



Palm oil extraction rate prediction based on the fruit ripeness levels using C4.5 algorithm

Wahyu Supriyatin

*Universitas Gunadarma, Jalan Margonda Raya No. 100 Depok 16424, Indonesia
ayu_ws@staff.gunadarma.ac.id*

Article history: Received January 01, 2021; Revised April 19, 2021; Accepted July 07, 2021; Available online August 07, 2021

Abstract

Oil palm plantations are one of the main keys in supporting Indonesia's economic growth. The rising consumption needs for palm oil products make it necessary to carry out data mining activities to increase CPO production. The maturity factor of palm fruit dramatically affects the quality of the oil extraction content (CPO yield) produced. This study aims to investigate the effect of fruit ripeness on the yield of CPO by using a data mining classification method with a decision tree. The algorithm used to generate decision tree classification is the C4.5 algorithm. The implementation of the C4.5 algorithm in the study was carried out using the Rapid Miner Studio 5.2 tools. The results show that the yield of CPO is influenced by the attributes of the condition of the long and ripe fruit, the condition of the long and overripe fruit, the normal condition of the fruit and the age of 3-6 years and the condition of the fruit of normal and age of 7-10 years. Decision tree C4.5 algorithm generates 8 rules with 4 rules showing a high production value, which means that the four rules affect the yield of CPO.

Keywords: Algorithm C4.5; Data Mining; Oil extraction rate; Fruit ripeness; Classification

Introduction

Indonesia is one of the countries that have the largest oil palm plantation area and is the number one producer of Crude Palm Oil (CPO) in the world [1] [2]. Oil palm is a plantation crop that has an essential role in the agricultural and plantation sectors because it produces oil (fat) [3]. Palm oil is one of the plantations Argo-industry businesses that has a significant role in the economy and people's incomes as well as industrial raw materials [4]. Oil palm production is influenced by three factors; environmental, genetic, and cultivation techniques. Environmental factors that affect the increase in oil palm production are climate and plantation land. Genetic factors are in the form of superior varieties of plant material. Cultivation factors are in the form of fertilization, plant care, and water [5]. Superior oil palm can produce from the planting age of 3.5 years marked by red fruit, influenced by the oil content [5]. If the fruit is ripe, free fatty acids (FFA) content will increase, and the fruit will drop naturally to the ground [5].

Harvesting palm oil fruit is an important activity to improve the quality of CPO [2]. The quality of CPO can be seen in terms of the quality and quantity of oil palm fruit that has an optimal mature condition [6]. Oil palm fruit must be harvested on time. If it is too ripe, the oil will contain high free fatty acids. However, if it is unripe, then the FFA and yield will be low [2]. The level of ripeness can influence the high and low quality of CPO at the time of fruit harvest [6]. Oil palm fruit that has good quality is a fruit that is harvested at the optimal level of ripeness, with color changes due to pigment concentration, appropriate moisture, and free fatty acid content [5]. Factors that affect the quality of the yield of palm oil (CPO) are poor quality fruit, late transported fruit (leftover), bruised fruit, and crop loss [1]. Good quantity fruit production will produce good CPO yield [1]. Good oil yield is obtained by processing ripe palm fruit. Ripe fruit quality has a high yield of palm oil extraction (CPO yield) [1].

This study aims to predict the level of palm oil extraction (CPO yield) from the ripeness of palm fruit by using one of the Knowledge Discovery in Database (KDD) techniques. The KDD technique used is data mining using the decision tree qualification method. The algorithm used in the qualification method is the C4.5 algorithm. The level of palm oil extraction (CPO yield) in this study was carried out by looking at the condition of the fruit, the condition of the bunch fruit, and the age of the fruit. The fruits used in the study are ripe, unripe, half-ripe and overripe. The conditions of the bunch fruit used in this study are long, short, and normal. The age categories used in the study are in the range of 3-6 years, 7-10 years, 11-15 years and >15 years. This research will produce a decision (goal) in the form of high production, which indicates the condition of the fruit that mainly influences the level of palm oil extraction. Implementation of data mining classification algorithm C4.5 is conducted using Rapid Miner Studio 5.2 tools.

Data mining is a process to draw knowledge that is in the database [7]. The application of data mining has been widely applied to convert data into useful information and gain knowledge in an extensive application [7]. Data mining is the process of analyzing data from different perspectives and summing it up into essential information that can be used to increase profits and reduce expenses [8]. Data mining is utilized to find the relationships, patterns, and meanings by filtering complex data using pattern recognition techniques [9]. It is a part of Knowledge Discovery in Database (KDD), and KDD acts as a database. Data mining can be used to predict the future [9]. KDD is a process of generating knowledge from existing patterns and applying the scientific method [7].

