

Application of the Association Method Using FP-Growth Algorithm to Find Pattern Of Medicine Purchasing Transactions at Pharmacy

Chanadya Yustia Purnamawati¹, Joanna Ardhyanti Mita Nugraha²(✉)

^{1,2} Informatics Department, Universitas Atma Jaya Yogyakarta, Daerah Istimewa Yogyakarta, Indonesia

¹180709844@students.uajy.ac.id, ²joanna.mita@uajy.ac.id

Article Info

Article history:

Received August 04, 2023

Revised August 14, 2023

Accepted November 15, 2023

Keywords:

Medicine

Association Method

FP-Growth

Procurement

ABSTRACT

This research was conducted with the aim of finding medicine sales transaction patterns at the Family Pharmacy so that it can be used as a reference for providing medicine stock so that service to consumers can improve. This research uses the association method by implementing the FP-Growth algorithm. Through the association method, medicine purchasing patterns will be found in the form of medicine purchased simultaneously. The data used is data from the second semester of 2021. This dataset consists of 20,266 transactions and 1,166 attributes in the form of medicine. The results of the research by setting different minimum support and minimum confidence values, obtained 7 association rules with a minimum support value of 10% and a minimum confidence of 90%. The detailed value of each rule formed is 2,683, which means the strength level of each rule is strong because it has a value of more than 1. After knowing the association rules that are formed, these rules will be used as a reference in the process of procuring medicine stock and medicine recommendations.

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



Corresponding Author:

Joanna Ardhyanti Mita Nugraha

Informatics Department, Universitas Atma Jaya Yogyakarta, Daerah Istimewa Yogyakarta, Indonesia

Email: joanna.mita@uajy.ac.id

1. INTRODUCTION

Health is a topic that cannot be separated from human life. Health itself is not only about the condition of the human body and mentality, but also things that can affect it [1]. When the body and mental state is sick, there are various ways that can be done to treat it, either by using therapy or using medicines. This depends on the disease suffered. In fact, to treat the disease, treatment can be done by combining the two methods.

Medicine is one of the health products produced by a company that aims both to prevent and treat patients [2]. Health products have various forms according to their use, namely in liquid and non-liquid forms (powder, tablets). Apart from being medicinal medicines, health products also include products that can increase stamina (supplements) to cosmetic products used by humans for purposes related to health [3]. These products can be purchased at hospitals and pharmacies either with a doctor's prescription or not. The need for medicines or other health products in each region is different because it depends on the population, types and quantities of certain medicines. If the medicine available is not the medicine that is needed, then this could have an adverse impact on the distributor and pharmacy owner. Medicine stocks will slowly decrease and the product's expiration date will get closer. If the medicine is not sold out until the expiration date, both the pharmacy and the distributor will suffer losses. Therefore, it is necessary to procure the right medicine.

Procurement is a process to meet operational needs as planned [4]. In the planning process, the activities carried out are making a list of medicines needed based on the pattern of the disease to the ability of the

community to buy medicine. Then, proceed with determining the amount of medicine stock that must be provided. After the planning process is completed, medicine procurement can then be carried out. This needs to be done so that the medicine procurement process can be directed and controlled. Good medicine procurement management will have a positive impact on pharmacies. Pharmacies can calculate the stock that is balanced with consumer needs so as to guarantee quality and control the capital issued by the pharmacy. In determining this plan, it is necessary to analyze medicine sales transactions with reference to previous sales transactions. However, if the data being analyzed is large and complex and is done manually, it will require greater effort [14].

Along with the development of technology, human work can run more effectively and efficiently. One of the rapidly growing technological developments is data mining. Data mining is a series of activities to collect, process, until information is found based on data originating from various sources [5]. The application of data mining in various sectors can make it easier to obtain information according to needs, for example finding patterns to increase sales, monitoring market direction and competition, and others. In processing data, there are several techniques or methods that have been developed in data mining, including association rules, grouping, classification, regression, prediction, and outlier detection. Each of these methods has a different function and is accompanied by the use of various algorithms.

In this study, the target object under study was medicine sales transactions at the Family Pharmacy. The transaction is examined with the aim of being able to determine the medicine stock that must be provided in the following month by utilizing the sales transaction. In order for this to be achieved, the authors use the association method in data mining. The association method or association rule is a method in data mining that is used to obtain relationships between items or attributes [6]. The relationship between these items is represented by a number of generated rules.

