Optimization Of The Decision Tree Algorithm Used Particle Swarm Optimization In The Selection Of Digital Payments

Optimization Of The Decision Tree Algorithm Used Particle Swarm Optimization In The Selection Of Digital Payments

I Ariyati[1]*, S Rosyida[2], K Ramanda[3], V Riyanto[4], S Faizah[5] and Ridwansyah[6]

1 Sistem Informasi, Universitas Bina Sarana Informatika
2,3,4,5,6 Teknik Informatika, Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri

E-mail: indah.ayi@bsi.ac.id

Abstract. One of the developments of information technology in Indonesia was the developments in Fintech (financial technology) that made it easy for people to access financial products, facilitated online transactions and also increased financial literacy. The development of fintech occurs when the use of cash was reduced to cashless when making payment transactions so that transactions would be more practical, easy, safe and comfortable. The purpose of this study to improve the quality of decision tree modeling and accuracy in the selection of digital payments. This research was focused on determining the fintech applications that were widely used by the people of Indonesia, namely Gopay or Ovo by comparing the advantages of applications and features of each fintech application. Optimization method of Decision Tree Algorithm (C4.5) and Particle Swarm Optimization by selecting several attributes including the level of ease of use, data security, application trust and convenience, maximum balance increase, discounts, cashback, ease of top-ups, range of existing merchants, return the money and customer complaint services. The results of the development of a decision tree algorithm based on particle swarm optimization provide a good classification and increase the validation value in the selection of digital payments.

1. Introduction

Internet and smartphone usage changed people's behavior when shopping (Andretti Abdillah et al., 2018). This happened because the use of cash began to decrease and preferred to make payment transactions with technology so that transactions were carried out more practically, easily, safely and comfortably (Prabaningtyas, Surjandari, & Laoh, 2019). The public was aware of e-money, transfers between accounts and digital wallets in the smartphone application. This transaction trend changed the world economy and Indonesia into a digital economy (Leon A. Abdillah, 2019) known as Financial Technology (Fintech). The fintech growth to continue to show significant results (Aristiana, Waluyo, & Muchtasib, 2019). This provided an opportunity for cellular payment service providers (Chandra, Kristin, Suhartono, Sutarto, & Sung, 2018) to develop in the Fintech business. With this fintech made it easier for people to make buying and selling transactions and bill payments, because cellular payments provided a lot of convenience in the transaction (Prabaningtyas et al., 2019) and reduced cost and safer because it could be done from home.

Gojek and Grab were a company engaged in the field of online transportation. The competition between the two companies was very tight because they had almost the same service in nearby locations. So that each company was required to maintain the quality of service so that customers continue to use it. This could be seen from the value given after completing the transaction and using the service, the payment method that offered non-cash payments made into the Fintech that was leading the way in facilitating daily activities (Susilo, Iksan Prabowo, Taman, Pustikaningsih, & Samlawi, 2019). To optimize digital payments provided by Gojek, Gopay. Gopay was one of the Fintech products brought by Gojek to support and implement financial technology in providing services (Danuarta & Darma, 2019). Ovo was also the same as Gopay, one of the Fintech products brought by Grab. Today, Gopay and Ovo transactions were not only in the Gojek and Grab applications, but by other offline and online merchants that could be used for digital payments. Transactions at all merchants marked with Gopay and Ovo we could make transactions quickly, safely and comfortably, in addition to that the system also provided points, cashback, discounts from every purchase.

Decision Tree algorithm (C4.5) had an appeal to be implemented in various applications because this algorithm was simple in its application, (Rosandy, 2016). The decision tree had advantages in choosing features from different internal nodes. The decision tree was quite flexible, in the same node, it would distinguish a criterion compared to other criteria for the selected feature (Syamsu and Wijaya, 2019). The advantage of PSO was that parameters could be adjusted and more easily applied (Widiastuti, Santosa and Supriyanto, 2014), and had more or even superior comparisons in the search for performance for many optimization problems with faster and more stable convergence rates (Park, Lee and Choi, 2009). PSO was proven to be an algorithm whose performance was consistent, good and effective for optimization problems because it was easier in applying the code (Dhanasaputra and Santosa, 2015).