Knowledge Discovery in Database is a method used to obtain database knowledge information [10]. The existing tables are interconnected and related. The results of the knowledge obtained will be used as a knowledge base in making decisions [10]. The stages in the KDD process are Data Selection, Preprocessing/Cleaning, Transformation, Data Mining and Interpretation/Evaluation [7] [10]. The data mining method in KDD is done by extracting existing patterns from the data depending on the applied data mining task [7].

Classification is the process of finding the same properties in a collection of objects in a database and classifying them into groups [11]. Classification is a type of data analysis that can assist in determining the label class of the sample you want to classify [12]. Classification is a method that finds the relationship between input attributes and target attributes with the aim of increasing the reliability of the results obtained from the data [12]. The purpose of classification is to find a different training set model in the appropriate category or class [11]. One of the methods used in data mining classification is decision tree classification.

The decision tree represents the tree structure where nodes represent attributes, branches represent attribute values, and leaves represent classes [11]. A decision tree is a structure used to divide an extensive data set into a small record set by applying decision rules [8]. A decision tree is a classification technique for objects or records consisting of a collection of decision nodes and connected to branches [9]. The decision tree forms a decision tree that produces conclusions in the form of classification rules, and one of the algorithms used is the C4.5 algorithm.

The C4.5 algorithm is a data mining algorithm used to create a decision tree [10] [11]. The C4.5 algorithm involves decision construction and a collection of decision nodes where each branch will lead to another node, either a decision node or an end node [12]. The C4.5 algorithm is a development algorithm from the ID3 algorithm created by J. Rose Quinlan [10]. The development that is in the C4.5 algorithm compared to the ID3 algorithm is that it can overcome missing values, can overcome continue data and pruning [11]. The C4.5 algorithm in building a decision tree uses the rule [10]:

- a. Select attribute as root.
- b. Create a branch for each value.
- c. Split cases in branches.
- d. Repeats the process for each branch until all cases in the branch have the same class.

Previous research related to CPO yield and ripeness was carried out by Pryo [1]. In his study, Pryo [1] analyzed the problem of CPO quality degradation caused by low palm oil (CPO) yield and high levels of Free Fatty Acid (FFA) in CPO. Low oil yields can be obtained due to poor fruit quality and yield loss. High CPO FFA content was obtained due to poor fruit quality, leftover fruit (late transportation) and bruised fruit.

Joko's study [5] used the morphological characteristics of oil palm fruit as an indicator of ripeness fruit harvest. The study was conducted randomly on several trees that were 7 years old. The results of the study [5] showed that the determination of ripeness fruit harvest could use the variables, fruit color, length of time after pollination, fruit bunch weight, fruit diameter, fruit mesocarp thickness, fruit fresh weight, fruit dry weight and fruit moisture content.

An ancient study [13] regarding the relationship between fruit maturity and bunch height on the number of loose fruits with the parameters observed were the number of loose fruits after harvest, free fatty acids, and percentage of fruit weight with bunch weight after harvest. The result of the study [13] indicated that the higher the ripeness, the higher the number of loose fruits, free fatty acids, and the percentage of fruit weight with bunch weight after harvest.

Rahmadhania [14], in her study, stated that the success of palm oil mills is determined by the maturity level of fresh fruit bunches at harvest and the content of CPO oil and free fatty acid (FFA) content. Research [14] was used to determine the effect of oil palm harvest fraction (palm FFB maturity) on yield (CPO content). The highest CPO yield is in fraction 3 with an average of 26.1%, and the lowest CPO yield in fraction 0 with an average of 22.0%.

Pradifta conducted his study [15] to determine the yield in crude palm oil and its minor composition. The parameters used [15] in his research were unripe fruit, ripe fruit, and overripe fruit. The study results [15] showed that palm fruit's ripeness level affects the yield, beta-carotene, crude palm oil compounds at an altitude of 650 masl and 850 masl.

Shiddiq's research [16] regarding the ripeness level of oil palm FFB is a determining factor for the quality of CPO produced by palm oil mills. The sorting of FFB after harvest was previously conducted manually by sight, but in the study [16], a laser-induced fluorescence imaging method was used to classify the ripeness level of oil palm FFB. Classification of the maturity level of FFB is carried out using K-mean Clustering. The study results [16] showed a potential laser-induced fluorescence imaging method could be used to classify the ripeness level of FFB.

Method

The method used in this study is a data mining classification method using the C4.5 algorithm. The algorithm is used to obtain the relationship between the ripeness of oil palm fruit and the level of extraction from palm oil. Dividing the level of fruit maturity into several conditions and comparing one condition to another, so that it will produce a conclusion of the fruit conditions that most influence the extraction of palm oil. The flowchart of the C4.5 algorithm can be seen in **Figure 1**. **Figure 2** is a research flow used in data mining classification.

