Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th October 2021. Vol. 1. No. 1

Implementation Data Mining of Employement Contract Extension at Indosat Using *Naïve Bayes*

Andini Fadila Sari[1*], M. Safii[2*], Dedi Suhendro[3], Irfan Sudahri Damanik[4]

[1,4]STIKOM Tunas Bangsa Pematangsiantar, Sumatera Utara, Indonesia*
[2,3]AMIK Tunas Bangsa Pematangsiantar, Sumatera Utara, Indonesia*
[1,2,3,4]Jln. Sudirman Blok A No. 1-3 Pematangsiantar, Sumatera Utara
**andinifdlsr123@gmail.com, *m.safii@amiktunasbangsa.ac.id*

Abstract

Contract employees are company resources in carrying out oprasional activities for a certain time based on an agreement or contract. Every company that uses a work contrak system every year, there must be employees who are extended and not renewed. Employees will get additional contracts if they have good performance. In this case to determine whether an employee is extended or not extended his work contract, there is difficulty in determining it and requires a long time and process. Therefore, this research was conducted to help guarantee the extension of the employee's work contract by classifier it into the labes "Eligble" and "Not Feasible" which has 4 variables for the process of employees who will be extended or not. The four variables are age, years of service, aspects of delay, achievement. In this study, the alternatives used as samples were employees at PT. Indosat Ooredoo. The number of data tested is 5 employees with two classes. From the results of the calculation of the Naïve Bayes Algorithm, it is obtained classification with 3 employees eligible class and 2 employees not eligible class. The results of this study found that the level of accuracy of 100.00%.

Keywords: Data Mining, Naive Bayes, Work Contract Extension.

1. Introduction

Employees are human resources (SDM) which is a very valuable company asset [1],[2]. Good management must be done to provide the best contribution. The more the company grows, the more human resources are needed. In a company there are two types of employees, namely contract employees and permanent employees. Contract employees is a time work agreement system (PKWT), in which the agreement system signed by the company has a period of 2 years and can only be extended once, with a maximum period of 1 year. PT Indosat Ooredo is a telecommunications giant that dominates and consistently monopolizes the cellular telecommunications industry in Indonesia [3]. The company was carried out in fighting over a very large number of potential Indonesian consumers in order to maintain and increase their respective markwt shares. In the problems so far, there is a reduction in employees every year because the work contract period has expired, with this it can motivate employees to be able to improve their performance. With the reduction of employees that occurs every year the company needs a system that can help determine the extension of work contracts that can improve human resources. In general, the evaluation of this system can help the manager in determining the extension of the employee's employment contract [4],[5]. The extension of the employee's employment contract is based on the employee's performance. If the employee's performance is considered good, the contract will be extended, otherwise if the employee's behavior is considered bad and does not meet company standards, the work contract will not be extended[6].

To determine the extension of the employment contract of employees at PT Indosat Ooredoo, the company has a standard employee assessment of each of these assessments in various aspects, namely Age, Working Period, Delay Aspects, Achievement. To solve the problem, a research was conducted using the Naive Bayes Algorithm [7]. The Naive Bayes algorithm is one of the statistical classifications[8],[9], where this classification can predict the probability of class membership of a data that will fall into a certain class, according to the probability calculation [10] with the title Data Mining Implementation for Predicting Employee Contract Status Using the Naive Bayes Algorithm Case Study of Kospin Jasa[11]. Conveying the results of his research regarding this research, he stated that the existence of a predictive system software for the status of an employee's employment contract can minimize the time required for the process of determining the status of an employee contract extension and is in accordance with existing procedures.

In this case [12] with the title Decision Support System for Extension of Employment Contracts at Grand Inna Daira Hotel Palembang with the Topsis Method, the results of his research regarding this study stated that decision support systems can assist general managers in making better decisions [13], [14]. From alternative decisions in a simple mathematical form and with this Topsis method can produce precise and accurate calculations through the process of existing calculation stages.

2. Results and Discussion

In this chapter the author describes the results of research and data analysis. By implementing the final result of the Naive Bayes Algorithm using 2 stages, namely by calculating Naive Bayes manually and testing using the RapidMiner 5.3 application.

