K-NEAREST NEIGHBOR METHOD FOR MONITORING OF PRODUCTION AND PRESERVATION INFORMATION (TREATMENT) OF RUBBER TREE PLANT

Fitria1), Hariyanto Wibowo2), Feven Indriyani3)

Informatics & Business Institute Darmajaya1), Informatics & Business Institute Darmajaya2), Informatics & Business Institute Darmajaya3)

*Email:*1fitria@darmajaya.ac.id1), hariyantowibowo@darmajaya.ac.id2)

ABSTRACT

The District in Karang Agung Village, the report of production results by the foreman is still done by making regular reports in the foreman's order, supervision of production by the assistant head of the region is still done by looking at reports from the foreman who is still in the form of regular records in the foreman's book and there are still many rubber plants that are not well developed due to lack of knowledge about how to preserve the rubber plants. The methods used have weaknesses, including foremen having difficulty reporting production data because of the distance of plantations and remote plantation offices. In addition, the assistant head of the region was able to find it difficult to monitor production results because the data provided was still in the form of records and is often late by the foreman. The algorithm analysis used the classification method, K-Nearest Neighbor algorithm. A classification method for a set of data based on previously classified data learning. The nearest Neighbor was an approach to finding a case by calculating the proximity between a new case and an old case. It was based on matching the weights of a number of features.

Keywords: District, production, K-Nearest Neighbor, Classification

INTRODUCTION

In the District in Karang Agung Village, the foreman's production report is still done by making regular reports for the foreman, supervision of production by the assistant head of the region is also still done by looking at the foreman's report which is still a regular record in the foreman's book and there are still many rubber plants that are not well developed due to lack of knowledge about how to preserve (care) the rubber plants. The method used has a weakness, for example, the foreman has difficulty reporting production data due to the distance of plantations and remote plantation offices, besides that assistant head of the region will find it difficult to monitor production results because the data provided is still in the form of notes and is often late given by the foreman. And the difficulty in accessing information if you are still searching manually on a website one to another website. Based on this, a system is needed that can input data on rubber latex production by the foreman, and the assistant head of the region can oversee the production of rubber plants. And the media needed to make it easier to access information about rubber plants. For this reason, an application for monitoring the results of production and information on preservation (maintenance) of rubber plants is built using the K-Nearest Neighbor method which is expected to help the performance in PTP Nusantara Vii in Karang Agung Village Unit.

LITERATURE REVIEW

Rubber plants (Hevea Brasiliensi) are important plantation crops both in the economic context of the community and non-oil and gas foreign exchange sources for the country. Rubber plants come from tropical regions of the Amazon Brazillia with rainfall from 2000 to 3000 mm/year and rainy days between 120-170 days/year [1]. Development of concentrated rubber in areas 10 LU and 10 LS Most areas of Indonesian rubber plantations are located in Sumatra (70%), Kalimantan (24%) and Java (4%) with rainfall of 1500-4000 mm / year with average dry months 0-4 months per year and is located at an elevation below 500 m above sea level. [1]

K-Nearest Neighbor (KNN) Algorithm

KNN is done by finding groups of objects in the closest (similar) training data to objects on new data or testing data. K-Nearest Neighbor Algorithm is a method for classifying objects based on learning data which is the closest distance to the object. Nearest Neighbor is an approach to look for cases by calculating the closeness between new cases and old cases which is based on matching weights of a number of features. The K-NN classification works based on an analogy, where test and training data are compiled and drawn conclusions based on the similarity of the data produced by the comparison.

$$d_{ik} = \sqrt{\sum_{j}^{m} (Cij - Ckj)^2}$$

Figure 1. Formula of Knn Algorithm

Where :
Cij = Sample data
Ckij = New data

K-Nearest Neighbor (KNN) is a method for classifying new objects based on (K) their closest neighbors. KNN includes a supervised learning algorithm, where the results of new query instances are classified based on the majority of the categories on the KNN. The most appearing classes will be the class of classification. [6]