Pseudocode Algorithm C4.5 in **Figure 1** which is used to build a decision tree is:

1. Start
2. Checking problem cases.
3. Determine what will be used as attributes in case of problems.
4. Perform testing of the specified attributes to obtain the highest Gain value by finding the Entropy value of each attribute. Equation (1) is the formula used to find the value of Entropy.

$$Entropy(S) = - \sum_{i=1}^m p_i \log_2(p_i) \quad (1)$$

Equation (2) is used to find the Entropy value for each variable or attribute.

$$Entropy_A(S) = \sum_v \frac{|S_v|}{|S|} Entropy(S_v) \quad (2)$$

Equation (1) and equation (2) are used to find the Gain value such as equation (3).

$$Gain(A) = Entropy(S) - Entropy_A(S) \quad (3)$$

Where (S) is the set of problem cases, A is the specified attribute, m is the number of partitions of the attribute A, $|S_v|$ is the number of cases on the vth- partition and |S| is the number of cases S.

5. If the highest Gain value has not been obtained from the problem, then recalculation is carried out to find the Entropy value of the attribute.
6. If the highest Gain value has been obtained, then partition the data according to the attributes that have been determined in step 3 based on the highest Gain value obtained.
7. The attribute with the highest Gain value will automatically become the initial root in a decision tree.
8. If the root has been obtained, then determine which attribute will be the branch by looking at the highest Gain value from each data partition result. This is done continuously for all the specified attributes to form a decision tree.
9. Finish.

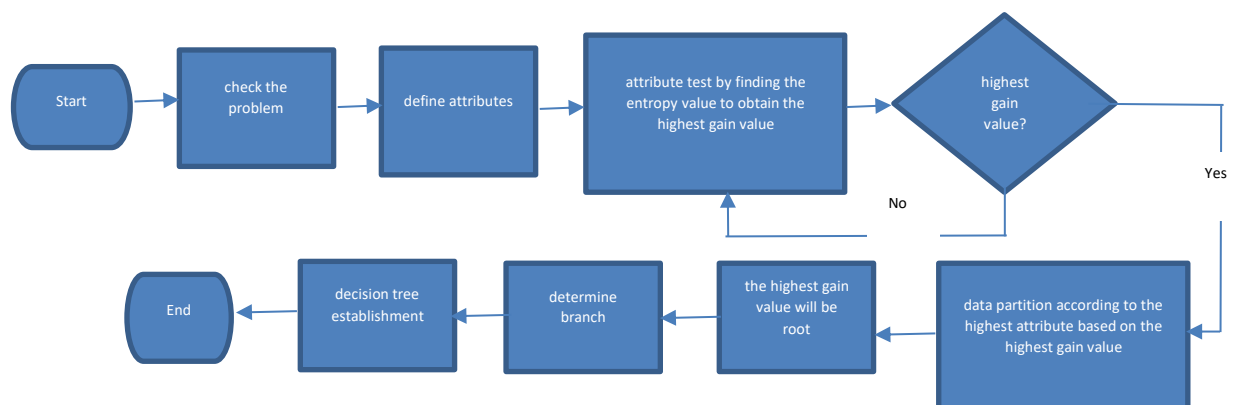


Figure 1. Flowchart of Algorithm C4.5



Figure 2. Research Flow

The steps in the classification of data mining with the C4.5 Algorithm in this study are:

A. Data Collection Stage

The data collection stage is the initial stage carried out before processing the data into a decision tree using the C4.5 Algorithm. The stages of data collection were carried out by taking sample data from oil palm plantation production data.

B. Data Cleanup Stage

This stage is the stage where the simplification of the data is obtained. Data that does not fit in this stage will be removed because it is considered invalid. Only necessary data attributes used in making the decision tree will be taken from the sample. Inappropriate data will be cleaned, checked, corrected, and eliminated to avoid duplicate data. So that from the data cleaning stage, it will produce the necessary data as condition attributes and decision attributes to be used in the following stages of data mining.

C. Data Integration Stage

This is the stages to combine some of the data to be used in research obtained from data sources. The condition attribute data used is adjusted so that it can produce a decision attribute data in the decision tree classification process.

D. Data Transformation Stage

This stage is the step that is carried out by simplifying or grouping data. The data used in the study are grouped into 2 variables; input variables and output variables. These variables will be classified into attributes as conditions and decisions as results. Input variables are input variables or condition variables in making a decision tree. Output variables are output variables or decision variables as a result of the decision tree.

