Data Mining Prediction of Oil Palm Fruit

Anisya[1],

[1]Informatics Engineering, Institut Teknologi Padang, Padang, Indonesia

Abstract

The development of information technology today is very meaningful for all circles. Currently, information technology has become a necessity in everyday life. The use of information technology is proven to facilitate human performance. Where the number of suppliers that supply palm oil fruit every year will affect the activities of companies engaged in palm oil production. So that currently the company needs a decision-making strategy in the procurement of oil palm fruit. Data mining is a technology that is very useful to help companies find very important information from data centers. Data mining predicts trends and characteristics of business behavior which are very useful to support important decision making. One of the techniques the writer uses is the Association Rule technique. Association Rule is a data mining technique to find association rules between item combinations. Using the Association Rule technique will help companies predict which suppliers will supply palm fruit in the following year. Meanwhile, to predict the load does not use the association rule method but uses existing data analysis.

1. Introduction

The development of information technology today is very meaningful for all circles. Currently, information technology has become a necessity in everyday life. The use of technology is proven to facilitate human performance. This is what causes information technology to be applied in all fields, including business. Large amounts of business data are one of the valuable assets that a company has, one of the objects of research here is that companies engaged in palm oil production have to think about how many suppliers will supply palm fruit each year. The number of suppliers that supply will affect production activities in the form of cooking oil.

Data mining is a technology that is very useful to help companies find very important information from data centers. Data mining predicts trends and traits of business behavior which is very useful to support important decision making. The availability of sufficient data, the need for information (or knowledge) to support decision making to make business solutions and support infrastructure in the field of information technology is the birth of a data mining technology. Data mining provides real solutions for decision makers in the business world to develop a business [1].

By using existing and relevant data, data mining create several models to identify patterns among the attributes that exist in the dataset. The model is a mathematical representation (simple linear equation / non-linear equation) that identifies patterns

Anisya
Email : nisa.anisya@gmail.com

between various object attributes (for example customers) in the dataset. Some of these patterns are descriptive (explaining the interrelationships / similarities between attributes). The second is predictive (predicting the "value" / result that will occur in certain attributes). There are four patterns identified by data mining, namely association to find groupings that occur simultaneously. Prediction to predict the future. Clusters to identify based on character. The last one is sequential relationships for finding a series of events.

The technique used by the author is by using the Association Rule technique. Association Rule is a data technique Mining to find association rules between item combinations. Using the Association Rule technique will help companies predict which suppliers will supply oil palm fruit in the following year. Many, at least, suppliers Supplying oil palm fruit every year will affect the yield of oil palm fruit. By using Technique Association Rule can find out a knowledge or information from the data center that is useful for companies in making decisions.

2. Research Methods

Previously, there was a lot of research on the application of Data Mining, using a variety of methods, and various implementations. Of course, with a few similarities and differences. In this study, the authors obtained several theories and literature that are used as guidelines and references. There are several studies related to the author's research.

Research [3] which explains how to predict the plan to increase fertilizer stock using the associatiation rule method mining (ARM). In this study, there are several similarities between the authors with research [3], one of which aims to predict an item. Research conducted by Bahrur Roji will determine the prediction of the plan to increase fertilizer stock using the association rule (ARM) method, from the test results will determine which fertilizers are often purchased by customers. The difference between the author's research and Bahrur Roji's research is that the case studies used are different.

Research [2] which explains how to apply Data Mining Using the Association Rules Method to support the marketing strategy of prospective new students (case study Ibi Darmajaya). Research [2] will be used as guidance in making further decisions about information sources as the right promotional media to support the marketing strategy in Darmajaya. The difference between the author's research and [2] is a different strategy and case study. Research [4] which explains how to apply data mining using the Market Basket Analysis method to sales data of book products using the Apriori Algorithm. And Frequent Pattern Growth (FP-Growth) with case studies Printing PT.Gramedia. There are several differences to the research [4] uses two algorithms, namely the Apriori algorithm and Frequent Pattern Growth (Fp-Growth). Research [4] will assist in the managerial decision-making process, especially those related to the making of marketing and sales strategies for PT.Gramedia's printed book products.

3. Results and Discussion

3.1 Analysis Data

There are three parts to analyze the data, namely input data, process data, and output data. Input data is data that is entered by the admin into the system. Process data is data that is processed to get the calculation of the association rule mining. While the output data is the result of data that has been entered and processed in the system.

3.2 Input Data

The data used to implement the device soft, in predicting suppliers who will supply oil palm fruit in the following year using the method of association rule mining, in the input data system that the writer displays, namely the data supply (purchase).

3.3 Diagram Context

3.3.1 Process Data

Calculation data with association rule mining This data contains the results of calculations carried out by determining minimum support and minimum confidence, to obtain an association relationship from oil palm fruit supply data. Supply data these are obtained from each supplier who supplies each year.

