

BUKTI KORESPONDENSI

Judul :

Ensemble Classification Method for Daily Return Stock Market"

- 1. Submit artikel (24 Juni 2020)**
 - Artikel yang di-submit**

- 2. Review artikel (6 September 2020)**
 - Tanggapan Review**
 - Revisi Paper**

- 3. Review ke 2 (22 November 2020)**

- 4. Artikel Accepted (22 November 2020)**

- 5. Pemberitahuan publish artikel (21 November 2021)**

- 1. Submit artikel (24 Juni 2020)
- Artikel yang di-submit**



Meilany Nonsi Tentua <meilany@upy.ac.id>

Songklanakarin Journal of Science and Technology - Manuscript ID SJST-2020-0281

1 message

Songklanakarin Journal of Science and Technology <onbehalf@manuscriptcentral.com>

Wed, Jun 24, 2020 at 10:47 AM

Reply-To: sjst@psu.ac.th

To: meilany@upy.ac.id

23-Jun-2020

Dear Mrs. Tentua:

Your manuscript entitled "Ensemble Classification Method for Daily Return Stock Market" **has been successfully submitted** online and is presently being given full consideration for publication in the Songklanakarin Journal of Science and Technology.

Your manuscript ID is SJST-2020-0281.

Please mention the above manuscript ID in all future correspondence or when calling the office for questions. If there are any changes in your street address or e-mail address, please log in to ScholarOne Manuscripts at <https://mc.manuscriptcentral.com/sjst> and edit your user information as appropriate.

You can also view the status of your manuscript at any time by checking your Author Center after logging in to <https://mc.manuscriptcentral.com/sjst>.

Thank you for submitting your manuscript to the Songklanakarin Journal of Science and Technology.

Sincerely,
Songklanakarin Journal of Science and Technology Editorial Office



Ensemble Classification Method for Daily Return Stock Market

Journal:	<i>Songklanakarín Journal of Science and Technology</i>
Manuscript ID	Draft
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Tentua, Meilany; Universitas PGRI Yogyakarta, informatic Rosadi, Dedi; Universitas Gadjah Mada, Department of Mathematics
Keyword:	Ensemble Method, Stock Market Returns, Classification

SCHOLARONE™
Manuscripts

Original Article

Ensemble Classification Method for Daily Return Stock Market

Meilany Nonsi Tentua¹, Dedi Rosadi²

¹Department of Informatics, Sains and Technology Fakultas, Universitas

PGRI Yogyakarta, Yogyakarta, 55182, Indonesia

²Universitas Gajah Mada, Yogyakarta, Indonesia, Sekip Utara Bulaksumur,
Yogyakarta, 55281, Indonesia

* Corresponding author, Email address: meilany@upy.ac.id

Abstract

Stock price movements in Indonesia are measured using indices, one of which is the Composite Stock Price Index (CSPI). CSPI is a stock index that measures a combination of all shares from various sectors listed on the Indonesia Stock Exchange.

The ensemble method is to build predictive models by combining the strengths of the classical classification method. In this research, the purpose of ensemble based on Boosting for Regression appeared to enhance simple tree analysis and deals with some of the weaknesses found in uncomplicated techniques. The ensemble tree combines the prediction values of many simple trees into a single prediction value. Based on the experiments that have been carried out, the ensemble method proved to have a better accuracy rate, which amounted to 82%. The model has been assumed by the ensemble model that can get the relationship between variables to more precise than the previous model.

Keywords: Ensemble Method, Stock Market Returns, Classification.

1. Introduction

Stock is one of the most popular financial market instruments. Stock can be defined as a sign of capital participation of a person or party (business entity) in a company or limited liability company (“PT Bursa Efek Indonesia,” n.d.). Share price movements in Indonesia are measured using indices, one of which is the Composite Stock Price Index (CSPI). CSPI is a stock index that measures a combination of all shares from various sectors listed on the Indonesia Stock Exchange. Investors usually use parameters in the CSPI to read price developments and make a reference to the portfolio.

A stock market index is a measurement of the value of a part of the stock market. This is often used to describe market aggregate trends. A basic financial problem will predict this trend. Such stochastic values are very difficult to predict (Bruni, 2017) . The difficulty in predicting the return of shares purchased by investors means that an investor is still guessing about the returns that investors will obtain in the future. There are four approaches to predicting stock price movements: technical analysis (Sedighi, Mohammadi, Fard, & Sedighi, 2019), fundamental analysis (Miralles-Quirós, Miralles-Quirós, & Oliveira, 2017), traditional time series forecasting (Park, Lee, & Lee, 2019), and statistical or machine learning methods (Kim & Han, 2016).

Statistical learning refers to a set of tools for modeling and understanding complex data sets. Statistical learning is a newly developed area in statistics and

1
2
3
4
5
6
7 integrates with parallel developments in computer science and, in particular,
8 machine learning (Gareth James, Daniela Witten, Trevor Hastie, 2013).

9
10
11 Predictive modeling using datasets is one of the tasks in statistical learning. An
12 approach to predict the qualitative response is a process known as
13 classification. Classification is the attribution of labels to notes according to
14 criteria that are automatically learned from a series of labeled record training.
15
16
17
18
19

20 Some algorithms classification of classics that are widely used today
21 (Trevor Hastie, Robert Tibshirani, 2017) like, linear discriminant analysis,
22 logistic regression (Gündüz, Çataltepe, & Yaslan, 2017)(Shah, Ismail, & Bin
23 Shahrin, 2019) , K-nearest neighbors (Tanuwijaya & Hansun, 2019), and trees. In
24 making predictions for the daily stock return market, an appropriate method is
25 needed to produce an accurate model. Classification methods such as random
26 forests (Jyothirmayee, Dilip Kumar, Someswara Rao, & Shiva Shankar, 2019),
27 support vector machines (Selakov, Cvijetinović, Milović, Mellon, & Bekut,
28 2014), boosting (Nayak, Pai, & Pai, 2016), bagging and stacking are methods that
29 have emerged recently. These methods are also called ensemble methods, which
30 are methods that can improve the accuracy of prediction models that are
31 formed. The idea of ensemble learning is to build predictive models by combining
32 the strengths of classical classification methods (Trevor Hastie, Robert Tibshirani,
33 2017).
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

52 The finding of the study (Kaur & Kalra, 2019) showed that using an
53 ensemble approach based on KNN and Naive Bayes can detect credit card fraud.
54 The use of the Naïve Bayes method in credit card fraud detection is appropriate
55
56
57
58
59
60

1
2
3
4
5
6
7 because the values in each data class are independent. So that there is no
8
9 dependency between one attribute with other attributes in each class. This is one
10
11 of the requirements for using the Naïve Bayes method in classification. Because
12
13 each attribute class must be independent, the ensemble approach with the
14
15 combination of the Naïve Bayes method would be inappropriate if used in the
16
17 daily stock market prediction. Because classes on the daily stock market data are
18
19 not independent.
20
21

22
23 In the study this time the proposed methods of classification boosting for
24
25 regression tree (ensemble) that used to predict daily stock market returns in the
26
27 movement of the stock Jakarta Composite Index (JCI).
28
29

30 31 32 **2. Materials and Methods**

33
34 Several studies have been conducted in the application
35
36 of statistical learning in the stock market domain. The use of the bootstrap method
37
38 is modified with Random Forest to predict the direction of the movement of stock
39
40 index prices (Kim & Han, 2016). Training sets generated by a modified bootstrap
41
42 m taking account of the impact of response variables simultaneously and applied
43
44 in a random forest. The data used is the Korea composite stock price index with
45
46 the variable close price, open price, high price, low price, trading volume, and
47
48 training amount.
49
50

51
52 A hierarchical Deep Neural Network (DNN) is a method that can be used
53
54 to predict stock market returns. This study predicts stock market returns within
55
56 the next five minutes (Lachiheb & Gouider, 2018). The data used are the main
57
58
59
60

1
2
3
4
5
6
7 index of the Tunisian stock market (TUNINDEX) with the variable TUNINDEX
8
9 current value, Number of up direction stocks, Number of down direction stocks
10
11 Number of unchanged stocks, Total stock values in Tunisian dinars. Other
12
13 research that uses DNN is a big comprehensive data analysis process to predict
14
15 the daily return direction of the S&P 500 SPDR ETF (ticker symbol: SPY) based
16
17 on 60 financial and economic features (Zhong & Enke, 2019).
18
19

20 The approach of a cross-sectional statistical estimation model for selecting
21
22 stocks on the Chinese stock market allows investors to identify stocks that are
23
24 likely to perform well and to build a suitable portfolio (Wu, Chen, Xu, He, &
25
26 Tindall, 2019). The data used is sourced from the Shanghai Composite Index.
27
28

29 Predicting stock market prices using Fuzzy Metagraph (FM) based on
30
31 stock market decision making, classification, and prediction are proposed for
32
33 short-term investors from the Indian stock market (Anbalagan & Maheswari,
34
35 2015). Simple Moving Average (SMA), Exponential Moving Average (EMA),
36
37 Moving Average Convergence Divergence (MACD), and Relative Strength Index
38
39 (RSI) are some Technical Indicators that are used as input to train an integrated
40
41 system with Fuzzy Meta graph. The data used in the Indian stock market data,
42
43 namely the Bombay Stock Exchange (BSE).
44
45
46
47

48 Artificial Neural Network (ANN) is a method in machine learning that can
49
50 be used for prediction and classification. Implementation of ANN back-
51
52 propagation is used to predict the closing price of the S & P 500 stock
53
54 exchange [19]. Historical data consists of five variables that are open, high, low,
55
56 close, and volume.
57
58
59
60

1
2
3
4
5
6
7 The Adaptive Neuro-Fuzzy Inference System (ANFIS) method is used to
8
9 determine the Jakarta Composite Index (JCI) value movement. This study
10
11 provides two outputs, namely prediction results and classification results based on
12
13 several relative error values. The classification method used is Fuzzy Kernel C-
14
15 Means (Fanita & Rustam, 2018).
16
17

18 This research is done to classify and predict the daily return of the stock
19
20 market in Indonesia. The undertaken method in this research can be seen in
21
22 Figure 1.
23

24
25 Step 1 : collect CSPI data from the Indonesia Composite Index through the
26
27 Yahoo Finance web.
28

29
30 Step 2: calculate the market stock daily return value from the data collected in
31
32 step 1 with the formula :
33

$$34 \quad \text{ReturnValue} = \frac{\text{currentvalue} - \text{originalvalue}}{\text{originalvalue}} * 100 \quad (1)$$

35
36
37 Step 3: Data Preprocessing will be done where missing values and excessive data
38
39 from the dataset will be handled.
40

41
42 Step 4: The training dataset must be large compared to the test
43
44 set. The training dataset is taken from 2015 to 2018 data, while the test
45
46 dataset is 2019 data.
47

48
49 Step 5: Applying the ensemble method, namely, 2) Boosting Regression Tree
50
51 (BoosRT). Steps for the ensemble method:
52

53
54 Step 6: The results of the ensemble model are used to analyze performance in
55
56 terms of accuracy, memory, accuracy, and specificity.
57
58
59
60

Boostered Regression Tree (BoosRT) models are a combination of two techniques: decision tree algorithms and boosting methods (Shin, 2015). In these algorithms, the BoosRT approximates the function (x) as an preservative growth of the base learner :

$$f(x) = f_0(x) + \beta_1 f_1(x) + \beta_2 f_2(x) + \dots + \beta_m f_m(x) \quad (2)$$

Single learning cannot predict enough using training data, so we must use the lowest residuals to improve predictive performance. BoosRT uses an iterative algorithm, so every iteration m , must have a new regression tree. To predicts a separate constant value in each one :

$$h(x; \{R_{lm}\}_l^L = \sum_{i=1}^L \overline{y_{lm}}(x \in R_{lm}) \quad (3)$$

where $y_{lm} = \text{mean } x_i \in R_{lm}$ (y_{lm}) is the mean of pseudo-residuals the m^{th} iteration :

$$\tilde{y}_{lm} = - \left[\frac{\partial \Psi(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x)=F_{m-1}(x)} \quad (4)$$

BoosRT algorithm for the generalized boosting of regression trees :

- a) Initialize (x), $F_0(x) = \arg \min_Y \sum_{i=1}^N \Psi(y_i, \gamma)$.
- b) For $m=1$ to M do
- c) From the full training dataset, select a subset randomly,

$$\{\pi(i)\}_i^N = \text{rand_perm}\{i\}_i^N \quad (7)$$

- d) Fit the base learner

$$\tilde{y}_{\pi(i)} = - \left[\frac{\partial \Psi(y_{\pi(i)}, F(x_i))}{\partial F(x(i))} \right]_{F(x)=f_{m-1}(x)} \quad , i = 1, \tilde{N} \quad (8)$$

- e) update model for the current iteration,

$$\{R_{lm}\}_l^L - \text{terminalmodetree} \left(\left\{ \widetilde{y}_{\pi(i)}, x_{\pi(i)} \right\}_l^N \right) \quad (9)$$

f) Choose a gradient descent step size as

$$Y_{lm} = \operatorname{argmin}_y \sum_{x(i) \in R_{lm}} \Psi(y_{\pi(i)}, F_{m-1}(x_{\pi(i)})) + Y \quad (10)$$

g) Update the estimate of (x) as

$$F_m(x) = F_{m-1}(x) + v \cdot Y_{lm} \mid (x \in R_{lm}) \quad (11)$$

3. Results and Discussion

The daily return stock market classification using the ensemble method is implemented using R. Some of the libraries contained in R are tools for statistical analysis and can also be used to create models with classification techniques (Gareth James, Daniela Witten, Trevor Hastie, 2013), (Rosadi, 2017). The dataset used in this study is JCI, PT Bank Negara Indonesia (Persero) Tbk, PT Perusahaan Perkebunan London Sumatra Indonesia Tbk, PT Unilever Indonesia Tbk, PT Astra International Tbk, PT Bank Mandiri (Persero) Tbk, PT Hanjaya Mandala Sampoerna Tbk. The daily data collected from the Jakarta Composite Index via the Yahoo Finance website. IHSB data were collected from December 25, 2014, to December 25, 2019. A data frame with 1193 observations on the following nine variables, such as

Year: The year that the data was recorded

X1: return value at t-1

X2: return value at t-2

X3: return value at t-3

X4: return value at t-4

X5: return value at t-5

Volume: Volume of shares traded (number of daily shares traded in billions)

Today: return value at today (t)

Direction: the label Down and Up, which indicates the market had a positive or negative return on a given day.

The observation data taken on Yahoo Finance consists of 7 (seven) variables that are, date, Open, High, Low, Close, Adj Close, and Volume. In preprocessing data, we will process the values in the observation data into a return value. The return value is calculated by using the close price value in the observed data. For example, data observation can be shown in table.1 below. The results of return value based on observation data table.1

$$Today = \frac{6.236,69 - 6.219,44}{6.219,44} * 100 = 0,277$$

$$X1 = \frac{6.219,44 - 6.334,84}{6.334,84} * 100 = -1,82$$

$$X2 = \frac{6.334,84 - 6.342,17}{6.342,17} * 100 = -0,12$$

$$X3 = \frac{6.342,17 - 6.381,95}{6.381,95} * 100 = -0,63$$

$$X4 = \frac{6.381,95 - 6.336,67}{6.336,67} * 100 = 0,715$$

$$X5 = \frac{6.336,67 - 6.326,21}{6.336,67} * 100 = 0,165$$

The return value of today becomes the label Direction, which indicating the market had a positive (Up) or negative (Down) return on a given day. The distribution of JKSE return value for 2015-2019 can be seen in figure.2

We compared the performance of the BoosRT method with the Logistic Regression, Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), k-Nearest Neighbors (kNN), and Naive Bayes methods.

The performance of the proposed method is validated by comparing it based on accuracy, recall, precision, and specificity. The analysis is done by taking different proportions of tests and training. In conducting training, data from 2015 to 2018 were used, namely as many as 951 data. Data testing used data in 2019, as many as 242 data.

Table 1 shows the results of the accuracy, recall, precision, and specificity of the classification method, where the BoosRT methods as ensemble methods. While the other method is the classical method.

In the first experiment, it appears that the logistic regression method has an accuracy of 54.32% or has $100 - 54,32 = 45,68\%$ training error rate. By removing variables that are not very influential in the formation of logistic regression models, an effective model is obtained. The accuracy level of 54.32% is the highest level of accuracy obtained by using very influential variables in the direction prediction.