2.1. Data Processing Using Naïve Bayes Algorithm

To obtain the the results of the research conducted, the following is a description of the manual calculation of the Naïve Bayes classification process in determining the extension of an employee's employment contract using the Nave Bayes method. The criteria used are 4, namely: Age, Working Period, Delay Aspects, and Achievement. The following sub-criteria of each criterion are as follows:

Table 1. Research Data

No	Alternative	Age	Years of Service	Lateness Aspect	Achievement	Classification
1	A1	50 years	33 years	60	200 Target	Not Feasible
2	A2	45 years	24 years	80	150 Target	Not Feasible
3	A3	45 years	24 years	60	250 Target	Not Feasible
4	A4	42 years	21 years	50	100 Target	Not Feasible
5	A5	38 years	15 years	80	200 Target	Worthy
6	A6	26 years	5 years	85	225 Target	Worthy
7	A7	44 years	23 years	80	40 Target	Not Feasible
8	A8	27 years	6 years	100	95 Target	Worthy
9	A9	33 years	10 years	85	270 Target	Worthy
10	A10	24 years	3 years	50	320 Target	Worthy
11	A11	22 years	1 years	100	200 Target	Worthy
12	A12	25 years	4 years	85	150 Target	Worthy
13	A13	25 years	4 years	50	200 Target	Worthy
14	A14	21 years	1 years	50	150 Target	Worthy
15	A15	21 years	1 years	100	150 Target	Worthy
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
40	A40	26 years	4 years	100	55 Target	Worthy
41	A41	22 years	1 years	50	150 Target	Worthy
42	A42	50 years	28 years	85	45 Target	Not Feasible
43	A43	33 years	11 years	100	60 Target	Worthy
44	A44	45 years	22 years	50	135 Target	Not Feasible
45	A45	26 years	5 years	60	155 Target	Worthy
46	A46	25 years	21 years	50	50 Target	?
47	A47	26 years	23 years	50	95 Target	?
48	A48	30 years	10 years	100	150 Target	?
49	A49	27 years	6 years	85	95 Target	?
50	A50	25 years	8 years	65	170 Target	?

After the data has been determined, the next step the author calculates the number of Eligible and Unfeasible based on Table 2.1, from 45 training data used, it is known that the Eligible class is 28 data, and the Inappropriate class is 17 data. Prior Probability calculation is likely to be feasible in determining the extension of the employee's employment contract, namely:

$$P(\text{Worthy}) = \frac{28}{45} = 0{,}62222$$

While the probability calculation is not feasible, namely:

$$P(\text{Not Feasible}) = \frac{17}{45} = 0{,}37778$$

After each criterion probability has been known, the next step is to calculate the value of one of the values given by the manager to determine the classification. Based on the training data in table 2.1, the alternative data 46 to 50 are classified into feasible classes. So to calculate the feasible value on alternative data 46 to 50 are as follows:

P(46|Worthy) = P(Age=Mature|Worthy) x P(Years of Service=TL|Worthy) x P(Lateness Aspect=K|Worthy) x P(Achievement =Low|Worthy)
= 0,82143 x 0 x 0,32143 x 0,28571
= 0,00000.
P(47|Worthy) = P(Age=Mature|Worthy) x P(Years of Service =TL|Worthy) x P(Lateness Aspect =K|Worthy) x P(Achievement =Low|Worthy)
= 0,82143 x 0 x 0,32143 x 0,28571
= 0,00000.
P(48|Worthy) = P(Age=Mature|Worthy) x P(Years of Service =L|Worthy) x P(Lateness Aspect =SB|Worthy) x P(Achievement =Height|Worthy)
= 0,82143 x 1 x 0,35714 x 0,71429
= 0,20955.
P(49|Worthy) = P(Age=Mature|Worthy) x P(Years of Service =L|Worthy) x P(Lateness Aspect =B|Worthy) x P(Achievement =Low|Worthy)
= 0,82143 x 1 x 0,28571 x 0,28571
= 0,06705.
P(50|Worthy) = P(Age=Mature|Worthy) x P(Years of Service =L|Worthy) x P(Lateness Aspect =C|Worthy) x P(Achievement =Height|Worthy)
= 0,82143 x 1 x 0,03571 x 0,71429
= 0,02095.

