

# A comparison between deep learning, naïve bayes and random forest for the application of data mining on the admission of new students

Nurhachita<sup>1</sup>, Edi Surya Negara<sup>2</sup>

<sup>1</sup>Universitas Islam Negeri Raden Fatah Palembang, Indonesia

<sup>2</sup>Universitas Bina Darma, Indonesia

---

## Article Info

### Article history:

Received Feb 21, 2020

Revised Jan 5, 2021

Accepted Mar 29, 2021

---

### Keywords:

Data mining  
Deep learning  
Naïve bayes  
New students  
Random forest

---

## ABSTRACT

The process of admitting new students at Universitas Islam Negeri Raden Fatah each year produces a lot of new student data. So that there is an accumulation of student data continuously. The purpose of this study is to compare deep learning, naïve bayes, and random forest on the admission of new students as well as being one of the bases for making decisions to determine the promotion strategy of each study program. The data mining method used is knowledge discovery in database (KDD). The tools used are rapid miner. The attributes used are student ID number, name, program study, faculty, gender, place of birth, date of birth, year of entry, school origin, national examination, type of payment, and nominal payment. The new student data used from 2016 to 2019 was an 18.930 item. The results of this study used deep learning bayes results resulted in an accuracy value of 52.65%, naïve bayes results resulted in an accuracy value of 99.79%, and random forest results resulted in an accuracy value of 44.65%.

*This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.*



---

## Corresponding Author:

Edi Surya Negara  
Data Science Interdisciplinary Research Center  
Computer Science Faculty  
Universitas Bina Darma, Indonesia  
Email: e.s.negara@binadarma.ac.id

---

## 1. INTRODUCTION

Information technology has an important role in most organization that manipulates and collects data in large databases. Decision making is generated by useful information from storing data. Data storage patterns that have been discovered by users to help the data analysis process is called data mining [1]. Along with the development of the internet, the data stored, both in the form of text, images, sound, and video also increased very quickly and significantly. In Indonesia, internet users in 1998 were only 500,000 users whereas by 2015 it was projected that internet users had reached 139 million [2]. information can identified by characteristics [3]. The large volume of data volume will become "garbage" in storage if it is not processed into useful information. Data mining technology provides a user-oriented approach to novel and hidden patterns in the data [4]. Mixed data models that have many topics can form a text data set model [5]. This is consistent with the definition of data that data is a fact that is recorded but has no meaning. Many universities have used information technology to support the admission process [6]. The application of information technology to education can also produce abundant student data and learning processes [7]. At universities, data can be obtained from databases, data will continue to grow, such as student data. Hope after this data mining technique can be used and useful and help analyze data in higher education institutions.

The process of admitting new students at Universitas Islam Negeri Raden Fatah every year produces a lot of new student data. This happens continuously so that there is an accumulation of student data which will continuously increase in the search for student information. Based on the amount of new student data, by managing the data, information that can be seen can be done by the University. Based on the number of new student data, by organizing the data so that information can be accessed and accepted by the university, for example, a compilation of university promotions or outreach and study programs in schools to accept new students, universities access schools for promotion. This causes a waste of budget because too many schools will be visited, and not time efficient. This research will classify and clarify data on the admission of new students at Universitas Islam Negeri Raden Fatah by utilizing the data mining process by applying clarification techniques. By comparing the three algorithms, deep learning, naïve bayes, and random forest. The tools used are rapid miner. The attributes used are student ID number, name, program study, faculty, place of birth, gender, date of birth, school origin, year of entry, national examination, type of payment, and nominal payment. Based on the results of deep learning, naïve bayes, and random forest can determine the promotion strategy of each study program. Based on the results of deep learning, naïve bayes, and random forest can see courses of interest in each school. The final results of the cluster can help the University [8].