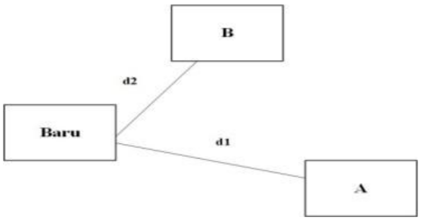


Figure 2. Illustration of K-Nearest Neighbor Algorithm

features. To find which cases to use, the closeness between new cases and all old cases is calculated. The old case with the greatest closeness will be the solution to be used in new cases. [7]

RESEARCH METHOD

Analysis of Algorithms used The steps of the K-Nearest Neighbor algorithm are as follows:

a) Determine the parameter k (number of closest neighbors).
b) Calculates the square of the Euclidean distance of an object against the training data provided
c) Sort results in number 2 in sequence.
d) Collect Y categories (The nearest neighbor classification is based on the value of k)
e) By using the most majority nearest neighbor category, object categories can be predicted.

Application Design (System And Software Design)
Application design is the stage of system design or application which includes the preparation of processes, data, process flow, and fulfillment of needs in accordance with the results of the needs analysis. Application

design documentation generated from this stage is a use case diagram and activity diagram. [8]

UML design

Draft the proposed application case diagram.

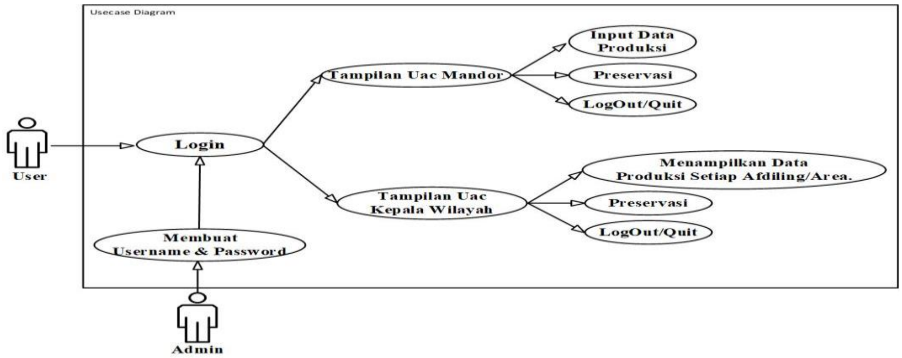


Figure 3. Use Case Proposed Application Chart

FINDINGS AND DISCUSSION

Preservation Page Views

On the preservation page it displays articles about preservation of rubber plants, and this preservation page can be accessed by the assistant head of region. Preservation page views can be seen in Figure 4.

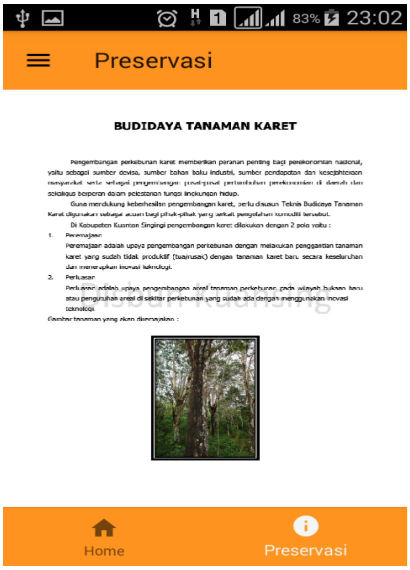


Figure 4. (Preservation Page Views)

Discussion of K-Nearest Neighbor Algorithm

K-Nearest Neighbor (K-NN) algorithm is a method of classification of a set of data based on data learning that has been previously classified. Included in supervised learning, where the results of querying new instances are classified based on the majority of the proximity of the categories in the K-NN. The following is a table of production results at PTP. Nusantara VII unit of Karang Agung Village from January to December 2017.