1
2
3
4
5
6
7 The use of the Linear Discriminant Analysis (LDA) method does not
8
9 differ greatly in its accuracy with the Logistic Regression method. The LDA
10
11 method only has an accuracy rate of 54.48%. Likewise, with the QDA method,
12
13 which only has an accuracy rate of 1% better than the logistic regression and LDA
14
15 methods.
16

17
18 For kNN method, models that created are used in the data testing level
19
20 accuracy achieved were 65 % with $k = 2$. This model better than LDA and QDA.
21

22
23 Training data using the Naïve Bayes method turns out to have a better
24
25 level of accuracy when applied to test data, compared to the previous method. The
26
27 accuracy of this method is 77 % or 12 % better for training error.
28

29
30 Moreover, when using the ensemble method, namely BoosRT it has a
31
32 perfect accuracy rate of 82 %. This is more impressive for stock market data,
33
34 which is known to be very difficult to model accurately. This shows that the
35
36 model assumed by BoosRT can capture the relationship between variables more
37
38 accurately than the previously used method.
39

40
41 Figure 2 shows the level of accuracy, recall, precision, and specificity of
42
43 each method used in the JCI daily data between 2015 and 2019.
44

45
46 This ensemble method is also tested on other data sets, which are

- 47
48 a) PT Kimia Farma (Persero) Tbk, produces and sells medicines and their
49
50 derivative products.
51
52 b) PT Bank Negara Indonesia (Persero) Tbk, provides various banking products
53
54 and services in Indonesia,
55
56
57
58
59
60

- 1
2
3
4
5
6
7 c) PT Perusahaan Perkebunan London Sumatra Indonesia Tbk is the plantation
8
9 business in Indonesia.
10
11 d) PT Unilever Indonesia Tbk, is engaged in manufacturing, marketing, and
12
13 distribution of consumer goods
14
15
16 e) PT Astra International Tbk, together with its subsidiaries, engages in the
17
18 automotive, financial services in Indonesia.
19
20 f) PT Indofarma (Persero) Tbk, manufactures and markets generic, over the
21
22 counter, and branded generic drugs in Indonesia.
23
24
25 g) PT Hanjaya Mandala Sampoerna Tbk, manufactures and trades in cigarettes
26
27 in Indonesia.
28
29
30 h) PT Telekomunikasi Indonesia Tbk provides telecommunications, information,
31
32 and media and edutainment services worldwide.
33

34 We used of the stock market in these companies, because the shares that
35
36 exist in these companies greatly affect changes in shares in JCI. This data set is
37
38 taken for a model that has been formed.. Table 3 shows the results of the
39
40 accuracy, recall, precision, and specificity of the classification method BoosRT.
41
42

43 The results of the classification using ensemble models that have been
44
45 obtained on 8 (eight) data sets show an accuracy level of 76% to 82%. The
46
47 average level of accuracy of the data set tested using the BoosRT model is 80.4%.
48
49 The average recall from the model is 79.2%, which means that the success rate of
50
51 the model in finding back a piece of information has better performance. While
52
53 the average precision model is 77.6%, which means that the level of prediction
54
55 matches the data with the predictions given by the model also has good
56
57
58
59
60

1
2
3
4
5
6
7 performance. From the measurement of accuracy, recall, and precision, the
8
9 BoosRT model used has a better average than the logistic Regression, LDA,
10
11 QDA, kNN, or Naïve Bayes methods. This shows that the ensemble model using
12
13 the BoosRT model is stable enough to be used in different data sets. The accuracy
14
15 of the BoosRT model in the test data set can be seen in figure.4.
16
17
18
19

20 **4. Conclusions**

21
22
23 Based on the experimental results, it can be seen that the logistic
24
25 regression method, LDA and QDA only have an accuracy level of 54% to 55%.
26
27 For the kNN and Naïve Bayes methods, each has an accuracy level of 65% and 77
28
29 %. The kNN and Naïve Bayes methods are relatively better than the previous
30
31 three methods.
32
33

34
35 The ensemble method turns out to have a perfect level of accuracy based
36
37 on the experiments that have been done, which is 82%. This is more impressive
38
39 for stock market data, which is known to be very difficult to model accurately.
40
41 This shows that the model assumed by BoosRT can capture the relationship
42
43 between variables more accurately than the method used before.
44
45

46
47 The classification using ensemble models that have been obtained on the
48
49 other data set. The model has shown an accuracy level of 76% to 82%. The
50
51 average level of accuracy of the data set tested using the BoosRT model is 80.4%.
52
53 The average recall from the model is 79.2%, which means that the success rate of
54
55 the model in finding back a piece of information has better performance. While
56
57 the average precision model is 77.6%, which means that the level of prediction
58
59
60

1
2
3
4
5
6
7 matches the data with the predictions given by the model also has good
8
9 performance. This shows that the ensemble model using the BossRT model is
10
11 stable enough to be used in different data sets.
12
13

14 15 16 **References**

- 17
18 Anbalagan, T., & Maheswari, S. U. (2015). Classification and prediction of stock
19
20 market index based on Fuzzy Metagraph. *Procedia Computer Science*,
21
22 47(C), 214–221. <https://doi.org/10.1016/j.procs.2015.03.200>
23
24
25 Bruni, R. (2017). Stock Market Index Data and indicators for Day Trading as a
26
27 Binary Classification problem. *Data in Brief*, 10, 569–575.
28
29 <https://doi.org/10.1016/j.dib.2016.12.044>
30
31
32 Fanita, F., & Rustam, Z. (2018). Predicting the Jakarta composite index price
33
34 using ANFIS and classifying prediction result based on relative error by
35
36 fuzzy Kernel C-Means. *AIP Conference Proceedings*, 2023.
37
38 <https://doi.org/10.1063/1.5064203>
39
40
41 Gareth James, Daniela Witten, Trevor Hastie, R. T. (2013). An Introduction to
42
43 Statistical Learning with Applications in R. In G. Casella S. Fienberg I.
44
45 Olkin (Ed.), *Springer Texts in Statistic*.
46
47 <https://doi.org/10.2174/0929867003374372>
48
49
50 Gündüz, H., Çataltepe, Z., & Yaslan, Y. (2017). Stock daily return prediction
51
52 using expanded features and feature selection. *Turkish Journal of Electrical*
53
54 *Engineering and Computer Sciences*, 25(6), 4829–4840.
55
56
57 <https://doi.org/10.3906/elk-1704-256>
58
59
60

- 1
2
3
4
5
6
7 Jyothirmayee, S., Dilip Kumar, V., Someswara Rao, C., & Shiva Shankar, R.
8
9 (2019). Predicting stock exchange using supervised learning algorithms.
10
11 *International Journal of Innovative Technology and Exploring Engineering*,
12
13 9(1), 4081–4090. <https://doi.org/10.35940/ijitee.A4144.119119>
14
15
16 Kaur, I., & Kalra, M. (2019). *Ensemble Classification Method for Credit Card*
17
18 *Fraud Detection*. (3), 423–427. <https://doi.org/10.35940/ijrte.C4213.098319>
19
20
21 Kim, H., & Han, S. T. (2016). The Enhanced Classification for the Stock Index
22
23 Prediction. *Procedia Computer Science*, 91(Itqm), 284–286.
24
25 <https://doi.org/10.1016/j.procs.2016.07.077>
26
27
28 Lachiheb, O., & Gouider, M. S. (2018). A hierarchical Deep neural network
29
30 design for stock returns prediction. *Procedia Computer Science*, 126, 264–
31
32 272. <https://doi.org/10.1016/j.procs.2018.07.260>
33
34
35 Miralles-Quirós, M. del M., Miralles-Quirós, J. L., & Oliveira, C. (2017). The role
36
37 of liquidity in asset pricing: the special case of the Portuguese Stock Market.
38
39 *Journal of Economics, Finance and Administrative Science*, 22(43), 191–
40
41 206. <https://doi.org/10.1108/JEFAS-12-2016-0001>
42
43
44 Nayak, A., Pai, M. M. M., & Pai, R. M. (2016). Prediction Models for Indian
45
46 Stock Market. *Procedia Computer Science*, 89, 441–449.
47
48 <https://doi.org/10.1016/j.procs.2016.06.096>
49
50
51 Park, M., Lee, M. L., & Lee, J. (2019). Predicting stock market indices using
52
53 classification tools. *Asian Economic and Financial Review*, 9(2), 243–256.
54
55 <https://doi.org/10.18488/journal.aefr.2019.92.243.256>
56
57
58 PT Bursa Efek Indonesia. (n.d.). Retrieved December 25, 2019, from
59
60

- 1
2
3
4
5
6
7 <https://www.idx.co.id/produk/saham/>
8
9 Rosadi, D. (2017). *Analisis statistika dengan R*. Retrieved from
10 <http://opac.perpusnas.go.id/DetailOpac.aspx?id=985555#>
11
12
13 Sedighi, M., Mohammadi, M., Fard, S. F., & Sedighi, M. (2019). The Nexus
14 between stock returns of oil companies and oil price fluctuations after heavy
15 oil upgrading: Toward theoretical progress. *Economies*, 7(3), 1–17.
16
17 <https://doi.org/10.3390/economies7030071>
18
19
20
21
22 Selakov, A., Cvijetinović, D., Milović, L., Mellon, S., & Bekut, D. (2014). Hybrid
23 PSO-SVM method for short-term load forecasting during periods with
24 significant temperature variations in city of Burbank. *Applied Soft*
25 *Computing Journal*, 16, 80–88. <https://doi.org/10.1016/j.asoc.2013.12.001>
26
27
28
29
30
31
32 Shah, S. Q., Ismail, I., & Bin Shahrin, A. R. (2019). Stock market manipulation:
33 A comparative analysis of East Asian emerging and developed financial
34 markets. *Management Science Letters*, 9(1), 183–192.
35
36 <https://doi.org/10.5267/j.msl.2018.10.006>
37
38
39
40
41
42 Shin, Y. (2015). *Application of Boosting Regression Trees to Preliminary Cost*.
43 2015.
44
45
46 Tanuwijaya, J., & Hansun, S. (2019). LQ45 stock index prediction using k-nearest
47 neighbors regression. *International Journal of Recent Technology and*
48 *Engineering*, 8(3), 2388–2391. <https://doi.org/10.35940/ijrte.C4663.098319>
49
50
51
52 Trevor Hastie, Robert Tibshirani, J. F. (2017). The Elements of Statistical
53 Learning: Data Mining, Inference, and Prediction, Second Edition. In
54 *SpringerPlus*. https://doi.org/10.1111/j.1751-5823.2009.00095_18.x
55
56
57
58
59
60

- 1
2
3
4
5
6
7 Wu, W., Chen, J., Xu, L., He, Q., & Tindall, M. L. (2019). A statistical learning
8
9 approach for stock selection in the Chinese stock market. *Financial*
10
11 *Innovation*, 5(1), 1–18. <https://doi.org/10.1186/s40854-019-0137-1>
12
13
14 Zhong, X., & Enke, D. (2019). Predicting the daily return direction of the stock
15
16 market using hybrid machine learning algorithms. *Financial Innovation*,
17
18 5(1). <https://doi.org/10.1186/s40854-019-0138-0>
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Review Only

1
2
3 **List of Figure.**
4

5 Figure 1. Steps were taken in the research
6

7 Figure 2. Return Value Today of JKSE 2015-2019
8

9 Figure 3. Analysis of Classification Methods with JKSE Data Set
10
11

12 Figure 4. Model accuracy in other data set
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

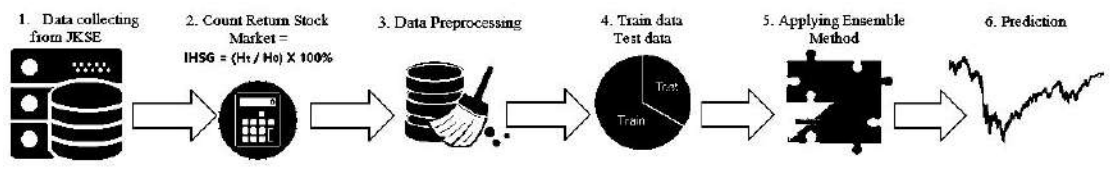


Figure 1. Steps were taken in the research

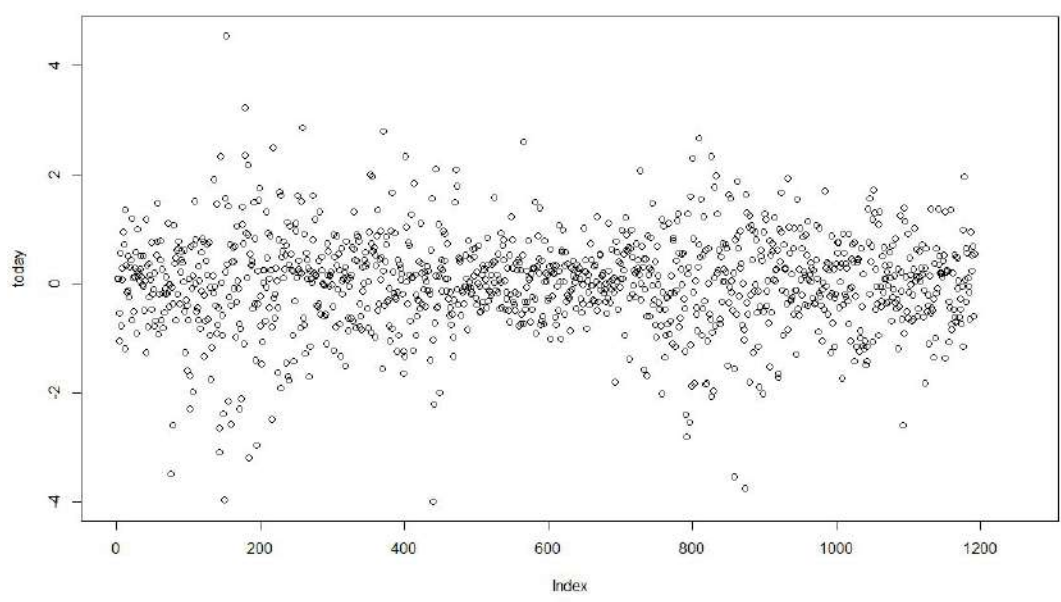


Figure 2. Return Value Today of JKSE 2015-2019

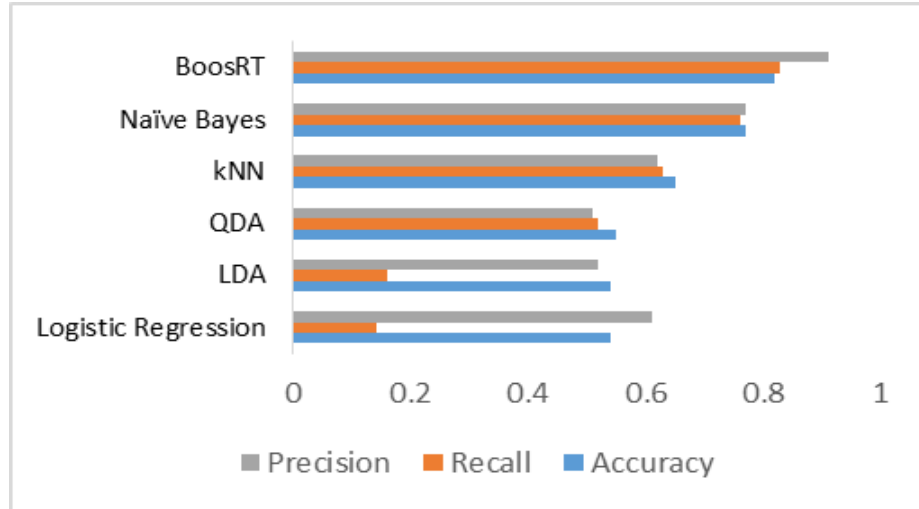


Figure 3. Analysis of Classification Methods with JKSE Data Set

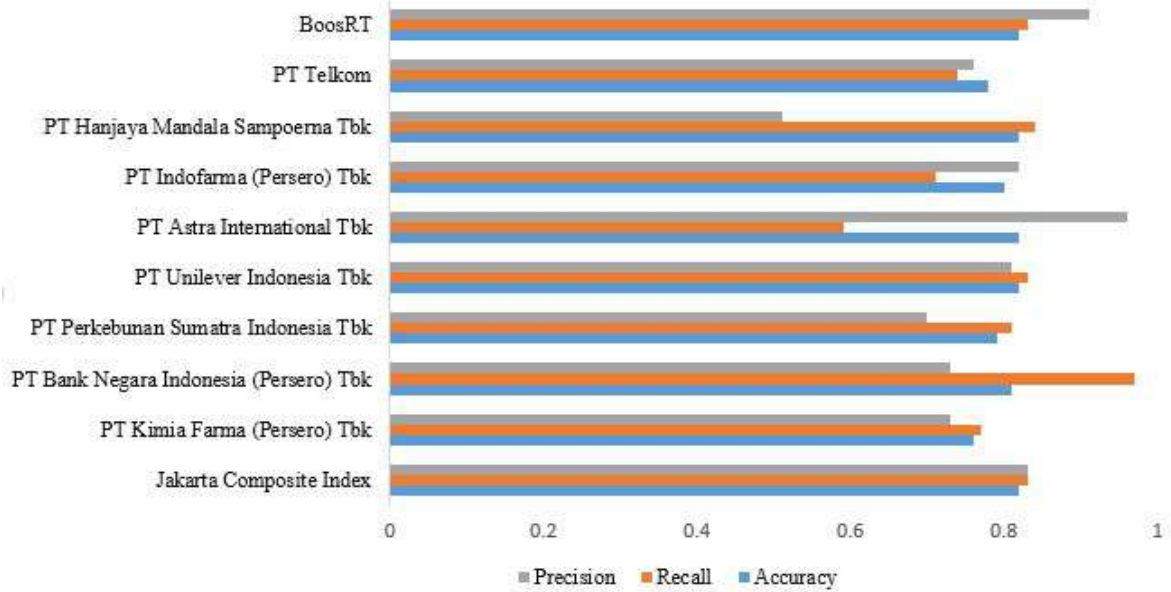


Figure 4. Model accuracy in other data set

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

List of Table.

