

Prediction Model of Production Completion Delay to Improve Service Quality Using Decision Tree Versus Multilayer Perceptron Method

Arko Djajadi^{*1}, Winarno², Abdullah Dwi Srenggini³

^{1,2}Lecture Master of Information System Department, Universitas Raharja

³Student Master of Information System Department, Universitas Raharja

E-mail: ^{*1}arkodjajadi@raharja.info, ²winarno1060@yahoo.com, ³abdullah.d@raharja.info

Abstract

Delays in the completion of pvd production can be caused by several factors. Including the actual experience in the production of the difficulty of each process and color type, even the difficulty of the product type can also be affected. In this study, the prediction of the delay in the completion of pvd production was carried out using the C4.5 decision tree and Multilayer Perceptron data mining method approach using Production Results data at PT. Surya Toto Indonesia, whose results are expected to provide information and input for the company in making production plans in the future. The data testing method was carried out with 5 (five) testing times with different amounts of data to determine the level of consistency of accuracy obtained. C4.5 gives the results of a decision tree where the root is the color type and as the leaf is the product category, type type and order period. The average value of accuracy generated in the C4.5 decision tree method is 87.15%. While the Multilayer Perceptron obtained an average accuracy of 87.98%, which is greater than the decision tree method with a difference of 0.83%.

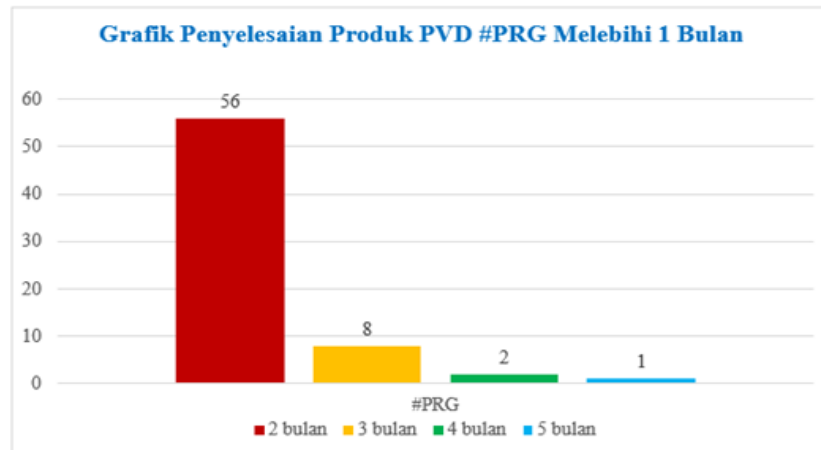
Keywords — Monitoring, Expenses, Operational

1. INTRODUCTION

In order to fulfill orders and to complete production from January 2020 to July 2021, this PVD category product is the category with the highest number of production that does not reach the target (gap) when compared to other categories. It can be seen in Table 1 and Figure 1. Table 1 Gap between Plans and Production Results by Category.

Table 1. Gap between Plans and Production Results by Category

No	Kategori Produk	Periode Jan'20 s/d Jul'21			
		Product Category	Production Target Amount	Gap	Gap Percentage
1	Chrome	2844	2732	112	4%
2	Painting	337	288	49	15%
3	Plastic	491	435	56	11%
4	PVD	736	545	191	26%
Total		4408	4000	408	9%

Figure 1. #PRG PVD Product Completion Chart

(Sumber : PT. Surya Toto Indonesia, Tbk)

From Table 1. it can be seen that the #PRG Polished Rose Gold color is the color category with the highest number of delays in completing production, resulting in delays in the assembly process or delivery process. For that PT. Surya Toto Indonesia, Tbk needs a system to be able to overcome these problems so that later they will not lose competitiveness with other companies or competitors in accelerating delivery.

Improve the performance of existing algorithms in terms of saving time by using rules and increasing the efficiency of the decision tree, but also improving the information generated [1]. Laksmi, (2013) in his research conducted a discussion by comparing the performance of ID3, C4.5 and CART algorithms [2]. Prediction is "thinking about a quantity, for example the demand for one or more products in the future period. In essence, the prediction is only an estimate (forecast) [3]. Prediction can be interpreted as an estimate which is a process in which an estimate is made about something that is most likely to happen in the future based on past or present information, so that errors (differences between something that has happened and the forecast results) can be minimized. [4].

