Journals Mantik
Volume 3 Number 3, November 2019 E-ISSN2685-4236
https://iocscience.org/ejournal/index.php/mantik/index

Naïve Bayes Algorithm Implementation To Predict Gum Production at PT. Sri Rahayu Court

Muhammad ibn Hawari, Bosker Sinaga

[1.2] program Information Engineering,
STMIK Pelita Nusantara, Jl, Iskandar Muda No. 1 Medan

E-mail: ibnuhawari5.@ gmail.com

ARTICLE INFO

Article history:
Received: 24 Aug2019
Revised: 24 Sep 2019
Accepted: 11 Oct 2019

***Keywords**:*
Prediction of the rubber;
Data Mining;
naïve Bayes

ABSTRACT

Latex is the production produced from rubber trees by means of tapping using a special knife, and in the plantations there are some plants produce processed in the plant itself becomes a basic ingredient or finished goods to consumers. During this time determine the outcome prediction latex production does not use a method so the results are not in accordance with what diharafkan. This is due to the lack of an objective method to decide the outcome prediction latex production, the rapid selection based on the data of workers tapping some rubber trees with some techniques and ways to get good results. With reference to the solutions provided Naïve Bayes algorithm to help predict the outcome of latex production, A leader collects sap yield of 3 items are Latex, Lump and Treelace. Calculations are done with the third item Naïve Bayes algorithm method with classifications that apply data mining and data probability. Data Mining is generally defined as a system capable of generating and handling of problem solving. Naïve Bayes algorithm can determine the value of the preferences of each alternative, and may be the best alternative from a number of alternatives. ,

1. Introduction

Science and technology developments are now served in various fields, one of them in plantation. Many technologies are applied to assist workers in the field perkebunanmisalnya predict the outcome of latex production and various other activities that facilitate workers or the public.Developments in science and technology is increasingly rapidly now this makes us want more open in the face of many changes that occur due to the progress and development of technology. Advances in technology can support activities in the field of plantation,

Improving the ability to use information effectively is highly beneficial.Data mining is a technique that utilizes large amounts of data to obtain valuable information that can be used for critical decision making. Data mining is an analysis step of the process of knowledge discovery in databases or knowledge discovery in databases at KDD lift. Data mining is the technique of data ataupenambatan a relatively quick and easy to find knowledge, patterns and willing siantar data automatically.

By combining the two purposes, descriptive and predictive. Descriptive data mining means used to search for patterns - patterns that can be understood humans that describes the characteristics of data, whereas the predictive data mining is used to form a model of knowledge that will be used to make predictions.. [1]

2. Research methods

The method used in data collection as a supporter of the research consists of observation, interviews, literature study method.

Journals Mantik
Volume 3 Number 3, November 2019 E-ISSN2685-4236
https://iocscience.org/ejournal/index.php/mantik/index

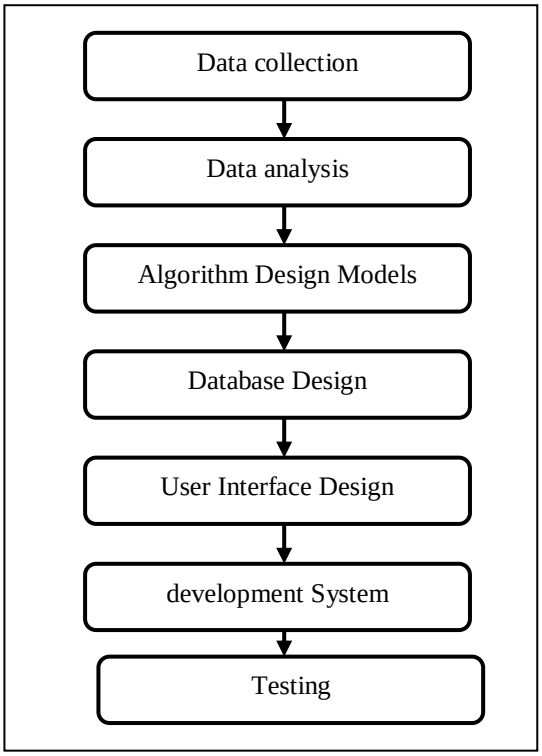

Figure 1. Research Framework

a) Data collection

In collecting the data, the data collection tool perananan very important because the tool is used as a guide or banister researchers for ongoing data collection. In addition there are various kinds of data used pengempulan tool according to the method selected researchers in the research process. In order to obtain the data that is complete, accurate, and reliable truth, researchers use data collection techniques as follows:

b) Data analysis

At this stage is the stage where an analysis of the issues identified are the predicted yield at PT. Sri Rahayu Supreme Plantation Kotarih using Naïve Bayes algorithm and classification.