Meanwhile, to calculate the inappropriate value in the 46th data, the 50 formula used is the same as the formula to determine the appropriate value. So to get the value is done as follows:
P(46| Not Feasble) = P(Age=Mature|Not Feasble) x P(Years of Service=TL|Not Feasble) x P(Lateness Aspect=K|Not Feasble) x P(Achievement=Low|Not Feasble)
= 0,11765 x 0,76471 x 0,70588 x 0,47059
= 0,02989
P(47| Not Feasble) = P(Age=Mature|Not Feasble) x P(Years of Service=TL|Not Feasble) x P(Lateness Aspect=K|Not Feasble) x P(Achievement=Low|Not Feasble)
= 0,11765 x 0,76471 x 0,70588 x 0,47059
= 0,02989
P(48| Not Feasble) = P(Age=Mature|Not Feasble) x P(Years of Service=L|Not Feasble) x P(Lateness Aspect=SB|Not Feasble) x P(Achievement=Height|Not Feasble)
= 0,11765 x 0,23529 x 0,05882 x 0,52941
= 0,00086
P(49| Not Feasble) = P(Age=Mature|Not Feasble) x P(Years of Service=L|Not Feasble) x P(Lateness Aspect=B|Not Feasble) x P(Achievement=Low|Not Feasble)
= 0,11765 x 0,23529 x 0,11765 x 0,47059
= 0,00153
P(50| Not Feasble) = P(Age=Mature|Not Feasble) x P(Years of Service=L|Not Feasble) x P(Lateness Aspect=C|Not Feasble) x P(Achievement=Height|Not Feasble)
= 0,11765 x 0,23529 x 0,11765 x 0,52941
= 0,00172

After the appropriate and unfeasible values in the data 46 to 50 are known, the writer then calculates the maximum for each classification. Alternative data calculation 46 to 50 to calculate the maximization of feasible values, namely:
P(Worthy|C) = P(Rn|C) * P(Worthy)
= P(46|C) * P(Worthy)
= 0,00000 x 0,62222
= 0,00000
P(Worthy|C) = P(Rn|C) * P(Worthy)
= P(47|C) * P(Worthy)
= 0,00000 x 0,62222
= 0,00000
P(Worthy|C) = P(Rn|C) * P(Worthy)
= P(48|C) * P(Worthy)
= 0,20955 x 0,62222
= 0,13039
P(Worthy|C) = P(Rn|C) * P(Worthy)
= P(49|C) * P(Worthy)
= 0,06705 x 0,62222
= 0,04172
P(Worthy|C) = P(Rn|C) * P(Worthy)
= P(50|C) * P(Worthy)
= 0,02095 x 0,62222
= 0,01304

While the calculation of the maximum value is not feasible in the data manager 46 to 50, namely:
P(Not Feasible|C) = P(Rn|C) * P(Not Feasible)
= P(46|C) * P(Not Feasible)
= 0,02989 x 0,37778
= 0,01129

P(Not Feasible|C) = P(Rn|C) * P(Not Feasible)
= P(47|C) * P(Not Feasible)
= 0,02989 x 0,37778
= 0,01129

P(Not Feasible|C) = P(Rn|C) * P(Not Feasible)
= P(48|C) * P(Not Feasible)
= 0,00086 x 0,37778
= 0,00032

P(Not Feasible|C) = P(Rn|C) * P(Not Feasible)
= P(49|C) * P(Not Feasible)
= 0,00153 x 0,37778
= 0,00058

P(Not Feasible|C) = P(Rn|C) * P(Not Feasible)
= P(50|C) * P(Not Feasible)
= 0,00172 x 0,37778
= 0,00065

After calculating the maximization of the feasible and unfeasible values, then the writer compares the feasible and unfeasible values. So that it can be seen that the employee is included in the eligible or ineligible category.
R46 =Worthy >= Not Feasible
= 0,00000 >= 0,01129
= 0,01129 (Not Feasible).

R47 =Worthy >= Not Feasible
= 0,00000 >= 0,01129
= 0,01129 (Not Feasible).

R48 =Worthy >= Not Feasible
= 0,13039 >= 0,00032
= 0,13039 (Worthy).

R49 =Worthy >= Not Feasible
= 0,04172 >= 0,00058
= 0,04172 (Worthy).

R50 =Worthy >= Not Feasible
= 0,01304 >= 0,00065
= 0,01304 (Worthy).

