removed. However the relative RMSE-e of the two rest actually are small enough, below the average, so, in these cases, the relative RMSE-e is not sufficient to become the reason of the removing; there is other reason, of course, to make the “two LQ45’s criteria” are hold, and the reason is beyond the scope of the research. Also relative RMSE-e analysis fails to explain why stock “39” still in LQ45 index.

Table 2:- Degree of the best polynomial (deg), relative RMSE-i (%), and relative RMSE-e (%) of each of 43 stocks. Average(deg) is 73.7209, average(relative RMSE-i) is 0.6719%, and average(relative RMSE-e) is 11.5455%. The bold text indicate the removed stock fro LQ45.

| No | deg | RMSE-i | RMSE-e | No | deg | RMSE-i | RMSE-e | No | deg | RMSE-i | RMSE-e |
|-----------|-----------|---------------|----------------|-----------|-----------|---------------|----------------|----|-----|--------|---------|
| 01 | 97 | 0.9205 | 6.3036 | 16 | 95 | 0.8179 | 7.3793 | 31 | 97 | 0.3743 | 13.5326 |
| 02 | 80 | 0.9227 | 9.0233 | 17 | 49 | 0.2248 | 9.4617 | 32 | 93 | 0.8519 | 17.4633 |
| 03 | 60 | 0.6751 | 15.9618 | 18 | 97 | 0.3895 | 4.4632 | 33 | 80 | 0.7535 | 18.7823 |
| 04 | 37 | 0.3909 | 7.3283 | 19 | 77 | 0.4072 | 7.0916 | 34 | 97 | 0.7672 | 8.5405 |
| 05 | 72 | 0.6073 | 29.7395 | 20 | 37 | 0.4764 | 8.3784 | 35 | 97 | 0.4414 | 4.6256 |
| 06 | 84 | 0.3459 | 2.1838 | 21 | 60 | 0.3285 | 4.5545 | 36 | 80 | 0.4336 | 9.1012 |
| 07 | 95 | 0.8046 | 15.9732 | 22 | 97 | 0.5881 | 2.9206 | 37 | 54 | 1.3913 | 10.6823 |
| 08 | 60 | 0.4150 | 6.5224 | 23 | 28 | 0.8960 | 48.2903 | 38 | 80 | 0.4396 | 8.9642 |
| 09 | 74 | 0.5415 | 6.7029 | 24 | 50 | 0.2561 | 3.8816 | 39 | 69 | 2.6804 | 56.9015 |
| 10 | 74 | 0.5673 | 7.7924 | 25 | 97 | 0.3932 | 6.5289 | 40 | 64 | 0.3363 | 3.7506 |
| 11 | 97 | 0.3502 | 8.5664 | 26 | 60 | 1.0949 | 5.1224 | 41 | 87 | 0.2665 | 8.1895 |
| 12 | 74 | 0.5637 | 3.6122 | 27 | 69 | 0.6812 | 13.0897 | 42 | 76 | 0.7175 | 10.5502 |
| 13 | 42 | 0.8210 | 18.2026 | 28 | 50 | 1.1610 | 6.2316 | 43 | 80 | 1.1097 | 17.4144 |
| 14 | 80 | 0.7973 | 5.3651 | 29 | 84 | 0.5055 | 11.8445 | | | | |
| 15 | 60 | 0.7457 | 13.3012 | 30 | 80 | 0.6384 | 12.1394 | | | | |