Data mining concept is to extract hidden patterns and to discover relationships between parameters in a vast amount of data [9]. Data mining is the process of extracting data (previously unknown, implicit, and considered useless) into information or knowledge or patterns from large amounts of data [10]. Data that is considered "garbage" because it is not patterned/not structured and is not useful, is processed (filter) so that it forms information or knowledge or new patterns that are useful [9]. Data mining is a process to explore the excess value of information that was not previously known to exist in a database. The patterns that are very useful and have more value than the data contained in the database are obtained by recognizing the information obtained [11]. From the explanation above it can be concluded that data mining is a step of analyzing the process of knowledge discovery in the database. Data mining is a process that employs one or more machine learning techniques (machine learning) to analyze and extract knowledge automatically [12].

A method based on learning from a feature that is not noticed is called deep learning [13]. Naïve bayes algorithm is one of the clarification algorithms based on the bayesian theorem in statistics. The availability of a class can be predicted by the naïve bayes algorithm [14]. The naïve bayes method is the beginning to build a method that has been designed using a corpus that has been formed [15]. Naïve bayes is often called the bayes' rule which is a prefix for data mining methods and machine learning. It builds a model with predictions. This is a new way to find out data and learn more [16]. Modeling data obtained by working with binary data, and is a category of data called random forest [17]. Clarification techniques on naïve bayes can be used at very large input dimensions. This is a simple algorithm but can produce very good results than other algorithms [18].

## 2. RESEARCH METHOD

The field of study that focuses on a methodology to add knowledge that is very much useful from the data is called knowledge discovery in database (KDD). The rapid development of online data on an ongoing basis due to the widespread use of the internet and databases has made a very large need from the KDD methodology [19]. Obstacles to adding knowledge from data to database research, machine learning, knowing patterns, statistics, and increasing performance to become smart and sophisticated business solutions [20]. In this study, the method used for data processing is the admission data by using the stages of knowledge discovery in database (KDD) as illustrated in Figure 1. Knowledge discovery in database (KDD) is the process of determining useful information and patterns in data. This information is contained in a large database that was previously unknown and potentially useful. Data mining is one step in a series of KDD iterative processes [21].

### 2.1. Data selection

In this process the selection of data sets is done, creating a target data set, or focusing on a subset of variables (data samples) where the discovery will be performed [22]. The results of the selection are stored in a separate file from the operational database. The attributes used are student ID number, name, program study, faculty, gender, school origin, year of entry, national examination, type of payment, and nominal payment. The data in this study were sourced from Universitas Islam Negeri Raden Fatah where this data is secondary data consisting of new student data for 2016 up to 2019. The amount of data obtained was 18,930 consisting of student ID number, name, program study, faculty, place of birth, gender, date of birth, school origin, year of entry, national examination, type of payment, and nominal payment. The following are some examples of new students data Table 1. The stages of the knowledge discovery in database (KDD) process consist of:

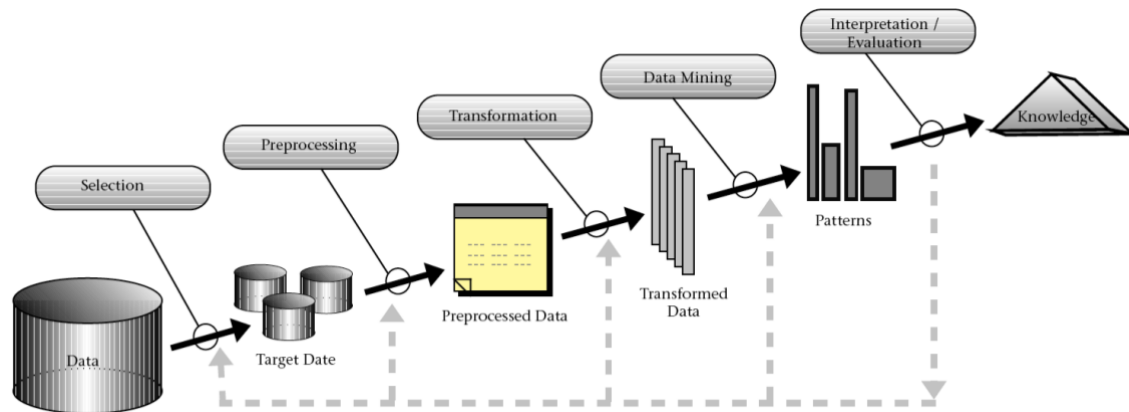


Figure 1. Stages in KDD

Table 1. New student data obtained

No	Student ID Number	Name	Study Program	Gender	Yea of Entry	School Origin	National Examination	Type of Payment	Nominal Payment
1	1683600004	Siska Apriyanti	Hadith Science	F	2016	Islamic school Nurul Hikmah	77	Group 3	Rp. 1.800.000
2	1683600003	Sifaul Hasanah	Quranic Sciences and Interpretation	M	2016	Islamic school Nurul Hikmah	67	Group 3	Rp. 1.800.000
3	1683600002	Maulidin Anam	Quranic Sciences and Interpretation	M	2016	Islamic school Nurul Hikmah	83	Group 3	Rp. 1.800.000
4	1683400008	Bayu Putra	Aqeedah and Islamic Philosophy	F	2016	Senior High School Puspita	85	Group 3	Rp. 1.800.000

## 2.2. Pre-processing and cleaning

Data pre-processing and data cleaning are done by removing inconsistent data and noise, duplicating data, correcting data errors, and can be enriched with relevant external data [23].

## 2.3. Transformation

This process transforms or combines data into a more appropriate way to do the mining process by summarizing (aggregation).

## 2.4. Data mining

A cycle to obtain a pattern or information that is very interested in data and used by a technique [24], methods, or algorithms under the objectives of the KDD process is called data mining process [10].

## 2.5. Interpretation/evaluation

The process for translating patterns generated from data mining. Evaluate (test) whether the patterns or information found are by or contradictory to previous facts or hypotheses. Knowledge obtained from the patterns formed is presented in the form of visualization.

## 3. RESULTS AND ANALYSIS

### 3.1. Deep learning

The data processing of new students using deep learning with rapidminer software is shown in Figure 2. Detail of the validation process in deep learning as shown in Figure 3. Using deep learning modeling as shown in Figure 2, with the amount of training data (new student admission data from 2016 to 2019) receiving 18.930 items. The accuracy of using deep learning is 52.65% as shown in Figure 4. Besides

producing an accuracy value, deep learning also produces a kappa value of 0.511, a correlation value of 0.804, and a cross-entropy value of 1.793 as shown in Figure 5.

