

INCREASE SALES FOR THE BEST SEED CLUSTERING USING RAPID MINER

Fauzan Azim Institute Technology and Science Mandala JI. Sumatra No. 118 - 120 Jember, Indonesia 082385584973 fauzan@stie-mandala.ac.id Yuli Hartati Institute Technology and Social Sciences Khatulistiwa Jl. Mayor Said Zamzam, Simpang Tiga Ophir, Kabupaten Pasaman Barat 082283073518 Yulihartati802.yh@gmail.com

Eko Afrianto Institute Technology and Science Mandala JI. Sumatra No. 118 - 120 Jember, Indonesia 082329319788 eko.afrianto@itsm.ac.id

ABSTRACT

The superiority of seeds is one of the main factors for farmers. UD. Tiara Bersaudara is a shop that sells seeds and agricultural needs. Poor sales of seeds and fertilizers can hurt UD. Tiara Brothers. To maintain the stock of seeds that are in demand by farmers, sellers must be able to analyze data on the sale of seeds that are in great demand by farmers so as not to cause losses in business. This process is difficult for UD Tiara Brothers to do because it has a lot of sales data. Existing problems can be solved by clustering seed sales data according to most farmers' interests. Clustering is the grouping of data into several clusters based on the level of data similarity. The purpose of the study was to group the best-selling seed data at UD. Tiara Brothers in increasing sales. The clustering method uses the K-Means algorithm by partitioning data into clusters based on the centroid closest to the data. Furthermore, testing by comparing the calculation results with RapidMiner studio software 9.7. The results of this study can later be used as a benchmark for decision support by UD. Tiara Brothers to set marketing strategies for increasing sales.

Keywords: Clustering, Data mining, RapidMiner, K-Means

1. INTRODUCTION

Trading Business (UD) is a form of business whose main activity is to buy and resell goods or services to make a profit. One form of a trading business is UD. Tiara Brothers are engaged in the sale of seeds and agricultural needs. In the process of selling UD. Tiara Brothers often experience obstacles, especially in managing seed supplies for farmers' needs. This obstacle causes a frequent shortage of seed stocks that are in demand because sales are very high. So there is a buildup of seeds that are not in demand in warehouses because sales are low. In this study, seed sales transaction data selection is used as an object for analysis. In sales transaction data, facts are stored that can be extracted and processed to become useful information for trading businesses.

Data mining is a technology that automates the process of finding interesting and sensitive patterns from large sets of data. This allows human understanding to be able to find patterns and scalability of techniques. Data mining techniques are used to perform descriptive mining (describing a grouping of common traits) or predictive mining (trying to predict based on data classification conclusions) on large volumes of data (Siregar, 2018). Data mining is often used in the form of wider applications and can be applied in various areas of life such as business, industry, science, research, and other fields (S, 2019).

Clustering is a technique of grouping several data on a large set of data based on certain criteria. The clustering results are given to the end user to be able to give an idea of what is happening to the database (Ali, 2019). The cluster technique has two methods of grouping data, namely hierarchical clustering and non-hierarchical clustering. Hierarchical clustering is a method of grouping data by grouping two or more data that have similarities or similarities, then forwarding to other objects that have close data, the process will last until a cluster forms a kind of tree where levels or clear hierarchy between objects from the least similar to the most similar (Alkhairi & Windarto, 2019).

Grouping data using the K-Means algorithm is very simple, at the initial stage determine in advance the number of data groups to be determined. After the data grouping is determined, select the document that will be used as a centroid point cluster. Next, iterate until there is stability for all groups of objects that have converged (Yunita, 2018).

In this study, seed sales data will be grouped by calculating the closest distance between the data and the midpoint (centroid) of a cluster. So that the grouping can produce several seed clusters, namely best-selling and nonbest-selling seeds. The application of the K-Means Clustering algorithm is expected to help UD. Tiara Brothers in increasing sales.

The application of the K-Means algorithm was once used to group sales data at one of the outdoor equipment retail stores by producing three clusters. The grouping of data is used to improve stock management and sales strategies (Liu et al., 2018). The application of the K-Means algorithm has also been used as a form of application of pre-consumption monthly load data analysis from electricity consumers from certain domestic areas in China (Bhargava, 2019)

2. RESEARCH METHODOLOGY

In this study, the research method used can be seen in Figure 1. The process will be described starting from data collection, and literature study until the research objectives are achieved and by the initial plan that has been determined. The purpose of this study is to analyze and group data using the K-Means algorithm to design strategies to increase sales.