Table 1. Production Data

Area	Production Results (Kg)											
	Januari	Febuari	Maret	April	Mei	Juni	Juli	Agustus	Sept	Okt	Nov	Des
AFD I	119.912	120.216	149.291	150.221	151.967	161.755	154..211	136.132	113.194	120.125	132.321	150.112
AFD II	100.512	101.912	135.998	137.210	141.003	142.725	138.159	124.250	94.239	101.211	122.150	137.210
AFD III	108.214	110.112	139.514	140.291	139.980	142.252	139.255	126.109	96.112	109.104	124.291	140.191
AFD IV	124.712	124.121	125.100	126.161	130.209	151.202	143.108	138.185	114.397	125.791	135.129	125.721
AFD V	114.292	113.891	172.001	175.161	175.999	172.407	163.204	133.176	101.008	115.921	130.192	172.216
AFD VI	110.912	111.132	147.412	150.126	155.461	161.222	159.301	124.341	98.143	110.213	124.912	146.421
AFD VII	121.319	120.192	163.555	163.692	109.300	170.125	165.112	134.102	104.107	121.111	133.291	164.912
AFD VIII	100.912	101.219	132.191	135.729	136.012	135.001	130.921	114.108	86.231	100.012	115.119	132.191

Production data from January-December 2017 are 96 data. Where in the 2017 production data, the lowest production results were found in the AFD II area in September with the production yield of 86,231 Kg and the highest yield was found in the AFD V area in May with a yield of 175,999 Kg. Classification of production is divided into 3 categories, namely:

a) Low production category: With production yield of ≤120,215 Kg
b) Medium production category: With production results of 20120,216 Kg up to 140,290 Kg
c) Category of high production yield: With production yield ≥ 140,300 Kg

Calculate the distance of new data with training data for each Area using the Euclidean distance formula

$$d_{ik} = \sqrt{\sum_{i}^{m}(Cij - Ckj)^2}$$

Information;
dik = distance
Cij = training data / sample
Ckj = New data
Manual calculation uses the K-nearest Neighbor algorithm with new data or data testing = 100,000 Kg with K = 4.

Table 2. Production Data Calculation Tables

No	Production Data	Calculate Euclidean distance	Smallest Distance	K=4	Category
1	86.231	(86.231-100.000)^2 = 13.769	20	No	Low
2	94.239	(94.239-100.000)^2 = 5.761	11	No	Low
3	96.112	(96.112-100.000)^2 = 3.888	9	No	Low
4	98.143	(98.143-100.000)^2 = 1.857	8	No	Low
5	100.012	(100.012-100.000)^2 = 12	1	Yes	Low
6	100.512	(100.512-100.000)^2 = 512	2	Yes	Low
7	100.912	(100.912-100.000)^2 = 912	3	Yes	Low
8	101.008	(101.008-100.000)^2 = 1.008	4	Yes	Low
9	101.211	(101.211-100.000)^2 = 1.211	5	No	Low
10	101.219	(101.219-100.000)^2 = 1.219	6	No	Low
11	101.219	(101.219-100.000)^2 = 1.219	7	No	Low
12	104.107	(104.107-100.000)^2 = 4.107	10	No	Low
13	108.214	(108.214-100.000)^2 = 8.214	12	No	Low
14	109.104	(109.104-100.000)^2 = 9.104	13	No	Low
15	109.300	(109.300-100.000)^2 = 9.300	14	No	Low
16	110.112	(110.112-100.000)^2 = 10.112	15	No	Low
17	110.213	(110.213-100.000)^2 = 10.213	16	No	Low
18	110.912	(110.912-100.000)^2 = 10.912	17	No	Low

19	111.132	$(111.132-100.000)^2$ = 11.132	18	No	Low
20	113.194	$(113.194-100.000)^2$ = 13.194	19	No	Low
21	113.891	$(113.891-100.000)^2$ = 13.891	21	No	Low
22	114.108	$(114.108-100.000)^2$ = 14.108	22	No	Low
23	114.292	$(114.292-100.000)^2$ = 14.292	23	No	Low
24	114.397	$(114.297-100.000)^2$ = 14.297	24	No	Low
25	115.119	$(115.119-100.000)^2$ = 15.119	25	No	Low
26	115.921	$(115.921-100.000)^2$ = 15.921	26	No	Low
27	119.912	$(119.912-100.000)^2$ = 19.912	27	No	Low
28	120.125	$(119.912-100.000)^2$ = 19.912	28	No	Low
29	120.192	$(120.192-100.000)^2$ =20.192	29	No	Low
30	120.216	$(120.216-100.000)^2$ = 20.216	30	No	Low
31	121.111	$(121.111-100.000)^2$ = 21.111	31	No	Medium
32	121.319	$(121.319-100.000)^2$ = 21.319	32	No	Medium
33	122.150	$(122.150-100.000)^2$ = 22.150	33	No	Medium
34	124.121	$(124.121-100.000)^2$ = 24.121	34	No	Medium
35	124.250	$(124.250-100.000)^2$ = 24.250	35	No	Medium
36	124.291	$(124.291-100.000)^2$	36	No	Medium