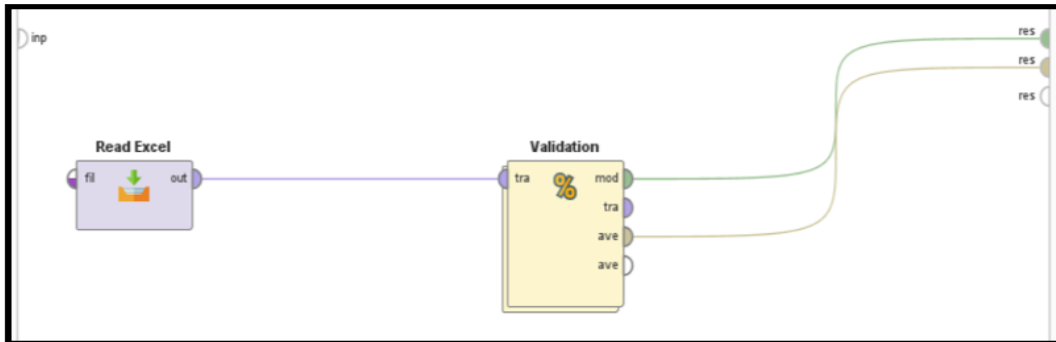


Figure 2. Deep learning modeling on rapidminer

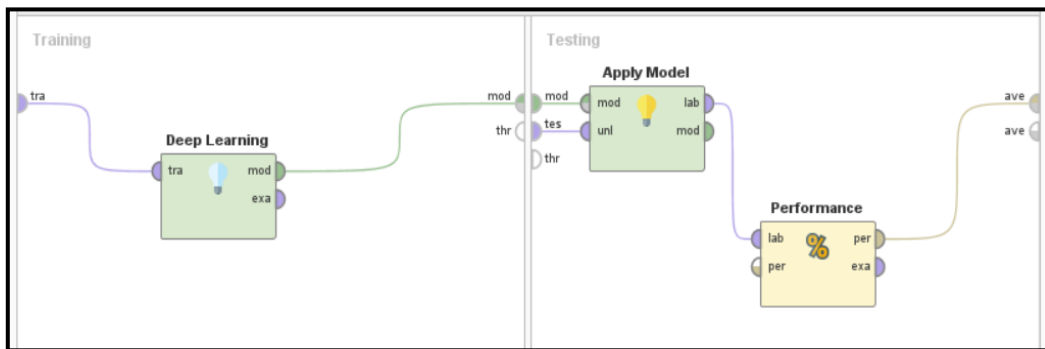


Figure 3. Validation deep learning

accuracy: 52.65%

	true S1 L...	true S1 L...	true S1 A...	true S1 S...	true S1 E...	true S1 P...	true S1 E...	true S1 S...	true S1 K...	true S1 P...	true S1 P...	true S1
pred. S1 ...	23	7	17	10	0	0	0	0	0	0	0	0
pred. S1 ...	13	43	12	12	0	0	0	0	0	0	0	0
pred. S1 ...	71	22	77	60	0	0	0	0	0	0	0	0
pred. S1 ...	10	1	10	10	0	0	0	0	0	0	0	0
pred. S1 ...	0	0	0	0	1	0	0	0	0	0	0	0
pred. S1 ...	0	0	0	0	0	296	193	0	0	0	0	0
pred. S1 ...	0	0	0	0	0	47	113	0	0	0	0	0
pred. S1 ...	0	0	0	0	14	0	0	105	0	0	0	0
pred. S1 ...	0	0	0	0	0	0	0	0	89	0	0	0
pred. S1 ...	0	0	0	0	0	0	0	0	0	174	0	0

Figure 4. View accuracy deep learning

### 3.2. Naïve bayes

The data processing of new students using naïve bayes with rapidminer software is shown in Figure 6. Using naïve bayes modeling as shown in Figure 6, with the amount of training data (new student admission data from 2016 to 2019) receiving 18.930 items and testing data using 2019 new student admission data with a total of 4762 items. The accuracy of using naïve bayes is 99.79% as shown in Figure 7. Besides producing an accuracy value, naïve bayes also produces a kappa value of 0.998, a correlation value of 0.998, and a cross-entropy value of 0.029 as shown in Figure 8.

### 3.3. Random forest

The data processing of new students using random forest with rapidminer software is shown in Figure 9. Using random forest modeling as shown in Figure 9, with the amount of training data (new student admission data from 2016 to 2019 receiving 18,930 and testing data using 2019 new student admission data with a total of 4762. The accuracy of using random forest 44.65% as shown in Figure 10.

kappa: 0.535

	true S1 L...	true S1 L...	true S1 A...	true S1 S...	true S1 B...	true S1 P...	true S1 E...	true S1 S...	true S1 K...	true S1 P...	true S1 P...	true S1 F...
pred. S1 ...	23	7	17	10	0	0	0	0	0	0	0	0
pred. S1 ...	13	43	12	12	0	0	0	0	0	0	0	0
pred. S1 ...	71	22	77	60	0	0	0	0	0	0	0	0
pred. S1 ...	10	1	10	10	0	0	0	0	0	0	0	0
pred. S1 ...	0	0	0	0	1	0	0	0	0	0	0	0
pred. S1 ...	0	0	0	0	0	296	193	0	0	0	0	0
pred. S1 ...	0	0	0	0	0	47	113	0	0	0	0	0
pred. S1 ...	0	0	0	0	14	0	0	105	0	0	0	0
pred. S1 ...	0	0	0	0	0	0	0	89	0	0	0	0
pred. S1 ...	0	0	0	0	0	0	0	0	174	0	0	0