In finding the relationship between these items, there are various algorithms that can be utilized, one of which is the FP-Growth algorithm. The FP-Growth algorithm is a part of the frequent pattern association method functionality which is the development of the a priori algorithm [7]. This algorithm only scans at the beginning so that the time required and memory usage is smaller, while the a priori algorithm performs scans repeatedly so that the time and memory required are greater. Based on this, research will be carried out to find patterns of medicine sales transactions in the form of which medicines are purchased together so as to make it easier for the pharmacy to determine the amount of medicine stock. The transaction data used in this study are medicine sales transactions at the Family Pharmacy.

2. METHOD

The type of research applied to this research is research with quantitative methods. The quantitative research method is a systematic method and the data is presented in the form of numbers and in processing it can apply knowledge in the fields of computing, statistics, or mathematics [12]. This is in line with the research conducted by the author. This research focuses on processing transaction data and looking at the number of transactions for a product. Then, this research was only carried out by the author, the data processing was sequential.

In this study, the method used is CRISP-DM. CRISP-DM has 6 phases starting from the business understanding phase to deployment. This method was chosen because many studies use CRISP-DM. Based on 24 studies sourced from ScienceDirect, IEEE, and ACM Digital Library that most authors define CRISP-DM as the de-facto standard for implementing models in data mining [13]. In addition, this method is considered structured. Therefore, in this study, the method used is CRISP-DM.

3. RESULTS AND DISCUSSION

This study implements the CRISP-DM methodology. Here are the stages:

3.1. Business Understanding

At this stage, the things that are done are understanding the needs and goals of the business. The Family Pharmacy as the object of research is one of the pharmacies in Sangasanga District. The selling rate of the items contained in this pharmacy was high. In order for the level of sales to be carried out optimally, it is necessary to do research to find patterns of sales. When a sales pattern has been found, the pharmacy can make decisions about the progress of its business.

3.2. Data Understanding

Sales patterns can be identified from sales transaction data. The sales transaction data studied is transaction data for 6 months (July-December) in 2021. This is done to determine the procurement of medicine stocks in the second semester of 2022. The limitations of the data studied are transactions that only contain

items in the form of medicines because the sales pattern you want to know is the pattern of medicine sales. The data provided by the pharmacy is in .xls format which is divided into 12 files with 1 month divided into 2 files. Therefore it is necessary to do the sorting to make it into 1 file. Then, because the format of the data provided is not suitable for testing purposes, it is necessary to adjust the format.

3.3. Data Preparation

Data that was originally divided into 12 files combined into 1 file. Then, the required attributes are selected. There were 1166 attributes used in this study which were divided into attributes consisting of medicine sales transactions and types of medicines sold and the number of transactions was 20266 transactions. It is necessary to clean data on sales transactions because data in the form of single data and transaction data that does not include drugs do not need to be included in the dataset file. The value of the medicine attribute is labeled 0 or 1 where the label is 0 if the transaction contains the medicine and 1 if the transaction does not contain the medicine. After the data has been transferred into 1 file, then it is saved in the .xlsx format. The following Figure 1 is a snippet of the dataset display used.