Figure 1 Method Research

2.1. Identify The Problem

Identifying problems is the initial stage in determining and formulating problems that exist in an object of research. The purpose of determining the problem formulation is so that the research carried out is directed and does not go out of the limits of the problem to be studied. This stage is the first step of a study

2.2. Analyze The Problem

Analyzing problems is a stage to understand problems in the object of research with a predetermined scope so that the results of the research can be as expected.

2.3. Data Collection

Collect the data needed to conduct research. The data collection process in this study is as follows:

a. Observation

Observation is the direct observation made on the object of research. The purpose of observation is so that the existing problems can be known.

b. Interview

The interview is a data collection technique carried out with direct questions and answers between the researcher collector and the resource person. In this study, researchers will conduct interviews with UD. Tiara Brothers. The purpose of the interview is to obtain information related to the existing problem.

c. Literature Study

A literature study is analyzing data and information by reading books, literature, journals, and articles that support research.

2.4. Processing Data Using The K-Means Clustering Algorithm

After all, the data is collected, the data will be analyzed and processed using the K-Means Clustering algorithm. The data to be processed at this stage is seed sales data from January-April 2022 at UD. Tiara Brothers. Data management is done by following the flow chart of the K-Means Algorithm process as seen in Figure 2:



Figure 2 K-Means Clustering Algorithm Flow Diagram

From the flow chart, you can describe the steps of data processing with the K-Means algorithm

- Determine the data to be in the cluster The data to be processed at this stage is seed sales data from January-April 2022 at UD. Tiara Bersaudara.
- b. Determine the number of clusters (k) to be formed to perform K-Means calculations
- c. Determine the center of the cluster (centroid) randomly.
- d. Calculates the distance of each data to the center of the cluster by using the Euclidean formula (1). $D(j,k) = \sqrt{(J1-Ki)^2 + (Ji-Ki)^2}....(1)$

Where:

D(J, K) = Data distance to j to cluster center k

Ji = Data Record

Yi = Data Centroid

- e. Grouping data into clusters with the shortest distance to the center of the cluster.
 - Calculate the new cluster center by using formula (2) :

Where:

f.

CI = Centroid data

J1 = Data to "i" on attribute data to "i"

 $\sum J = Amount of data$

g. If the calculation of the new cluster center has been completed, then some data moves from the results of the previous iteration and repeat steps 2 to 4 until no more data is moving to another cluster.

2.5. Implementing The K-Means Clustering Algorithm On Rapidminer

After the manual calculation process using K-Mean is implemented. The next step is to apply the manual calculation using RapidMiner software. Here's an overview of the calculation steps using the Rapid Miner application:

2.6. Evaluation

The evaluation is carried out to see whether the results of manual calculations with the system testing process that has been made are by user requests. The purpose of the evaluation is to see to what extent the system is functioning and to identify problems that occur

3. RESULT AND DISCUSSION

In this discussion, several steps were taken to analyze the data with software that aims to determine the best seed sales data clustering at UD. Tiara Brothers using Rapid Miner software.

3.1 Data Collection

After the manual calculation process using K-Mean is implemented. The next step is to apply the manual calculation using RapidMiner software. Here's an overview of the calculation steps using the Rapid Miner application.