Figure 5. View kappa deep learning

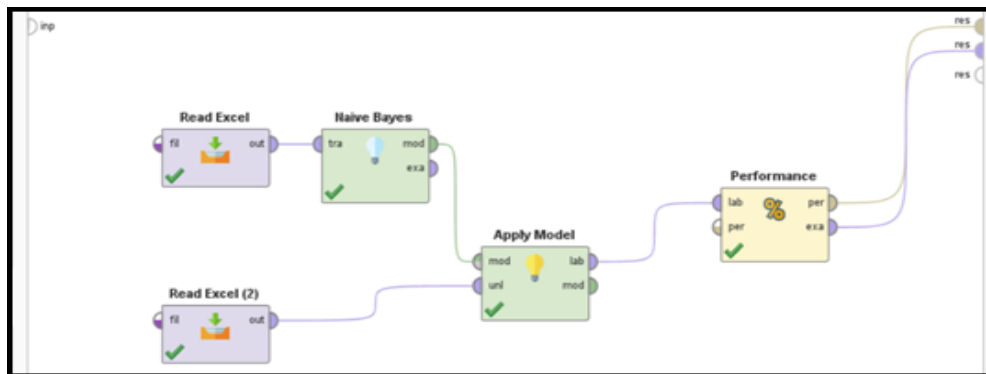


Figure 6. Naïve bayes modeling on rapidminer

accuracy: 99.79%

	true S1 B...	true S1 K...	true S1 P...	true S1 P...	true S1 P...	true S1 ...	true S1 P...	true S1 S...	true S1 K...	true S1 B...	true S1 L...	true S1 L...
pred. S1 ...	126	0	0	0	0	0	0	0	0	0	0	0
pred. S1 ...	0	124	0	0	0	0	0	0	0	0	0	0
pred. S1 ...	0	0	141	0	0	0	0	0	0	0	0	0
pred. S1 ...	0	0	0	105	0	0	0	0	0	0	0	0
pred. S1 ...	0	0	0	0	274	0	0	0	0	0	0	0
pred. S1 ...	0	0	0	0	0	108	0	0	0	0	0	0
pred. S1 ...	0	0	0	0	0	0	187	0	0	0	0	0
pred. S1 ...	0	0	0	0	0	0	0	75	0	0	0	0
pred. S1 ...	0	0	0	0	0	0	0	1	38	0	0	0
pred. S1 ...	0	0	0	0	0	0	0	0	0	35	0	0
pred. S1 ...	0	0	0	0	0	0	0	0	0	0	90	0

Figure 7. View accuracy naïve bayes

kappa: 0.998

	true S1 B...	true S1 K...	true S1 P...	true S1 P...	true S1 P...	true S1 ...	true S1 P...	true S1 S...	true S1 K...	true S1 B...	true S1 I...	true S1
pred. S1 ...	126	0	0	0	0	0	0	0	0	0	0	0
pred. S1 ...	0	124	0	0	0	0	0	0	0	0	0	0
pred. S1 ...	0	0	141	0	0	0	0	0	0	0	0	0
pred. S1 ...	0	0	0	105	0	0	0	0	0	0	0	0
pred. S1 ...	0	0	0	0	274	0	0	0	0	0	0	0
pred. S1 ...	0	0	0	0	0	108	0	0	0	0	0	0
pred. S1 ...	0	0	0	0	0	0	187	0	0	0	0	0
pred. S1 ...	0	0	0	0	0	0	0	75	0	0	0	0
pred. S1 ...	0	0	0	0	0	0	0	1	38	0	0	0
pred. S1 ...	0	0	0	0	0	0	0	0	0	35	0	0