		= 24.291			
37	124.341	$(124.341-100.000)^2$ = 24.341	37	No	Medium
38	124.712	$(124.712-100.000)^2$ = 24.712	38	No	Medium
39	124.912	$(124.912-100.000)^2$ = 24.912	39	No	Medium
41	125.100	$(125.100-100.000)^2$ = 25.100	40	No	Medium
41	125.721	$(125.721-100.000)^2$ = 25.721	41	No	Medium
42	125.791	$(125.791-100.000)^2$ = 25.791	42	No	Medium
43	126.109	$(126.109-100.000)^2$ = 26.109	43	No	Medium
44	126.161	$(126.161-100.000)^2$ = 26.161	44	No	Medium
45	130.192	$(130.192-100.000)^2$ = 30.192	45	No	Medium
46	130.209	$(130.209-100.000)^2$ = 30.209	46	No	Medium
47	130.921	$(130.921-100.000)^2$ = 30.921	47	No	Medium
48	132.191	$(132.191-100.000)^2$ = 32.191	48	No	Medium
49	132.191	$(132.191-100.000)^2$ = 32.191	49	No	Medium
50	132.321	$(132.321-100.000)^2$ = 32.321	50	No	Medium
51	133.176	$(133.176-100.000)^2$ = 33.176	51	No	Medium
52	133.291	$(133.291-100.000)^2$ = 33.291	52	No	Medium
53	134.102	$(134.102-100.000)^2$ = 34.102	53	No	Medium

54	135.001	$(135.001-100.000)^2$ = 35.001	54	No	Medium
55	135.129	$(135.129-100.000)^2$ = 35.129	55	No	Medium
56	135.729	$(135.729-100.000)^2$ = 35.729	56	No	Medium
57	135.998	$(135.998-100.000)^2$ = 35.998	57	No	Medium
58	136.012	$(136.012-100.000)^2$ = 36.012	58	No	Medium
59	136.132	$(136.132-100.000)^2$ = 36.132	59	No	Medium
60	137.210	$(137.210-100.000)^2$ = 37.210	60	No	Medium
61	137.210	$(137.210-100.000)^2$ = 37.210	61	No	Medium
62	138.159	$(138.159-100.000)^2$ = 38.159	62	No	Medium
63	138.185	$(138.185-100.000)^2$ = 38.185	63	No	Medium
64	139.255	$(139.255-100.000)^2$ = 39.255	64	No	Medium
65	139.514	$(139.514-100.000)^2$ = 39.514	65	No	Medium
66	139.980	$(139.980-100.000)^2$ = 39.980	66	No	Medium
67	140.191	$(140.191-100.000)^2$ = 40.191	67	No	High
68	140.291	$(140.291-100.000)^2$ = 40.291	68	No	High
69	141.003	$(141.003-100.000)^2$ = 41.003	69	No	High
70	142.252	$(142.252-100.000)^2$ = 42.252	70	No	High
71	142.725	$(142.725-100.000)^2$	71	No	High