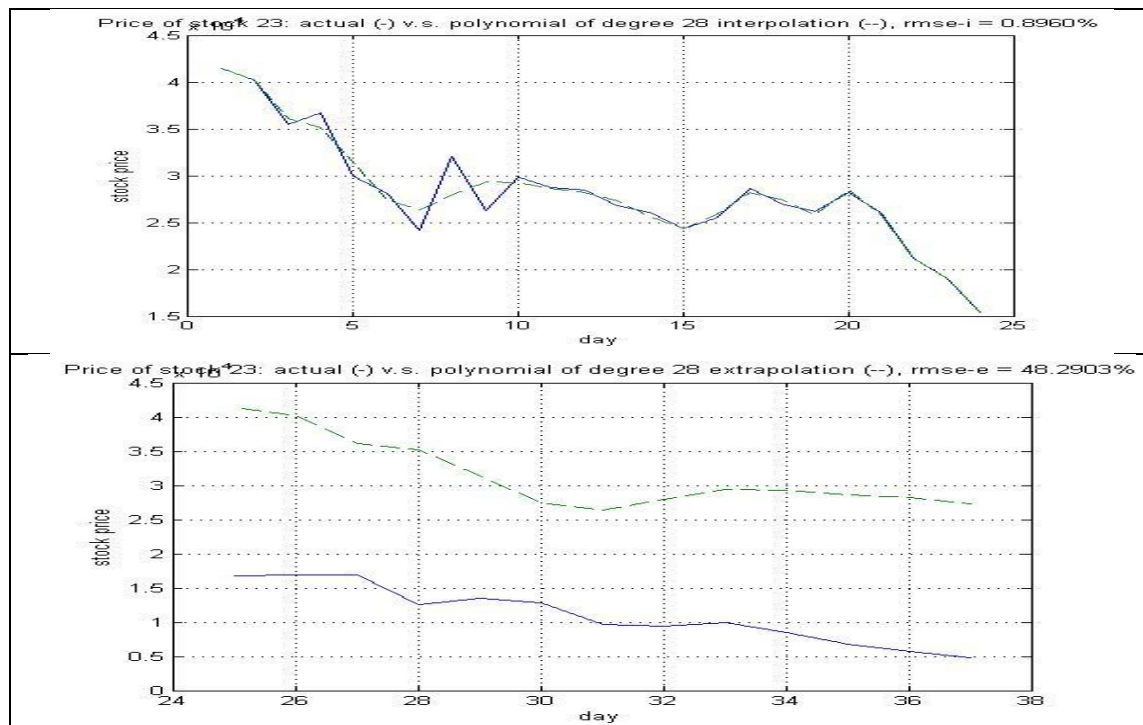


Figure 1:- Stock “23” (ITMG, Indo Tambangraya Megah Tbk) has the smallest degree of interpolation, i.e. 28, but its relative RMSE-e almost 50%. This stock is removed from the LQ45 in February 2016.

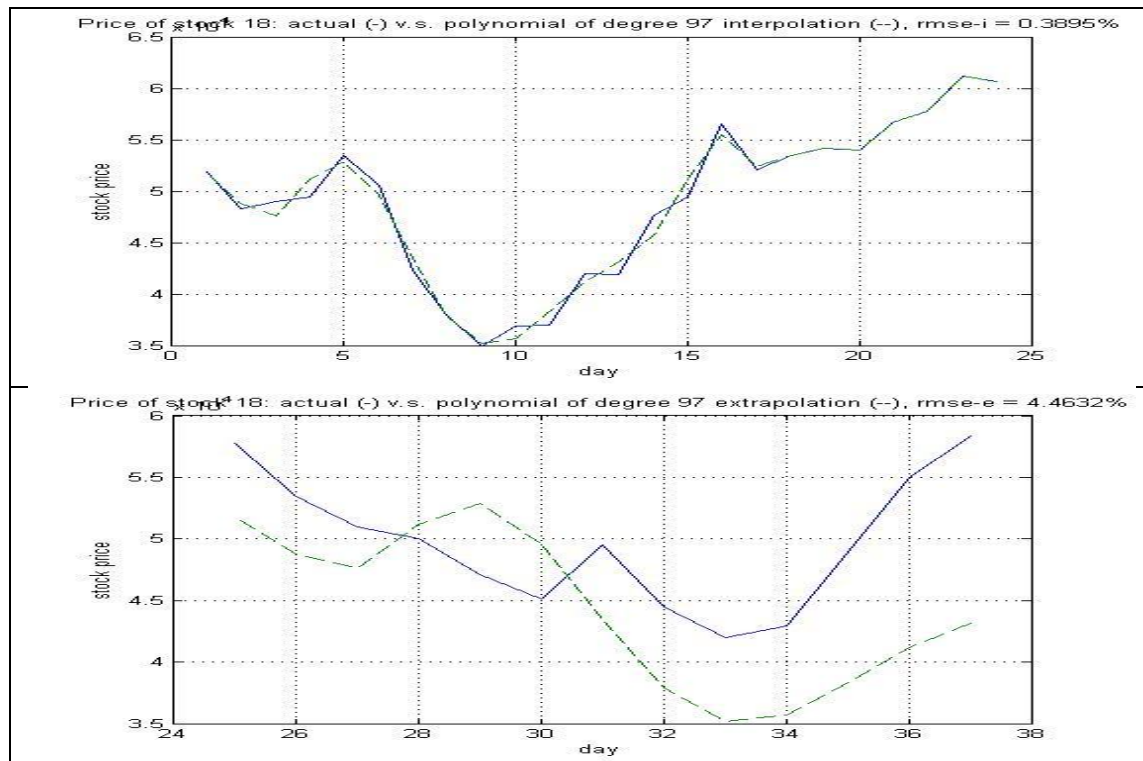


Figure 2:- Stock “18” (GGRM, Gudam Garam Tbk) is the stock having the largest degree of polynomial and relative RMSE-e less than 5%. This stock is saved in LQ45 in February 2016.