c) Formulation of the problem

This stage is the process for the elaboration of problem identification and problem limitation. The formulation of this problem is a complete and detailed questions about the scope of the problem that will be tested and designated as the final results will be used to reference the target production of rubber latex production at PT. Sri Rahayu Supreme Plantation Kotarih.

d) Algorithm Design Models

1) Perancanan Model Algorithm performed by stage - following stage
2) Calculate P (Ci) for each class
3) Calculate P (X | Ci) for each criterion and each class
4) Find P (X | Ci) that most be concluded

As the case is predicting the results of latex production would do the work manually by taking three items, namely the production of Latex, Lump and Treelace different sample.

e) Database Design

Database Design is the process to determine the content and regulation of data required to support a variety of system designs. Database design is done in the following manner;

1) Make a table that serves to store data latex production.
2) Queries, functions to manipulate the production data that would predict.
3) Designing Form as a place to enter production data into the database.
4) Switcboard which serves to design the Main Menu to be more attractive to be applied.

f) User Interface Design

Journals Mantik
Volume 3 Number 3, November 2019 E-ISSN2685-4236
https://iocscience.org/ejournal/index.php/mantik/index

Designing Interfaces is the most important part of designing the system. Usually it is also the most difficult part, because in designing the interface must meet three requirements: an interface should be simple, an interface must be complete, and an interface needs to have a fast performance. The main reason why the interface is difficult to plan for each interface is a small programming language: interface describes a set of objects and operations that can be used to manipulate objects.

g) System planning

Designing a system that describes the stages of building design the system. The system design consists of designing Use Case Diagrams, Activity Diagrams and Class Diagrams. This design is based on a direct survey.

h) Implementation

Application has been made to be tested will be implemented at PT. Sri Rahayu Supreme Plantation Kotarih.

i) examination

Tests will be conducted trials in PT. Sri Rahayu Supreme Plantation Kotarih to find bugs or errors in the application.

3. Analysis And Design

At the stage of the analysis that will be built starting with collecting sap production data on methods of predicting outcome in particular latex production Naive Bayes algorithm method. Naive Bayes algorithm method is used to view the data from existing latex production and refers to research that has been done.

Table 1.
Data Training Rubber Production

N0	Month	Rainfall (C1)	Cleanliness (C2)	Expertise Tapper (C3)	Production (C4)
1	January	moderate	Clean	moderate	No.
2	February	moderate	Clean	moderate	Yes
3	March	moderate	Clean	moderate	Yes
4	April	moderate	Clean	Well	Yes
5	May	moderate	Clean	Well	No.
6	June	moderate	Dirty	Well	No.
7	July	moderate	Dirty	Well	No.
8	August	swift	Dirty	Well	Yes
9	September	swift	Dirty	moderate	Yes
10	October	swift	Clean	moderate	Yes
11	November	moderate	Clean	Well	No.
12	December	swift	Dirty	moderate	Yes

Class: 1: Indications Production = Yes
Class: 2: Indications Production of Rubber = No

To determine what the no and yes from the data above, the calculation of the value as below:

Rubber Production indication yes = 7/12 = 0.58
Rubber Production indication no = 5/12 = 0.42

a. Process Methods Naive Bayes classifier

In deciding which areas are often flooded then be calculated by the method of Naïve Bayes classifier to calculate P (x | i) For each class i with the following formula:

Journals Mantik
Volume 3 Number 3, November 2019 E-ISSN2685-4236
https://iocscience.org/ejournal/index.php/mantik/index

$$p\ (x \mid i) = \frac{p(i)\ .p(x|i)}{\sum_{j=1}^{c} p(j).p(x|j)}$$

If : **P (Rainfall (C1) = "Running" | Production (C4) = "Yes") = 4/7 = 0:57**
P (Rainfall (C1) = "Running" | **Production** (C4) = "No") = 4/5 = 0.8
P (Rainfall (C1) = "Medium" | **Production** (C4) = "Yes") = 8/7 = 1:14
P (Rainfall (C1) = "Medium" | **Production** (C4) = "No") = 4/5 = 0.8
(Hygiene (C2) = "Net" | Production(C4) = "Yes") = 7/7 = 1
P (Cleanliness (C2) = "Net" | Production (C4) = "No") = 7/5 = 1.4
P (Cleanliness (C2) = "Dirty" | Production (C4) = "Yes") = 5/7 = 0.71
P (Cleanliness (C2) = "Dirty" | Production (C4) = "No") = 5/5 = 1
P (Expertise(C3) = "Medium" | Production (C4) = "Yes") = 6/7 = 0.85
P (Expertise (C3) = "Medium" | Production (C4) = "No") = 6/5 = 1.2
P (Expertise (C3) = "Good" | Production (C4) = "Yes") = 6/7 = 0.85
P (Expertise (C3) = "Good" | Production (C4) = "No") = 6/5 = 1.2

After calculating each class i in the data table above, the continued production of rubber calculation by determining a class in how good rubber production for the calculation of the probability value

b. Decision Methods Naive Bayes classifier

If the value of p in substitution into x independent unconnected, so he found a new formula as follows

$$P\ (x \mid i) = \prod_{k=1}^{p} p(x^k \mid i)|$$

To find the value of the parameter probability of each hypothesis then obtained class i in determining each kuputusan that panel production is good, then from the above formula dihitunglah probability value, X = (Rainfall = "Cleanliness Rubber", Expertise Tapper) then to calculate P (x | i) for Class "Yes" and "No." Then:

P (X | Production (C4) = "Yes") = 0.57 * 1.14 * 1 * 0.71 * 0.85 * 0.85 * 0.55 = 0.043
P (X | Production (C4) = "No") = 0.8 * 0.8 * 1.4 * 1 * 1.2 * 1.2 * 0.45 = 0.058

From the above results, it appears that the highest probability is in class P (X | Production = No) so that it can be concluded that the production is classified as a "No" means that each month the production of rubber improved where the results of rubber production can be divided into three types of rubber, latex , Lump, Treelace.

c. Implementation

1) Login Form

Login form is the form for entry into the program the system before using the program must first enter your user id and password, if verified then the admin will be entered into the system that has been designed. Display login form can be seen in the picture below.

Figure 2. Display Form Login

2) Display Form Main Menu

Display main menu page is an early look at the time the application is run. The shape of the main menu display can be seen in the picture below.

Journals Mantik
Volume 3 Number 3, November 2019 E-ISSN2685-4236
https://iocscience.org/ejournal/index.php/mantik/index

Figure 3. Main Menu Display

Display main menu has several sub-menus to display other menus, among others, the File Menu, sub menu file consists of Rubber Production data, the data criteria and probability, whereas the decision menuhasil there is a sub-menu that is the decision of the last probabilities exit menu.

3) Form Data Production

Form data input output 5.5 rubber on the image serves to mengimput rubber production results data. This form has five buttons, namely: the save button serves to store data every rubber production, the Change button is used to store data that is converted into data form the results of rubber production, the Delete button is used to delete the data output, the button Clean function to clear data is in the form of production, the search button is used to find the data criteria for the production of rubber and the exit button is used to exit from a form of rubber production data and will return to the main menu.