		= 42.725			
72	143.108	$(143.108-100.000)^2$ = 43.108	72	No	High
73	146.421	$(146.421-100.000)^2$ = 46.421	73	No	High
74	147.412	$(147.412-100.000)^2$ = 47.412	74	No	High
75	149.291	$(149.291-100.000)^2$ = 49.291	75	No	High
76	150.112	$(150.112-100.000)^2$ = 50.112	76	No	High
77	150.126	$(150.126-100.000)^2$ = 50.126	77	No	High
78	150.221	$(150.221-100.000)^2$ = 50.221	78	No	High
79	151.202	$(151.202-100.000)^2$ = 51.202	79	No	High
80	151.967	$(151.967-100.000)^2$ = 51.967	80	No	High
81	154.211	$(154.211-100.000)^2$ = 54.211	81	No	High
82	155.461	$(155.461-100.000)^2$ = 55.461	82	No	High
83	159.301	$(159.301-100.000)^2$ = 59.301	83	No	High
84	161.222	$(161.222-100.000)^2$ = 61.222	84	No	High
85	161.755	$(161.755-100.000)^2$ = 61.755	85	No	High
86	163.204	$(163.204-100.000)^2$ = 63.204	86	No	High
87	163.555	$(163.555-100.000)^2$ = 63.555	87	No	High
88	163.692	$(163.692-100.000)^2$ = 63.692	88	No	High

89	164.912	$(164.912\text{-}100.000)^2$ = 64.912	89	No	High
90	165.112	$(165.112\text{-}100.000)^2$ = 65.112	90	No	High
91	170.125	$(170.125\text{-}100.000)^2$ = 70.125	91	No	High
92	172.001	$(172.001\text{-}100.000)^2$ = 72.001	92	No	High
93	172.216	$(172.216\text{-}100.000)^2$ = 72.216	93	No	High
94	172.407	$(172.407\text{-}100.000)^2$ = 72.407	94	No	High
95	175.161	$(175.161\text{-}100.000)^2$ = 75.161	96	No	High
96	172.999	$(172.999\text{-}100.000)^2$ = 72.999	95	No	High

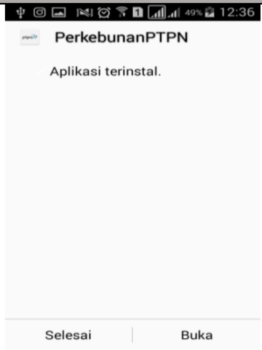
By sorting the smallest distance, K = 4 is taken, and the 4 closest distance is included in the low category, it can be concluded that the production rate of 100,000 Kg includes low production yield.

Software System Testing

Installation Testing

Installation testing is done whether the built-in application can run on the Android operating system.

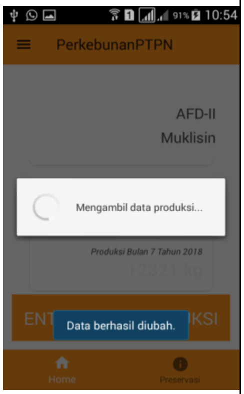
Table 3. Installation testing

No.	Brand	Specification	Test result	Information
1.	Samsung J1 Ace	RAM 0.75GB Android 4.4.4 KitKat Speed HSPA 21.1/5.76 Mbps, LTE Ca t4 150/50 Mbps. Ukuran Layar 4.3 inches		SUCCESS
2	Xiomi Redmi 4x			SUCCESS

Usage Testing
At the stage of testing the use, will be tested on several smartphones as

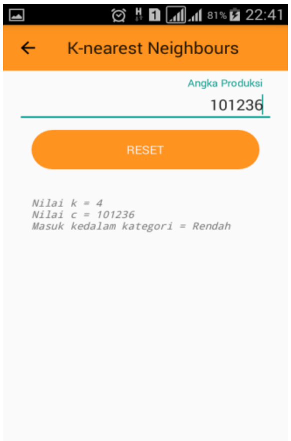
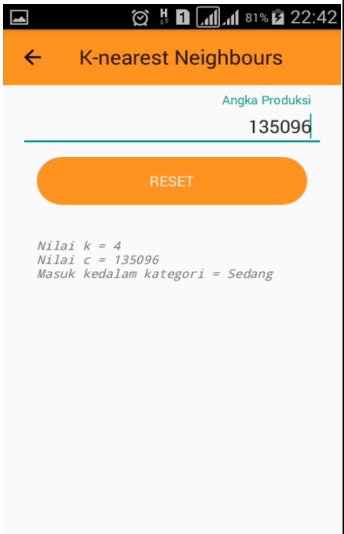
follows: The results of the test can be seen in table 4.