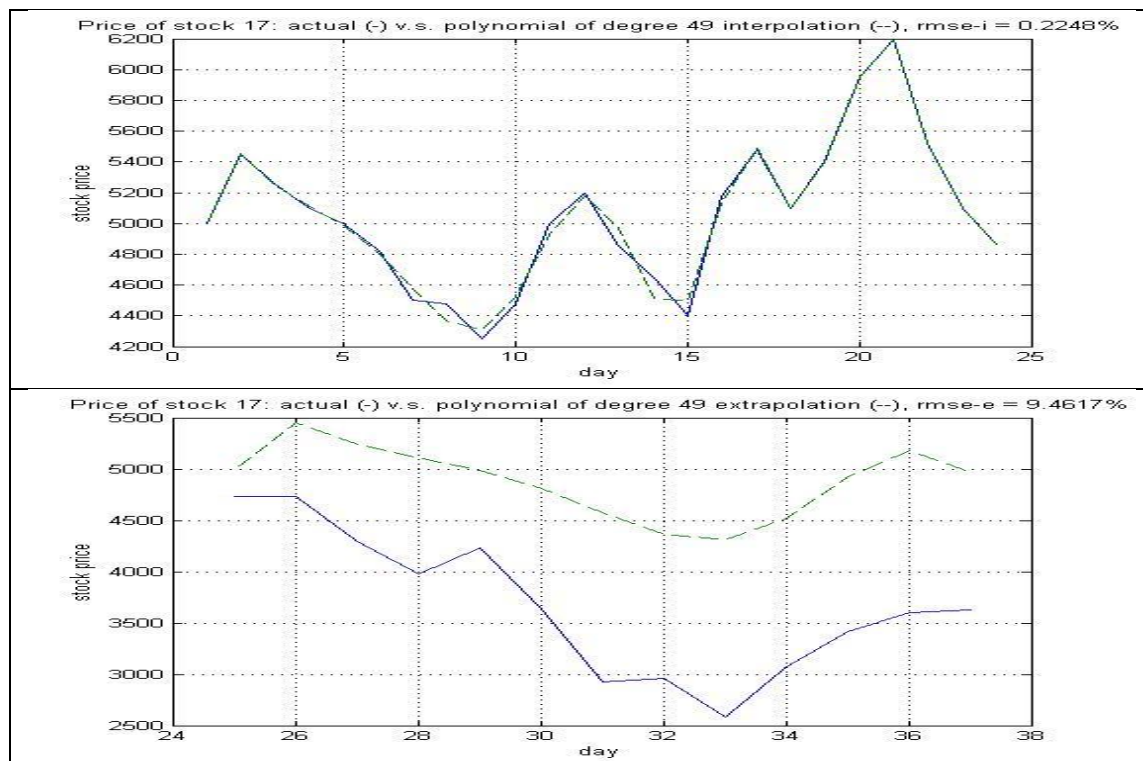


Figure 3:- Stock “17” (EXCL, XL Axiata Tbk) has smallest relative RMSE-i and relative RMSE-e small enough, below 10%. However this stock is removed from LQ45 in February 2016.