3.3.2 Output Data

Application of association rule mining in software will produce the output of the data analysis of the purchase of oil palm fruit which will produce knowledge and information in the form of a base rule pattern used in the supply of oil palm fruit.

3.4 Test Results

In this purchase data displays each data supply from each supplier. The data used as data analysis is data on the supply of oil palm fruit in 2016. Supply data from each supplier in one year will be used to predict which suppliers will supply in the following year.

3.4.1 Support for Candidate Itemset 2

$$Support = \frac{Value\ A + B}{Total\ Purchase} x 100\% \qquad (1)$$

$$confidence = \frac{Value\ A}{Value\ A+B} x 100\% \qquad (2)$$

Table 1. Support for Candidate Itemset 2

Rules	Support (Value A + B / Total Purchase x 100%)	Confidence (Value A / A + B x 100%)	Best Rule Support x Confidence / 100 %
If buy a Khairunas supplier then buy Idris	273+234/12937*100 % =3.91 %	273/507x100% =53.84%	3.91x53.84/100% =2.10%
If buy a Idris supplier then buy Khairunas	234+273/12937*100 % =3.91 %	234/507x100% =46.15%	3.91x46.15/100% =1.80%
If Buy A Khairunas Supplier Then Buy Dt Musa	273+367/12937*100 % =4.94 %	273/640x100% =42,65%	4.94x42,65/100% =2.10%
If Buy Dt Musa Supplier Then Buy Khairunas	367+273/12937*100 % =4.94 %	367/640x100% =57.34%	4.94x57.34/100% =2.83%
If Buy A Khairunas Supplier, Then Buy PTP	273+3528/12937*100 % =29.38 %	273/3801x100% =7,18%	29.38x7.18/100% =2.10%
If Buy A PTP Supplier Then Buying Khairunas	3528+273/12937*100 % =29.38 %	3528/3801x100% =92.81%	29.38x92.81/100% =27.26%
If Buy A Khairunas Supplier, Then Buy Talao	273+4341/12937*100 % =35.66 %	273/4614x100% =5.91%	35.66x5.91/100% =2.10%

9

If buy a Talao supplier then buy Khairunas	4341+273/12937*1000 % =35.66 %	4341/4614x100% =94.08%	35.66x94.08/100% =33.54%
If buy a Khairunas supplier then buy Sei Kunyit	273+4194/12937*1000 % =34.52 %	273/4467x100% =6.11%	34.52x6.11/100% =2.10%
If buy a Sei Kunyit supplier then buy Sei Khairunas	4194+273/12937*1000 % =34.52	4194/4467x100% =93.88%	34.52x93.88/100% =32,40%
If buy a Idris supplier then buy Dt Musa	234+367/12937*1000 %	234/601x100% =38.93%	4.64x38.93/100% =1.80%
If buy a Dt Musa supplier then buy Idris	367+234/12937*1000 % =4.64 %	367/601x100% =61,06%	4.64x61.06/100% =2.83%
If buy a Idris supplier then buy PTP	234+3528/12937*1000 % =29.07 %	234/3762x100% =6.22%	29.07x6.22/100% =1.80%
If buy a PTP supplier then buy Idris	3528+234/12937*1000 % =29.07 %	3528/3762x100% =93.77%	29.07x93.77/100% =27.25%
If buy a Idris supplier then buy Talao	234+4341/12937*1000 % =35.36 %	234/4575x100% =5.11%	35.36x5.11/100% =1.80%
If buy a Talao supplier then buy Idris	4341+234/12937*1000 % =35.36 %	4341/4575x100% =94.88%	35.36x94.88/100% =33.54%
If buy a Idris supplier then buy Sei Kunyit	234+4194/12937*1000 % =34.22 %	234/4428x100% =5.28%	34.22x5.28/100% =1.80%
If buy a Sei Kunyit supplier then buy Idris	4194+234/12937*1000 % =34.22 %	4194/4428x100% =94.71%	34.22x94.71/100% =32.40%
If buy a Dt Musa supplier then buy PTP	367+3528/12937*1000 % =30.10 %	367/3895x100% =9.42%	30.10x9.42/100% =2.83%
If buy a PTP supplier then buy Dt. Musa	3528+367/12937*1000 % =30.10 %	3528/3895x100% =90.57%	30.10x90.57/100% =27.26%
If buy a Dt Musa supplier then buy Talao	367+4341/12937*1000 % =36.39 %	367/4708x100% =7.79%	36.39x7.79/100% =2.83%
If buy a Talao supplier then buy Dt Musa	4341+367/12937*1000 % =36.39 %	4341/4708x100% =92.20%	36.39x92.20/100% =33.55%