E. Data Mining Application

This stage is a stage of determining the data mining classification algorithm that will be used in research. The stage is to produce a decision tree so that a conclusion or result is obtained. Decision tree classification is done using the C4.5 Algorithm.

F. Pattern Evaluation Stage

This stage is a stages of grouping patterns from the attributes of the condition of the level of ripeness of oil palm fruit. This pattern will be evaluated so as to produce an appropriate decision in the form of whether or not the level of maturity of oil palm fruit on the level of palm oil extraction.

G. Pattern Presentation Stage

This stage is the final stage in the data mining process in research. This stage is the stage of obtaining the decision tree results from the analysis that has been done using the Rapid Miner Studio 5.2 tools. Processing of oil palm fruit ripeness level data using criterion provisions, minimum gain, maximum depth, and confidence from the C4.5 algorithm used. The decision tree results will be used to determine the level of ripeness of the oil palm fruit, which will affect the level of palm oil.

Results and Discussion

Implementation of data mining classification uses the C4.5 algorithm with processing tools Rapid Miner Studio 5.2 to obtain a decision tree. The decision tree is used to see how the ripeness level of oil palm fruit is used to predict the level of palm oil produced by an oil palm plantation. The steps taken from the data collection stage until the decision result are:

A. Data Collection Stages

This stage is the stage of collecting data that will be used as samples in the study. The research sample data was obtained from data on the ripeness level of oil palm fruit in one month on an oil palm plantation. The data used as the sample in the study is 50 samples. The data obtained from oil palm plantations include the attribute data of area, mill name, mill code, estate name, estate code, date, fruit condition, fruit bunch condition and fruit age.

B. Stages of Data Cleaning and Data Integration

The data cleaning stage is carried out on the sample data that has been collected for the process of deleting unnecessary attributes. So that the existing attributes can be simplified or minimized and can produce a decision tree as the searched result. The sample data collected was then simplified and three attributes were obtained to be processed as conditions, which are the attribute of "Condition of Fruit", the attribute "Condition of Fruit bunch" and the attribute of "Age of Fruit".

C. Data Transformation Stages

The stages of grouping data are divided into two variables called input variables and output variables. The input variable becomes the input (condition attribute) while the output variable becomes the output (result attribute). The input variables or conditions in the study are:

1. Attribute "Condition of Fruit" with records "Condition of Ripe Fruit", "Condition of Half Ripe Fruit". "Unripe Fruit Condition" and "Overripe Fruit Condition".
2. Attribute "Condition of Fruit Bunch" with records "Empty Fruit Bunch", Long Fruit Bunch " and "Normal Fruit".
3. Attribute "Age of Fruit" with records "3-6 years", "7-10 years", "11-15 years" and "> 15 years".

The output variable or result in the study is the "Extraction" attribute with the result of the decision "High Production" or "Low Production".

No	Kondisi Bua	Kondisi Janj	Usia Buah	Ekstraksi
integer	polyno...	polyno...	polyno...	binomi...
id	attribute	attribute	attribute	label
1	Buah Masak	Buah Janjan	3-6	Produksi Re
2	Buah Setenç	Buah Janjan	11-15	Produksi Re
3	Buah Mental	Buah Janjan	3-6	Produksi Re
4	Buah Setenç	Buah Janjan	>15	Produksi Re
5	Buah Masak	Buah Norma	7-10	Produksi Tin
6	Buah Terlalu	Buah Janjan	3-6	Produksi Tin
7	Buah Masak	Buah Janjan	>15	Produksi Tin
8	Buah Masak	Buah Janjan	11-15	Produksi Tin
9	Buah Setenç	Buah Janjan	3-6	Produksi Re
10	Buah Terlalu	Buah Janjan	>15	Produksi Tin
11	Buah Setenç	Buah Janjan	3-6	Produksi Re
12	Buah Masak	Buah Janjan	3-6	Produksi Re
13	Buah Mental	Buah Janjan	7-10	Produksi Re
14	Buah Masak	Buah Janjan	11-15	Produksi Tin
15	Buah Masak	Buah Janjan	3-6	Produksi Tin

Figure 3. Data Transformation Steps

Figure 3 is the stage of data transformation after the data cleaning process from the data collection process was carried out. The output variable or result, which is the "Extraction" attribute, is created in the binomial data type. The binomial data type is a data type that only has 2 types of data, which are the "Low Production" and "High Production". The attribute "Extraction" because it is a decision is made in the form of "Label".

The condition attributes, which are "Fruit Condition", "Fruit Bunch Condition" and "Fruit Age" use polynomial data types. The polynomial data type is a data type that has more than two data types. While the attribute "No" is used as an id with an integer data type.