"Products can also be defined as physical, psychological, and symbolic attributes that can produce satisfaction and benefits for consumers" [5]. "Product is anything that can fulfill and satisfy human needs and wants, both tangible or intangible or services" [6]. According to Pandensia et al.(2017) in Liviona and Sinatra,(2018) service quality is "an effort to meet customer expectations in accordance with their needs and desires. [7] So it can be concluded that the quality of service is an action taken to fulfill the customer's desire for the commitment that we give to the customer. Physical Vapor Deposition (PVD), where the coating material is solid (Solid) using a High vacuum. And the manufacture of metal atoms by evaporation, sputter and ion bombardment methods, at a temperature of 500°C. known as the "Cold Process". Physical Vapor Deposition (PVD) is a subset of Vacuum coating technology, a general term used to describe the various methods for depositing thin films by condensation that is evaporated from the material onto various surfaces.

"Data mining is a series of processes, data mining can be divided into several stages. These stages are interactive in which the user is directly involved or through a knowledge base" [8]. "Data mining is a process that uses statistical, mathematical, artificial intelligence,

and machine learning techniques to extract and identify useful information and related knowledge from large databases” [9]. Some of the classification methods commonly used in data mining are: decision tree classifier, Bayesian classifier, classifying k-nearest neighbors, case-based reasoning, genetic algorithms and fuzzy logic techniques [10]. The C4.5 algorithm is one of the classification techniques in machine learning that is used in the data mining process by forming a decision tree which is represented in the form of rules [11]. Meanwhile, according to (Kamagi and Seng, 2014) is a classification algorithm that can produce a decision tree that is easy to interpret, has an acceptable level of accuracy, and is efficient in handling discrete and numeric type attributes [12]. Decision tree is a classification methodology, where the classification process is modeled using a set of hierarchical decisions on feature variables, arranged in a tree-like structure.” [13]. This method can divide (training data) into a smaller set of records by applying a series of decision rules (rules), so that the members of the result set become similar to each other by paying attention to the target variable [14].

Decision tree is one method that can be used to classify a team of objects or data to produce a decision [15]. The C4.5 algorithm is an algorithm used to form a decision tree. The advantage of the C4.5 algorithm is that it can produce a decision tree that is easy to interpret, can handle numeric (continuous) and discrete data [16]. Supervised Learning is a widely used technique. This technique is the same as "programming by example". This technique involves a training phase in which historical training data whose characters are assigned to known results and processed into data mining algorithms [17]. The multilayer neural network method is a backpropagation algorithm that uses a learning rule gradient descent. This algorithm is very useful, quite reliable and easy to understand. This algorithm is also capable of fixing the weights on the hidden layer [18].

2. RESEARCH METHOD

Classification Data mining is a learning method, to predict the value of a group of attributes in describing and distinguishing data classes or concepts that aim to predict the class of objects whose class label is unknown [19]. Data mining is the process of employing one or more machine learning techniques to analyze and extract knowledge automatically. In generating knowledge from existing patterns, it is necessary to apply a scientific method called Knowledge Discovery in Database (KDD) [20].

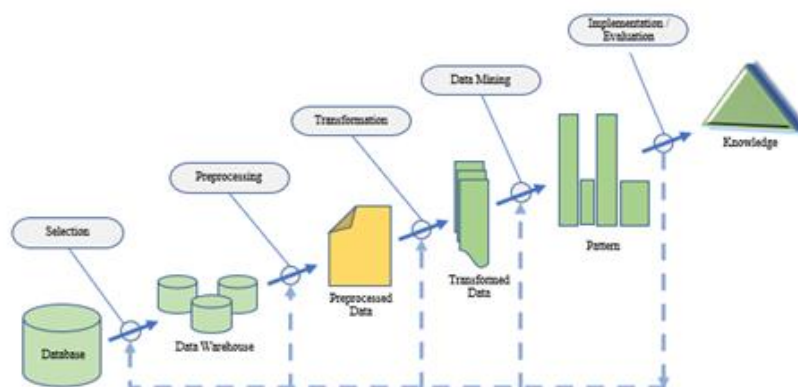


Figure 2. Data Mining Stages

Han's opinion in the publication of his book entitled Data Mining Concepts and Techniques Third Edition, there are several points in the classification process:

1. Learning model: In the early stages, the training data is analyzed by a classification algorithm, the class attributes are concluded as loan decisions and the classifier is represented as a form of classification rule.
2. Classification: This phase is a classifier that has been designed to be used to classify data.