If buy a Dt Musa supplier then buy Sei Kunyit	367+4194/12937*100 % =35.25 %	367/4561x100% =8.04%	35.25x8.04/100% =28.34%
If buy a Sei Kunyit supplier then buy Dt Musa	4194+367/12937*100 % =35.25 %	4194/4561x100% =91.95%	35.25x91.95/100% =32.41%
If buy a PTP supplier then buy Talao	3528+4341/12937*100 % =60.82 %	3528/7869x100% =44.83%	60.82x44.83/100% =27.26%
If buy a Talao supplier then buy PTP	4341+3528/12937*100 % =60.82 %	4341/7869x100% =55.16%	60.82x55.16/100% =33.54%
If buy a PTP supplier then buy Sei Kunyit	3528+4194/12937*100 % =59.68 %	3528/7722x100% =45.68%	59.68x45.68/100% =27.26%
If buy a Sei Kunyit supplier then buy PTP	4194+3528/12937*100 % =59.68 %	4194/7722x100% =54.31%	59.68x54.31/100% =32.41%
If buy a Talao supplier then buy Sei Kunyit	4341+4194/12937*100 % =65.97 %	4341/8535x100% =50.86%	65.97x50.86/100% =33.55%
If buy a Sei Kunyit supplier then buy Talao	4194+4341/12937*100 % =65.97 %	4194/8535x100% =49.13%	65.97x49.13/100% =32.41%

The data above is the result of data processing using a minimum support value of 10%. and a minimum value of 10% confidence. Itemset that meet the specified minimum support and minimum confidence values are the best rules.

3.4.2 Support for Candidate Itemset 3

$$Support = \frac{Value\ A + B + C}{Total\ Purchase} x100\% \qquad (3)$$

$$confidence = \frac{Value\ A}{Value\ A+B+C} x100\% \qquad (4)$$

The formula above is the formula used in calculating the association rule (ARM). As shown in the table below:

Table 2. Support for Candidate Itemset 3

Aturan	Support (Value A + B + C / Total Purchase x 100%)	Confidence (Value A / A + B + C x 100%)	Best Rule Support x Confidence / 100%
If you buy PTP and khairunas you will buy Talao	3528+273+4341/12937 x 100% =62.93%	4341/3528+273+4341x100% =53.31%	62.93×53.31/100% =33.54%
If you buy PTP and khairunas you will buy Sei Kunyit	3528+273+4194/12937 x 100% =62.93%	4194/3528+273+4194x100% =52.45%	62.93×52.45/100% =33.00%
If you buy PTP and Talao you will buy Sei Kunyit	3528+4341+4194/12937 x 100% =93.24%	4194/3528+4341+4194x100% =34.76%	93.24×34.76/100%= 32.41%
If you buy Khairunas and Talao you will buy PTP	273+4341+4194/12937 x 100% =68.03%	4194/273+4341+4194x100% =47.61%	68.03×47.61/100% =32.38%

The data above is a calculation of 3 itemset, data that meets the minimum value of support and minimum confidence, then it is declared the best rule. If the value does not meet the minimum support and minimum confidence, which is 10%, a prune will be carried out. A value that meets minimum

support and minimum confidence is the best rule.

4. Conclusion

The conclusion of this research is that the application can predicting the prediction of the supplier along with the load that will be supplied in the following year. The data used to support this application uses supplier supply data from 2017 to 2021.

For further development, this application can not only predict suppliers and supply expenses in the following year, but also can obtain expenditure predictions the company uses other data mining methods.

5. Reference

[1] Sutrisno, Dkk (2013), Penerapan Data Mining Pada Penjualan Menggunakan Metode Clustering Study Kasus PT.Indomarco Palembang.

[2] Nurjoko, Darmawan Abdi, (2015), Penerapan Data Mining Menggunakan Association

[3] Bahrur Roji (2013), Penerapan Metode Association Rule Mining (ARM) Untuk Memprediksi Rencana Penambahan Stok Pupuk Berdasarkan Kebiasaan Pelanggan (Studi Kasus: CV. Tani Makmur Jaya).

[4] Gunadi Goldie, Dkk (2012), Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (Fp-Growth) : Studi Kasus Percetakkan PT.Gramedia.

[5] Nurdin dan astika dewi, (2015), Penerapan Data Mining Untuk Menganalisis Penjualan Barang Dengan Menggunakan Metode Apriori Pada Supermarket Sejahtera Lhokseumawe.

[6] Tumpubolon Kannedi, Dkk (2013), Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat – Alat Kesehatan.

[7] Asa Verano Dwi, (2016), Assosiasi Rules Dan Moving Average Untuk Memprediksi Persediaan Bahan Buku Produksi.

[8] Widodo Prabowo Pudjo, Dkk. 2013. Penerapan data mining dengan matlab. Bandung : Penerbit Pasar Buku Palasari.

[9] Fathansyah. 2012. Basis Data. Bandung: Penerbit Informatika Bandung.