D. Stages of Decision Tree Formation (Data Mining Application Stage)

The decision tree formation stage is the final stage in implementing data mining classification using the C4.5 algorithm. The 50-sample data in Figure 3 will be processed using the Rapid Miner Studio 5.2 tools. The sample data in this study were obtained from one of the oil palm plantations in the Kalimantan area for December 2020. The results obtained from the decision tree are 3 condition attributes used with 1 decision attribute. Figure 4 is a decision tree obtained from 50 samples using the C4.5 algorithm.

Figure 4 shows the rules generated from the attributes of "Fruit Condition", "Fruit Bunch Condition" and "Fruit Age". By using the criterion gain ratio, it appears that the "Fruit Bunch Condition" has the most significant influence on the level of palm oil extraction. If only "Fruit Bunch Condition" is used to determine palm oil extraction it will result in "Low Production", so it must be influenced by other attributes. The combination of attributes of "Fruit Bunch Condition" and "Condition of Fruit" as well as attributes of "Fruit Bunch Condition" and "Age of Fruit" will result in "High Production" in influencing the level of palm oil extraction.

The rules obtained from the decision tree in Figure 4 obtained 8 rules with 4 "High Production" rules and 4 "Low Production" rules. The "High Production" rule that affects the extraction rate of palm oil is obtained for the following conditions:

1. "Fruit Bunch Condition = Long Fruit Bunch" and "Condition of Fruit = Ripe Fruit"
2. "Fruit Bunch Condition = Long Fruit Bunch" and "Condition of Fruit = Overripe fruit"
3. "Fruit Bunch Condition = Long Fruit Bunch" and "Fruit Age = 3 – 6 years"
4. "Fruit Bunch Condition = Long Fruit Bunch" and "Fruit Age = 7 – 10 years"

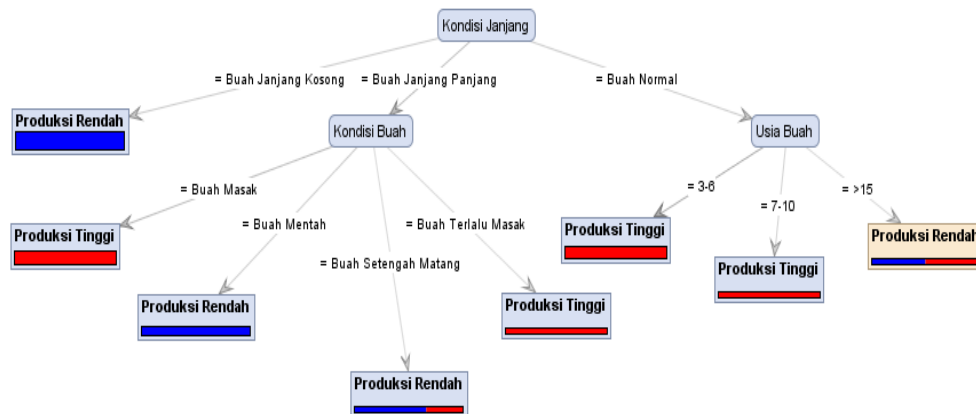


Figure 4. C4.5. Algorithm Decision Tree

Figure 5 is the rule of the rule used to generate a decision tree in Figure 4. From Figure 5 it can be seen that the number of rules that exist in the conditions of "High Production" and "Low Production". The "High Production" rule that affects the level of palm oil extraction is obtained by the rule:

1. If "Fruit Bunch Condition = Long Fruit Bunch" and "Fruit Condition = Ripe Fruit", then "High Production" produces 10 samples
2. If " Fruit Bunch Condition = Long Fruit Bunch " and "Fruit Condition = Half Ripe Fruit", then "High Production" produces 1 sample
3. If " Fruit Bunch Condition = Long Fruit Bunch " and "Fruit Condition = Overripe Fruit", then "High Production" produces 3 samples
4. If " Fruit Bunch Condition = Normal Fruit" and "Fruit Age = 3 - 6 years", then "High Production" produces 8 samples
5. If " Fruit Bunch Condition = Normal Fruit" and "Fruit Age = 7 - 10 years", then "High Production" produces 3 samples
6. If " Fruit Bunch Condition = Normal Fruit" and "Fruit Age = > 15 years", then "High Production" produces 1 sample