Figure 8. View kappa naïve bayes

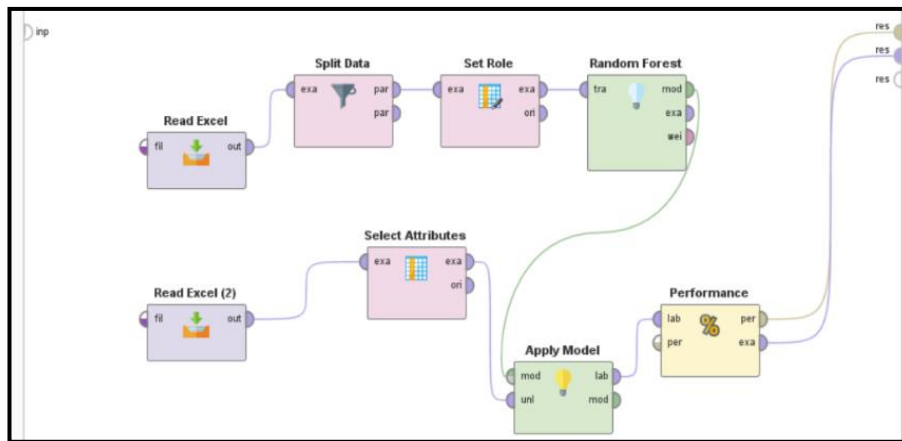


Figure 9. Random forest modeling on rapidminer

accuracy: 44.65%

	true S1 I...	true S1 I...	true S1 A...	true S1 S...	true S1 B...	true S1 P...	true S1 E...	true S1 S...	true S1 K...	true S1 P...	true S1 P...	true S1 P...
pred. S1 ...	256	78	86	70	0	0	0	0	0	0	0	0
pred. S1 ...	3	37	2	2	0	0	0	0	0	0	0	0
pred. S1 ...	115	99	282	143	0	0	0	0	0	0	0	0
pred. S1 ...	20	29	39	94	0	0	0	0	0	0	0	0
pred. S1 ...	0	0	0	0	66	0	0	2	0	0	0	0
pred. S1 ...	0	0	0	0	0	1121	837	0	0	0	0	0
pred. S1 ...	0	0	0	0	0	34	192	0	0	0	0	0
pred. S1 ...	0	0	0	0	86	0	0	348	0	0	0	0
pred. S1 ...	0	0	0	0	0	0	0	0	231	0	0	0
pred. S1 ...	0	0	0	0	0	0	0	0	0	579	0	0
pred. S1 ...	0	0	0	0	0	0	0	0	0	0	18	0

Figure 10. View accuracy random forest

Besides producing an accuracy value, random forest also produces a kappa value of 0.421, a correlation value of 0.729, and a cross-entropy value of 2.003, as shown in Figure 11. The parameter results of accuracy, precision, recall can be said to produce good classification results or not by using the classification result parameter guidelines shown in Table 2 [25]. The results of the comparison between the three algorithms namely deep learning, naïve bayes, and random forest, as shown in Table 2.

The highest accuracy value is naïve bayes with a value of 99.79%, while for the highest kappa value is naïve bayes with a value of 0.998, while for the highest correlation value is naïve bayes with a value of 0.998, and for the highest cross-entropy value that is random forest with a value of 2.003.

kappa: 0.421												
	true S1 L...	true S1 I...	true S1 A...	true S1 S...	true S1 B...	true S1 P...	true S1 E...	true S1 S...	true S1 K...	true S1 P...	true S1 P...	true S1 P...
pred. S1 ...	256	78	86	70	0	0	0	0	0	0	0	0
pred. S1 ...	3	37	2	2	0	0	0	0	0	0	0	0
pred. S1 ...	115	99	262	143	0	0	0	0	0	0	0	0
pred. S1 ...	20	29	39	94	0	0	0	0	0	0	0	0
pred. S1 ...	0	0	0	0	66	0	0	2	0	0	0	0
pred. S1 ...	0	0	0	0	0	1121	837	0	0	0	0	0
pred. S1 ...	0	0	0	0	0	34	192	0	0	0	0	0
pred. S1 ...	0	0	0	0	86	0	0	348	0	0	0	0
pred. S1 ...	0	0	0	0	0	0	0	0	231	0	0	0
pred. S1 ...	0	0	0	0	0	0	0	0	0	579	0	0
pred. S1 ...	0	0	0	0	0	0	0	0	0	0	18	0
pred. S1 ...	0	0	0	0	0	0	0	0	0	0	450	1268

Figure 11. View accuracy random forest

Table 2. Results comparison of all three algorithms

Result	The algorithm		
	Deep learning	Naïve bayes	Random forest
Accuracy	52.65%	99.79%	44.65%
Kappa	0.511	0.998	0.421
Correlation	0.804	0.998	0.729
Cross-entropy	1.793	0.029	2.003