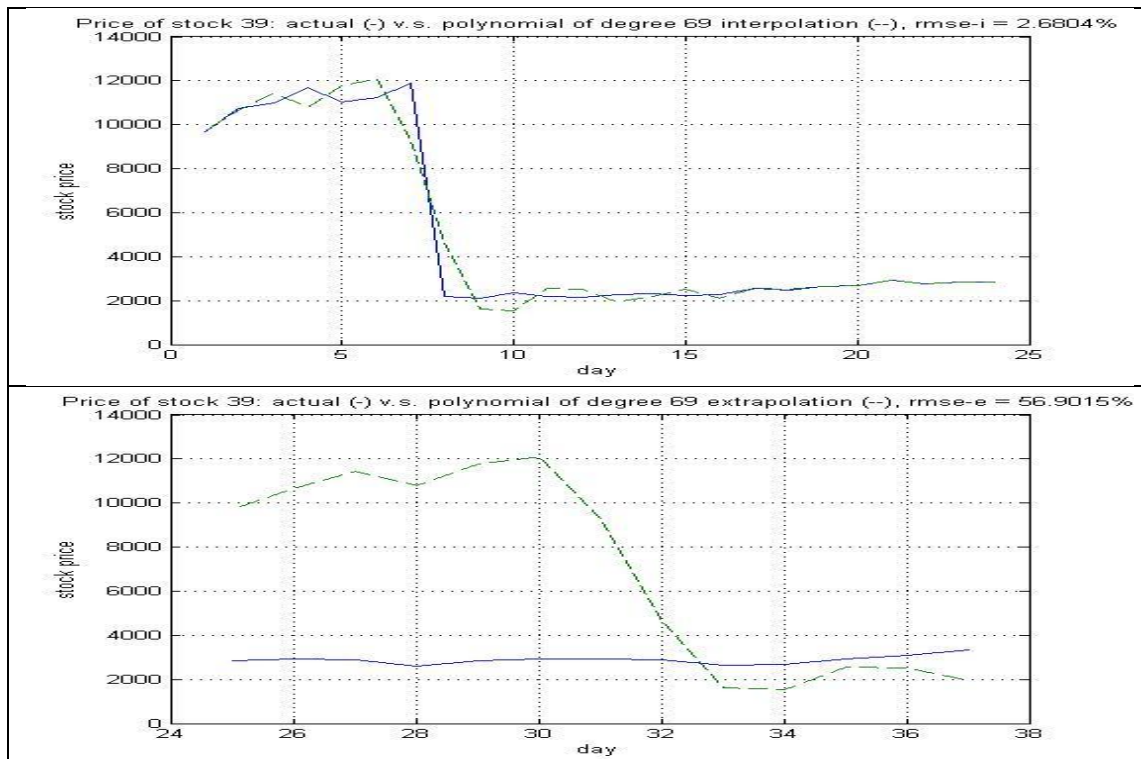


Figure 4:- Stock “39” (TLKM, Telekomunikasi Indonesia (Persero) Tbk) has largest both relative RMSE-i and relative RMSE-e; however this stock save in LQ45 in February 2016.

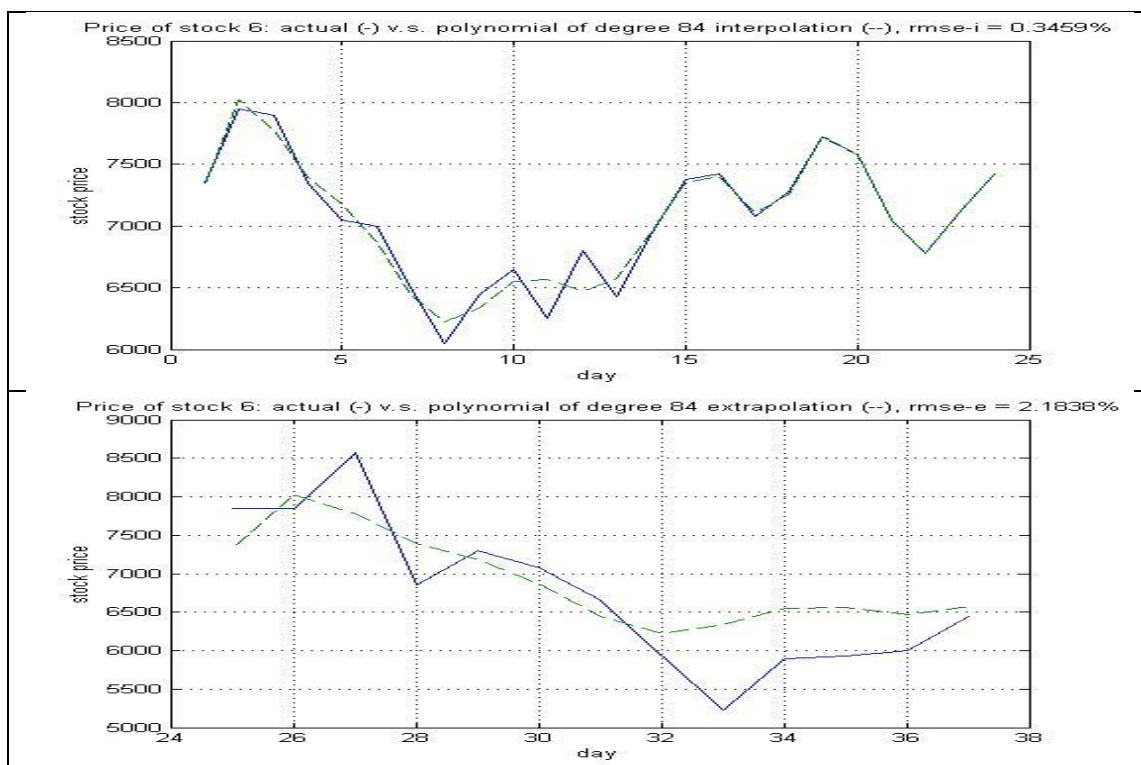


Figure 5:- Stock “6” (ASII, Astra International Tbk) has smallest relative RMSE-e. This stock is save in LQ45 in February 2016.