- Table 1. Data Observation
- Table 2. Evaluation of Classification Methods
- Table 3. Ensemble Classification With the other Data Set

For Review Only

Table 1. Data Observation

Date	Open	High	Low	Close
Sep 05, 2019	6,294.28	6,307.35	6,281.95	6,306.80
Sep 06, 2019	6,329.41	6,336.91	6,305.02	6,308.95
Sep 09, 2019	6,328.28	6,333.90	6,306.74	6,326.21
Sep 10, 2019	6,331.73	6,342.01	6,311.34	6,336.67
Sep 11, 2019	6,334.59	6,381.95	6,328.69	6,381.95
Sep 12, 2019	6,399.00	6,414.48	6,337.52	6,342.17
Sep 13, 2019	6,369.42	6,375.80	6,318.92	6,334.84
Sep 16, 2019	6,262.29	6,266.14	6,193.51	6,219.44
Sep 17, 2019	6,215.24	6,240.35	6,205.30	6,236.69

Table 2. Evaluation of Classification Methods

Metode	Accuracy	Recall	Precision
Logistic Regression	0.54	0.14	0.61
Linear Discriminant Analysis (LDA)	0.54	0.16	0.52
Quadratic Discriminant Analysis (QDA)	0.55	0.52	0.51
kNN	0.65	0.63	0.62
Naïve Bayes	0.77	0.76	0.77
BoosRT	0.82	0.83	0.71

Table 3. Ensemble Classification With the other Data Set

Data Set	Accuracy	Recall	Precision
PT Kimia Farma (Persero) Tbk	0,76	0,77	0,73
PT Bank Negara Indonesia (Persero) Tbk	0,81	0,97	0,73
PT Perusahaan Perkebunan London Sumatra Indonesia Tbk	0,79	0,81	0,70
PT Unilever Indonesia Tbk	0,82	0,83	0,81
PT Astra International Tbk	0,82	0,59	0,96
PT Indofarma (Persero) Tbk	0,8	0,71	0,82
PT Hanjaya Mandala Sampoerna Tbk	0,82	0,84	0,71
PT Telkom	0,78	0,74	0,76

- 2. Review artikel (6 September 2020)**
 - Tanggapan Review**
 - Revisi Paper**



Meilany Nonsi Tentua <meilany@upy.ac.id>

Songklanakarín Journal of Science and Technology - Decision on Manuscript ID SJST-2020-0281

1 message

Songklanakarín Journal of Science and Technology <onbehalf@manuscriptcentral.com> Sun, Sep 6, 2020 at 5:46 PM
Reply-To: proespichaya.k@psu.ac.th
To: meilany@upy.ac.id

06-Sep-2020

Dear Mrs. Tentua:

Manuscript ID SJST-2020-0281 entitled "Ensemble Classification Method for Daily Return Stock Market" which you submitted to the Songklanakarín Journal of Science and Technology, **has been reviewed**. The comments of the reviewer(s) are included at the bottom of this letter.

The reviewer(s) have recommended publication, but also suggest some revisions to your manuscript. Therefore, I invite you to respond to the reviewer(s)' comments and revise your manuscript.

To revise your manuscript, log into <https://mc.manuscriptcentral.com/sjst> and enter your Author Center, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision.

You may also click the below link to start the revision process (or continue the process if you have already started your revision) for your manuscript. If you use the below link you will not be required to login to ScholarOne Manuscripts.

*** PLEASE NOTE: This is a two-step process. After clicking on the link, you will be directed to a webpage to confirm. ***

https://mc.manuscriptcentral.com/sjst?URL_MASK=18795cdb9e52441ab54199d09828bad1

You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript using a word processing program and save it on your computer. Please also highlight the changes to your manuscript within the document in MS Word or by using bold or colored text.

Once the revised manuscript is prepared, you can upload it and submit it through your Author Center.

When submitting your revised manuscript, you will be able to respond to the comments made by the reviewer(s) in the space provided. You can use this space to document any changes you make to the original manuscript. In order to expedite the processing of the revised manuscript, please be as specific as possible in your response to the reviewer(s).

IMPORTANT: Your original files are available to you when you upload your revised manuscript. Please delete any redundant files before completing the submission.

Because we are trying to facilitate timely publication of manuscripts submitted to the Songklanakarín Journal of Science and Technology, your revised manuscript should be submitted by 05-Nov-2020. If it is not possible for you to submit your revision by this date, we may have to consider your paper as a new submission.

Once again, thank you for submitting your manuscript to the Songklanakarín Journal of Science and Technology and I look forward to receiving your revision.

Sincerely,
Assoc. Prof. Dr. Proespichaya Kanatharana
Editor in Chief, Songklanakarín Journal of Science and Technology
proespichaya.k@psu.ac.th

Reviewer(s)' Comments to Author:

Reviewer: 1**Comments to the Author**

check line 23 in page 4, rephrase the aim of the study.

check line 55 [19]

check step 5 in page 6

check line 35 page 12.

Your results are not compared with other studies

the paper should be well formatted

any policy implication?

any limitation?

check for grammatical errors.

more mathematical expressions of all the methods used.

More work needed.

Reviewer: 2**Comments to the Author****Review Comments to authors:**

This article used the Ensemble Classification Method for the prediction of stock prices.

Although the authors made huge effort on this paper, while the authors can still improve this research work.

I share the following few observations, which should be improved before the paper is accepted.

1. In introduction section, the study should mention what has already discovered. What is the research gap, and how this paper covers that gap?
2. Introduction section should include contribution of paper.
3. The authors should try to reduce the typos and grammar errors in the paper.
4. The methodology section is explained well, but the author has to provide equations in proper way.
5. In my view, discussion should be a sub-section (3.1) and need more details about it.
6. The conclusion section is a summary of results, instead of policy implications. I recommend that the author have to write the specific (not general) implications of this study.

Associate Editor

Comments to the Author:

(There are no comments.)

 **Review comments.pdf**
284K



Ensemble Classification Method for Daily Return Stock Market

Journal:	<i>Songklanakarín Journal of Science and Technology</i>
Manuscript ID	SJST-2020-0281.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	26-Oct-2020
Complete List of Authors:	Tentua, Meilany; Universitas PGRI Yogyakarta, informatic Rosadi, Dedi; Universitas Gadjah Mada, Department of Mathematics
Keyword:	Ensemble Method, Stock Market Returns, Classification

SCHOLARONE™
Manuscripts

Author's Response

-----Reviewer: 1-----

check line 23 in page 4, rephrase the aim of the study.

answer : I already make the aim of study in this line, which i write in introduction like below :

In this study, we aim to predict daily stock market returns in the movement of the stock Jakarta Composite Index (JCI). We also use the most regulated daily share price, such as PT Bank Negara Indonesia (Persero) Tbk, PT Perusahaan Perkebunan London Sumatra Indonesia Tbk, PT Unilever Indonesia Tbk, PT Astra International Tbk, PT Bank Mandiri (Persero) Tbk, and PT Hanjaya Mandala Sampoerna Tbk. To predict the daily stock market, we use the Boosted Regression Tree (Boost RT) method. Besides, we also made comparisons with other classical methods such as LDA, Logistic regression, KNN, and Naïve Bayes.

check line 55 [19]

answer : I already delete

check step 5 in page 6

answer : I corected in writing

check line 35 page 12.

Your results are not compared with other studies

answer : I already compared the result with other studies which use ensemble method. I explain that in sub. secssion Discussion.

In this study, we use the return value in the past five days as a feature. This feature considers that the stock price may be affected by stock movements from a few preceding days. This feature makes it easy for a new investor to predict the value of shares. Our research's use of features is different from research conducted by (Jyothirmayee et al., 2019) using the random forest method; they use the data feature taken from Yahoo Finance. It will be a weakness of the model because the prediction will only consider the stock's movement on the previous day to lower the accuracy value.

Besides, a study by (Nayak et al., 2016) shows that the use of features and stock value also should considers the sentiment analysis of the company. It will not be easy for a new investor in deciding to buy shares if he has ti know the company's sentiment analysis.

the paper should be well formated

answer : I alredy done with proofread.

any policy implication?

answer : yes, i add the policy implication in conclusion

The policy implications in this study can use from two sides, investors and companies. For investors, predictions of the stock price can determine their investment strategy. It will

1
2
3 relate to the rate of return they expect in investing. For companies, this prediction will be
4 considered for the company to determine policies related to the company's stock price.
5 any limitation?
6

7
8 check for gramatical errors.

9 answer : yes, i already done
10

11
12 more mathematical expressions of all the methods used.

13 answer : ya, I already add mathematical expressions in methodology for evaluate the model.
14
15
16
17

18 **----- Reviewer: 2-----**
19

20 I share the following few observations, which should be improved before the paper is
21 accepted.
22

23 1. In introduction section, the study should mention what has already discovered. What is
24 the research gap, and how this paper covers that gap?

25 answer : yes, I already add in introduction the gap and how this papers covers the gap, like i
26 write below:
27

28 However, using a relatively small amount of data in stock market data, models built with
29 Linear Discriminant Analysis, Logistic Regression, K-Nearest Neighbors, Naive Bayes, and
30 Tress will easily over-fit. Their variance also tends to increase. Thus it has an impact on the
31 low accuracy achieved.
32

33 In making predictions for the daily stock return market, an appropriate method is needed to
34 produce an accurate model. Classification using the Random Forests method in the study of
35 (Jyothirmayee et al., 2019) and the boosting process in the research of (Nayak, Pai, & Pai,
36 2016) have provided a slightly better level of accuracy compared to previous classical
37 methods. These methods are also called ensemble methods. They can improve the accuracy
38 of prediction models formed. However, in their research, they used analytical sentiment
39 variables that would be difficult for new investors.
40
41
42
43

44 2. Introduction section should include contribution of paper.

45 answer : yes, i already write the contribution of paper in introduction,
46

47 The contributions of our study is the use of simple features, only by calculating the stock
48 closing price's return value from the five previous days. The classification method that we
49 use is the Boosted Regression Tree (BoostRT) method, which is an ensemble method that
50 combines the Boosting method and the Decision Tree method. This method is rarely used
51 for predicting stock prices in Indonesia. We use accuracy techniques in evaluating the
52 classification model.
53
54

55
56 3. The authors should try to reduce the typos and grammar errors in the paper.

57 answer : yes, I alredy done it, with proof read
58
59
60

1
2
3 4. The methodology section is explained well, but the author has to provide equations in
4 proper way.

5
6 answer : yes, I already done this part. with add the formula to evaluate model in
7 methodology
8

9
10 5. In my view, discussion should be a sub-section (3.1) and need more details about it.

11 answer : yes, I make sub-section : 3.1 Result and 3.2 as Discussion
12

13
14 6. The conclusion section is a summary of results, instead of policy implications. I
15 recommend that the author have to write the specific (not general) implications of this
16 study.
17

18 answer : yes, i already add this in conclusion, like I write below :

19 According to the experiments, the ensemble method has a high accuracy (82%). It is
20 significant for stock market data, which is known to be difficult to model accurately. It
21 shows that the model assumed by BoosRT can capture the relationship between variables
22 more accurately than the method used before. The classification using ensemble models
23 have been obtained on the other data set. The model has shown an accuracy level of 76% to
24 82%. It shows that the ensemble model using the BossRT model is stable to be used and
25 generalized in different data sets.
26

27
28 The policy implications in this study can use from two sides, investors and companies. For
29 investors, predictions of the stock price can determine their investment strategy. It will
30 relate to the rate of return they expect in investing. For companies, this prediction will be
31 considered for the company to determine policies related to the company's stock price.
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Original Article

Ensemble Classification Method for Daily Return Stock Market

Meilany Nonsi Tentua¹, Dedi Rosadi²

¹Department of Informatics, Sains and Technology Fakultas, Universitas PGRI Yogyakarta, Yogyakarta, 55182, Indonesia

²Universitas Gajah Mada, Yogyakarta, Indonesia, Sekip Utara Bulaksumur, Yogyakarta, 55281, Indonesia

* Corresponding author, Email address: meilany@upy.ac.id

Abstract

Stock price movements in Indonesia are measured using indices, one of which is the Composite Stock Price Index (CSPI). CSPI is a stock index that measures a combination of all shares from various sectors listed on the Indonesia Stock Exchange.

The ensemble method is to build predictive models by combining the strengths of the classical classification method. In this research, the purpose of ensemble based on Boosting for Regression appeared to enhance simple tree analysis and deals with some of the weaknesses found in uncomplicated techniques. The ensemble tree combines the prediction values of many simple trees into a single prediction value. Based on the experiments that have been carried out, the ensemble method proved to have a better accuracy rate, which amounted to 82%. The model has been assumed by the ensemble model that can get the relationship between variables to more precise than the previous model.

Keywords: Ensemble Method, Stock Market Returns, Classification.

1. Introduction

Stock is one of the most popular financial market instruments. Stock can be interpreted as a sign of capital participation of a person or party (business entity) in a company or limited liability company (“PT Bursa Efek Indonesia,” n.d.). Share price movements in Indonesia are measured using indices, one of which is the Composite Stock Price Index (CSPI). CSPI is a stock index that measures a combination of all shares from various sectors listed on the Indonesia Stock Exchange. Investors usually use parameters in the CSPI to read price developments and make a reference to the portfolio.

A stock market index is a measurement of the value of a part of the stock market. It is often used to describe market aggregate trends that will be predicted by a basic financial issue. Such stochastic value-share is very difficult to predict (Bruni, 2017). The difficulty in predicting the return of shares purchased by investors means that an investor is still guessing about the returns that investors will obtain in the future. There are four approaches used in predicting stock price movements such as technical analysis (Sedighi, Mohammadi, Fard, & Sedighi, 2019), fundamental analysis (Miralles-Quirós, Miralles-Quirós, & Oliveira, 2017), traditional time series forecasting (Park, Lee, & Lee, 2019), and statistical or machine learning methods (Kim & Han, 2016).

Statistical learning refers to a set of tools for modeling and understanding complex data sets. Statistical learning is a newly developed area in statistics and integrates with parallel developments in computer science and, in particular, machine learning (Gareth James, Daniela Witten, Trevor Hastie, 2013).

1
2
3
4
5
6
7 Predictive modeling using data-sets is one of the tasks in statistical learning. An
8 approach to predict the qualitative response is a process known as
9 classification. Classification is the attribution of labels to notes according to
10 criteria that are automatically learned from a series of labeled record training.
11
12
13
14

15
16 Some algorithms classification of classics is widely used today (Trevor
17 Hastie, Robert Tibshirani, 2017) such as Linear Discriminant Analysis,
18 Logistic Regression (Gündüz, Çataltepe, & Yaslan, 2017) (Shah, Ismail, & Bin
19 Shahrin, 2019), K-Nearest Neighbors (Tanuwijaya & Hansun, 2019), Naïve Bayes
20 (Marianto & Maruddani, 2020), Support Vector Machines (Selakov, Cvijetinović,
21 Milović, Mellon, & Bekut, 2014), and Trees (Sitorus & Tarihoran, 2018).
22
23
24
25
26
27
28

29 However, using a relatively small amount of data in stock market data, models
30 built with Linear Discriminant Analysis, Logistic Regression, K-Nearest
31 Neighbors, Naive Bayes, and Tress will easily over-fit. Their variance also tends
32 to increase. Thus it has an impact on the low accuracy achieved.
33
34
35
36
37
38

39 In making predictions for the daily stock return market, an appropriate
40 method is needed to produce an accurate model. Classification using the Random
41 Forests method in the study of (Jyothirmayee et al., 2019) and the boosting
42 process in the research of (Nayak, Pai, & Pai, 2016) have provided a slightly
43 better level of accuracy compared to previous classical methods. These methods
44 are also called ensemble methods. They can improve the accuracy of prediction
45 models formed. However, in their research, they used analytical sentiment
46 variables that would be difficult for new investors. The idea of ensemble learning
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7 is to build predictive models by combining the strengths of classical classification
8 methods (Trevor Hastie, Robert Tibshirani, 2017).
9

10
11 In this study, we aim to predict daily stock market returns in the movement
12 of the stock Jakarta Composite Index (JCI). We also use the most regulated daily
13 share price, such as PT Bank Negara Indonesia (Persero) Tbk, PT Perusahaan
14 Perkebunan London Sumatra Indonesia Tbk, PT Unilever Indonesia Tbk, PT
15 Astra International Tbk, PT Bank Mandiri (Persero) Tbk, and PT Hanjaya
16 Mandala Sampoerna Tbk. To predict the daily stock market, we use the Boosted
17 Regression Tree (Boost RT) method. Besides, we also made comparisons with
18 other classical methods such as LDA, Logistic regression, KNN, and Naïve
19 Bayes.
20
21
22
23
24
25
26
27
28
29
30
31

32 The contributions of our study is the use of simple features, only by
33 calculating the stock closing price's return value from the five previous days. The
34 classification method that we use is the Boosted Regression Tree (BoostRT)
35 method, which is an ensemble method that combines the Boosting method and the
36 Decision Tree method. This method is rarely used for predicting stock prices in
37 Indonesia. We use accuracy techniques in evaluating the classification model.
38
39
40
41
42
43
44
45
46
47

48 **2. Materials and Methods**

49
50 Several studies have been conducted in the application
51 of statistical learning in the stock market domain. The use of the bootstrap method
52 is modified with Random Forest to predict the direction of the movement of stock
53 index prices (Kim & Han, 2016). Training sets are generated by a modified
54
55
56
57
58
59
60

1
2
3
4
5
6
7 bootstrap considering the impact of response variables simultaneously and applied
8
9 in a random forest. The data used is the Korea composite stock price index with
10
11 the variable close price, open price, high price, low price, trading volume, and
12
13 training amount.
14

15
16 A hierarchical Deep Neural Network (DNN) is a method that can be
17
18 applied to predict stock market returns. A study applying this method had been
19
20 conducted to predict the stock market returns within the next five minutes
21
22 (Lachiheb & Gouider, 2018). The data used were the main index of the Tunisian
23
24 stock market (TUNINDEX) with the variables of TUNINDEX current value,
25
26 Number of up direction stocks, Number of down direction stocks, Number of
27
28 unchanged stocks, and Total stock values in Tunisian dinars. Other research that
29
30 uses DNN is a large comprehensive data analysis process to predict the daily
31
32 return direction of the S&P 500 SPDR ETF (ticker symbol: SPY) based on 60
33
34 financial and economic features (Zhong & Enke, 2019).
35
36
37

38
39 The approach of a cross-sectional statistical estimation model for selecting
40
41 stocks on the Chinese stock market allows investors to identify stocks that are
42
43 likely to perform well and to build a suitable portfolio (Wu, Chen, Xu, He, &
44
45 Tindall, 2019). The data used is sourced from the Shanghai Composite Index.
46
47

48 Predicting stock market prices using Fuzzy Metagraph (FM) based on
49
50 stock market decision making, classification, and prediction are proposed for
51
52 short-term investors from the Indian stock market (Anbalagan & Maheswari,
53
54 2015). Simple Moving Average (SMA), Exponential Moving Average (EMA),
55
56 Moving Average Convergence Divergence (MACD), and Relative Strength Index
57
58
59
60

(RSI) are some technical indicators that are used as input to train an integrated system with Fuzzy Meta graph. The data used in the Indian stock market data is the Bombay Stock Exchange (BSE).