The analysis of 6 rules that results in the condition of "High Production" is only 4 rules with a large sample value that affects the level of palm oil extraction. For 2 rules are in a "Low Production" condition where 1 rule has a 1:1 balance condition stated as "Low Production" and 1 rule has a "Low Production" condition with a value of 2:1.

```
Kondisi Janjang = Buah Janjang Kosong: Produksi Rendah {Produksi Rendah=15, Produksi Tinggi=0}
Kondisi Janjang = Buah Janjang Panjang
| Kondisi Buah = Buah Masak: Produksi Tinggi {Produksi Rendah=0, Produksi Tinggi=10}
| Kondisi Buah = Buah Mentah: Produksi Rendah {Produksi Rendah=6, Produksi Tinggi=0}
| Kondisi Buah = Buah Setengah Matang: Produksi Rendah {Produksi Rendah=2, Produksi Tinggi=1}
| Kondisi Buah = Buah Terlalu Masak: Produksi Tinggi {Produksi Rendah=0, Produksi Tinggi=3}
Kondisi Janjang = Buah Normal
| Usia Buah = 3-6: Produksi Tinggi {Produksi Rendah=0, Produksi Tinggi=8}
| Usia Buah = 7-10: Produksi Tinggi {Produksi Rendah=0, Produksi Tinggi=3}
| Usia Buah = >15: Produksi Rendah {Produksi Rendah=1, Produksi Tinggi=1}
```

Figure 5. Decision Tree Description Rules

E. Stages of Evaluation and Presentation

The evaluation and presentation stages are the documentation stages in processing research data using the Rapid Miner Studio 5.2 tools. Figure 6 is the process for importing data into Rapid Miner Studio 5.2 from 50 samples used in the study. The 50 samples used were checked, so that they can be processed and produce a decision tree.

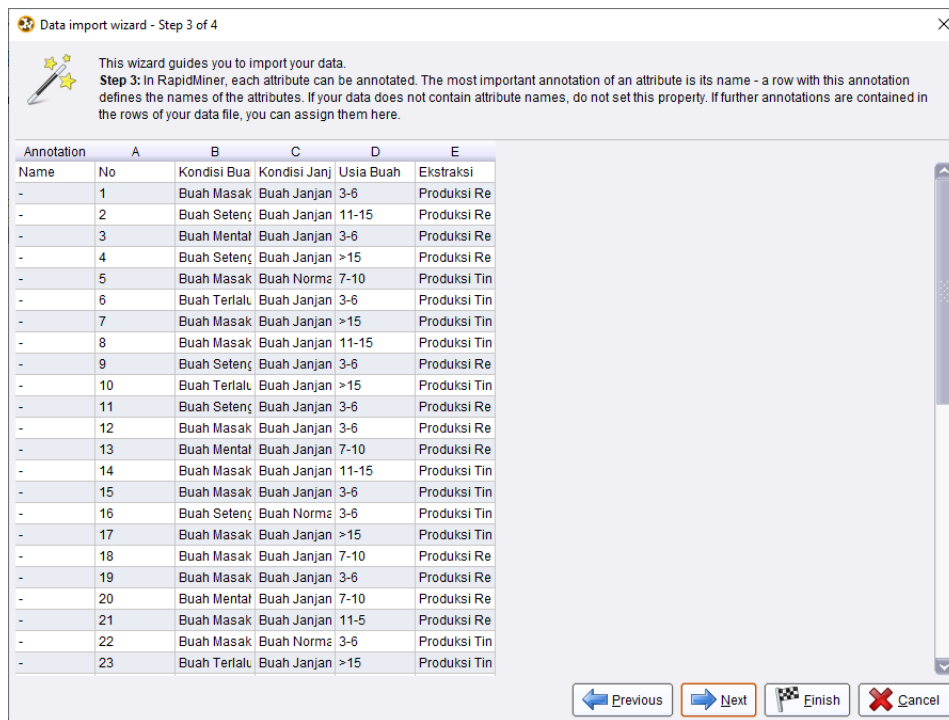


Figure 6. Data Import Process

After the data import process in **Figure 6** is carried out, the process of determining the condition and result attributes is conducted as shown in **Figure 7**. **Figure 7** determines the data type for the condition and result attributes.