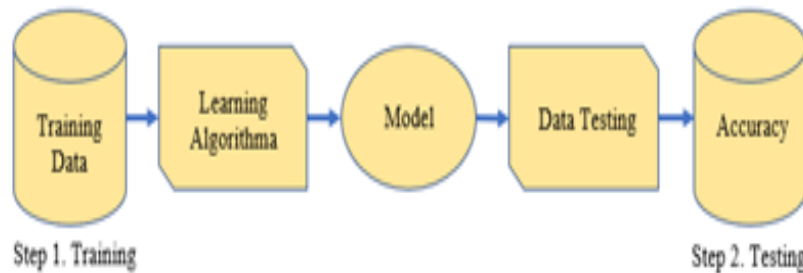


Figure 3. Stages in Classification in General

Likewise in this case, the process is carried out to build a more efficient classification segmentation process than before. The segmentation process of the decision tree and multilayer perceptron algorithms predicts the data that will be the decision. Figure 5 is an image of the prediction process flow, which is a top-level diagram that describes the incoming and outgoing data flows from the pvd production completion prediction aid program using the decision tree and multilayer perceptron methods. Can be seen in the image below:

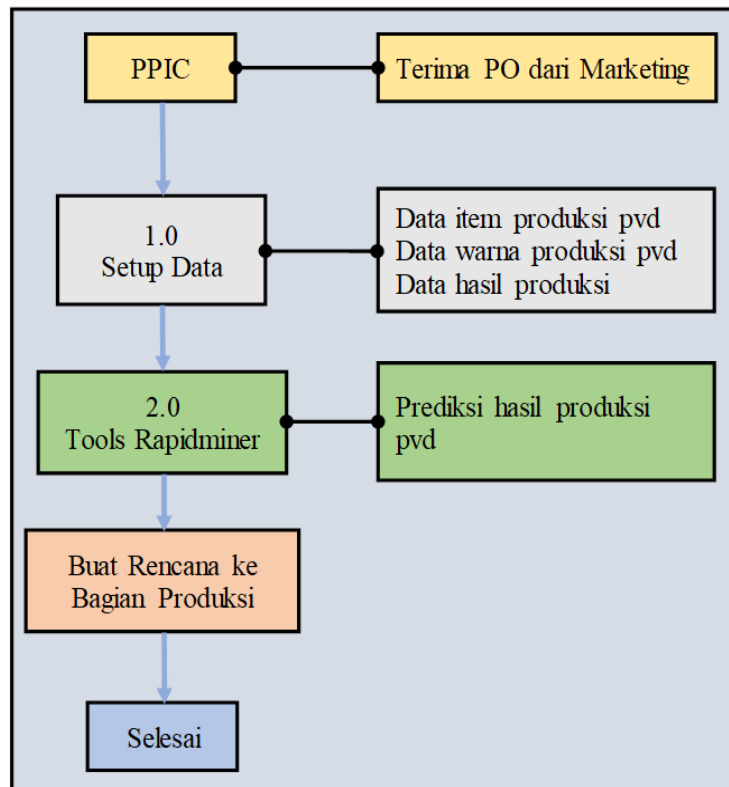


Figure 4. Prediction Process Flow

3. RESEARCH RESULTS AND DISCUSSION

This study uses data from the completion of PVD Coating production data as much as 584 data as training data and 256 data as testing data. The purpose of this research is to find a rule tree as a reference for predicting the completion of pvd production so that it is on time according to the production plan and can minimize delays in completing production. This study tested the data with 5 (five) times of testing with different amounts of data to determine the level of consistency of accuracy obtained. C4.5 gives the results of a decision tree where the root is the color type and as the leaf is the product category, type type and order period. Furthermore, correlation test is used to measure which level of accuracy is the highest between the decision tree and multilayer perceptron algorithms. The following is the testing data used:

Table 2. Testing Data Table

No	Nama Tipe	Kategori Produk	Jenis Tipe	Periode Order / Bulan	Periode Order / Tahun	Nama Produk	Jenis Produk	Kode Produk	Material	Asal Produk	Masalah Kualitas	Kesulitan Nilai La,b	Planning	Warna	Keterangan
1	NON SERIES	BATH & SHOWER	TIPE B 6	January	2020	Triangle Handle	Part Sedang	S32034	Metal	LOCAL	Ya	Sedang	1 bulan	#PRG	terlambat
2	NON SERIES	BATH & SHOWER	TIPE B 6	January	2020	Cap Nut	Part Sedang	S23018	Metal	LOCAL	Ya	Sedang	1 bulan	#PRG	terlambat
3	NON SERIES	BATH & SHOWER	TIPE B 6	January	2020	Spindle	Part Sedang	S31084	Metal	LOCAL	Tidak	Sedang	1 bulan	#PRG	terlambat
4	NON SERIES	BATH & SHOWER	TIPE B 6	January	2020	Flexible	Part Sedang	N40012	Plastik	IMPORT	Ya	Ya	1 bulan	#PRG	terlambat
5	NON SERIES	BATH & SHOWER	TIPE B 6	January	2020	Body	Part Sedang	S11017	Metal	LOCAL	Ya	Sedang	1 bulan	#PRG	terlambat
6	NON SERIES	BATH & SHOWER	TIPE B 6	January	2020	Character Bis	Part Sedang	S62050	Metal	LOCAL	Ya	Sedang	1 bulan	#PRG	terlambat
7	NON SERIES	BATH & SHOWER	TIPE B 6	January	2020	Shower Spray	Part Sedang	S90030	Plastik	LOCAL	Ya	Ya	1 bulan	#PRG	terlambat
8	NON SERIES	BATH & SHOWER	TIPE B 6	January	2020	Cover	Part Kecil	S90108	Plastik	LOCAL	Ya	Ya	1 bulan	#PRG	terlambat

After analyzing the data requirements and methods, the next step is to test the C4.5 decision tree and multilayer perceptron methods. The testing process using the C4.5 decision tree method can be shown in Figure 5. While the testing process using the multilayer perceptron neural network method can be shown in Figure 6.

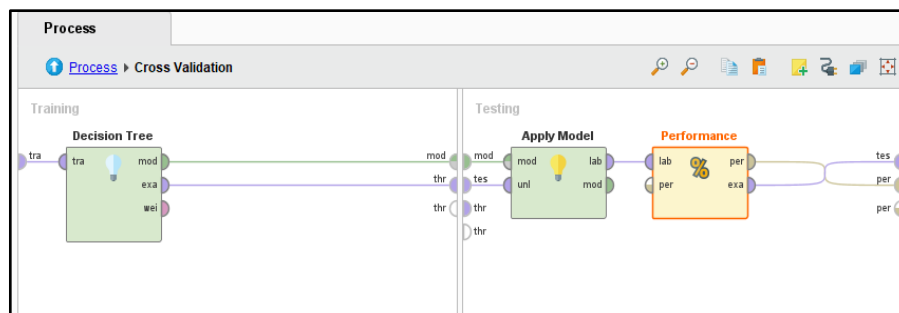


Figure 5. Testing process using decision tree C4.5

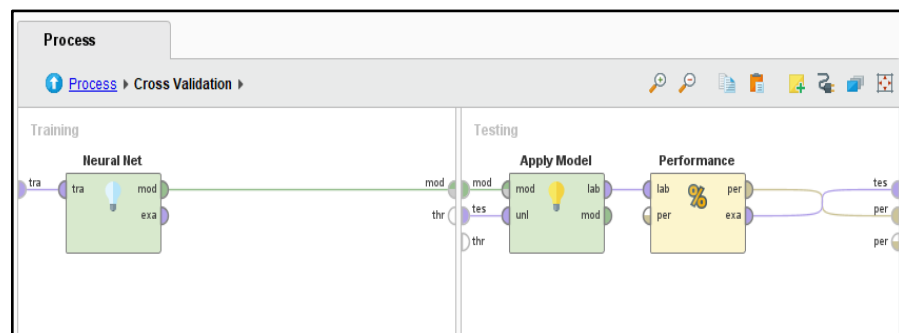


Figure 6. Testing Process using Neural Network Multilayer Perceptron

After the testing process has been carried out, the results obtained can be seen in Figures 8 and 11. In Figure 7 the results of the decision tree show that the root variable is the color type, as for the leaf is the product category, type type, and order period. This shows that the type of color affects the outcome of the decision. Figure 8 is the result of calculating the accuracy of the C4.5 method and multilayer perceptron with criteria using the confusion matrix.