Artificial Neural Network (ANN) is a method in machine learning that can be applied for prediction and classification. Implementation of ANN back-propagation is used to predict the closing price of the S & P 500 stock exchange. Historical data consists of five variables, such as open, high, low, close, and volume.

The Adaptive Neuro-Fuzzy Inference System (ANFIS) method is used to determine the Jakarta Composite Index (JCI) value movement. This study provides two outputs, namely prediction results and classification results based on several relative error values. The classification method used is Fuzzy Kernel C-Means (Fanita & Rustam, 2018).

This research is done to classify and predict the daily return of the stock market in Indonesia. The undertaken method in this research can be seen in Figure 1.

Step 1 : Collecting CSPI data from the Indonesia Composite Index through the Yahoo Finance web.

Step 2: Calculating the market stock daily return value from the data collected in step 1 using the formula :

$$ReturnValue = \frac{currentvalue - originalvalue}{originalvalue} * 100 \quad (1)$$

Step 3: Pre-processing data by handling the missing values and excessive data from the data-set.

Step 4: Making sure that the training data-set is larger compared to the test set. The training data-set is taken from 2015 to 2018 data, while the test data-set is taken from 2019 data.

Step 5: Applying the ensemble Boosting Regression Tree (BoosRT).

Step 6: Analyzing the results of the ensemble model performance in terms of its accuracy, memory, accuracy, and specificity.

2.1. Boosted Regression Tree (BoosRT)

Boosted Regression Tree (BoosRT) models are a combination of two techniques: decision tree algorithms and boosting methods (Shin, 2015). In these algorithms, the BoosRT approximates the function (x) as an preservative growth of the base learner :

$$f(x) = f_0(x) + \beta_1 f_1(x) + \beta_2 f_2(x) + \dots + \beta_m f_m(x) \quad (2)$$

A single study cannot predict enough using training data, so we must use the lowest residuals to improve predictive performance. BoosRT uses an iterative algorithm, so every iteration m , and a new regression tree. The formula to predict a separate constant value is:

$$h(x; \{R_{lm}\}_l^L = \sum_{i=1}^L \bar{y}_{lm}(x \in R_{lm}) \quad (3)$$

where $y_{lm} = \text{mean } x_i \in R_{lm}$ (y_{lm}) is the mean of pseudo-residuals the m^{th} iteration :

$$\tilde{y}_{lm} = - \left[\frac{\partial \Psi(y_i, F(x_i))}{\partial f(x_i)} \right]_{f(x)=F_{m-1}(x)} \quad (4)$$

After that the approximation obtained will be updated with the following equation

$$f_m(x) = f_{m-1}(x) + v \cdot Y_{lm} | (x \in R_{lm}), \quad (5)$$

Where

$$Y_{lm} = \arg \min_y \sum_{x_i \in R_{lm}} \Psi(y_i, F_{m-1}(x_i) + Y) \quad (6)$$

BoosRT algorithm for the generalized boosting of regression trees are:

a) Initialize (x) , $F_0(x) = \arg \min_Y \sum_{i=1}^N \Psi(y_i, \gamma)$.

b) For $m=1$ to M do

c) From the full training dataset, select a subset randomly,

$$\{\pi(i)\}_i^N = \text{rand_perm}\{i\}_i^N \quad (7)$$

d) Fit the base learner

$$\tilde{y}_{\pi(i)} = - \left[\frac{\partial \Psi(y_{\pi(i)}, F_{m-1}(x_{\pi(i)}))}{\partial F(x_{\pi(i)})} \right]_{F(x)=F_{m-1}(x)}, i = 1, \tilde{N} \quad (8)$$

e) Update model for the current iteration,

$$\{R_{lm}\}_i^L = \text{terminalmodetree}(\{\tilde{y}_{\pi(i)}, x_{\pi(i)}\}_i^N) \quad (9)$$

f) Choose a gradient descent step size as

$$Y_{lm} = \arg \min_y \sum_{x(i) \in R_{lm}} \Psi(y_{\pi(i)}, F_{m-1}(x_{\pi(i)}) + Y) \quad (10)$$

g) Update the estimate of (x) as

$$F_m(x) = F_{m-1}(x) + v \cdot Y_{lm} \mid (x \in R_{lm}) \quad (11)$$

2.2. Performance Evaluation

The performance of the classification model can be measured by counting the number of classes that were predicted correctly (true positive), the predicted numbers that were not included in the class and were true (true negative), and those that were wrongly predicted (false positive or false negative). The formulation of accuracy, precision, and recall is used in this study using the following formula:

$$\text{Accuracy} = \frac{tp + tn}{tp + fn + fp + tn} \quad (12)$$

$$Precision = \frac{tp + tn}{tp + fp} \quad (13)$$

$$Recall = \frac{tp}{tp + fn} \quad (14)$$

Where, tp = true positive, tn = true negative, fp = false positive, and fn = false negative.

3. Results and Discussion

3.1. Result

The daily return stock market classification using the ensemble method is implemented using R. Some of the libraries contained in R are tools for statistical analysis and can also be used to create models with classification techniques (Gareth James, Daniela Witten, Trevor Hastie, 2013),(Rosadi, 2017). The data-set used in this study is JCI, PT Bank Negara Indonesia (Persero) Tbk, PT Perusahaan Perkebunan London Sumatra Indonesia Tbk, PT Unilever Indonesia Tbk, PT Astra International Tbk, PT Bank Mandiri (Persero) Tbk, and PT Hanjaya Mandala Sampoerna Tbk. The daily data were collected from the Jakarta Composite Index via the Yahoo Finance website. IHS data were collected from December 25, 2014 to December 25, 2019. A data frame with 1193 observations on the following nine variables, such as:

Year: The year that the data was recorded

X1: return value at t-1

X2: return value at t-2

X3: return value at t-3

X4: return value at t-4

X5: return value at t-5

Volume: Volume of shares traded (number of daily shares traded in billions)

Today: return value at today (t)

Direction: the label Down and Up, indicating that the market had a positive or negative return on a given day.

The observation data taken from Yahoo Finance consists of 7 (seven) variables such as Date, Open, High, Low, Close, Adj Close, and Volume. In pre-processing data, we will process the values in the observation data into a return value. The return value is calculated by using the close price value in the observed data. The results of return value based on observation data is shown in table.1:

$$Today = \frac{6.236,69 - 6.219,44}{6.219,44} * 100 = 0,277$$

$$X1 = \frac{6.219,44 - 6.334,84}{6.334,84} * 100 = -1,82$$

$$X2 = \frac{6.334,84 - 6.342,17}{6.342,17} * 100 = -0,12$$

$$X3 = \frac{6.342,17 - 6.381,95}{6.381,95} * 100 = -0,63$$

$$X4 = \frac{6.381,95 - 6.336,67}{6.336,67} * 100 = 0,715$$

$$X5 = \frac{6.336,67 - 6.326,21}{6.336,67} * 100 = 0,165$$

1
2
3
4
5
6
7 The return value of a day becomes the label Direction, which indicating
8
9 the market had a positive (Up) or negative (Down) return on a given day. The
10
11 distribution of JKSE return value for 2015-2019 can be seen in figure.2.
12

13 We compared the performance of the BoosRT method with the Logistic
14
15 Regression, Linear Discriminant Analysis (LDA), Quadratic Discriminant
16
17 Analysis (QDA), k-Nearest Neighbors(kNN), and Naive Bayes methods.
18
19

20 The performance of the proposed method is validated by comparing it in
21
22 terms of its accuracy, recall, precision, and specificity. The analysis is done by
23
24 taking different proportions of tests and training. In conducting the training, data
25
26 from 2015 to 2018 were used, as many as 951 data. Data testing used data from
27
28 2019, as many as 242 data.
29
30

31 Table 1 shows the results of the accuracy, recall, precision, and specificity
32
33 of the classification method, where the BoosRT methods as ensemble
34
35 methods. While the other method is the classical method.
36
37

38 Figure 2 shows the level of accuracy, recall, precision, and specificity of
39
40 each method used in the JCI daily data between 2015 and 2019.
41
42

43 This ensemble method is also tested on other data sets, such as:

- 44
45 a) PT Kimia Farma (Persero) Tbk, produces and sells medicines and their
46
47 derivative products.
48
49
50 b) PT Bank Negara Indonesia (Persero) Tbk, provides various banking products
51
52 and services in Indonesia.
53
54
55 c) PT Perusahaan Perkebunan London Sumatra Indonesia Tbk, a plantation
56
57 business in Indonesia.
58
59
60

- 1
2
3
4
5
6
7 d) PT Unilever Indonesia Tbk, engages in manufacturing, marketing, and
8 distribution of consumer goods.
9
10
11 e) PT Astra International Tbk, together with its subsidiaries, engages in the
12 automotive and financial services in Indonesia.
13
14
15 f) PT Indofarma (Persero) Tbk, manufactures and markets generic, over the
16 counter, and branded generic drugs in Indonesia.
17
18
19
20 g) PT Hanjaya Mandala Sampoerna Tbk, manufactures and trades cigarettes in
21 Indonesia.
22
23
24
25 h) PT Telekomunikasi Indonesia Tbk, provides telecommunications,
26 information, media, and edutainment services worldwide.
27
28
29

30 We used the stock market data from these companies because their shares
31 greatly affect JCI. This data set is taken from a model that has been formed. Table
32 3 shows the results of the accuracy, recall, precision, and specificity of the
33 BoosRT classification method.
34
35
36
37

38 3.2. Discussion

39
40
41 In the first experiment we used IHSG data, it appeared that the logistic
42 regression method had an accuracy of 54.32% or had $100 - 54,32 = 45,68\%$
43 training error rate. By removing variables that were not very influential in the
44 formation of logistic regression models, an effective model was obtained. The
45 accuracy level of 54.32% was the highest level of accuracy obtained by using very
46 influential variables in the direction prediction.
47
48
49
50
51
52
53

54 The use of the Linear Discriminant Analysis (LDA) method did not
55 significantly differ its accuracy with the Logistic Regression method. The LDA
56
57
58
59
60

1
2
3
4
5
6
7 method had an accuracy rate of 54.48%. Likewise, with the QDA method, it only
8
9 had an accuracy rate of 1% better than the logistic regression and LDA methods.

10
11 For kNN method, the models created are used in the data testing. The level
12
13 of accuracy achieved were 65 % with $k = 2$. This model was better than LDA and
14
15 QDA. In using this method, we tried k values from 1 to 15. When $k = 2$ the
16
17 experiment achieved the highest accuracy values.
18
19

20
21 Training data using the Naïve Bayes method turned out to have a better
22
23 level of accuracy when applied to test data, compared to the previous method. The
24
25 accuracy of this method was 77 % or 12 % better for training error.
26
27

28
29 Moreover, by using the ensemble method called BoosRT, it had an
30
31 accuracy rate of 82 %. This is significant for stock market data, which is known to
32
33 be very difficult to be predicted accurately. It shows that the model assumed by
34
35 BoosRT can capture the relationship between variables more accurately than the
36
37 previously used method.
38

39
40 In the second experiment we used the BoosRT method on the return value
41
42 of eight companies, such as PT Bank Negara Indonesia (Persero) Tbk, PT
43
44 Perusahaan Perkebunan London Sumatra Indonesia Tbk, PT Unilever Indonesia
45
46 Tbk, PT Astra International Tbk, PT Bank Mandiri (Persero) Tbk, and T Hanjaya
47
48 Mandala Sampoerna Tbk. The results of the classification using ensemble models
49
50 that have been obtained on 8 (eight) data sets showed an accuracy level of 76% to
51
52 82%. The average level of accuracy of the data set tested using the BoosRT model
53
54 was 80.4%. The average recall from the model was 79.2%, which means that the
55
56 success rate of the model in finding information had a better performance. The
57
58
59
60

1
2
3
4
5
6
7 average precision model of 77.6% means that the level of prediction matches the
8 data. It also meand that the model has a good performance. From the
9 measurement of accuracy, recall, and precision, the BoosRT model used has a
10 better average than the logistic Regression, LDA, QDA, kNN, or Naïve Bayes
11 methods. This shows that the ensemble model using the BossRT model is stable to
12 be used in different data sets. The accuracy of the BoosRT model in the test data
13 set can be seen in figure.4.
14
15
16
17
18
19
20
21
22

23 In this study, we use the return value in the past five days as a feature. This
24 feature considers that the stock price may be affected by stock movements from a
25 few preceding days. This feature makes it easy for a new investor to predict the
26 value of shares. Our research's use of features is different from research conducted
27 by (Jyothirmayee et al., 2019) using the random forest method; they use the data
28 feature taken from Yahoo Finance. It will be a weakness of the model because the
29 prediction will only consider the stock's movement on the previous day to lower
30 the accuracy value.
31
32
33
34
35
36
37
38
39
40

41 Besides, a study by (Nayak et al., 2016) shows that the use of features and
42 stock value also should considers the sentiment analysis of the company. It will
43 not be easy for a new investor in deciding to buy shares if he has ti know the
44 company's sentiment analysis.
45
46
47
48
49

50 The limitation of this study is that it only uses five return value variables
51 and does not pay attention to the socio-political and security conditions that occur
52 in Indonesia which also affect the stock market performance.
53
54
55
56
57
58
59
60

5. Conclusions

According to the experiments, the ensemble method has a high accuracy (82%). It is significant for stock market data, which is known to be difficult to model accurately. It shows that the model assumed by BoosRT can capture the relationship between variables more accurately than the method used before. The classification using ensemble models have been obtained on the other data set. The model has shown an accuracy level of 76% to 82%. It shows that the ensemble model using the BossRT model is stable to be used and generalized in different data sets.

The policy implications in this study can use from two sides, investors and companies. For investors, predictions of the stock price can determine their investment strategy. It will relate to the rate of return they expect in investing. For companies, this prediction will be considered for the company to determine policies related to the company's stock price.

References

- Anbalagan, T., & Maheswari, S. U. (2015). Classification and prediction of stock market index based on Fuzzy Metagraph. *Procedia Computer Science*, 47(C), 214–221. <https://doi.org/10.1016/j.procs.2015.03.200>
- Bruni, R. (2017). Stock Market Index Data and indicators for Day Trading as a Binary Classification problem. *Data in Brief*, 10, 569–575. <https://doi.org/10.1016/j.dib.2016.12.044>
- Fanita, F., & Rustam, Z. (2018). Predicting the Jakarta composite index price

1
2
3
4
5
6
7 using ANFIS and classifying prediction result based on relative error by
8 fuzzy Kernel C-Means. *AIP Conference Proceedings*, 2023.
9
10 <https://doi.org/10.1063/1.5064203>
11
12

13
14 Gareth James, Daniela Witten, Trevor Hastie, R. T. (2013). An Introduction to
15 Statistical Learning with Applications in R. In G. Casella S. Fienberg I.
16 Olkin (Ed.), *Springer Texts in Statistic*.
17
18 <https://doi.org/10.2174/0929867003374372>
19
20
21

22
23 Gündüz, H., Çataltepe, Z., & Yaslan, Y. (2017). Stock daily return prediction
24 using expanded features and feature selection. *Turkish Journal of Electrical*
25 *Engineering and Computer Sciences*, 25(6), 4829–4840.
26
27 <https://doi.org/10.3906/elk-1704-256>
28
29
30

31
32 Jyothirmayee, S., Dilip Kumar, V., Someswara Rao, C., & Shiva Shankar, R.
33 (2019). Predicting stock exchange using supervised learning algorithms.
34 *International Journal of Innovative Technology and Exploring Engineering*,
35 9(1), 4081–4090. <https://doi.org/10.35940/ijitee.A4144.119119>
36
37
38
39

40
41 Kim, H., & Han, S. T. (2016). The Enhanced Classification for the Stock Index
42 Prediction. *Procedia Computer Science*, 91(Itqm), 284–286.
43
44 <https://doi.org/10.1016/j.procs.2016.07.077>
45
46
47

48
49 Lachiheb, O., & Gouider, M. S. (2018). A hierarchical Deep neural network
50 design for stock returns prediction. *Procedia Computer Science*, 126, 264–
51 272. <https://doi.org/10.1016/j.procS.2018.07.260>
52
53
54

55
56 Marianto, F. Y., & Maruddani, D. A. I. (2020). *Perbandingan Metode Naive*
57 *Bayes dan Bayesian Regularization Neural Network (BRNN) untuk*
58
59
60

1
2
3
4
5
6
7 *Klasifikasi Sinyal Palsu pada Indikator Stochastic Oscillator*. 9, 16–25.