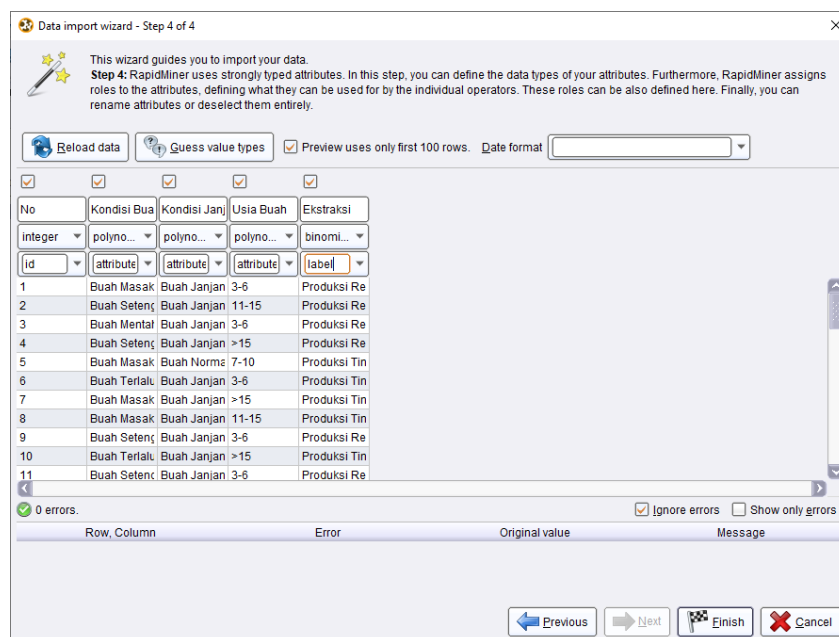


Figure 7. The process of determining condition attributes and result attributes

Figure 8 is the process of applying the C4.5 algorithm to produce a decision tree. Determination of the decision tree in the Rapid Miner Studio 5.2 tools uses the criterion provisions, minimum gain, maximum depth and confidence from the C4.5 algorithm used. The conditions used are:

1. *Criterion* = gain_ratio
2. *Minimal Gain* = 0,1
3. *Maksimal Depth* = 20
4. *Confidence* = 0,25

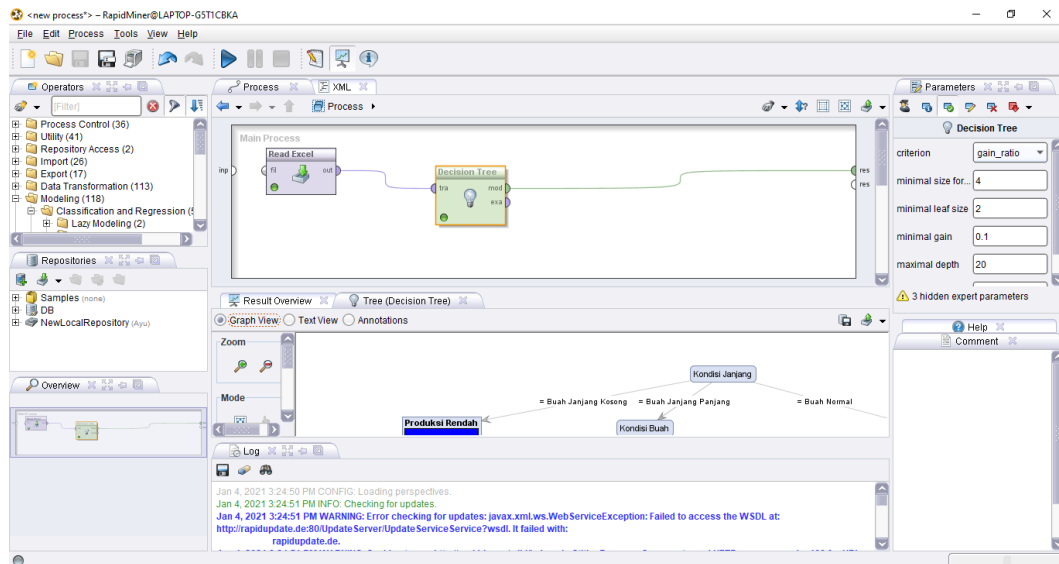


Figure 8. The C4.5 . Algorithm Application Process

Conclusion

The data mining classification method using the C4.5 algorithm is used to predict the content of palm oil-based on fruit ripeness to obtain the appropriate decision tree. The implementation is carried out using Rapid Miner Studio 5.2 tools. The decision tree produces 8 rules with 4 rules that produce high production attribute values. High production means fruit ripeness affects the level of palm oil extraction. Attributes of conditions that affect high production are the condition of long fruit bunches and ripe fruits, the condition of long fruit bunches and overripe fruits, the condition of normal fruit bunches and 3-6 years old fruit, and the condition of normal fruit bunches and 7-10 years old fruit.

Development for further research can use other data mining classification algorithms such as the C5.0 Algorithm, ID3 Algorithm, Classification and Regression Trees Algorithm, Naive Bayes Algorithm, or KNN Algorithm. The development of using other algorithms is used to see how the highest Gain value is generated for the same case, whether it will produce the same root or a better root. Apart from the algorithm side, the attributes used can also be developed, among others, determining the wasted extraction losses based on the condition of the fruit bunch or the quality of the CPO produced based on the freshness of the fruit.