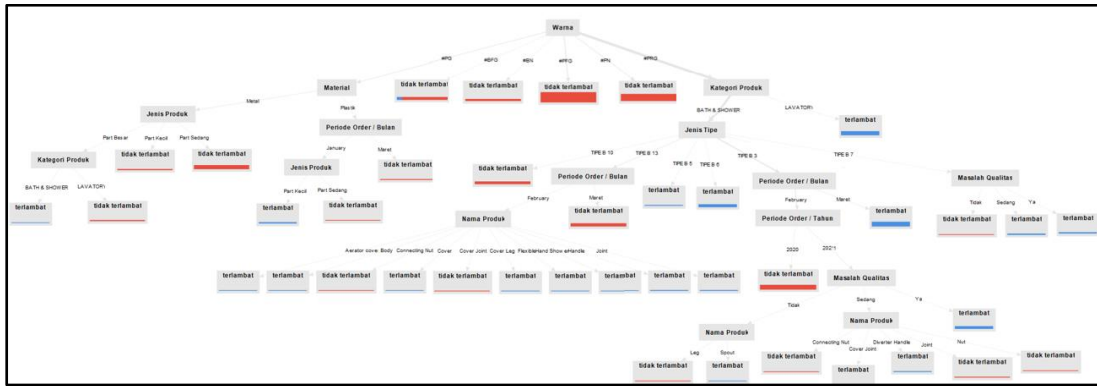


Figure 7. Decision Tree Results

The final result obtained from the process of forming this decision tree is a rule base that can be read from the path of the decision tree starting from the root to the leaf which is the class label of the conditions arranged by the attributes above the leaf. The basic rules obtained are:

1. If Color is #BFG: then the result is not late.
2. If Color is #PFG: then the result is not late.
3. If Color is #BN: then the result is not late.
4. If Color is #PN: then the result is not late.
5. If Color #PG: material = Metal, product type = Large Parts and product category = Bath & Shower: then the result is late.
6. If Color #PG: material = Metal, product type = Large Parts and product category = Lavatory: then the result is not late.
7. If Color #PG: material = Metal, product type = Small or Medium Part and product category = Bath & Shower and Lavatory: then the result is not late.
8. If Color #PG: material = Plastic, order period month = January, product type = Small Part and product category = Bath & Shower and Lavatory: then the result is late.
9. If Color #PG: material = Plastic, order period month = January, product type = Part Medium and product category = Bath & Shower and Lavatory: then the results are not late.
10. If Color = #PG: material = Plastic, order period month = March and product category = Bath & Shower and Lavatory: then the results are not late.
11. If Color #PRG: product category = Bath & Shower and type = TYPE B 10: then the result is not late.
12. If Color #PRG: product category = Bath & Shower type type = TYPE B 13 order period month = February and product name = Aerator Cover, Body, Cover, Cover Leg, Flexible, Hand & Shower: then the result is late.

13. If Color #PRG: product category = Bath & Shower type type = TYPE B 13 order period month = February and product name = Connecting Nut, Cover Joint, Handle, Joint: then it's not too late.
14. If Color #PRG: product category = Bath & Shower type type = TYPE B 13 month order period = March: then the results are not late.
15. If Color #PRG: product category = Bath & Shower type type = TYPE B 3 order period month = March: and order period year = 2020: then the results are not late.
16. If Color #PRG: product category = Bath & Shower type type = TYPE B 3 order period month = March: order period year = 2021 quality problem = Medium and product name = Connecting Nut, Joint, Nut: then the results are not late.
17. If Color #PRG: product category = Bath & Shower type type = TYPE B 3 order period month = March: order period year = 2021 quality problem = Medium and product name = Cover Joint, Diverter Handle: then the result is late.
18. If Color #PRG: product category = Bath & Shower type type = TYPE B 3 month order period = March: year order period = 2021 quality problem = No and product name = Leg: then the result is not late.
19. If Color #PRG: product category = Bath & Shower type type = TYPE B 3 month order period = March: year order period = 2021 quality problem = No and product name = Spout: then the result is late.
20. If Color #PRG: product category = Bath & Shower type type = TYPE B 3 month order period = March: year order period = 2021 quality problem = Yes: then the result is late.
21. If Color #PRG: product category = Bath & Shower and type type = TYPE B 5 and 6: then the result is late.
22. If Color #PRG: product category = Bath & Shower and type type = TYPE B 7 quality problems = Moderate and Yes: then results are late.
23. If Color #PRG: product category = Bath & Shower and type type = TYPE B 7 quality problems = No: then the result is not late.
24. If Color #PRG: product category = Lavatory: then results are late.
25. From the search for the rule base obtained from the decision tree, it can be seen that the overall training data record (256 records) can be taken and the biggest influence in the completion of pvd production is from the color type, material type, and product category.