| 1 | Transaksi | 19 Cair | 19 Salep | Acarbose | Acdat Cr | Acetylsist | Acifar 400 | Acifar Cr | Acnes Gel | Acnes Gel | Acnol Gel | Acnol Loti | Actifed Hi | Actifed M | Actifed O | Acyclovir | Acycl |
|----|--|---------|----------|----------|----------|------------|------------|-----------|-----------|-----------|-----------|------------|------------|-----------|-----------|-----------|-------|
| 2 | BROCHIFAR TAB, MASKER ALL, RHEMAFAR 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | DEXAHARSEN 0.75, NATRIUM DIC NOVEL | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | AMOXICILIN TAB, PARACETAMOL TAB | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | DEMACOLIN SYR, DEMACOLIN TAB, VITACIMIN DKK | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | FASIDOL TAB, IFARSYL TAB | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | AMOXICILIN TAB, OBHC PLUS DWS MENTOL 100, PAR, 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | ANTIMO TAB, HISTIGO TAB, MITU REFIL | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | GLIKOS 500, PARACETAMOL TAB, VITAMIN B COMP KT 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | GPU MINYAK 30, MKP LANG IS, TOLAK LINU ALL | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | DEXAHARSEN 0.5, MASKER DUCKBILL SENSI, MASKER : 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | COUNTERPAIN 30, DEXANTA TAB | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | IFARSYL TAB, SAFE CARE AROMATHERAPI | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | LACTOGEN 2 180G, SCOPMA PLUS, SHENLING, THROM 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | ANTANGIN CAIR, DULCOLAX ISI 4, TONIKUM SYR 330 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | FITKOM TAB JERUK, LERZIN TAB, MADU TJ SUPER 150, 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | ANACETINE, BYE FEVER BAYI | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 18 | MILORIN, WINATIN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 19 | BIO 7, CDR 10, OVISKIN N, VITAQUIN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 20 | FARIDEXON 0.5, FAXIDEN 20, PARAMEX, SANMOL FOR 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 21 | PLOSSA, VENTOLIN INH | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 22 | AMLODIPIN 10 HJ, AMLODIPIN 5 HJ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 23 | BODREX FLUBA HIJAU, BODREX TAB, GANDAPURA 60, 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 24 | GRATHAZONE TAB, MINYAK BANTENG K, PROMAG TA 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 25 | ORA CACINGAN, SANADRYL DMP 120, YUSIMOX TAB 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 26 | DEXAHARSEN 0.5, VOLTADEX | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 27 | AMLODIPIN 10 MEDIKON, CEDOCARD 10, MYLANTA T/0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 28 | BUFACARYL TAB, FARSIFEN TAB 400, FAXIDEN 20, SAM 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Figure 1. Medicine Sales Transaction Dataset

3.4. Modeling

After the data is ready, do the modeling using the RapidMiner application. Model the dataset using a predetermined algorithm, namely FP-Growth. Before that, it is necessary to add a numerical to binomial operator to change the attribute data type.

3.4.1. Prepare and import dataset

Implementation of the FP-Growth algorithm association method, the number of transactions used is 20266 transactions with a total of 1166 medicine items. In conducting research, the tools or applications used are RapidMiner Studio.

3.4.2. Design

At this stage, drag DatasetTA as the dataset that will be used to the view process section and add the necessary operators. Then set minimum support and minimum confidence. Figure 2 shows the process view on Rapidminer.

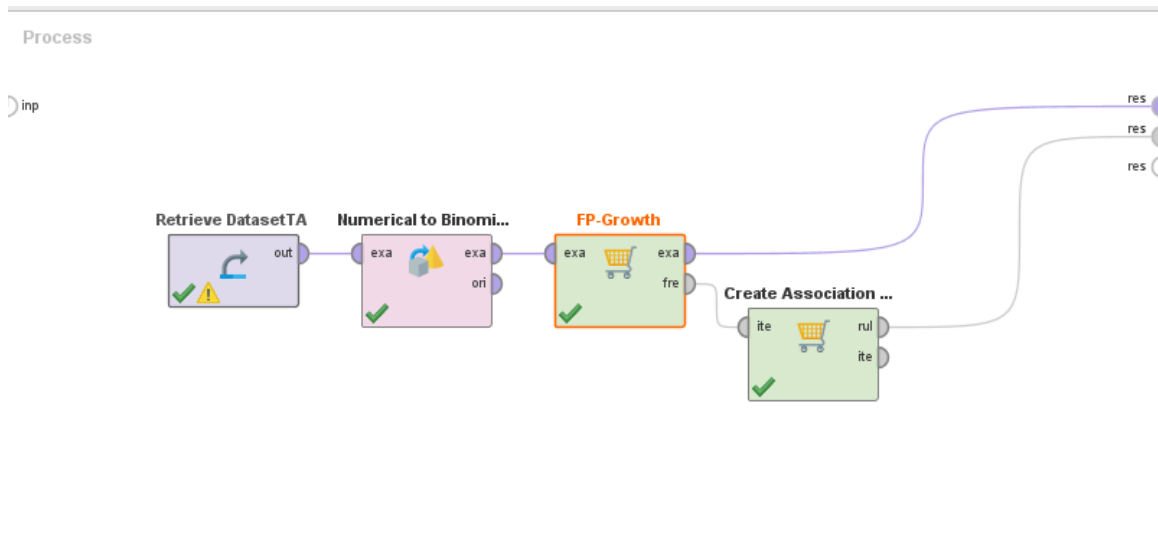


Figure 2. Process display on RapidMiner

3.4.3. Implementation Results

Click the run button then click result to see the results. The following is Figure 3 which displays the 7 rules produced.