8
9 Miralles-Quirós, M. del M., Miralles-Quirós, J. L., & Oliveira, C. (2017). The role
10 of liquidity in asset pricing: the special case of the Portuguese Stock Market.
11 *Journal of Economics, Finance and Administrative Science*, 22(43), 191–
12 206. <https://doi.org/10.1108/JEFAS-12-2016-0001>

13
14
15
16
17
18 Nayak, A., Pai, M. M. M., & Pai, R. M. (2016). Prediction Models for Indian
19 Stock Market. *Procedia Computer Science*, 89, 441–449.
20 <https://doi.org/10.1016/j.procs.2016.06.096>

21
22
23
24
25 Park, M., Lee, M. L., & Lee, J. (2019). Predicting stock market indices using
26 classification tools. *Asian Economic and Financial Review*, 9(2), 243–256.
27 <https://doi.org/10.18488/journal.aefr.2019.92.243.256>

28
29
30
31
32 PT Bursa Efek Indonesia. (n.d.). Retrieved December 25, 2019, from
33 <https://www.idx.co.id/produk/saham/>

34
35
36
37 Rosadi, D. (2017). *Analisis statistika dengan R*. Retrieved from
38 <http://opac.perpusnas.go.id/DetailOpac.aspx?id=985555#>

39
40
41 Sedighi, M., Mohammadi, M., Fard, S. F., & Sedighi, M. (2019). The Nexus
42 between stock returns of oil companies and oil price fluctuations after heavy
43 oil upgrading: Toward theoretical progress. *Economies*, 7(3), 1–17.
44 <https://doi.org/10.3390/economies7030071>

45
46
47
48
49
50 Selakov, A., Cvijetinović, D., Milović, L., Mellon, S., & Bekut, D. (2014). Hybrid
51 PSO-SVM method for short-term load forecasting during periods with
52 significant temperature variations in city of Burbank. *Applied Soft*
53 *Computing Journal*, 16, 80–88. <https://doi.org/10.1016/j.asoc.2013.12.001>

- 1
2
3
4
5
6
7 Shah, S. Q., Ismail, I., & Bin Shahrin, A. R. (2019). Stock market manipulation:
8
9 A comparative analysis of East Asian emerging and developed financial
10
11 markets. *Management Science Letters*, 9(1), 183–192.
12
13 <https://doi.org/10.5267/j.msl.2018.10.006>
14
15
16 Shin, Y. (2015). *Application of Boosting Regression Trees to Preliminary Cost*.
17
18 2015.
19
20 Sitorus, H., & Tarihoran, Y. (2018). Analisis Harga Saham PT Astra Internasional
21
22 Tbk Menggunakan Data Dari Bursa Efek Indonesia dalam Jangka Waktu
23
24 Pendek Menggunakan Metode Naïve Bayes dan Decision Tree-J48. *TeIKa*,
25
26 8(1), 21–33. <https://doi.org/10.36342/teika.v8i1.2239>
27
28
29 Tanuwijaya, J., & Hansun, S. (2019). LQ45 stock index prediction using k-nearest
30
31 neighbors regression. *International Journal of Recent Technology and*
32
33 *Engineering*, 8(3), 2388–2391. <https://doi.org/10.35940/ijrte.C4663.098319>
34
35
36 Trevor Hastie, Robert Tibshirani, J. F. (2017). The Elements of Statistical
37
38 Learning: Data Mining, Inference, and Prediction, Second Edition. In
39
40 *SpringerPlus*. https://doi.org/10.1111/j.1751-5823.2009.00095_18.x
41
42
43 Wu, W., Chen, J., Xu, L., He, Q., & Tindall, M. L. (2019). A statistical learning
44
45 approach for stock selection in the Chinese stock market. *Financial*
46
47 *Innovation*, 5(1), 1–18. <https://doi.org/10.1186/s40854-019-0137-1>
48
49
50 Zhong, X., & Enke, D. (2019). Predicting the daily return direction of the stock
51
52 market using hybrid machine learning algorithms. *Financial Innovation*,
53
54 5(1). <https://doi.org/10.1186/s40854-019-0138-0>
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

List of Table.

- Table 1. Data Observation
- Table 2. Evaluation of Classification Methods
- Table 3. Ensemble Classification With the other Data Set

For Review Only

Table 1. Data Observation

Date	Open	High	Low	Close
Sep 05, 2019	6,294.28	6,307.35	6,281.95	6,306.80
Sep 06, 2019	6,329.41	6,336.91	6,305.02	6,308.95
Sep 09, 2019	6,328.28	6,333.90	6,306.74	6,326.21
Sep 10, 2019	6,331.73	6,342.01	6,311.34	6,336.67
Sep 11, 2019	6,334.59	6,381.95	6,328.69	6,381.95
Sep 12, 2019	6,399.00	6,414.48	6,337.52	6,342.17
Sep 13, 2019	6,369.42	6,375.80	6,318.92	6,334.84
Sep 16, 2019	6,262.29	6,266.14	6,193.51	6,219.44
Sep 17, 2019	6,215.24	6,240.35	6,205.30	6,236.69

Table 2. Evaluation of Classification Methods

Metode	Accuracy	Recall	Precision
Logistic Regression	0.54	0.14	0.61
Linear Discriminant Analysis (LDA)	0.54	0.16	0.52
Quadratic Discriminant Analysis (QDA)	0.55	0.52	0.51
kNN	0.65	0.63	0.62
Naïve Bayes	0.77	0.76	0.77
BoosRT	0.82	0.83	0.71

Table 3. Ensemble Classification With the other Data Set

Data Set	Accuracy	Recall	Precision
PT Kimia Farma (Persero) Tbk	0,76	0,77	0,73
PT Bank Negara Indonesia (Persero) Tbk	0,81	0,97	0,73
PT Perusahaan Perkebunan London Sumatra Indonesia Tbk	0,79	0,81	0,70
PT Unilever Indonesia Tbk	0,82	0,83	0,81
PT Astra International Tbk	0,82	0,59	0,96
PT Indofarma (Persero) Tbk	0,8	0,71	0,82
PT Hanjaya Mandala Sampoerna Tbk	0,82	0,84	0,71
PT Telkom	0,78	0,74	0,76

List of Figure.

Figure 1. Steps were taken in the research

Figure 2. Return Value Today of JKSE 2015-2019

Figure 3. Analysis of Classification Methods with JKSE Data Set

Figure 4. Model accuracy in other data set

For Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

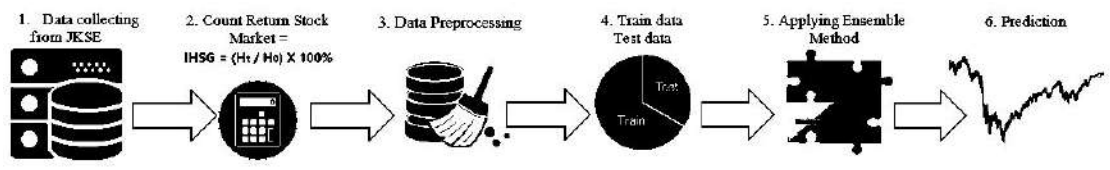


Figure 1. Steps were taken in the research

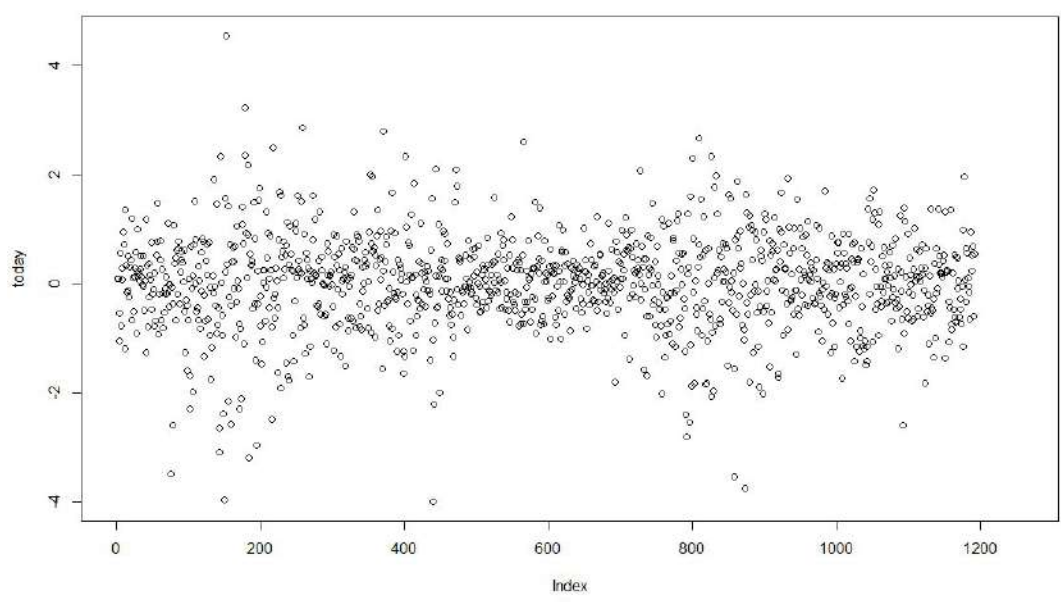


Figure 2. Return Value Today of JKSE 2015-2019

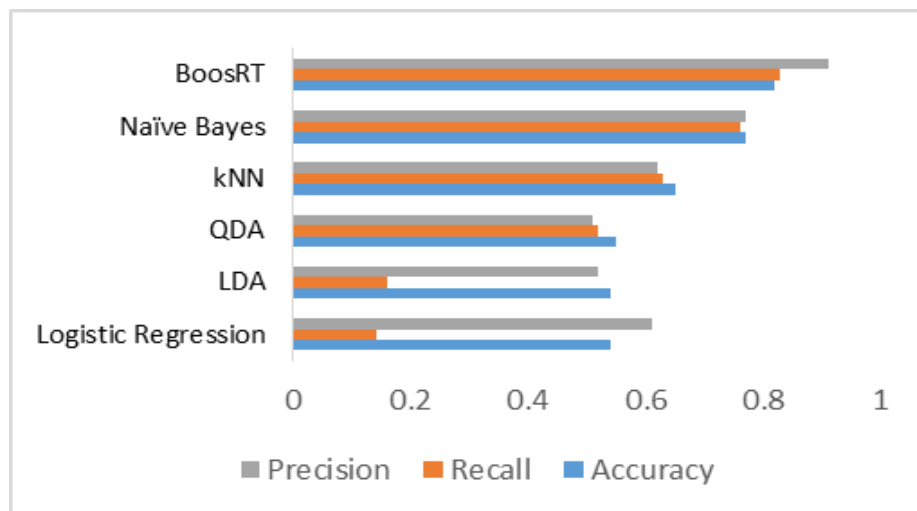


Figure 3. Analysis of Classification Methods with JKSE Data Set

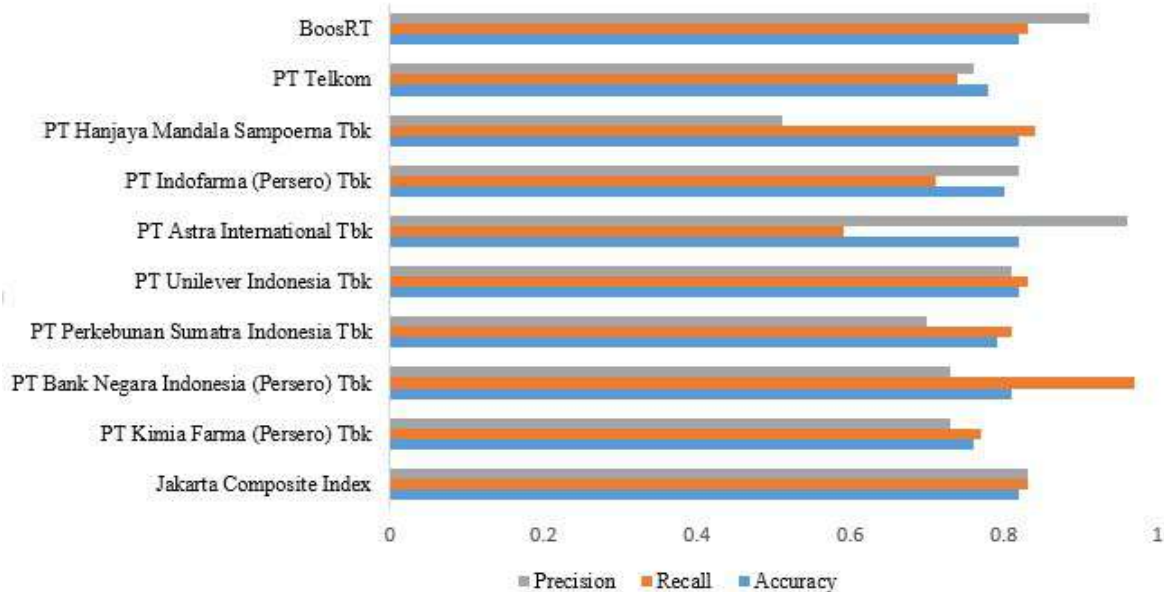


Figure 4. Model accuracy in other data set

3. Review ke 2 (23 November 2020)



Meilany Nonsi Tentua <meilany@upy.ac.id>

Songklanakar Journal of Science and Technology - Decision on Manuscript ID SJST-2020-0281.R1

2 messages

Songklanakar Journal of Science and Technology <onbehalf@manuscriptcentral.com>

Mon, Nov 23, 2020 at 8:02 AM

Reply-To: proespichaya.k@psu.ac.th

To: meilany@upy.ac.id

22-Nov-2020

Dear Mrs. Tentua:

It is a pleasure to accept your manuscript entitled "Ensemble Classification Method for Daily Return Stock Market" in its current form for publication in the Songklanakar Journal of Science and Technology. The comments of the reviewer(s) who reviewed your manuscript are included at the foot of this letter.

Thank you for your fine contribution. On behalf of the Editors of the Songklanakar Journal of Science and Technology, we look forward to your continued contributions to the Journal.

Sincerely,
Assoc. Prof. Dr. Proespichaya Kanatharana
Editor in Chief, Songklanakar Journal of Science and Technology
proespichaya.k@psu.ac.th

Reviewer(s)' Comments to Author:**Reviewer: 1****Comments to the Author**
the author has addressed all the comments.**Reviewer: 2****Comments to the Author**
I think you have attempted to the the query i raised.

Associate Editor
Comments to the Author:
(There are no comments.)

Meilany Nonsi Tentua <meilany@upy.ac.id>

To: proespichaya.k@psu.ac.th

Sun, Mar 7, 2021 at 3:50 PM

dear Editor in Chief, Songklanakarin Journal of Science and Technology,

thank you for receiving my article entitled "Ensemble Classification Method for Daily Return Stock Market"

I want to ask when my article will be included in the current issue?

Thank you

Sincerely

Meilany Nonsi Tentua

[Quoted text hidden]

4. Artikel Accepted (22 November 2020)

Author Dashboard

- 2 Manuscripts with Decisions >
- Start New Submission >
- Legacy Instructions >
- 5 Most Recent E-mails >
- English Language Editing Service >

Manuscripts with Decisions

ACTION	STATUS	ID	TITLE	SUBMITTED	DECISIONED
	<ul style="list-style-type: none"> Accept (22-Nov-2020) Awaiting Production Checklist 	SJST-2020-0281.R1	Ensemble Classification Method for Daily Return Stock Market View Submission	26-Oct-2020	22-Nov-2020
a revision has been submitted (SJST-2020-0281.R1)	<ul style="list-style-type: none"> Major Revision (06-Sep-2020) a revision has been submitted 	SJST-2020-0281	Ensemble Classification Method for Daily Return Stock Market View Submission	27-Jun-2020	06-Sep-2020

[view decision letter](#)

[view decision letter](#)

5. Pemberitahuan publish artikel (21 November 2021)



Meilany Nonsi Tentua <meilany@upy.ac.id>

Checking Artwork for Publication in SJST

2 messages

Songklanakarín Journal of Science and Technology SJST <sjst@psu.ac.th>
To: Meilany Nonsi Tentua <meilany@upy.ac.id>

Sun, Nov 21, 2021 at 10:24 PM

Dear Author,

I'm pleased to inform you that your manuscript will be available online at the Songklanakarín Journal of Science and Technology website. The processes of English Correction and Artwork Production have been done. Please find the attachments and carefully check the correction of your artwork. Are there any mistakes?


If there are any mistakes, please reply to me within November 22, 2021. As well, please specify the categorization for your manuscript:

- Agricultural and Biological Sciences
- Biotechnology and Agro-Industry
- Chemistry and Pharmaceutical Sciences
- Engineering and Industrial Research
- Environmental and Natural Science
- Physical Sciences and Mathematics

If you need any further information, please feel free to let me know.

Yours sincerely,

Aussani Sribenchapon (Nok)
Editorial Assistant
Songklanakarín Journal of Science and Technology
Tel: +66 (0) 62-227-6757
Email: sjst@psu.ac.th or aussani.s@psu.ac.th

 **26.Meilany_1428-1434.pdf**
699K

Meilany Nonsi Tentua <meilany@upy.ac.id>
To: Songklanakarín Journal of Science and Technology SJST <sjst@psu.ac.th>

Tue, Nov 23, 2021 at 7:31 PM

thank you for being my manuscript will be available online at the Songklanakarín Journal of Science and Technology website.
and i inform that my manuscript all correct
Yes, everything is correct.