2.2. Testing Process With RapidMiner

From the results of the above probability, 5 data will be tested and resolved using the RapidMiner tool so that it is produced with the classification results as shown in Figure 1.

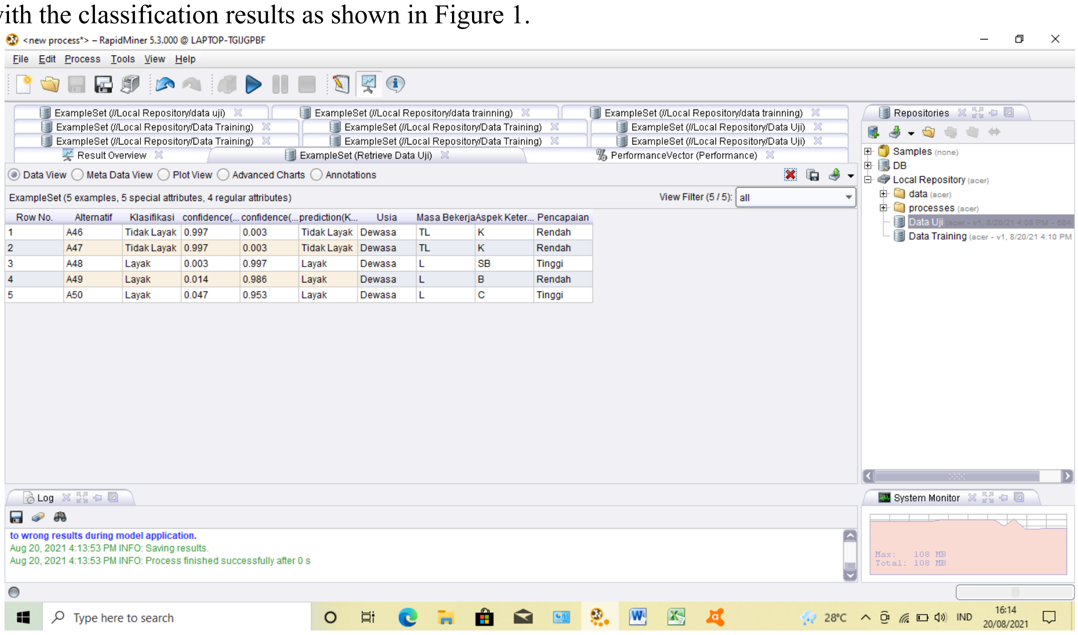

Figure 1. Calculation Results Using RapidMiner Tools

2.3. Validation Data

In conducting data validation, there are things that must be considered including: manual calculation algorithms must display the final result in the form of a decision tree, and the data used must be valid and the same as those used in the tools. The test results of the Naïve Bayes algorithm model are shown as follows:

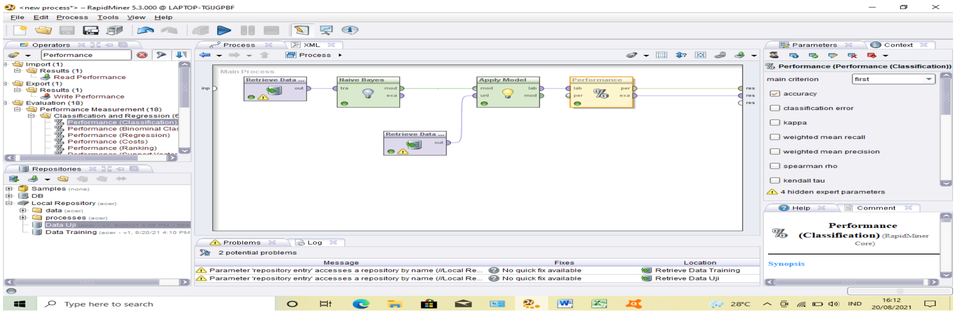

Figure 2. ***Classification Design of Training Data and Test Data***

Testing on the verification and validation side of the application, using the help of RapidMiner version 5. In processing and testing accuracy with the Naïve Bayes algorithm, the RapidMiner version 5 application can be used. After forming 5 rules where there are 3 rules that have been successfully classified with decent values, and the rest of the rules after being classified, there are as many as 2 rules with inappropriate values.

2.4. Accuracy

Accuracy is the result of how well the model correlates the results with the attributes in the data provided.