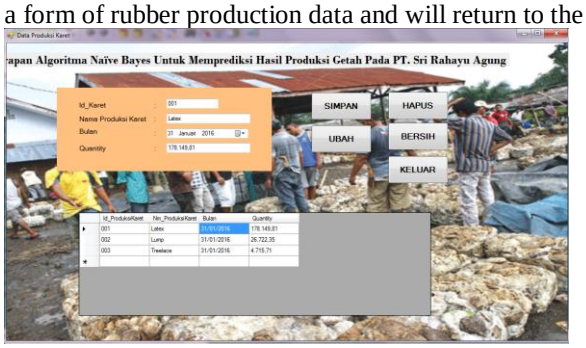

Figure 4. Data Input Form Rubber Production

4) Display Form Sub Decision

See the results in menu form has a sub menu is a sub menu result NBC decision that serves to determine the outcome of the algorithm memprediksi NBC in the production of rubber, the display sub-menu can be seen in the image below:

Figure 5. Display Sub Menu Sub Decision

Journals Mantik
Volume 3 Number 3, November 2019 E-ISSN2685-4236
https://iocscience.org/ejournal/index.php/mantik/index

5) Form Decision NBC

Form data results to see NBC Decision serves to calculate the value of each production class on the rubber, which in this case has some value calculation to determine the class of production to determine the probability value. In this form have a 5 button for that; The button functions to process the attribute values to obtain the results in the method of NBC, the save button serves to store data values from the process, the delete button is used to delete the values that have been stored in the database, the clean button serves to clear existing values on the form and the exit button serves to get out of the form and return the results of NBC's decision on the main menu.

Figure 6. Display Form NBC Decision

4. Conclusion

The following can be concluded from Naive Bayes Algorithm Implementation To Predict Production of the rubber at. Sri Rahayu Supreme running based on the description that has been described in the fifth chapter. The conclusion is a program that has been completed designed and constructed in the form of Naive Bayes Algorithm Implementation To Predict Production of the rubber at. Sri Rahayu Court can run properly. Software used is Visual Basic 2008 and Microsoft Access successfully designed and can be used to predict the outcome of latex production in accordance with the identification of the problem. PT. Sri Rahayu Supreme easier in determining the outcome of latex production precisely and accurately.

5. References

[1] Alfina, T., Santosa, B., & Barakbah, A. R. (2012). Analisa Perbandingan Metode Hierarchical Clustering, K-Means dan Gabungan Keduanya dalam Cluster Data (Studi Kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS). Jurnal Teknik ITS , 1, A521- A525.
[2] Arista, R. (2009). Analisis Sistem Penggajian Pada PMI Cabang Surakarta. Surakarta: Universitas Sebelas Maret.
[3] Fadlina. (2014). Data Mining Untuk Analisa Tingkat Kejahatan Jalanan Dengan Algoritma Association Rule Metode Apriori. Informasi dan Teknologi (INTI) , 3 (1), 144-154.
[4] Fais, S. N., D, A. M., & I, S. M. (2014). KLasifikasi Calon Pendonor Darah Dengan Metode Naive Bayes Classifier. Malang: Universitas Brawijaya Malang.
[5] Gumilang, I. C., Sudjalwo, & Rakhmadi, A. (2014). Prediksi Persediaan Obat Dengan Metode Naive Bayes (Studi Kasus: Apotek Saputra). Surakarta: Universitas Muhammadiyah Surakarta.
[6] Hermawati, F. A. (2013). Data Mining. Yogyakarta: Penerbit Andi.
[7] Junanto, A. (2013). Algoritma Naive Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa. Jurnal Informatika DINAMIK , 18 (1), 9-16.
[8] Meilani, B. D., & Susanti, N. (2014). Aplikasi Data Mining Untuk Menghasilkan Pola Kelulusan Siswa Dengan Metode Naive Bayes. Jurnal LINK , 21 (2), 1-6.
[9] Nugroho, A., & Subanar. (2013). Klasifikasi Naive Bayes untuk Prediksi Kelahiran pada Data Ibu Hamil. Berkala MIPA , 23 (3), 297-308.
[10] Nurani, A., Susanto, B., & Proboyekti, U. (2007). Implementasi Naive Bayes Classifier Pada Program Bantu Penentuan Buku Referensi Mata Kuliah. Jurnal Informatika , 3 (2), 32-36.
[11] Panjaitan, M. R., Maas, L. T., & Tukiman. (2012). Faktor - Faktor Yang Mempengaruhi Anggota Organisasi Kepemudaan Alumni BUdi Mulia (ALBUM-MEDAN) Terhadap Donor Darah di PMI Medan Tahun 2012. Medan: Universitas Sumatera.