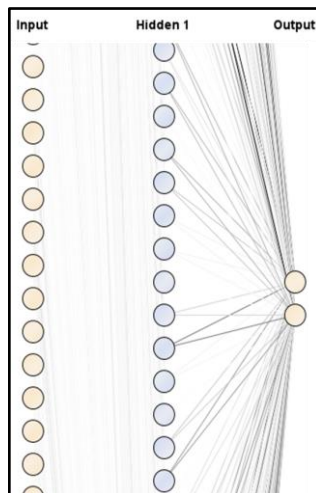


Figure 8. Architecture of Neural Networks Multilayer Perceptron

The network architecture in Figure 8 above is the best network architecture that was formed to be used to predict the completion of late and not late pvd production. The architecture consists of three layers, namely the input layer (179 nodes), hidden layer one (92 nodes), and the output layer (2 nodes). Each link or edge from one layer to another assigns a weight to each node. Based on the results of the two previous tests, the author has tested the dataset with 5 (five) testing scenarios to determine the level of accuracy obtained. The following is a recap of the comparison between the decision tree and multilayer perceptron methods.

Table 3. Comparison of Test Results

No Testing	Number of Records	Comparison	Algorithm	
			Decision Tree	Multilayer Perceptron
1	256	Akurasi	94.57%	92.98%
2	200		89.50%	89.00%
3	150		86.67%	90.00%
4	100		87.00%	92.00%
5	50		78.00%	76.00%
Average			87.15%	87.98%
Prediction Decision			The decision tree is well formed	-

Based on table 3 with the results of testing as much as 5 (five) times and the results of the highest accuracy value obtained by the multilayer perceptron algorithm of 87.98% and the decision tree algorithm of 87.15%. 0.83% difference.

4. CONCLUSION

Based on the results of the research and discussion, it can be concluded that the use of data mining with the decision tree method and the neural network multilayer perceptron can be applied in predicting the completion of pvd production. Here are some conclusions as follows:

1. The accuracy results obtained using the decision tree method with testing 5 (five) times with the number of datasets with different values obtained an average of 87.15%. Meanwhile, in the multilayer perceptron method, the average value obtained is 87.98%. Based on the results of the accuracy of the difference between the two methods, the multilayer perceptron method has a better and consistent level of accuracy at 90% in each testing test.
2. For the pvd production category which is often late, it needs to be re-evaluated, therefore (PPIC) Production Planning and Inventory Control in making the PRG color pvd production plan, the lead time is advanced faster than before to anticipate delays in the completion of color pvd (PRG) production. example: there is a request for an order in

March, therefore the schedule made is moved forward in February to anticipate a delay in the completion of production.

3. The positioning of the results of this study can be said to be in line with previous research that to perform predictive data mining using a decision tree or multilayer perceptron algorithm is very supportive. It is proven from several experiments that both algorithms provide a good level of accuracy. So it can be concluded that the accuracy of the multilayer perceptron algorithm is more accurate in predicting the delay in the completion of pvd production at PT. Surya Toto Indonesia when compared with the decision tree algorithm. Of course, these two methods can help the PPIC division in making future pvd production planning in order to minimize the occurrence of delays in the completion of production which can affect consumers.

5. SUGGESTED

1. With the results obtained from this study, the accuracy can be improved by increasing the number of parameters and the number of input nodes and layers in the next study. The discussion and application of a comprehensive feature model is expected to increase the accuracy value which is getting closer to 100% in subsequent studies.
2. For further research, it is more important to look at or take into account additional parameters in the production completion process to assist in decision making.
3. For the stages of the testing process, it is necessary to add more scenarios of more than 5 (five) times of testing with a larger number of datasets for each test to determine the level of consistency of accuracy obtained.