#### 4. CONCLUSION

Based on the research and discussion that has been carried out, it can be concluded that from the three methods of deep learning, naïve bayes, and random forest in determining the best student recruitment promotion strategy at the Raden Fatah State Islamic University in Palembang and referring to the original data. Data of new students used from 2016 to 2019 were 18.930 items as data training and data testing used data of new students from 2019 were 4762 items. The results of this study used deep learning results resulted in an accuracy value of 52.65%, naïve bayes results resulted in an accuracy value of 99.79%, and random forest results resulted in an accuracy value of 44.65%. So of the three algorithms that show the best results for the promotion strategy of incoming new students that are using naïve bayes with the highest accuracy value of 99.79%.

#### REFERENCES

- [1] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From data mining to knowledge discovery in databases," *AI Magazine*, vol. 17, no. 3, pp. 37-53, 1996.
- [2] Nurhachita and E. S. Negara, "A Comparison Between Naïve Bayes and the K-Means Clustering Algorithm for The Application of Data Mining on The Admission of New Students," *Jurnal Intelektualita: Keislaman, Sosial, dan Sains*, vol. 9, no. 1, pp. 51-62, 2020, <https://doi.org/10.19109/intelektualita.v9i1.5574>.
- [3] D. F. Brianna, E. Surya Negara, and Y. N. Kunang, "Network Centralization Analysis Approach in the Spread of Hoax News on Social Media," *2019 International Conference on Electrical Engineering and Computer Science (ICECOS)*, Batam, Indonesia, 2019, pp. 303-308, doi: 10.1109/ICECOS47637.2019.8984526.
- [4] J. Soni, U. Ansari, D. Sharma, and S. Soni, "Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction," *International Journal of Computer Applications*, vol. 17, no. 8, pp. 43-48, 2011, doi: 10.5120/2237-2860.
- [5] E. S. Negara, D. Triadi, and R. Andryani, "Topic Modelling Twitter Data with Latent Dirichlet Allocation Method," *2019 International Conference on Electrical Engineering and Computer Science (ICECOS)*, Batam, Indonesia, 2019, pp. 386-390, doi: 10.1109/icecos47637.2019.8984523.
- [6] W. Yathongchai, C. Yathongchai, K. Kerdprasop, and N. Kerdprasop, "Factor analysis with data mining technique in higher educational student drop out," *Latest Advances in Educational Technologies*, pp. 111-116, 2012.
- [7] A. P. Wibawa, et al., "Naïve Bayes Classifier for Journal Quartile Classification," *International Journal of Recent Contributions from Engineering, Science & IT (iJES)*, vol. 7, no. 2, pp. 91-99, 2019.
- [8] A. Cutler, D. R. Cutler, and J. R. Stevens, "Random Forests," *Springer Science+Business Media*, pp. 157-175, 2012, doi: 10.1007/978-1-4419-9326-7.
- [9] F. Ahmad, N. H. Ismail, and A. A. Aziz, "The prediction of students' academic performance using classification data mining techniques," *Applied Mathematical Sciences*, vol. 9, no. 129, pp. 6415-6426, 2015, doi: 10.12988/ams.2015.53289.