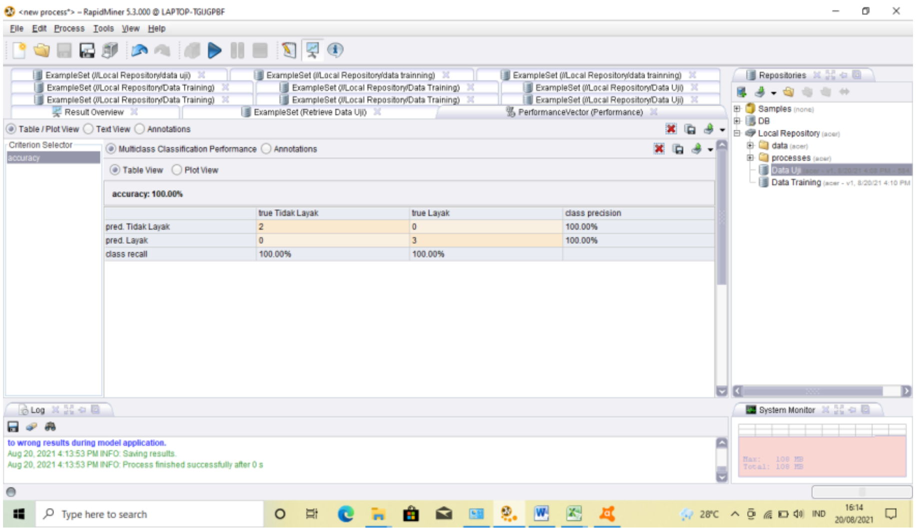

Figure 3. Accuracy Performance Value

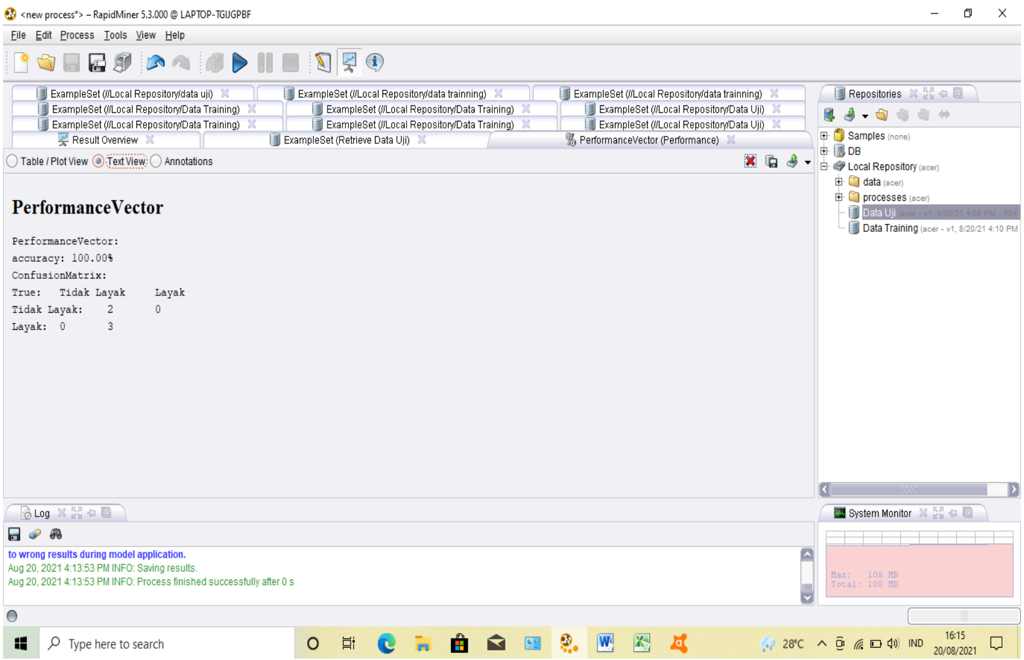

Figure 4. Detail *Performance Vector*

Based on the image that has been applied above, it can be seen that the data testing was carried out using the apply model and the % performance obtained 100% accuracy results can be categorized as a suitable method in solving security performance evaluation problems using the Naïve Bayes method.