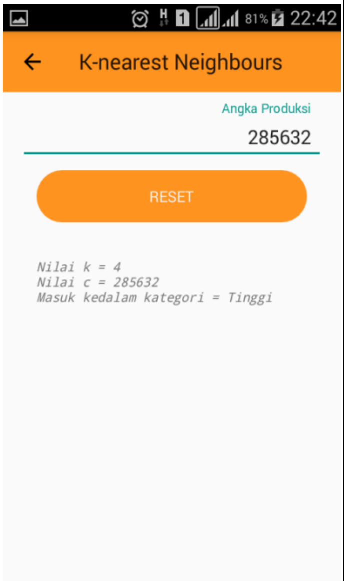
Table 4. Testing on inputting production data

No.	Brand	Specification	Test result	Information
1.	Samsung J1 Ace	RAM 0.75GB Android 4.4.4 KitKat Speed HSPA 21.1/5.76 Mbps, LTE Cat4 150/50 Mbps. Ukuran Layar 4.3 inches		SUCCESS
2	Xiomi Redmi 4x			SUCCESS

Testing of K-Nearest Neighbor Algorithm
K-Nearest Neighbor software testing can be seen in table 5.

Table 5. Software algorithms

No	Brand	Specification	Test result	Information
1	Samsung J1 Ace	RAM 0.75GB Android 4.4.4 KitKat Speed HSPA 21.1/5.76 Mbps, LTE Ca t4 150/50 Mbps.		New data is included in the low category based on the calculation of the algorithm by finding the closest distance to the old data that has been classified into the low, medium and high categories.
2	Samsung J1 Ace	RAM 0.75GB		New data is included in the medium category based on the calculation of the algorithm by finding the closest distance to the old data that has been classified into the low, medium and high categories.

2	Samsung J1 Ace	RAM 0.75GB		New data is included in the High category based on the calculation of the algorithm by finding the closest distance to the old data that has been classified into the low, medium and high categories.

CONCLUSIONS

Using the K-Nearest Neighbor method can help performance at PTP Nusantara VII in Karang Agung Village Unit. This system is proposed to make it easier for the foreman and assistant head of region and foreman to no longer carry out conventional/procedural manuals and can store the database as a production data archive.

REFERENCES

[1] Sabarman Damanik. 2012. Pengembangan Karet *(Havea Brasiliensis)* Berkelanjutan di Indonesia. Bogor : Perspektif Vol.11 No.1 Hlm 91-102 ISSN:1412-8004.

[2] Nazruddin, S. H. (2015). Pemprograman Aplikasi Mobile Smartphone dan Tablet Pc Berbasis Android. Bandung: Informatika

[3] Mudjahidin & Nyoman Dita Pahang Putra. 2010. Rancang Bangun Sistem Informasi Monitoring Perkembangan Proyek Berbasis Web Studi Kasus di Dinas Bina Marga dan Pemantusan. Surabaya : Jurnal Teknik Industri, Vol. 11, No. 1, Febuari 2010: 75-83.

[4] Moh. Rochman Wahid Maulana. 2017. Pengembangan Aplikasi Untuk Studi Bahasa Carakan Madura : PROSIDING ISSN: 2549 – 869X.

[5] Angga Setiyadi & Tati Hariyanti. 2016. Penerapan Sqlite Pada Aplikasi Pengaturan Waktu Ujian dan Presentasi. Bandung : Jurnal Teknik Informatika, Vol.13 No.2.
[6] Nuraini, R. (2015). Desain Algoritma Perkalian Matriks Menggunakan Metode Flowchart. *Jurnal Teknik Komputer Amik BSI,* 144-151.
[7] Toto Andri Puspito. 2017. Menentukan Pilihan Sekolah Didalam Penerimaan Peserta Didik Baru Dengan Menggunakan Metode Nave bayes dan K-Nearest Neighbor (Studi Kasus: PPDB Online Jenjang SMP Kota Metro). Bandar Lampung : Jurnal Informatika, Vol. 17, No. 1.
[8] Roger S.Pressman, Ph.D. (2012). Rekayasa Perangkat Lunak (pendekatan praktisi edisi 7). Yogyakarta: penerbit C.V ANDI.