- [10] T. Marnoto, "Drying of Rosella (*Hibiscus sabdariffa*) Flower Petals using Solar Dryer with Double Glass Cover Collector," *International Journal of Science and Engineering*, vol. 7, no. 2, pp. 155-160, 2014, doi: 10.12777/ijse.7.2.150-154.
- [11] F. A. Rahman, M. I. Desa, A. Wibowo, and N. A. Haris, "Knowledge discovery database (KDD)-data mining application in transportation," *Proceeding of International Conference on Electrical Engineering, Computer Science and Informatics (EECSI 2014)*, Yogyakarta, pp. 116-119, 2014, doi: 10.11591/eecsi.1.357.
- [12] B. Kolukisa, *et al.*, "Diagnosis of Coronary Heart Disease via Classification Algorithms and a New Feature Selection Methodology," *International Journal of Data Mining Science (IJDAT)*, vol. 1, no. 1, pp. 1-8, 2019.
- [13] F. Zhou, Y. Gao, and C. Wen, "A Novel Multimode Fault Classification Method Based on Deep Learning," *Journal of Control Science and Engineering*, vol. 2017, no. 1, 2017, doi: 10.1155/2017/3583610.
- [14] B. Kolukisa, *et al.*, "Evaluation of Classification Algorithms, Linear Discriminant Analysis and a New Hybrid Feature Selection Methodology for the Diagnosis of Coronary Artery Disease," *2018 IEEE International Conference on Big Data (Big Data)*, Seattle, WA, USA, 2018, pp. 2232-2238, doi: 10.1109/BigData.2018.8622609.
- [15] T. Sutabri, A. Suryatno, D. Setiadi, and E. S. Negara, "Improving naïve bayes in sentiment analysis for hotel industry in Indonesia," *2018 Third International Conference on Informatics and Computing (ICIC)*, Palembang, Indonesia, 2018, pp. 1-6, doi: 10.1109/IAC.2018.8780444.
- [16] S. Mujawar and H. P. R. Devale, "Prediction of Hear t Disease using Modified K-means and by using Naive Bayes," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 3, no. 11, pp. 039-0400, 2015, doi: 10.15680/IJIRCCCE.2015.
- [17] Jehad A., Khan, Rehanullah, N. Ahmad, I. Maqsood, "Random Forests and Decision Trees," *IJCSI International Journal of Computer Science Issues*, vol. 9, no. 5, pp. 272-278, 2012.
- [18] M. Doshi and S. K. Chaturvedi, "Correlation Based Feature Selection (CFS) Technique to Predict Student Perfomance," *International journal of Computer Networks & Communications*, vol. 6, no. 3, pp. 197-206, 2014, doi: 10.5121/ijcnc.2014.6315.
- [19] M. Ramageri, Bharati, M., "Data Mining Techniques and Applications," *Indian Journal of Computer Science and Engineering*, vol. 1, no. 4, pp. 301-305, 2010.
- [20] P. V. P. Sundar, "A Comparative Study for Predicting Student's Academic Performance Using Bayesian Network Classifiers," *IOSR Journal of Engineering*, vol. 3, no. 2, pp. 37-42, 2013, doi: 10.9790/3021-03213742.
- [21] N. Kumar, S. Jain, and K. Chauhan, "Knowledge Discovery from Data Mining Techniques," *International Journal of Engineering Research & Technology (IJERT)*, vol. 7, no. 12, pp. 1-3, 2019.
- [22] E. S. Egara, D. Kerami, I. M. Wiryana, and T. B. Maulana Kusuma, "Researchgate data analysis to measure the strength of Indonesian research," *Far East Journal of Electronics and Communications*, vol. 17, no. 5, pp. 1177-1183, 2017, doi: 10.17654/EC017051177.
- [23] R. Amanda, E. S. Negara, A. Info, D. Mining, E. Method, and Y. D. Classification, "Analysis and Implementation Machine Learning for YouTube Data Classification by Comparing the Performance of Classification Algorithms," *Jurnal Online Informatika* vol. 5, no. 1, pp. 61-72, 2020, doi: 10.15575/join.v5i1.505.
- [24] R. Dieng, O. Corby, and S. Lapalut, "Acquisition of gradual knowledge," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 723 LNAI, no. 3, pp. 407-426, 1993, doi: 10.1007/3-540-57253-8\_65.
- [25] Florin Gorunescu, "Data mining techniques and models," *Intelligent Systems Reference Library*, vol. 12, pp. 185-317, 2011, doi: 10.1007/978-3-642-19721-5\_5.

## BIOGRAPHIES OF AUTHORS



**Nurhachita** has obtained her bachelor's of Information System from Universitas Islam Negeri Raden Fatah Palembang. She is a master student majoring in Informatics Engineering at Universitas Bina Darma. And then, she works at Universitas Islam Negeri Raden Fatah Palembang.



**Edi Surya Negara** has obtained his bachelor's and master of informatics from Universitas Bina Darma and Doctor of Information Technology from Gunadarma University. He has 10 years of teaching and research experience. He publishes 9 research papers at the international level.