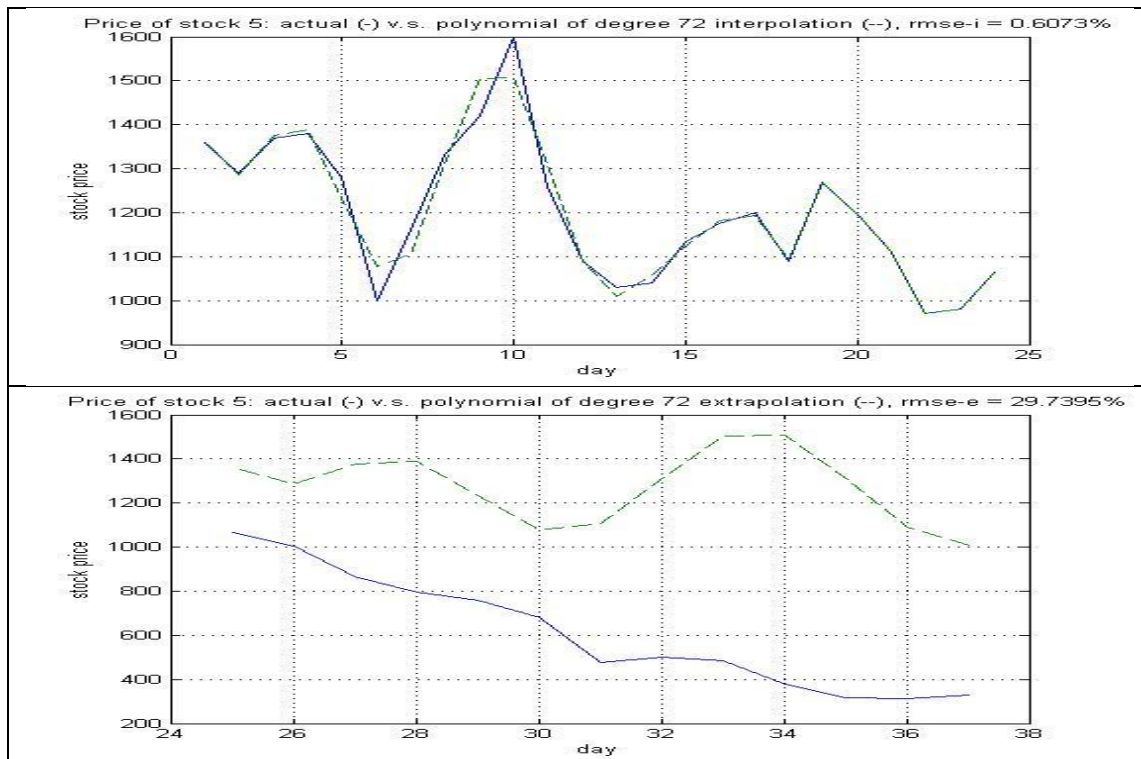


Figure 6:- Stock “5” (ANTM, Aneka Tambang (Persero) Tbk) has large relative RMSE-e. This stock is removed from LQ45 in August 2015.