| Premises | Conclusion | Support | Confidence | Lift |
|--------------------------------------|---------------------------|---------|------------|-------|
| Platogrix | Santagesik Tab | 0.373 | 1 | 2.683 |
| Santagesik Tab | Platogrix | 0.373 | 1 | 2.683 |
| Amlodipin 10 Medikon | Platogrix | 0.025 | 1 | 2.683 |
| Amlodipin 10 Medikon | Santagesik Tab | 0.025 | 1 | 2.683 |
| Amlodipin 10 Medikon | Platogrix, Santagesik Tab | 0.025 | 1 | 2.683 |
| Platogrix, Amlodipin 10 Medikon | Santagesik Tab | 0.025 | 1 | 2.683 |
| Santagesik Tab, Amlodipin 10 Medikon | Platogrix | 0.025 | 1 | 2.683 |

Figure 3. Frequent Item Set Results

3.5. Evaluation

The results of the association rules are formed by setting a support value of 10% and confidence of 90%. The confidence value for each rule formed is 1. This means that the level of confidence that the drug items will be purchased simultaneously is 100%. However, to see the level of strength of the association rules, you can see the lift ratio value in table 1. The lift ratio value for all the rules formed is 2.683, meaning that the level of association rules is strong because the value is more than 1. The lift ratio value itself is influenced by the confidence value and support.

Table 1. Association Rules Results

| Association Rules | Support | Confidence | Lift |
|--|---------|------------|-------|
| Platogrix → Santagesik Tab | 37.3% | 100% | 2.683 |
| Santagesik Tab → Platogrix | 37.3% | 100% | 2.683 |
| Amlodipin 10 Medikon → Platogrix | 2.5% | 100% | 2.683 |
| Amlodipin 10 Medikon → Santagesik Tab | 2.5% | 100% | 2.683 |
| Amlodipin 10 Medikon → Platogrix, Santagesik Tab | 2.5% | 100% | 2.683 |
| Platogrix, Amlodipin 10 Medikon → Santagesik Tab | 2.5% | 100% | 2.683 |
| Santagesik Tab, Amlodipin 10 Medikon → Platogrix | 2.5% | 100% | 2.683 |

3.6. Deployments

The results that come out are in the form of rules which can then be used as a reference for decision making. After the results come out, the rules that are formed are then presented in the form of a report containing an explanation of the knowledge gained from these rules. Here are 7 generated rules that can be applied by pharmacy owners in making decisions:

1. If consumers buy Platogrix, it will be recommended to buy Santagesik Tab.
2. If consumers buy Santagesik Tab, then it will be recommended to buy Platogrix.
3. If consumers buy Amlodipine 10 Medikon, it will be recommended to buy Platogrix.
4. If you buy Amlodipine 10 Medikon, it will be recommended to buy Santagesik Tab.
5. If consumers buy Amlodipine 10 Medikon, it will be recommended to buy Platogrix and Santagesik Tab.
6. If consumers buy Platogrix and Amlodipine 10 Medikon, it will be recommended to buy Santagesik Tab.
7. If consumers buy Santagesik Tabs and Amlodipine 10 Medikon, it will be recommended to buy Platogrix.

4. CONCLUSION

Based on the research obtained from testing the dataset at the Family Pharmacy, it can be concluded that: (1) Transaction patterns that are formed based on testing by implementing the FP-Growth algorithm using the RapidMiner application produce 7 rules that can help in finding relationships or relationships with medicines purchased by consumers so that the owner pharmacies can find out which medicines have a high level of sales for the purposes of medicine procurement. (2) The strength of the rules that are formed can be seen through the lift ratio value in each rule. In this study, the lift ratio value obtained for each rule is 2,683, meaning that the strength level of each rule is strong because it has a value of more than 1.