Our research focused on Gopay and Ovo as fintech applications that are often used by Indonesians. This study compared the advantages of applications and features of each application, using the Decision Tree (C4.5) algorithm and Particle Swarm Optimization researchers compared the level of ease of use, data security, application confidence and comfort, increased maximum balance, discounts and cashback, ease in topping up, the range of existing merchants, refunds, services to complaints from customers.

Previous studies did not use the data mining process of digital payment data and the absence of optimization to maximize the model of digital payment data. Therefore this study used a data mining process with Decision Tree and Particle Swarm Optimization models. The novelty of this research was in the dataset used. The dataset had never been used in data mining research with decision tree algorithms and particle swarm optimization.

2. Methodology

Decision Tree Algorithm (C4.5) process flow contained in the data mining processes to make a decision tree by selecting attributes as the root node and making branches for each value after that divided cases in branches and doing the same process for each branch until all cases in the branch had the same class. The Gain value calculation was used as the formula in Equation 1, while the Entropy value calculation could be seen in Equation 3.

$$Gain(S.A) = \frac{Gain(S.A)}{Si(S.A)} \quad \text{..... (1)}$$

$$Si = -\sum_{1}^{n} -pi\ log_2 pi \quad \text{..... (2)}$$

$$Entropy(S) = -Sin \sum_{i=1}^{n} Si/Slog_2 pi\ Si\ s\ c\ i = 1 \quad \text{..... (3)}$$

Particle Swarm Optimization (PSO) updated iteration to iteration by using a population of particles denoted as p size. To find the optimal solution, each particle was changed by direction in its search based on two factors: pbest and gbest. Each particle represented the position of the candidate (solution). A particle was considered as a point in D-dimensional space, and its status was marked based on its position and velocity. To find the optimal solution, each particle changed its speed according to the cognitive and social parts.

The subjects in this study were users of Ovo and Go-Pay digital money. The number of questionnaires was 750 respondents with objects of digital payment service research through Ovo and Go-Pay. Data collection techniques in the form of techniques answered the forms of structured questions by the criteria through questionnaires. The results of processing the questionnaire as a source of data to produce accuracy in the selection of electronic payments from various attributes. The research used in data processing focuses on determining the Fintech application used, namely Gopay or Ovo. The form of questionnaire questions by comparing the advantages of the application and the features of each application are used as fintech attributes, that are the level of ease of use of the application, security of user data, trust and convenience of the application, the maximum balance increased during application updates, transaction discounts and cashback transactions, ease in topping up, reach of collaborating merchants, refunds, service to complaints customer. Based on the questionnaire distributed, we processed and tested it with the particle swarm optimization algorithm which was used to improve accuracy. The algorithm proposed by the decision tree model in optimization with particle swarm optimization could be seen in Fig 1.