6. REFERENCES

- [1] Agrawal, Gaurav L., Hitesh Gupta, 2013, Optimization of C4.5 Decision Tree Algorithm for Data Mining Application , IJETAE (3):34-346
- [2] Laksmi, M, Begum, R.M, Venkatesan,VP., 2013, An Analysis on Performance of Decision Tree Algorithms using Student's Qualitative Data, International Journal of Modern Education and Computer Science 5: 5-18.
- [3] Hay's, R. N., Anharudin, & Adrean, R. (2017). Sistem Informasi Inventory Berdasarkan Prediksi Data Penjualan Barang Menggunakan Metode Single Moving Average Pada CV.Agung Youanda. Jurnal ProTekInfo, 29-33.
- [4] Herdianto. (2013). Prediksi Kerusakan Motor Induksi Menggunakan Metode Jaringan Syaraf Tiruan Backpropagation. Tesis, Fakultas Teknik, Universitas Sumatera Utara.
- [5] Ardianto, H., Anshori, A. N. M., & Hamidah, H. (2018). Potensi Larvasida dari Ekstrak Daun Jeruk Bali Citrus maxima terhadap dan Aedes aegypti dan Culex quinquefasciatus. Jurnal Vektor Penyakit, 12(1), 19-24.
- [6] Rahayu, D. W. S. (2020). Persepsi Merek Dalam Pemilihan Air Minum Dalam Kemasan (Studi Pada Mahasiswa Universitas Islam Balitar Blitar). AKUNTABILITAS: Jurnal Ilmiah Ilmu-Ilmu Ekonomi, 13(1), 81-89.

- [7] Curatman, A. (2015). MEMBANGUN KEPUASAN KONSUMEN MELALUI KEWAJARAN HARGA DAN MEREK PRIBADI. LOGIKA Jurnal Ilmiah Lemlit Unswagati Cirebon, 14(2), 42-49.
- [8] Azwanti. (2018). Analisa Algoritma C4.5 Untuk Memprediksi Penjualan Motor Pada PT.Capella Dinamik Nusantara Cabang Muka Kuning. Jurnal Ilmiah Ilmu Komputer, 33-38.
- [9] Harahap, F. (2015). Penerapan Data Mining Dalam Memprediksi Pembelian Cat. Konferensi Nasional Sistem & Informatika, 856-862.
- [10] Azwanti. (2018). Analisa Algoritma C4.5 Untuk Memprediksi Penjualan Motor Pada PT.Capella Dinamik Nusantara Cabang Muka Kuning. Jurnal Ilmiah Ilmu Komputer, 33-38.
- [11] R. Soewono, R. Hernowo, and P. S. Sasongko, "Sistem Pakar Identifikasi Modalitas Belajar Siswa Dengan Implementasi Algoritma C4.5," J. Sist. Inf. Bisnis, vol. 4, no. 1, 2014.
- [12] Hartanto Kamagi, David, Hansun, Seng. 2014. Implementasi Data Mining dengan Algoritma C4.5 untuk Memprediksi Tingkat Kelulusan Mahasiswa. ISSN 2085-4552. Ultimatix: Vol. VI. No. 1 Juni 2014.
- [13] Aggarwal. (2015). Data Mining The TextBook. New York: Springer International oPublishing Switzerland.
- [14] Hermanto, Bambang. 2017, Analisis Kinerja Decision Tree C4.5 dalam Prediksi Potensi Pelunasan Kredit Calon Debitur, JURNAL INNOTEK POLBENG - SERI INFORMATIKA, VOL. 2, NO. 2, ISSN : 2527-9866
- [15] Achmad, B. D. M., Slammat, F., & ITATS, F. T. I. (2012).Klasifikasi data karyawan untuk menentukan jadwal kerja menggunakan metode decision tree.Jurnal IPTEK, 16(1).
- [16] Prasetyo, Eko. Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab. Yogyakarta: CV.Andi Offset, 2014.
- [17] Faisal, M. R. (2017). Seri Belajar Data Science: Klasifikasi dengan Bahasa Pemrograman R.
- [18] Purnomo, M. H., & Kurniawan, A. (2006). Supervised Neural Network dan Aplikasinya. Yogyakarta: Graha Ilmu.
- [19] D. Widiastuti, "Analisa Perbandingan Algoritma SVM, Naïve Bayes , dan Decision Tree dalam Mengklasifikasi Serangan (Attacks) pada Sistem Pendeteksi Intrusi," 2012.
- [20] Jayawardana, Ivana Herliana and Seng Hansun. 2015. Rancang Bangun Sistem Pakar Untuk Deteksi Dini Katarak Menggunakan Algoritma C4.5. Ultima Computing. Vol. VII, No.2, ISSN: 2355-3286. 48-58.