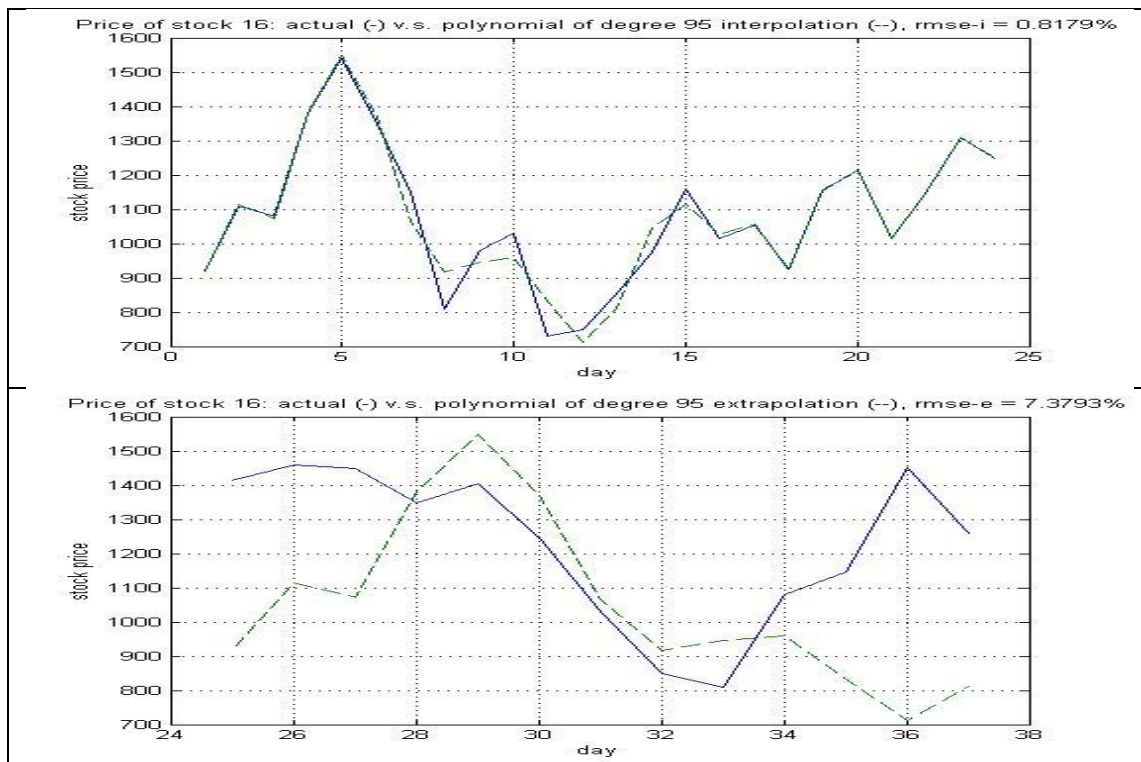


Figure 7:- Stock 16 (CTRA, Ciputra Development Tbk) has relative RMSE-e large enough. This stock is removed from LQ45 in August 2015.

Figure 1 until 7 shows seven selected stocks with variation of degree of the best polynomial, value of relative RMSE-i, and value of relative RMSE-i, represent all of 43 stocks. Figure 1, Figure 2, Figure 5, Figure 6, and Figure 7 show the stocks of which their relative RMSE-i determine their membership in LQ45 while Figure 3 and Figure 4 has anomaly in the sense of value of relative RMSE-i.

Conclusion:-

1. The average of relative RMSE-e of all 43 stocks is small enough; it indicates that *investment in the stocks for 13 months on average is save enough*.
2. With the average of relative RMSE-e which is 11.5455%, 40 of 43 stocks can be explained why a stock is removed from LQ45 or still save in it, i.e. because the RMSE-much larger than the average or about the average, respectively.

Acknowledgment:-

This work is funded by Directorate of Higher Education, Department of Education and Culture, Republic of Indonesia, under Grant No. 790/K3/KM/SPK.LT/2016, June 14-th, 2016.

References:-

1. Brigham, Eugene F, dan Weston. 2004. *Dasar-Dasar Manajemen Keuangan*. Erlangga. Jakarta.
2. Fakhruddin, Hendy M. 2006. *Pasar Modal Indonesia*. Salemba Empat. Jakarta.
3. Febri Kurnia. 2016, 26 Januari. *Ini Penyebab XL Terpentak dari LQ45*. Warta Ekonomi.
4. Investopedia. *What is the benefit of investing in blue chip stocks?*. <http://www.investopedia.com/ask/answers/031915/what-benefit-investing-blue-chip-stocks.asp>
5. Jianfeng Si, Arjun Mukherje, Bing Liu, Sinno Jialin Pan, Qing Li, dan Huay Li. 2014. *Exploiting Social Relations and Sentiment for Stock Prediction*. Dalam Proceeding of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Association for Computational Linguistics.
6. Kabar Bisnis. 2016, 25 Januari. *Sampoerna, Antam, dan Hanson masuk LG45, tiga saham ini terpentak*.
7. Xiao Deng, Yue Zhang, Ting Liu, and Junwen Duan. *Using Structured Events to Predict Stock Price Movement: An Empirical Investigation*. In Proceeding of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Association for Computational Linguistics.