References

- [1] P. A. Lukito and Sudrajat, "The Effect of Palm Oil Fruit Branch Injury to Free Fatty Acid Content and CPO Rendemen at Talisayan 1 Estate Berau," *Bul. Agrohorti*, vol. 5, no. 1, pp. 37-44, 2017.
- [2] N. Fitriya, S. P. Wirman and W. Fitri, "Identifikasi Karakteristik Buah Kelapa Sawit Siap Panen Dengan Metode Laser Spekel Imaging (LSI)," *Jurnal Photon*, vol. 9, no. 1, pp. 139-142, 2018.
- [3] F. Insani, I. Harani, S. Sanjaya and Yusra, "Peramalan Produksi Tandan Buah Segar (TBS) Kelapa Sawit Dengan Regresi Linear Dan Algoritma Genetik (Studi Kasus : PT. Peputra Masterindo)," in *Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI) 11*, Pekanbaru, 2019.
- [4] P. Imam and I. B. A. K. Santosa, "The Use of Multivariate Linear Regression to Analyse The Accuracy of Determination of Model Prediction of Free Fatty Acid and DOBI at Sterilization Fresh Fruit Bunch," *Jurnal Teknologi dan Industri Pertanian Indonesia*, vol. 9, no. 2, pp. 55-64, 2017.
- [5] J. Supriadi, G. Tabrani and Isnaini, "The Criteria of Palm Oil Ripe Level (*Elaeis Guineensis* Jacq.) as the Reap Indicator Observed from Morphological Characteristics," *JOMFAPERTA*, vol. 6, no. 1, pp. 1-9, 2019.
- [6] R. Salambue and M. Shiddiq, "Klasifikasi Kematangan Buah Sawit Menggunakan Model Warba RGB," in *Seminar Nasional APTIKOM (SEMNASTIK)*, Semarang, 2019.
- [7] B. Sugara, D. Widyatmoko, B. S. Prakoso and D. M. Saputro, "Penerapan Algoritma C4.5 Untuk Deteksi Dini Gangguan Autisme Pada Anak," in *Seminar Nasional Teknologi Informasi dan Komunikasi (SENTIKA)*, Yogyakarta, 2018.
- [8] H. D. Honesqi, "Klasifikasi Data Mining Untuk Menentukan Tingkat Persetujuan Kartu Kredit," *Jurnal TEKNOIF*, vol. 5, no. 2, pp. 57-62, 2017.
- [9] R. H. Pambudi, B. D. Setiawan and Indriati, "Penerapan Algoritma C4.5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternal," *Jurnal pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 7, pp. 2637-2643, 2018.
- [10] Y. Mardi, "Data Mining : Kalsifikasi Menggunakan Algoritma C4.5," *Jurnal Edik Informatika*, vol. 2, no. 2, pp. 213-219, 2017.

-
- [11] N. Azwanti, "Analisa Algoritma C4.5 Untuk Memprediksi Penjualan Motor pada PT. Capella Dinamik Nusantara Cabang Muka Kuning," *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, vol. 13, no. 1, pp. 33-38, 2018.
- [12] S. Hendrian, "Algoritma Klasifikasi Data Mining Untuk Memprediksi Siswa Dalam Memperoleh Bantuan Dana Pendidikan," *Faktor Exacta*, vol. 11, no. 3, pp. 266-274, 2018.
- [13] I. R. Purba, Irsal and Meiriani, "The Relationship between Fraction of Fruit Maturity and Height of Fruit Bunches to the Number of Oil Palm Fruit at Harvesting Oil Palm (*Elaeis guineensis* Jacq)nin Palm Plantation Rambutan PT. Perkebunan Nusantara III," *Jurnal Agroekoteknologi FP USU*, vol. 5, no. 2, pp. 315-328, 2017.
- [14] F. Rahmadhania, P. Sembiring and M. A. Sinaga, "The Effect of Palm Oil Fruit Maturity Dxp Bah Lias Variety to the Crude Palm Oil Content," *Jurnal Argo Estate*, vol. 3, no. 1, pp. 1-9, 2019.
- [15] I. U. P. Rangkuti, "Oil Extraction Rate and Minor Components Based on Level Ripeness of Crude Palm Oil at High Altitudes," *Argotekma : Jurnal Agroteknologi dan Ilmu Pertanian*, vol. 3, no. 1, pp. 10-16, 2018.
- [16] H. Ishak, M. Shiddiq, R. H. Fitra and N. Z. Yasmin, "Ripeness Level Classification of Oil Palm Fresh Fruit Bunch Using Laser Induced Fluorescence Imaging," *J. Aceh Phys. Soc.*, vol. 8, no. 3, pp. 84-89, 2019.