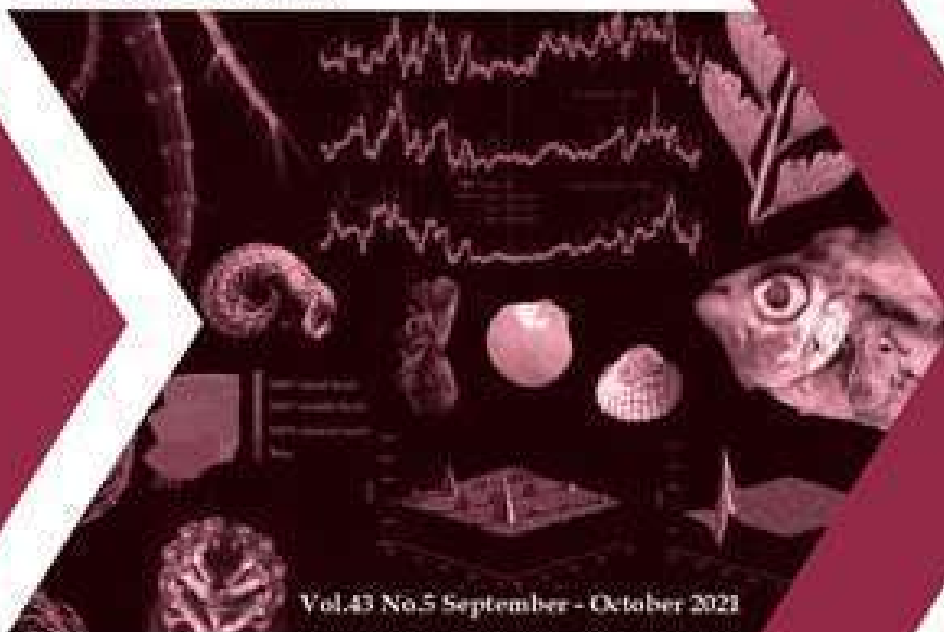
the categorization for my manuscript is Physical Sciences and Mathematics

best regards

meilany

[Quoted text hidden]

ISSN 0125-3395 (print)
ISSN 2408-1779 (online)



Vol.43 No.5 September - October 2021



SONGKLANAKARIN
JOURNAL OF
SCIENCE AND
TECHNOLOGY

Indexed in
SCOPUS, Zoological Record,
ACI, DOAJ and EBSCO

www.sjst.psu.ac.th

Published by Research and Development Office,
Prasert of Songklanakarin University, Thailand

Editorial Board

Executive Editor

Sunton Wongsiri, M.D.,
Prince of Songkla University, Thailand

Associate Executive Editor

Suppasil Maneerat,
Prince of Songkla University, Thailand

Editor in Chief

Proespichaya Kanatharana,
Prince of Songkla University, Thailand

Associate Editors

Benjamas Cheirsilpa,
Prince of Songkla University, Thailand

Chongdee Buranachai,
Prince of Songkla University, Thailand

Kuaanan Techato,
Prince of Songkla University, Thailand

Panote Thavarungkul,
Prince of Songkla University, Thailand

Phadungsak Rattanadecho,
Thammasat University, Thailand

Rawee Chiarawipa,
Prince of Songkla University, Thailand

Editorial Board

Ajcharaporn Piumsomboon,
Chulalongkorn University, Thailand

Alan Frederik Geater,
Prince of Songkla University, Thailand

Anurag Sunpapao,
Prince of Songkla University, Thailand

Apiradee Saelim,
Prince of Songkla University, Thailand

Nitsara Karoonuthaisiri,
National Center for Genetic Engineering and Biotechnology, Thailand

Oramas Suttinun,
Prince of Songkla University, Thailand

Pakamas Chetpattananondh,
Prince of Songkla University, Thailand

Pichaya Tandayya,
Prince of Songkla University, Thailand

Apon Numnuam,
Prince of Songkla University, Thailand

Athassawat Kammanee,
Prince of Songkla University, Thailand

Avirut Chinkulkijniwat,
Suranaree University of Technology, Thailand

Barry Noller,
The University of Queensland, Australia

Byeang Hyeon Kim,
Pohang University of Science and Technology, Republic of Korea

Chaiyawan Wattanachant,
Prince of Songkla University, Thailand

Che-Chen Chang,
National Taiwan University, Taiwan, R.O.C.

Chittanon Buranachai,
Prince of Songkla University, Thailand

Duangrat Thongpak,
Khon Kaen University, Thailand

George A. Gale,
King Mongkut's University of Technology Thonburi, Thailand

Helmut Duerrast,
Prince of Songkla University, Thailand

Hiroshi Kanzaki,
Okayama University, Japan

Hitoshi Watarai,
Osaka University, Japan

Igor Meglinski,
University of Otago, New Zealand

John A.N. Parnell,
The University of Dublin, Ireland

Juraithip Wungsintaweekul,
Prince of Songkla University, Thailand

Khamphe Phoungthong,
Prince of Songkla University, Thailand

Klairung Samart,
Prince of Songkla University, Thailand

Kornkanok Ingkaninan,
Naresuan University, Thailand

Manat Chaijan,
Walailak University, Thailand

Martin Wing Cheung Mak,
Linköping University, Sweden

Mizuo Maeda,
RIKEN Institute, Japan

Nisaudah Radenahmad,
Prince of Songkla University, Thailand

Piti Sukontasukkul,
King Mongkut's University of Technology North Bangkok, Thailand

Pongsaton Amornpitoksuk,
Prince of Songkla University, Thailand

Pornchai Phukpattaranont,
Prince of Songkla University, Thailand

Pornsak Sriamornsak,
Silpakorn University, Thailand

Prapas Patchanee,
Chiang Mai University, Thailand

Ray Chambers,
University of Wollongong, Australia

Ronnason Chinram,
Prince of Songkla University, Thailand

Sappasith Klomklao,
Thaksin University, Thailand

Sawasdee Yordkhayhun,
Prince of Songkla University, Thailand

Shoji Motomizu,
Okayama University, Japan

Shyue-Win Wei,
National Chi Nan University, Taiwan, R.O.C.

Sinchai Kamolphiwong,
Prince of Songkla University, Thailand

Sumetha Suwanboon,
Prince of Songkla University, Thailand

Supinya Tewtrakul,
Prince of Songkla University, Thailand

Surasak Sangkhathat, M.D.,
Prince of Songkla University, Thailand

Teerapong Senjuntichai,
Chulalongkorn University, Thailand

Timothy Wiedmann,
University of Minnesota, United States of America

Tipayaratn Musikacharoen,
Prince of Songkla University, Thailand

Tirayut Vilaivan,
Chulalongkorn University, Thailand

Warakorn Limbut,
Prince of Songkla University, Thailand

Woi Pei Meng,
University of Malaya, Malaysia

Yeong Yik Sung,
Universiti Malaysia Terengganu, Malaysia

Yoshio Otani,
Kanazawa University, Japan

Editorial Assistants

Journal Online Officer

Naliwan Heemham,
Prince of Songkla University, Thailand

About

 [About Us](#)

 [Aim & Scope](#)

 [Editorial Board](#)

 [Peer Review Process](#)

 [Reviewer Acknowledgements](#)

 [Contact](#)

Search

Editorial Board Search



Search by First name, Last name, Affiliation, Country



Songklanakarin Journal of Science and Technology (SJST),

Research and Development Office (RDO),
Prince of Songkla University,
Hat Yai, Songkhla, 90112 Thailand

Tel: +6674-286-959

Mobile: +6666-0483836 / +6662-2276757



Research and Development Office (RDO)



Email: sjst@psu.ac.th



Songklanakarin Journal of Science and Technology (SJST)

Volume 43 No 5 September - October 2021

pp. 1237 - 1541

Alternative matrix in cryptography with Hadamard matrix

Nilobol Kamyun, Supanee Sompong, Nichakan Kamtalang, Mongkoltong Sroypachr, and Kannika Khompongson

DOI: 10.14456/sjst-psu.2021.161

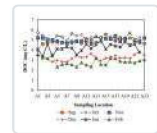
pp. 1237 - 1243 [PDF](#) [PDF w/Links](#)

Tracking seasonal changes of natural organic matter in a water treatment plant in Bangkok using EEM-PARAFAC approach

Kitiyot Tungsudjawong, Suchat Leungprasert, and Panitan Jutaporn

DOI: 10.14456/sjst-psu.2021.162

pp. 1244 - 1251 [PDF](#) [PDF w/Links](#)

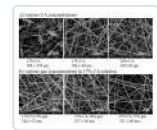


Preparation and characterization of gac loaded cellulose acetate nanofiber patch

Natyada Khammanewanitwong, Rattapol Rangkupan, and Pornpen Werawatganone

DOI: 10.14456/sjst-psu.2021.163

pp. 1252 - 1258 [PDF](#) [PDF w/Links](#)



The comparison of bilateral strength of thigh muscles in long-distance runners with history of anterior knee pain

Yosawin Sakunkaruna, Sutida Sakunkaruna, Prasert Sakulsriprasert, and Kamonwan Praopritpirom

DOI: 10.14456/sjst-psu.2021.164

pp. 1259 - 1263 [PDF](#) [PDF w/Links](#)

Phytochemical and pharmacological activity of *Tiliacora triandra* (Colebr.) Diels

Thanutchaporn Nutmakul

DOI: 10.14456/sjst-psu.2021.165

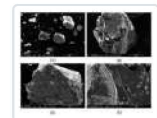
pp. 1264 - 1274 [PDF](#) [PDF w/Links](#)



Effect of pullulanase debranching on the physical and chemical properties of instant Jasmine rice porridges

Jirayus Woraratphoka, Chompoonuch Khongla, Sumalee Musika, Artit Ausavasukhi, Plengpin Pianpumpong, Dusadee Teimtes Boonmasongsung, and Worawikunya Kiatponglangr

DOI: 10.14456/sjst-psu.2021.166



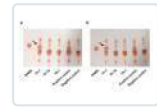
pp. 1275 - 1282 [PDF](#) [PDF w/Links](#)

Cholesterol-lowering activity and functional characterization of lactic acid bacteria isolated from traditional Thai foods for their potential used as probiotics

Narathip Puttarat, Boonyarut Ladda, Anongnard Kasorn, Somboon Tanasupawat, and Malai Tawechotipatr

DOI: 10.14456/sjst-psu.2021.167

pp. 1283 - 1291 [PDF](#) [PDF w/Links](#)

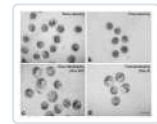


Proteomic analyses in fast and slow-cleaving bovine embryos – a search for potential markers of healthy pre-implantation embryos

Preeyalak Konongbua, Atthaboon Watthammawut, Sittiruk Roytrakul, Monsicha Somrit, Somluk Asuvapongpatana, and Wattana Weerachayanukul

DOI: 10.14456/sjst-psu.2021.168

pp. 1292 - 1298 [PDF](#) [PDF w/Links](#)

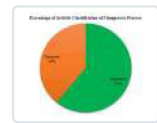


Validation of improvement in changeover of die bond process using basicmost technique

Puvanasvaran A.Perumal, Karmegam Karuppiah, and Lim Eu Sheng

DOI: 10.14456/sjst-psu.2021.169

pp. 1299 - 1304 [PDF](#) [PDF w/Links](#)

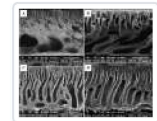


Identifying the potential activity of *Azolla pinnata* through *in vitro* assay and sem

Jerine Peter S., Sanjay V., Muruganandham L., and Evan Prince Sabina

DOI: 10.14456/sjst-psu.2021.170

pp. 1305 - 1311 [PDF](#) [PDF w/Links](#)



Improved growth performance of hybrid grouper (*Epinephelus fuscoguttatus* ♀ × *E. lanceolatus* ♂) fed with green pea meal based diets

M. Azfar-Ismail, Eka Rosyida, and Rossita Shapawi

DOI: 10.14456/sjst-psu.2021.171

pp. 1312 - 1316 [PDF](#) [PDF w/Links](#)

Shoulder Angle Measurement (SAM) system for home-based rehabilitation using computer vision with a web camera

Sunthorn Rungruangbaiyok, Rakkrit Duangsoithong, and Kanadit Chetpattananondh

DOI: 10.14456/sjst-psu.2021.172

pp. 1317 - 1328 [PDF](#) [PDF w/Links](#)

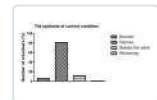


Monitoring adverse drug reactions to Thai herbal remedies (Ya Satee) in Thailand: Observational study

Pornthip Termwiset, Kusuma Sriyakul, Nanthiya Srikaew, and Parunkul Tungsukruthai

DOI: 10.14456/sjst-psu.2021.173

pp. 1329 - 1334 [PDF](#) [PDF w/Links](#)

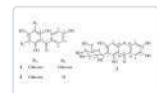


The fabric facial mask enhanced skin permeation of hydrophilic bioactive compounds in *Aquilaria crassna* leaf extract

Eakkaluk Wongwad, Kornkanok Ingkaninan, Wudtichai Wisuitiprot, and Neti Waranuch

DOI: 10.14456/sjst-psu.2021.174

pp. 1335 - 1341 [PDF](#) [PDF w/Links](#)

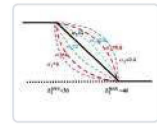


Multi-objective periodic maintenance scheduling and optimization via krill herd algorithm

 Pasura Aungkulanon, Pongchanun Luangpaiboon, and Roberto Montemanni

 DOI: 10.14456/sjst-psu.2021.175

 pp. 1342 - 1350  [PDF](#)  [PDF w/Links](#)



Isotherm, kinetics and thermodynamic studies of Cd²⁺ adsorption from aqueous solution onto cellulose nanocrystals obtained from cassava peel

 C. V. Abiazim, C. T. Onwordi, L. F. Petrik, and A. B. Williams

 DOI: 10.14456/sjst-psu.2021.176

 pp. 1351 - 1359  [PDF](#)  [PDF w/Links](#)

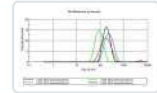


Development of chitosan edible coatings incorporated with clove essential oil nanoemulsions and its effect on shelf life of fresh-cut mangoes

 Pooja Singh, Tooba Khan, Farhan Jalees Ahmad, Gaurav Kumar Jain, and Jinku Bora

 DOI: 10.14456/sjst-psu.2021.177

 pp. 1360 - 1366  [PDF](#)  [PDF w/Links](#)

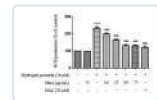


The effects of ethanolic extract of Okra fruit, *Abelmoschus Esculentus* (L.) Moench on cellular senescence in aging neuron

 Pichaya Jumnonprakhon, Krittalak Pasakawee, Panida Banjongsinsiri, Nowwapan Donrung, Supawadee Daodee, and Pennapa Chonpathompikunlert


 DOI: 10.14456/sjst-psu.2021.178

 pp. 1367 - 1373  [PDF](#)  [PDF w/Links](#)

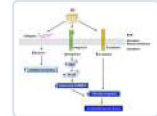


In vitro evaluation of wound healing potential of sulfated galactans from red alga *Gracilaria fisheri* in fibroblast cells

 Choowadee Pariwatthanakun, Tawut Rudtanatip, Boonyakorn Boonsri, Benjamart Pratoomthai, and Kanokpan Wongprasert

 DOI: 10.14456/sjst-psu.2021.179

 pp. 1374 - 1381  [PDF](#)  [PDF w/Links](#)

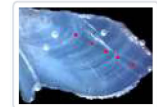


Length-weight relationships and growth parameters of five freshwater fishes in Raban Lake, Perak, Malaysia

 Rumeaida Mat Piah, Nur Atiana Syawani Zakaria, Nur Fathiah Jusoh, Nur Laila Noor Ghani, Nur Syafarrah Zamri, Nurin Zalikha Azman, Nur Syahirah Mohd Nasir, Siti Azrina Kamaruddin, Nuralif Fakhruallah Mohd Nur, and Muhammad Hanif Fadzli

 DOI: 10.14456/sjst-psu.2021.180

 pp. 1382 - 1386  [PDF](#)  [PDF w/Links](#)



Does ambulatory surgery work for open appendectomy?: A review of paediatric cases with appendicitis treated in a single institution

 Paiboon Sookpotarom and Kobkool Chakrapan Na Ayudhya

 DOI: 10.14456/sjst-psu.2021.181

 pp. 1387 - 1391  [PDF](#)  [PDF w/Links](#)



Structural characteristics and physiological responses along the salinity gradients of *Ceriops tagal* (Perr.) C.B. Rob. dominated mangrove associations

 Trittep Vichkovitten

 DOI: 10.14456/sjst-psu.2021.182



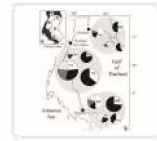
pp. 1392 - 1399 [PDF](#) [PDF w/Links](#)

Spatial variations in biological aspects of *Hippocampus* spp. in the Gulf of Thailand

Parichart Laksanawimol, Mareeya Bahe, and Haruthai Lamthaisong

DOI: 10.14456/sjst-psu.2021.183

pp. 1400 - 1407 [PDF](#) [PDF w/Links](#)

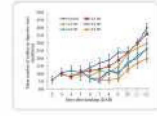


Survival, growth, and feeding ability of marble goby, *Oxyleotris marmorata* (Bleeker, 1852) Larvae under delayed initial feeding

Siti Fairus Mohamed Yusoff, Ching Fui Fui, and Shigeharu Senoo

DOI: 10.14456/sjst-psu.2021.184

pp. 1408 - 1413 [PDF](#) [PDF w/Links](#)

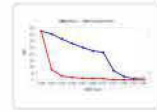


Explicit analytical solutions for the average run length of modified EWMA control chart for ARX(p,r) processes

Korakoch Silpakob, Yupaporn Areepong, Saowanit Sukparungsee, and Rapin Sunthornwat

DOI: 10.14456/sjst-psu.2021.185

pp. 1414 - 1427 [PDF](#) [PDF w/Links](#)



Ensemble classification method for daily return stock market

Meilany Nonsi Tentua and Dedi Rosadi

DOI: 10.14456/sjst-psu.2021.186

pp. 1428 - 1434 [PDF](#) [PDF w/Links](#)



Modified Thiel and saturated salt solutions for rabbit soft cadaveric embalming: A preliminary study

Jirawat Saetan, Jongdee Nopparat, Chittipong Tipbunjong, Wipapan Khimmaktong, Uraporn Vongvatcharanon, Kosol Puriwattanakul, Boonsin Tangtrakulwanich, Chongdee Thammakhet-Buranachai, and Wandee Udomuksorn

DOI: 10.14456/sjst-psu.2021.187

pp. 1435 - 1442 [PDF](#) [PDF w/Links](#)

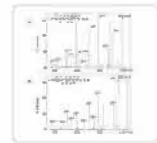


Separation of diethylpyrocarbonate-modified histidine-containing peptides using strong cation exchange chromatography