3. Conclusion

Based on all the results of research that has been carried out on the Naïve Bayes Method in determining the extension of the employee employment contract, it can be concluded as follows The problem in determining the extension of the employee's employment contract can be solved using a data mining technique, namely the Naïve Bayes Algorithm. Based on the employee contract extension data used as training data, the Naïve Bayes method succeeded in classifying 45 data from 50 data tested. So that the Naïve Bayes method is successful in classifying employment contract extensions with an accuracy percentage of 100% and the Naïve Bayes method utilizes training data to generate the probability of each criterion for a different class, so that the probability values of these criteria can be optimized to extend the employee contract based on the classification process carried out by the Nave Bayes method itself.

References

[1] J. A. Winarsih and S. Bahri, "Sistem Pendukung Keputusan Penilaian Kinerja Karyawan Untuk Kenaikan Jabatan Pada PT. SMS Cengkareng Barat Dengan Metode Profile Matching," *Indones. J. Softw. Eng.*, vol. 4, no. 1, pp. 85–92, 2018, doi: 10.31294/ijse.v4i1.6392.

[2] Y. Kasmawati, "Human Capital Dan Kinerja Karyawan," *J. Appl. Bus. Econ.*, vol. 3, no. 4, pp. 265–280, 2017.

[3] N. P. N. E. Sari, I. N. P. Budiartha, and D. G. D. Arini, "Perlindungan Hukum terhadap Pekerja dalam Perjanjian Kerja Waktu Tertentu Menurut Undang-Undang no 13 Tahun 2003," *J. Analog. Huk.*, vol. 2, no. 1, pp. 124–128, 2020, doi: 10.22225/ah.2.1.1613.124-128.

[4] "TERMINATION OF EMPLOYMENT RELATIONSHIP Zde ň ka Gregorová," pp. 1–87.

[5] T. Winarti, D. Priyanto, and V. Vydia, "Penerapan Data Mining Menggunakan Algoritma Naive Bayes ntuk Klasifikasi Perpanjangan Kontrak Kerja Karyawan," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2013.

[6] V. N. Sari, L. Y. Astri, and E. Rasywir, "Analisis Dan Penerapan Algoritma Naive Bayes Untuk Evaluasi," *J. Ilm. Mhs. Tek. Inform.*, vol. 2, no. 1, pp. 53–68, 2020.

[7] A. W. Syaputri, E. Irwandi, and M. Mustakim, "Naïve Bayes Algorithm for Classification of Student Major's Specialization," *J. Intell. Comput. Heal. Informatics*, vol. 1, no. 1, p. 17, 2020, doi: 10.26714/jichi.v1i1.5570.

[8] A. P. Wibawa *et al.*, "Naïve Bayes Classifier for Journal Quartile Classification," vol. 7, no. 2, pp. 91–99, 2019.

[9] R. Y. Hayuningtyas, "Penerapan Algoritma Naïve Bayes untuk Rekomendasi Pakaian Wanita," vol. 6, no. 1, pp. 18–22, 2019.

[10] A. Z. A. M. Bajabir, "Penerapan metode naive bayes untuk prediksi menentukan karyawan tetap pada pt. ysp industries indonesia," 2018.

[11] M. K. Wahyu Nurjaya WK, S.T. and F. Da Cintia, "Implementasi Data Mining Untuk Prediksi Status Kontrak Kerja Karyawan Menggunakan Algoritma Naive Bayes Studi Kasus Kospin Jasa," *Lpkia*, vol. 11, no. 2, pp. 23–32, 2018.

[12] F. D. Surury, M. . Nita Novita, S.E., and M. K. Ienda Meiriska, S.Kom., "Sistem Pendukung Keputusan Perpanjangan Kontrak Kerja Karyawan Hotel Grand Inna Daira Palembang dengan Metode Topsis," *Sist. Pendukung Keputusan Perpanjangan Kontrak Kerja Karyawan Hotel Gd. Inna Daira Palembang Dengan Metod. TOPSIS*, pp. 1–8, 2018.

[13] R. Rahim *et al.*, "TOPSIS Method Application for Decision Support System in Internal Control for Selecting Best Employees," *J. Phys. Conf. Ser.*, vol. 1028, no. 1, 2018, doi: 10.1088/1742-6596/1028/1/012052.

[14] K. Fatmawati and A. P. Windarto, "Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi," *Comput. Eng. Sci. Syst. J.*, vol. 3, no. 2, p. 173, 2018, doi: 10.24114/cess.v3i2.9661.