REFERENCES

- [1] S. D. Krisna Triyono and Y. K. Herdiyanto, "Konsep Sehat Dan Sakit Pada Individu Dengan Urolithiasis (Kencing Batu) Di Kabupaten Klungkung, Bali," *J. Psikol. Udayana*, vol. 4, no. 02, p. 263, 2018, doi: 10.24843/jpu.2017.v04.i02.p04.
- [2] Anonim, "Kesehatan," *Nano Natura*. <https://nanotechnatura.com/kesehatan/>.
- [3] Y. Xu *et al.*, "Retrospective study of reported adverse events due to complementary health products in Singapore from 2010 to 2016," *Front. Med.*, vol. 5, no. JUN, pp. 1–14, 2018, doi: 10.3389/fmed.2018.00167.
- [4] A. Drs. Medianto Henky Saputra, "Pengelolaan Persediaan Obat di Apotek," *Supply Chain Indonesia*, 2020. <https://supplychainindonesia.com/pengelolaan-persediaan-obat-di-apotek/>.
- [5] B. Anggraeni, Hapsari Dita, Saputra, Ragisl, Noranita, "Data Mining Data mining," *Min. Massive Datasets*, vol. 2, no. January 2013, pp. 5–20, 2005, [Online]. Available: https://www.cambridge.org/core/product/identifier/CBO9781139058452A007/type/book_part.
- [6] R. R. Ramadina, T. H. Pudjiantoro, and I. Santikarama, "Pembangunan Sistem Customer Relationship Management (CRM) Menggunakan Metode Asosiasi Algoritma Apriori," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 2, no. 2, p. 78, 2020, doi: 10.36499/jinrpl.v2i2.3539.
- [7] R. Fitria, W. Nengsih, and D. H. Qudsi, "Implementasi Algoritma FP-Growth Dalam Penentuan Pola Hubungan Kecelakaan Lalu Lintas," *J. Sist. Inf.*, vol. 13, no. 2, p. 118, 2017, doi: 10.21609/jsi.v13i2.551.
- [8] M. Badrul, "Algoritma Asosiasi Dengan Algoritma Apriori Untuk Analisa Data Penjualan," *J. Pilar Nusa Mandiri*, vol. XII, no. 2, pp. 121–129, 2016, [Online]. Available: <http://ejournal.nusamandiri.ac.id/index.php/pilar/article/view/266>.
- [9] D. Fitriati and M. Hardiyanto, "Perbandingan Algoritma Apriori Dan Algoritma Fp-Growth Untuk Mengetahui Pola Penggunaan Transportasi Online," *Snatif*, p. 9, 2018.
- [10] E. Muningsih, "Penentuan rekomendasi produk dengan metode data mining asosiasi generalized sequence pattern (gsp)," pp. 218–224, 2016.
- [11] D. Saadah, Rabiatus, "Analisa Asosiasi Data Mining Penjualan Meubel Menggunakan Algoritma Apriori Pada Master Borneo Pontianak Selatan," vol. 13, no. 2, pp. 31–39, 2020.
- [12] Syafnidawaty, "PENELITIAN KUALITATIF," *raharja.ac.id*, 2020. <https://raharja.ac.id/2020/10/29/penelitian-kualitatif/> (accessed Apr. 21, 2020).
- [13] C. Schröer, F. Kruse, and J. M. Gómez, "A systematic literature review on applying CRISP-DM process model," *Procedia Comput. Sci.*, vol. 181, no. 2019, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.
- [14] Yunarto, Holy Icutan dan Martinus Getty Santika, "Business Concepts Implementation Series in Inventory Management", Jakarta: Elex Media Komputindo, 2005.

BIOGRAPHIES OF AUTHORS

Chanadya Yustia Purnamawati was born in 2000 and received her Bachelor of Computer Science in Informatics Universitas Atma Jaya Yogyakarta, Indonesia in 2022. Her area of interest includes artificial intelligence and data mining.



Joanna Ardhyanti Mita Nugraha was born in 1991 and received her Master of Computer in Information System Universitas Diponegoro Semarang, Indonesia in 2017. Her area of interest includes artificial intelligence and data mining. She is a lecturer of Information Engineering Department in Universitas Atma Jaya Yogyakarta since 2018.