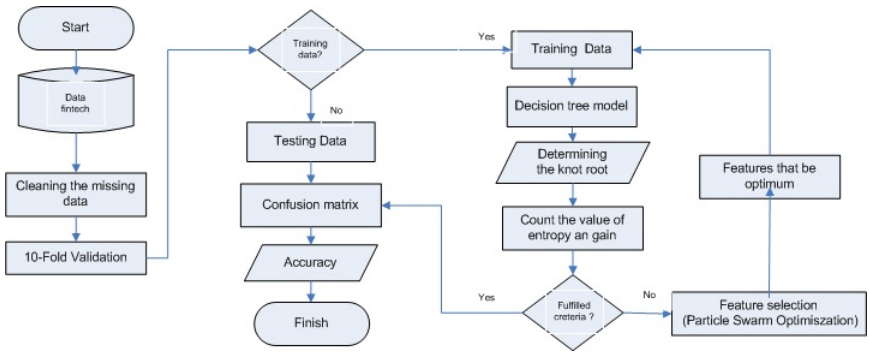

Figure 1. Proposed Model in Choosing Digital Payments

Fig 1 was a research model that started from collecting data in the form of attributes used in the process of selecting digital payments. Attributes were collected into a single fintech dataset, then performed data cleaning by eliminating blank data or data that do not fit the criteria. The model was built with Particle Swarm Optimization. Validation would be performed immediately using 10-fold cross-validation. The fintech dataset was divided into 2, namely testing data and training data. To produce the best model, testing was done by adding the decision tree model C.45 algorithm. Adding a decision tree model would determine the defining attribute as the root node of the other attributes. The determination of the root node was based on the calculation of the entropy value and the gain value of each attribute. The results of the dataset validation by adding the decision tree model were fulfilled then the accuracy was met by producing an efficient and effective feature but, if the accuracy criteria were not met, split validation would be performed to optimize the features of each attribute

3. Result and Analysis

3.1. Decision Tree Algorithm

The C45 decision tree algorithm was a data mining technique that converts very large facts into a decision tree that represents the rules. The process in decision trees changed the shape of data (tables) into tree models, changed the tree model to the rules, and simplify rules. In the preparation of the first decision tree that must be done was determining the attributes to be the root node and other attributes that would be the next node with steps that could be calculated entropy value and the gain value of each attribute, the results of the calculation of entropy and gain values could be seen from the table 1.

Table 1. Gain Values

Knot	Gain
Gender	1,14882E-05
Job Category	1,18954645
Ease of Application	0,11034921
Data Security	0,055174552
Trust in Comfort	0,10016227
Increased Maximum Balance	0,859414434
Discount	1,167752198
Cashback	1,181146456
Topup Facilities	0,149257111
Merchant Coverage	1,259037132
Refunds	0,903045941
Customer Service	1,148634633

From the calculation of entropy and the highest gain was found in the merchant coverage attribute with a value of 1.259037132 as a root in the selection of digital payments. Then obtained a decision tree that could be seen in Fig 2.

The results of the decision tree obtained confusion matrix which could be seen in table 2. From table 2 it could be calculated with the calculation formula of accuracy, precision and recall, and obtained an accuracy value of 80.00%, a precision of 81.99% and a recall value of 91.70% with the rule that the predicted choice of Ovo is 36 people, but from the results of the DT model were predicted to choose Gopay and vice versa which was predicted by Gopay 114 but Ovo prediction results. After obtaining the results of accuracy, precision and recall, the AUC value of 0.851 was obtained.

3.2. Particle Swarm Optimization

From the results of the decision tree, the method was used to optimize the decision tree model by using particle swarm optimization, after the feature was selected, we obtained attributes that had no weight that could be seen in Fig 3.