Pongsakorn Thawonpan, Supinya Thanapongpichat, Watthanachai umpathong, Aung Win Tun, Churat Weeraphan, Worrawit Wanitsuwan, and Hansuk Buncherd

DOI: 10.14456/sjst-psu.2021.188

pp. 1443 - 1448 [PDF](#) [PDF w/Links](#)



Efficient method to compute search directions of infeasible primal-dual path-following interior-point method for large scale block diagonal quadratic programming

Duangpen Jetpipattanapong and Gun Srijuntongsiri

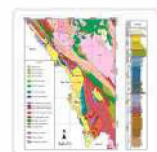
DOI: 10.14456/sjst-psu.2021.189

pp. 1449 - 1455 [PDF](#) [PDF w/Links](#)

Facies analysis of fluvial deposits of Cenozoic Doi Ton Formation in Mae Sot Basin, Western Thailand

Atiwut Bunlam, Yupa Thasod, Rattapanorn Fongngern, and Pitaksit Ditbanjong

DOI: 10.14456/sjst-psu.2021.190



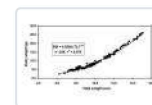
pp. 1456 - 1465 [PDF](#) [PDF w/Links](#)

Life-history traits of Indian river shad *Gudusia chapra* (Hamilton, 1822) in the Mahananda River (tributary of the Ganges River) of northwestern Bangladesh

Md. Yeamin Hossain, Asma Afroz Chowdhury, Zannatul Mawa, Md. Rabiul Hasan, Md. Akhtarul Islam, Md. Aatur Rahman, Obaidur Rahman, Dalia Khatun, Most. Shakila Sarmin, Sabana Nasrin, Md. Ashekur Rahman, Akhery Nima, and Farida Parvin

DOI: 10.14456/sjst-psu.2021.191

pp. 1466 - 1472 [PDF](#) [PDF w/Links](#)



An algorithm for packing postal items into roll pallets: A case study on a local branch of Thailand Post

Pakwan Riyapan, Sofar Tehroh, Fareeda Maeroh, and Arthit Intarasit

DOI: 10.14456/sjst-psu.2021.192

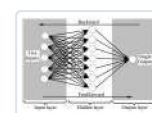
pp. 1473 - 1481 [PDF](#) [PDF w/Links](#)

Automating plastic bearing fault diagnosis using continuous wavelet transform and neural networks

Saipul Azmi Mohd Hashim, N. Abdul Razak, and N. S. Sholahuddin

DOI: 10.14456/sjst-psu.2021.193

pp. 1482 - 1490 [PDF](#) [PDF w/Links](#)

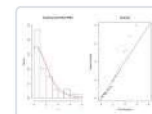


Repetitive acceptance sampling plan for type-ii generalized half logistic distribution based on truncated life test

Gadde Srinivasa, Sd. Jilani, and A. Vasudeva

DOI: 10.14456/sjst-psu.2021.194

pp. 1491 - 1498 [PDF](#) [PDF w/Links](#)



Phytochemical profile, antioxidant activity, and inhibition of α -amylase and α -glucosidase for banana central pseudo-stem juice

Intira Koodkaew, Khwanruedee Chaiyawan, and Pimchanok Satapoomin

DOI: 10.14456/sjst-psu.2021.195

pp. 1499 - 1506 [PDF](#) [PDF w/Links](#)

Biologically active compounds in *Delonix regia* (Flame of forest) seed oil extracted by hexane

Abdulrasaq Olalekan Oyedeji, Bibilomo Abiola Odeyemi Oluwatobi Cynthia Omoyeni, and Luqmon Azeez

DOI: 10.14456/sjst-psu.2021.196

pp. 1507 - 1511 [PDF](#) [PDF w/Links](#)

Broad activity against fungal phytopathogens of marine *Astrophaeriella cf. striatispora* BUCS 055-1

Sudarat Suanjit, Watanalai Panbangred, and Apiradee Pilantanapak

DOI: 10.14456/sjst-psu.2021.197

pp. 1512 - 1518 [PDF](#) [PDF w/Links](#)



The effects of dietary supplement with sodium citrate and probiotic on bacterial density in gastrointestinal tract and growth performance of Asian swamp eel (*Monopterus albus*)


Huynh Thanh Toi, Mashyaka Anicet, and Nguyen Thi Hong Van

DOI: 10.14456/sjst-psu.2021.198

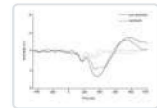
pp. 1519 - 1526 [PDF](#) [PDF w/Links](#)

Comparisons between recidivists and non-recidivists on affective, behavioral, and P300 brain activity

 Naruemon Intamuen, Apitchaya Chaiwutikomwanich, and Vorasith Siripornpanich

 DOI: 10.14456/sjst-psu.2021.199

 pp. 1527 - 1534  [PDF](#)  [PDF w/Links](#)

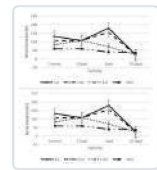


Salinity tolerance of freshwater-derived *Aedes aegypti* and *Aedes albopictus* from the artificial containers of Dumaguete, Philippines

 Venus Bantoto-Kinamot and Jaydon Soringa Lomoljo

 DOI: 10.14456/sjst-psu.2021.200

 pp. 1535 - 1541  [PDF](#)  [PDF w/Links](#)



Archives

 [Latest Issue](#)

 [Article in Press](#)

 [Back Issues](#)

Document Search

Article Search

Search by Article title, Abstract, Keywords, DOI

Author Search

Search by First name, Last name

Search 



Songklanakarin Journal of Science and Technology (SJST),

Research and Development Office (RDO),
Prince of Songkla University,
Hat Yai, Songkhla, 90112 Thailand

Tel: +6674-286-959

Mobile: +6666-0483836 / +6662-2276757



Research and Development Office (RDO)



Email: sjst@psu.ac.th

Copyright © Since 2021, All Rights Reserved by [Research and Development Office](#)

*Original Article***Ensemble classification method for daily return stock market**Meilany Nonsi Tentua^{1*} and Dedi Rosadi²¹ *Department of Informatics, Faculty of Science and Technology,
Universitas PGRI Yogyakarta, Yogyakarta, 55182 Indonesia*² *Department of Statistics, Gadjah Mada University, Yogyakarta, 55281 Indonesia*

Received: 27 June 2020; Revised: 26 October 2020; Accepted: 22 November 2020

Abstract

Stock price movements in Indonesia are measured using indices, one of which is the Composite Stock Price Index (CSPI). CSPI is a stock index that measures a combination of all shares from various sectors listed on the Indonesia Stock Exchange.

The ensemble method is to build predictive models by combining the strengths of the classical classification method. In this research, the purpose of ensemble based on Boosting for Regression appeared to enhance simple tree analysis and deals with some of the weaknesses found in uncomplicated techniques. The ensemble tree combines the prediction values of many simple trees into a single prediction value. Based on the experiments that have been carried out, the ensemble method proved to have a better accuracy rate, which amounted to 82%. It is assumed that the ensemble model can obtain the relationships between variables that are more precise than the previous model.

Keywords: ensemble method, stock market returns, classification

1. Introduction

Stock is one of the most popular financial market instruments. Stock can be interpreted as a sign of capital participation of a person or party (business entity) in a company or limited liability company ("PT Bursa Efek Indonesia," n.d.). Share price movements in Indonesia are measured using indices, one of which is the Composite Stock Price Index (CSPI). CSPI is a stock index that measures a combination of all shares from various sectors listed on the Indonesia Stock Exchange. Investors usually use parameters in the CSPI to read price developments and make a reference to the portfolio.

A stock market index is a measurement of the value of a part of the stock market. It is often used to describe market aggregate trends that will be predicted by a basic financial issue. Such stochastic value-share is very difficult to predict (Bruni, 2017). The difficulty in predicting the return of

shares purchased by investors means that an investor is still guessing about the returns that investors will obtain in the future. There are four approaches used in predicting stock price movements such as technical analysis (Sedighi, Mohammadi, Fard, & Sedighi, 2019), fundamental analysis (Miralles-Quirós, Miralles-Quirós, & Oliveira, 2017), traditional time series forecasting (Park, Lee, & Lee, 2019), and statistical or machine learning methods (Kim & Han, 2016).

Statistical learning refers to a set of tools for modeling and understanding complex data sets. Statistical learning is a newly developed area in statistics and integrated with parallel developments in computer science and, in particular, machine learning (Gareth James, Daniela Witten, Trevor Hastie, 2013). Predictive modeling using data-sets is one of the tasks in statistical learning. An approach to predict the qualitative response is a process known as classification. Classification is the attribution of labels to notes according to criteria that are automatically learned from a series of labeled record training.

Some classical classification algorithms are widely used today (Trevor Hastie, Robert Tibshirani, 2017) such as

*Corresponding author
Email address: meilany@upy.ac.id

Linear Discriminant Analysis, Logistic Regression (Gündüz, Çataltepe, & Yaslan, 2017) (Shah, Ismail, & Bin Shahrin, 2019), K-Nearest Neighbors (Tanuwijaya & Hansun, 2019), Naïve Bayes (Mariano & Maruddani, 2020), Support Vector Machines (Selakov, Cvijetinović, Milović, Mellon, & Bekut, 2014), and Trees (Sitorus & Tarihoran, 2018). However, using a relatively small amount of data in stock market data, models built with Linear Discriminant Analysis, Logistic Regression, K-Nearest Neighbors, Naive Bayes, and Tress will easily over-fit. Their variance also tends to increase. Thus, it has an impact on the low accuracy achieved.

In making predictions for the daily stock return market, an appropriate method is needed to produce an accurate model. Classification using the Random Forests method in the study of Jyothirmayee *et al.* (2019), and the boosting process in the research of Nayak, Pai, & Pai (2016), have provided a slightly better level of accuracy compared to previous classical methods. These methods are also called ensemble methods. They can improve the accuracy of prediction models formed. However, in their research, they used analytical sentiment variables that would be difficult for new investors. The idea of ensemble learning is to build predictive models by combining the strengths of classical classification methods (Trevor Hastie, Robert Tibshirani, 2017).

In this study, we aim to predict daily stock market returns in the movement of the stock Jakarta Composite Index (JCI). We also use the most regulated daily share price, such as PT Bank Negara Indonesia (Persero) Tbk, PT Perusahaan Perkebunan London Sumatra Indonesia Tbk, PT Unilever Indonesia Tbk, PT Astra International Tbk, PT Bank Mandiri (Persero) Tbk, and PT Hanjaya Mandala Sampoerna Tbk. To predict the daily stock market, we use the Boosted Regression Tree (Boost RT) method. Besides, we also made comparisons with other classical methods such as LDA, Logistic regression, KNN, and Naïve Bayes.

The contributions of our study is the use of simple features, only by calculating the stock closing price's return value from the five previous days. The classification method that we use is the Boosted Regression Tree (BoostRT) method, which is an ensemble method that combines the Boosting method and the Decision Tree method. This method is rarely used for predicting stock prices in Indonesia. We use accuracy techniques in evaluating the classification model.

2. Materials and Methods

Several studies have been conducted in the application of statistical learning in the stock market domain. The use of the bootstrap method is modified with Random Forest to predict the direction of the movement of stock index prices (Kim & Han, 2016). Training sets are generated by a modified bootstrap considering the impact of response variables simultaneously and applied in a random forest. The data used are the Korea composite stock price index with the variable close price, open price, high price, low price, trading volume, and training amount.

A hierarchical Deep Neural Network (DNN) is a method that can be applied to predict stock market returns. A study applying this method had been conducted to predict the stock market returns within the next five minutes (Lachiheb &

Gouider, 2018). The data used were the main index of the Tunisian stock market (TUNINDEX) with the variables of TUNINDEX current value, Number of up direction stocks, Number of down direction stocks, Number of unchanged stocks, and Total stock values in Tunisian dinars. Other research that uses DNN is a large comprehensive data analysis process to predict the daily return direction of the S&P 500 SPDR ETF (ticker symbol: SPY) based on 60 financial and economic features (Zhong & Enke, 2019).

The approach of a cross-sectional statistical estimation model for selecting stocks on the Chinese stock market allows investors to identify stocks that are likely to perform well and to build a suitable portfolio (Wu, Chen, Xu, He, & Tindall, 2019). The data used are sourced from the Shanghai Composite Index.

Predicting stock market prices using Fuzzy Metagraph (FM) based on stock market decision making, classification, and prediction are proposed for short-term investors from the Indian stock market (Anbalagan & Maheswari, 2015). Simple Moving Average (SMA), Exponential Moving Average (EMA), Moving Average Convergence Divergence (MACD), and Relative Strength Index (RSI) are some technical indicators that are used as input to train an integrated system with Fuzzy Meta graph. The data used in the Indian stock market data are from the Bombay Stock Exchange (BSE).

Artificial Neural Network (ANN) is a method in machine learning that can be applied for prediction and classification. Implementation of ANN back-propagation is used to predict the closing price of the S & P 500 stock exchange. Historical data consist of five variables, namely open, high, low, close, and volume.

The Adaptive Neuro-Fuzzy Inference System (ANFIS) method is used to determine the Jakarta Composite Index (JCI) value movement. This study provides two outputs, namely prediction results and classification results based on several relative error values. The classification method used is Fuzzy Kernel C-Means (Fanita & Rustam, 2018).

This research is done to classify and predict the daily return of the stock market in Indonesia. The method undertaken in this research can be seen in Figure 1.

Step 1: Collecting CSPI data from the Indonesia Composite Index through the Yahoo Finance web.

Step 2: Calculating the market stock daily return value from the data collected in step 1 using the formula:

$$\text{Return Value} = \frac{\text{Currentvalue} - \text{Originalvalue}}{\text{Originalvalue}} \times 100 \quad (1)$$

Step 3: Pre-processing data by handling the missing values and excessive data from the data-set.

Step 4: Making sure that the training data-set is larger compared to the test set. The training data-set is taken from 2015 to 2018 data, while the test data-set is taken from 2019 data.

Step 5: Applying the ensemble Boosting Regression Tree (BoosRT).

Step 6: Analyzing the results of the ensemble model performance in terms of its accuracy, memory, accuracy, and specificity.

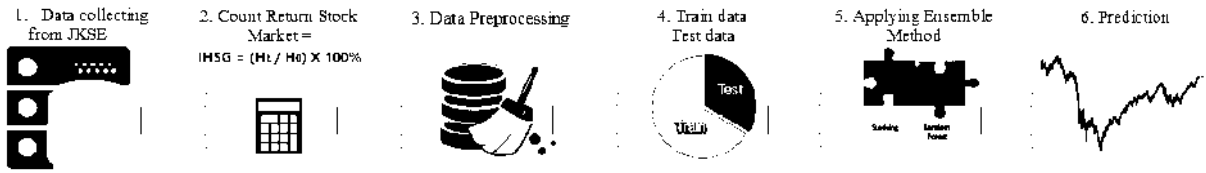


Figure 1. Steps under taken in the research

2.1 Boosted Regression Tree (BoosRT)

Boosted Regression Tree (BoosRT) models are a combination of two techniques: decision tree algorithms and boosting methods (Shin, 2015). In these algorithms, the BoosRT approximates the function (x) as preservative growth of the base learner:

$$f(x) = f_0(x) + \beta_1 f_1(x) + \beta_2 f_2(x) + \dots + \beta_m f_m(x) \tag{2}$$

A single study cannot predict enough using training data, so we must use the lowest residuals to improve predictive performance. BoosRT uses an iterative algorithm, so every iteration *m*, and a new regression tree. The formula to predict a separate constant value is:

$$h(x; \{R_{lm}\}_l^L = \sum_{l=1}^L \overline{y_{lm}}(x \in R_{lm}) \tag{3}$$

where y_{lm} = mean $x_i \in R_{lm}$ (\tilde{y}_{lm}) is the mean of pseudo-residuals the *m*th iteration:

$$\tilde{y}_{lm} = - \left[\frac{\partial \Psi(y_i, F(x_i))}{\partial f(x_i)} \right]_{f(x)=F_{m-1}(x)} \tag{4}$$

After that the approximation obtained will be updated with the following equation

$$f_m(x) = f_{m-1}(x) + v \cdot Y_{lm} \mid (x \in R_{lm}), \tag{5}$$

Where

$$Y_{lm} = \arg \min_Y \sum_{x_i \in R_{lm}} \Psi(y_i, F_{m-1}(x_i) + Y) \tag{6}$$

BoosRT algorithm for the generalized boosting of regression trees are:

- a) Initialize (x), $F_0(x) = \arg \min_Y \sum_{i=1}^N \Psi(y_i, \gamma)$.
- b) For $m=1$ to *M* do
- c) From the full training dataset, select a subset randomly,

$$\{\pi(i)\}_i^N = \text{rand_perm}\{i\}_i^N \tag{7}$$

- d) Fit the base learner

$$\check{y}_{\pi(i)} = - \left[\frac{\partial \Psi(y_{\pi(i)}, F(x_i))}{\partial F(x(i))} \right]_{F(x)=f_{m-1}(x)}, i = 1, \tilde{N} \tag{8}$$

- e) Update model for the current iteration,

$$\{R_{lm}\}_l^L - \text{terminalmodetree}(\{\overline{y_{\pi(i)}}, x_{\pi(i)}\}_i^N) \tag{9}$$

- f) Choose a gradient descent step size as

$$Y_{lm} = \arg \min_Y \sum_{x(i) \in R_{lm}} \Psi(y_{\pi(i)}, F_{m-1}(x_{\pi(i)}) + Y) \tag{10}$$

- g) Update the estimate of (x) as

$$F_m(x) = F_{m-1}(x) + v \cdot Y_{lm} \mid (x \in R_{lm}) \tag{11}$$

2.2. Performance evaluation

The performance of the classification model can be measured by counting the number of classes that were predicted correctly (true positive), the predicted numbers that were not included in the class and were true (true negative), and those that were wrongly predicted (false positive or false negative). The formulation of accuracy, precision, and recall is used in this study using the following formula:

$$\text{Accuracy} = \frac{tp + tn}{tp + fn + fp + tn} \tag{12}$$

$$\text{Precision} = \frac{tp + tn}{tp + fp} \tag{13}$$

$$\text{Recall} = \frac{tp}{tp + fn} \tag{14}$$

where, *tp* = true positive, *tn* = true negative, *fp* = false positive, and *fn* = false negative.