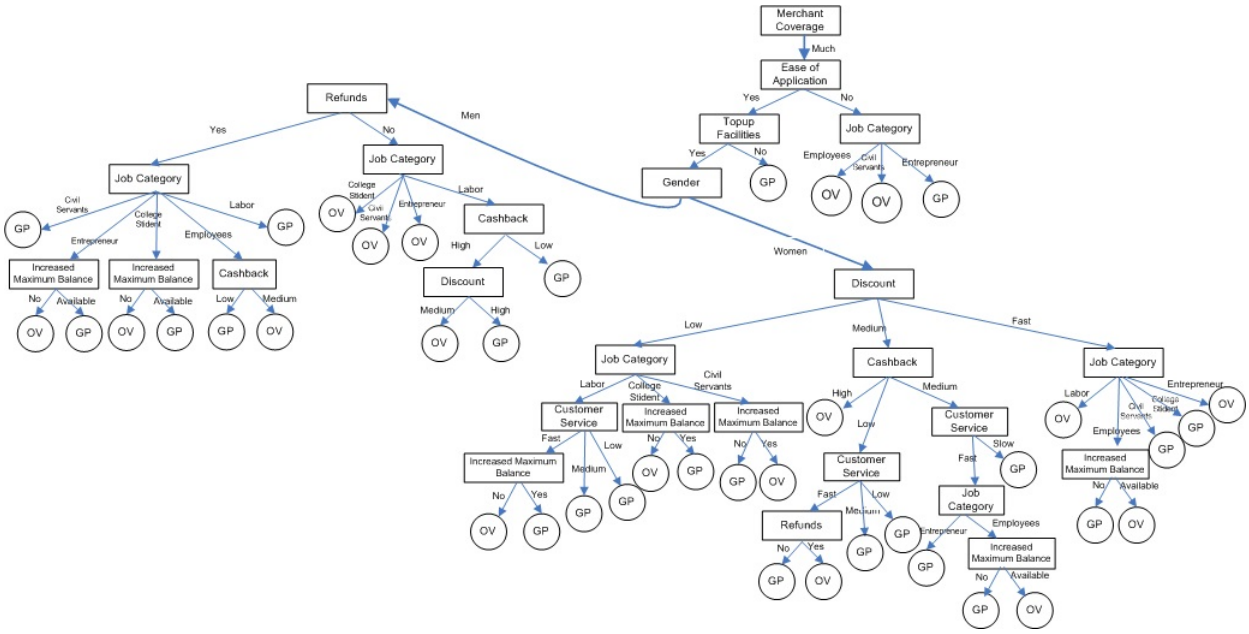

Figure 2. Decision tree model in The Digital Payment Selection

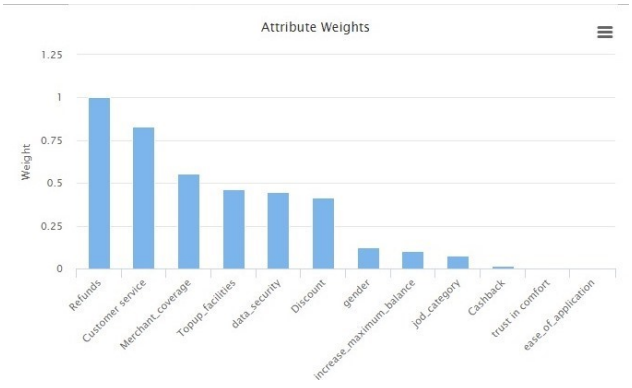

Figure 3. Weight of Attributes of Decision Tree Method and Particle Swarm Optimization

From Fig 3 it could be seen that among the attributes there were several attributes that could influence and there were two attributes that did not affect including trust and ease of application. After finding the attributes that did not affect the confusion matrix obtained from the DT and PSO models which could be seen in table 3.

Table 2. Confusion Matrix DT

Information	Ovo	Gopay
Prediction Yes	204	36
Prediction No	114	396

Table 3. Confusion Matrix DT used PSO

Information	Ovo	Gopay
Prediction Yes	283	40
Prediction No	35	392

From table 3 it could be calculated with the calculation formula of accuracy, precision and recall, and obtained an accuracy value of 90.00%, a precision of 92.07% and a recall value of 90.76% with the rule that the predicted choice of Ovo 40 people, but from The results of the DT and PSO models were predicted to choose Gopay and vice versa which were predicted in Gopay 35 but Ovo prediction results. After obtaining the results of accuracy, precision and recall, the AUC value is 0.971

4. Conclusion

The results of the decision tree algorithm based on particle swarm optimization provided a good classification and increased the validation value so that it could be used as a digital payment selection. The results of the accuracy got an accuracy rate of 90% increased by 10%, precision by 92.07% increased by 0.37% and obtained an AUC value of 0.971 increased by 0.12. In this study, we did not take sides between Gopay or Ovo because both were fintech as a result of technological developments and innovations in Indonesia to meet the needs of the public to carry out financial activities online and facilitate public access to these services. As users, we could not choose the best fintech between the two because the two Fintech complemented each other.