3. Results and Discussion

3.1. Results

The daily return stock market classification using the ensemble method is implemented using R. Some of the libraries contained in R are tools for statistical analysis and can also be used to create models with classification techniques (James, Witten, & Hastie, 2013), (Rosadi, 2017). The data-sets used in this study are JCI, PT Bank Negara Indonesia (Persero) Tbk, PT Perusahaan Perkebunan London Sumatra Indonesia Tbk, PT Unilever Indonesia Tbk, PT Astra International Tbk, PT Bank Mandiri (Persero) Tbk, and PT Hanjaya Mandala Sampoerna Tbk. The daily data were collected from the Jakarta Composite Index via the Yahoo Finance website. IHSIG data were collected from December

25, 2014 to December 25, 2019. A data frame with 1193 observations comprises the following nine variables, namely:

Year: The year that the data was recorded

X1: return value at t-1

X2: return value at t-2

X3: return value at t-3

X4: return value at t-4

X5: return value at t-5

Volume: Volume of shares traded (number of daily shares traded in billions)

Today: return value at today (t)

Direction: the label Down and Up, indicating that the market had a positive or negative return on a given day.

The observation data taken from Yahoo Finance consists of 7 (seven) variables: Date, Open, High, Low, Close, Adj Close, and Volume. In pre-processing data, we will process the values in the observation data into a return value. The return value is calculated by using the close price value in the observed data. The results of return value based on observation data is shown in Table 1.

$$Today = \frac{6.236,69 - 6.219,44}{6.219,44} * 100 = 0,277$$

$$X1 = \frac{6.219,44 - 6.334,84}{6.334,84} * 100 = -1,82$$

$$X2 = \frac{6.334,84 - 6.342,17}{6.342,17} * 100 = -0,12$$

$$X3 = \frac{6.342,17 - 6.381,95}{6.381,95} * 100 = -0,63$$

$$X4 = \frac{6.381,95 - 6.336,67}{6.336,67} * 100 = 0,715$$

$$X5 = \frac{6.336,67 - 6.326,21}{6.336,67} * 100 = 0,165$$

The return value of a day becomes the label Direction, which indicating the market had a positive (Up) or negative (Down) return on a given day. The distribution of JKSE return value for 2015-2019 can be seen in Figure 2.

We compared the performance of the BoosRT method with the Logistic Regression, Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), k-Nearest Neighbors (kNN), and Naive Bayes methods.

The performance of the proposed method is validated by comparing it in terms of its accuracy, recall, precision, and specificity. The analysis is done by taking different proportions of tests and training. In conducting the training, data from 2015 to 2018 were used, as many as 951 data. Data testing used data from 2019, as many as 242 data.

Table 1 shows the results of the accuracy, recall, precision, and specificity of the classification method, where the BoosRT methods as ensemble methods. While the other method is the classical method.

Figure 2 shows the level of accuracy, recall, precision, and specificity of each method used in the JCI daily data between 2015 and 2019.

Table 1. Data observation

Date	Open	High	Low	Close
Sep 05, 2019	6,294.28	6,307.35	6,281.95	6,306.80
Sep 06, 2019	6,329.41	6,336.91	6,305.02	6,308.95
Sep 09, 2019	6,328.28	6,333.90	6,306.74	6,326.21
Sep 10, 2019	6,331.73	6,342.01	6,311.34	6,336.67
Sep 11, 2019	6,334.59	6,381.95	6,328.69	6,381.95
Sep 12, 2019	6,399.00	6,414.48	6,337.52	6,342.17
Sep 13, 2019	6,369.42	6,375.80	6,318.92	6,334.84
Sep 16, 2019	6,262.29	6,266.14	6,193.51	6,219.44
Sep 17, 2019	6,215.24	6,240.35	6,205.30	6,236.69

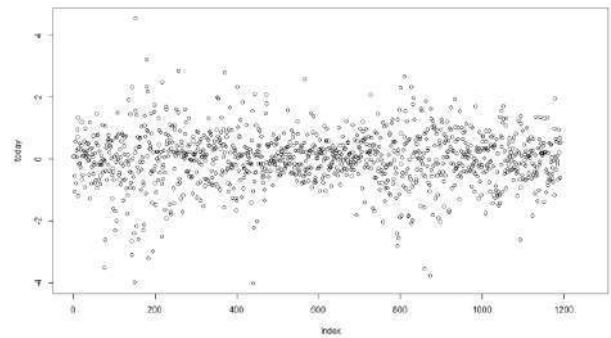


Figure 2. Return value today of JKSE 2015-2019

This ensemble method is also tested on other data sets, such as:

- PT Kimia Farma (Persero) Tbk, produces and sells medicines and their derivative products.
- PT Bank Negara Indonesia (Persero) Tbk, provides various banking products and services in Indonesia.
- PT Perusahaan Perkebunan London Sumatra Indonesia Tbk, a plantation business in Indonesia.
- PT Unilever Indonesia Tbk, engages in manufacturing, marketing, and distribution of consumer goods.
- PT Astra International Tbk, together with its subsidiaries, engages in the automotive and financial services in Indonesia.
- PT Indofarma (Persero) Tbk, manufactures and markets generic, over the counter, and branded generic drugs in Indonesia.
- PT Hanjaya Mandala Sampoerna Tbk, manufactures and trades cigarettes in Indonesia.
- PT Telekomunikasi Indonesia Tbk, provides telecommunications, information, media, and edutainment services worldwide.

We used the stock market data from these companies because their shares greatly affect JCI. This data set is taken from a model that has been formed. Table 3 shows the results of the accuracy, recall, precision, and specificity of the BoosRT classification method.

3.2 Discussion

In the first experiment we used IHSG data; it appeared that the logistic regression method had an accuracy of 54.32% or had $100 - 54,32 = 45,68\%$ training error rate. By removing variables that were not very influential in the

Table 2. Evaluation of classification methods

Methods	Accuracy	Recall	Precision
Logistic Regression	0.54	0.14	0.61
Linear Discriminant Analysis (LDA)	0.54	0.16	0.52
Quadratic Discriminant Analysis (QDA)	0.55	0.52	0.51
kNN	0.65	0.63	0.62
Naïve Bayes	0.77	0.76	0.77
BoosRT	0.82	0.83	0.71

Table 3. Ensemble classification with the other data set

Data set	Accuracy	Recall	Precision
PT Kimia Farma (Persero) Tbk	0,76	0,77	0,73
PT Bank Negara Indonesia (Persero) Tbk	0,81	0,97	0,73
PT Perusahaan Perkebunan London Sumatra Indonesia Tbk	0,79	0,81	0,70
PT Unilever Indonesia Tbk	0,82	0,83	0,81
PT Astra International Tbk	0,82	0,59	0,96
PT Indofarma (Persero) Tbk	0,8	0,71	0,82
PT Hanjaya Mandala Sampoerna Tbk	0,82	0,84	0,71
PT Telkom	0,78	0,74	0,76

formation of logistic regression models, an effective model was obtained. The accuracy level of 54.32% was the highest level of accuracy obtained by using very influential variables in the direction prediction.

The use of the Linear Discriminant Analysis (LDA) method did not significantly differ in accuracy compared with the Logistic Regression method. The LDA method had an accuracy rate of 54.48%. Likewise, the QDA method had an accuracy rate only 1% better than the Logistic Regression and LDA methods.

For kNN method, the models created were used in the data testing. The level of accuracy achieved was 65 % with $k = 2$. This model was better than LDA and QDA. In using this method, we tried k values from 1 to 15. When $k = 2$ the experiment achieved the highest accuracy values.

Training data using the Naïve Bayes method turned out to have a better level of accuracy when applied to test data, compared to the previous method. The accuracy of this method was 77 %, or 12 % better for training error.

Moreover, by using the ensemble method called BoosRT, it had an accuracy rate of 82 %. This is significant for stock market data, which is known to be very difficult to be predicted accurately. It shows that the model assumed by BoosRT can capture the relationship between variables more accurately than the previously used methods.

In the second experiment we used the BoosRT method on the return value of eight companies, namely PT Bank Negara Indonesia (Persero) Tbk, PT Perusahaan Perkebunan London Sumatra Indonesia Tbk, PT Unilever Indonesia Tbk, PT Astra International Tbk, PT Bank Mandiri (Persero) Tbk, and T Hanjaya Mandala Sampoerna Tbk. The

results of the classification using ensemble models that have been obtained on 8 (eight) data sets showed an accuracy level of 76% to 82%. The average level of accuracy of the data set tested using the BoosRT model was 80.4%. The average recall from the model was 79.2%, which means that the success rate of the model in finding information had a better performance. The average precision model of 77.6% means that the level of prediction matches the data. It also means that the model has a good performance. From the measurement of accuracy, recall, and precision, the BoosRT model used had a better average than the logistic Regression, LDA, QDA, kNN, or Naïve Bayes methods. This shows that the ensemble model using the BoosRT model is stable to be used in different data sets. The accuracy of the BoosRT model in the test data set can be seen in Figure 4.

In this study, we used the return value in the past five days as a feature. This feature considers that the stock price may be affected by stock movements from a few preceding days. This feature makes it easy for a new investor to predict the value of shares. Our research's use of features is different from research conducted by Jyothirmayee *et al.* (2019) using the random forest method; they used the data feature taken from Yahoo Finance. It is a weakness of the model because the prediction only considers the stock's movement on the previous day to lower accuracy.

Besides, a study by Nayak *et al.* (2016), shows that the use of features and stock value should also consider the sentiment analysis of the company. It will not be easy for a new investor in deciding to buy shares if he has to know the company's sentiment analysis.

The limitation of this study is that it uses only five return value variables and does not pay attention to the socio-political and security conditions that occur in Indonesia which also affect the stock market performance.

5. Conclusions

According to the experiments, the ensemble method has a high accuracy (82%). It is significant for stock market data, which is known to be difficult to model accurately. It shows that the model assumed by BoosRT can capture the relationship between variables more accurately than the methods used before. The classification using ensemble models had been obtained on the other data sets. The model has shown an accuracy level of 76% to 82%. It shows that the ensemble model using the BoosRT model is stable to be used and generalized in different data sets.

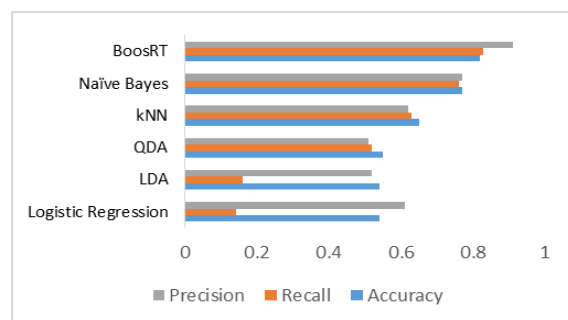


Figure 3. Analysis of classification methods with JKSE data set

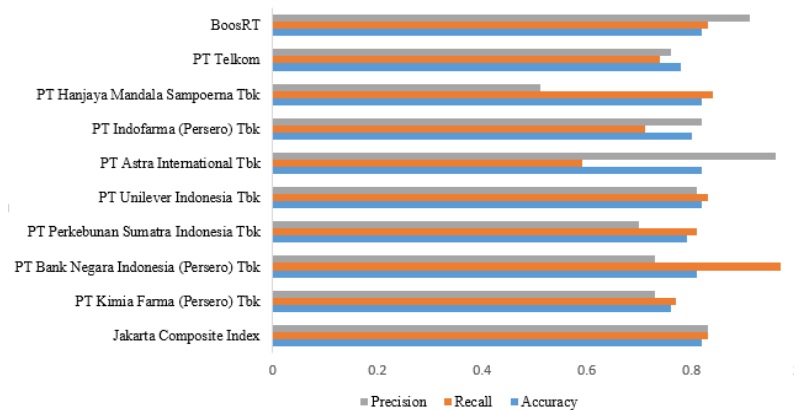


Figure 4. Model accuracy in other data sets

The policy implications in this study can use from two sides, investors and companies. For investors, predictions of the stock price can determine their investment strategy. It will relate to the rate of return they expect in investing. For companies, this prediction will be considered for the company to determine policies related to the company's stock price.

References

- Anbalagan, T., & Maheswari, S. U. (2015). Classification and prediction of stock market index based on Fuzzy Metagraph. *Procedia Computer Science*, 47(C), 214–221. doi:10.1016/j.procs.2015.03.200
- Bruni, R. (2017). Stock market index data and indicators for day trading as a binary classification problem. *Data in Brief*, 10, 569–575. doi:10.1016/j.dib.2016.12.044
- Fanita, F., & Rustam, Z. (2018). Predicting the Jakarta composite index price using ANFIS and classifying prediction result based on relative error by fuzzy Kernel C-Means. *AIP Conference Proceedings*, 2023. doi:10.1063/1.5064203
- Gareth James, Daniela Witten, Trevor Hastie, R. T. (2013). An introduction to statistical learning with applications in R. In G. Casella S. Fienberg I. Olkin (Ed.), *Springer Texts in Statistic*. doi:10.2174/0929867003374372
- Gündüz, H., Çataltepe, Z., & Yaslan, Y. (2017). Stock daily return prediction using expanded features and feature selection. *Turkish Journal of Electrical Engineering and Computer Sciences*, 25(6), 4829–4840. doi:10.3906/elk-1704-256
- Jyothirmayee, S., Dilip Kumar, V., Someswara Rao, C., & Shiva Shankar, R. (2019). Predicting stock exchange using supervised learning algorithms. *International Journal of Innovative Technology and Exploring Engineering*, 9(1), 4081–4090. doi:10.35940/ijitee.A4144.119119
- Kim, H., & Han, S. T. (2016). The enhanced classification for the stock index prediction. *Procedia Computer Science*, 91(Itqm), 284–286. doi:10.1016/j.procs.2016.07.077
- Lachiheb, O., & Gouider, M. S. (2018). A hierarchical Deep neural network design for stock returns prediction. *Procedia Computer Science*, 126, 264–272. doi:10.1016/j.procs.2018.07.260
- Mariato, F. Y., & Maruddani, D. A. I. (2020). *Perbandingan Metode Naive Bayes dan Bayesian Regularization Neural Network (BRNN) untuk Klasifikasi Sinyal Palsu pada Indikator Stochastic Oscillator*, 9, 16–25.
- Miralles-Quirós, M. del M., Miralles-Quirós, J. L., & Oliveira, C. (2017). The role of liquidity in asset pricing: the special case of the Portuguese Stock Market. *Journal of Economics, Finance and Administrative Science*, 22(43), 191–206. doi:10.1108/JEFAS-12-2016-0001
- Nayak, A., Pai, M. M. M., & Pai, R. M. (2016). Prediction models for indian stock market. *Procedia Computer Science*, 89, 441–449. doi:10.1016/j.procs.2016.06.096
- Park, M., Lee, M. L., & Lee, J. (2019). Predicting stock market indices using classification tools. *Asian Economic and Financial Review*, 9(2), 243–256. doi:10.18488/journal.aefr.2019.92.243.256
- Rosadi, D. (2017). *Analisis statistika dengan R*. Retrieved from <http://opac.perpusnas.go.id/DetailOpac.aspx?id=985555#>
- Sedighi, M., Mohammadi, M., Fard, S. F., & Sedighi, M. (2019). The Nexus between stock returns of oil companies and oil price fluctuations after heavy oil upgrading: Toward theoretical progress. *Economies*, 7(3), 1–17. doi:10.3390/economies7030071
- Selakov, A., Cvijetinović, D., Milović, L., Mellon, S., & Bekut, D. (2014). Hybrid PSO-SVM method for short-term load forecasting during periods with significant temperature variations in city of Burbank. *Applied Soft Computing Journal*, 16, 80–88. doi:10.1016/j.asoc.2013.12.001
- Shah, S. Q., Ismail, I., & Bin Shahrin, A. R. (2019). Stock market manipulation: A comparative analysis of East Asian emerging and developed financial markets. *Management Science Letters*, 9(1), 183–192. doi:10.5267/j.msl.2018.10.006
- Shin, Y. (2015). *Application of boosting regression trees to preliminary cost*, 2015.

- Sitorus, H., & Tarihoran, Y. (2018). Analisis Harga Saham PT Astra Internasional Tbk Menggunakan Data Dari Bursa Efek Indonesia dalam Jangka Waktu Pendek Menggunakan Metode Naïve Bayes dan Decision Tree-J48. *TeIKA*, 8(1), 21–33. doi:10.36342/teika.v8i1.2239
- Tanuwijaya, J., & Hansun, S. (2019). LQ45 stock index prediction using k-nearest neighbors regression. *International Journal of Recent Technology and Engineering*, 8(3), 2388–2391. doi:10.35940/ijrte.C4663.098319
- Trevor Hastie, Robert Tibshirani, J. F. (2017). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). *SpringerPlus*. doi:10.1111/j.1751-5823.2009.00095_18.x
- Wu, W., Chen, J., Xu, L., He, Q., & Tindall, M. L. (2019). A statistical learning approach for stock selection in the Chinese stock market. *Financial Innovation*, 5(1), 1–18. doi:10.1186/s40854-019-0137-1
- Zhong, X., & Enke, D. (2019). Predicting the daily return direction of the stock market using hybrid machine learning algorithms. *Financial Innovation*, 5(1). doi:10.1186/s40854-019-0138-0