References

[1] L. Andretti Abdillah, D. Napitupulu, J. Simarmata, R. Rahim, D. Abdullah, and A. Pranolo, "The Role of Big Data and Cloud Computing for Technology Based Business," vol. 144, pp. 183–187, 2018.

[2] N. I. Prabaningtyas, I. Surjandari, and E. Laoh, "Mining customers opinion on services and applications of mobile payment companies in Indonesia using sentiment analysis approach," 2019 16th Int. Conf. Serv. Syst. Serv. Manag. ICSSSM 2019, no. 18, pp. 1–5, 2019.

[3] Leon A. Abdillah, "An Overview of Indonesian Fintech Application," 1st Int. Conf. Commun. Inf. Technol. Youth Study, pp. 8–16, 2019.

[4] M. Aristiana, B. Waluyo, and A. B. Muchtasib, "Factors that Influence People Interest in Using a Digital Platform as a ZIS Payment (Case Study Go-Pay)," Proceedings, Int. Conf. Zakat 2019, no. 2017, pp. 50–59, 2019.

[5] Y. U. Chandra, D. M. Kristin, J. Suhartono, F. S. Sutarto, and M. Sung, "Analysis of Determinant Factors of User Acceptance of Mobile Payment System in Indonesia (A Case Study of Go-Pay Mobile Payment)," Proc. 2018 Int. Conf. Inf. Manag. Technol. ICIMTech 2018, no. September, pp. 454–459, 2018

[6] A. Z. Susilo, M. Iksan Prabowo, A. Taman, A. Pustikaningsih, and A. Samlawi, "A comparative study of factors affecting user acceptance of go-pay and OVo as a feature of Fintech application," Procedia Comput. Sci., vol. 161, pp. 876–884, 2019.

[7] G. L. N. Danuarta and G. S. Darma, "Determinants of Using Go-Pay and its Impact on Net Benefits," Int. J. Innov. Sci. Res. Technol., vol. 4, no. 11, pp. 173–182, 2019.

[8] T. Rosandy, "PERBANDINGAN METODE NAIVE BAYES CLASSIFIER DENGAN METODE DECISION TREE (C4.5) UNTUK MENGANALISA KELANCARAN PEMBIAYAAN (Study Kasus : KSPP / BMT AL-FADHILA)",J. Teknol. Inf. Magister Darmajaya, vol. 2, no. 01, pp. 52–62, 2016.

[9] S. Syamsu, M. Muhajirin, and N. S. Wijaya, "Rules Generation Untuk Klasifikasi Data Bakat dan Minat Berdasarkan Rumpun Ilmu Dengan Decision Tree," Inspir. J. Teknol. Inf. dan Komun., vol. 9, no. 1, p. 40, 2019.

[10] N. A. Widiastuti, S. Santoso, and C. Supriyanto, "ALGORITMA KLASIFIKASI DATA MINING NAÏVE BAYES BERBASIS PARTICLE SWARM OPTIMIZATION UNTUK DETEKSI PENYAKIT JANTUNG PARTICLE SWARM OPTIMIZATION UNTUK DETEKSI PENYAKIT JANTUNG," J. Pseudocode, vol. 1, no. 1, p. 11 14, 2014.

[11] T. S. Park, J. H. Lee, and B. Choi, "Optimization for artificial neural network with adaptive inertial weight of particle swarm optimization," Proc. 2009 8th IEEE Int. Conf. Cogn. Informatics, ICCI 2009, pp. 481–485, 2009.

[12] N. Dhanasaputra, B. Santosa, and J. T. Industri, "Pengembangan Algoritma Cat Swarm Optimization (Cso) Untuk Klasifikasi," no. Cat Swarm Optimization, Klasifikasi, Multiple Regression Linear Mode, p. 18, 2015.