1. The display of sales data to be tested can be seen in Figure 3:

)	۱ 🖬 ۱	*7 * (** -) =						Rekap	Data Pe	njualan Bibi	t UD.Tiar	ra Bers
	Ho	me Insert Page Layout	Form	ulas	Data	R	eview	View	Der	veloper	Load Te	st
orma	Pag Layo	e Page Break ut Preview Custom Full Views Screen	✓ Ru ✓ GI	uler ridlines essage	Bar	Formu Headir	ila Bar ngs	Zoom	100%	Zoom to Selection	New Windo	Arra w 4
		Workbook Views		2	now/Hi	de			Zoon	n		
-			Te	tal Penius	lan Perhul	-						
	No	Bibit	Januari	Februari	Maret	April						
	1	Bibit Jagung Pioneer P36	60	14	21	20						
	2	Bibit Tomat Ultima F1	6	4	3	8						
	3	Bibit Tomat Sayur Lazata	3	1	0	6						
	4	Bibit Terong Ungu Yuvita F1	5	2	3	1						
	5	Bibit Terong Ungu Yumi F1	1	7	5	8						
-	6	Bibit Terong Ungu Mustang F1	1	4	7	2						
	-	Ribit leaves NK 212	70	122		116						
	0	Bibit Terope Hilps Hitssi F1	4	2		1						
	10	Bibit Sawit polone Super	35	80	25	90						
	11	Bibit Sawit Biasa	70	18	47	28						
	12	Bibit Pare Raden F1	9	3	2	0						
	13	Bibit Pare Pondok Oktan F1	5	9	1	9						
	14	Bibit Pare Pondok	7	9	1	0						
	15	Bibit Onyong Prima F1	9	8	11	2						
	16	Bibit Jagung NK 7328	178	156	149	188						
-	17	Bibit Mentimun Panjang Batara F1	5	4	8	4						
-	18	Bibit Mentimun Hijau Calista F1	6			4						
- H	19	Bible Melon Putin Heni F1	11		10	- 2						
	20	Bibit Jahu Bulat Kuruma F1	3	2	19	1						
	22	Bibit Kanskuns Darat Lesend (Daun Lebar)	17	30	6	-						
	23	Bibit Kanekung Achroporia	19	6	6	9						
	24	Bibit Kacang Panjang Rumana	3	1	6	9						
	25	Bibit Kacang Pajang Kanto Tavi	3	6	4	6						
	26	Bibit Jagung Pionner 23	40	70	80	1						
	27	Bibit Jagung Pioneer 32	210	180	250	198						
	28	Bibit Mentimun Pondok Baresta F1	6	3	0	1						
-	29	Bibit Jagung NK 6172	170	90	88	64						
-	30	Bibit Terong Manggis Provita F1	6	14	11	2						
	51	Bibit Jagung Manis Donanza Bibit Jagung Manis Bill Dutih New Lorenza E1	12		19	21 6						
-	33	Bibit Jagung Manis Diji Putin New Lorenza P1	15	18	35	11						
	34	Bibit Cabe Rawit Seta OP Putih	7	10	18	23						
	35	Bibit Cabe Rawit Pelita F1	12	22	8	1						
	36	Bibit Cabe Rawit Taruna	15	7	30	26						
	37	Bibit Cabe Keriting Kiyu F1	10	9	14	7						
	38	Bibit Cabe Keriting Castilo	6	2	1	1						
	39	Bibit Cabe keriting Baja F1	1	2	0	3						
	40	Bibit Bayam Merah Mira	4	3	5	1						
	41	Bibit Bayam Hijau Maestro	19	32	10	13						
	42	Bibit Bayam Hijau Kartika	1	9	3	9					-	
+ +	E I	ATA AWAL TTERAST 1	TERAST	2 / T	TERASI	3 / 1	TERAS	та / н	ASTL DA	AN KESTMO	IL A N	/ \$

Figure 3 Prepared Sales Data

2. Import repository on Rapid Miner software

After the data in Microsoft Excel is prepared, the next step is to test the results by importing the repository into the RapidMiner software. Import repository in RapidMiner software can be done by selecting Menu import data as seen in Figure 4:

<new process*=""> - RapidMiner Study</new>	o Educational 9.7.002 (0 ayu-PC								(C) (0) (0)
e Edit Brocess View Conne	ctions Bettings Egt	ensions Help								
	 • 	Views De	nige .	Results	Turbo Prep	Auto Modeli	Nore =	Frid Arm, ope	10110.001	P 48 Studio
spository ×	Process							Parameters	20	
OimportData = -	Process				a a	3 E 0	a 🖝 🗵	Process		
Training Rest Import data to a	repository							logverboally	Inc	
Camples								logfie		
Community Samples	0.48						444.0			
DB (caperal)										
naralara -								-	out parameters	
								A Change con	and the read of the	001
ending character										2001
Blending (82)								Belp ×		
Cleansing (29)								Proces	ss	
Bandariana (1988)								KapidMa	ser Situidio Core	
and an and a second sec				man falloss h	ALLAN ON MONTHING	an destand		Synopsis		
Scoring (14) Validation (30)	Leasans the Villager	PROVIDE A COMPANY OF THE PART OF			density and hidden Process			The root operation	for which is the	Report Add and the Report of the
Scoreg (14) Validation (30)	Leverage the Vilado	m of Crowds to get o	1					operator of eve	ry process.	e batter mora.

Figure 4 Import Data

3. Processing Import Data

After selecting the Import Data menu option, then select the data file that has been prepared previously as seen in Figure 5. The Sales Data file is taken From the" SALES DATA" folder with the file name UD Seed Sales Data Recap. Tiara Brothers

	Import Data - Salact the data	instant and a second			-		and, he	1	el thole •
speakery =	-	Select the data locati	on.					10	
O report Data = +	DATE PERLONAN						a -		
Tetrang Researces Annuals Services Community Samples Annuals Community Samples Annuals Community Samples Community Community	Ballmana	199 Yanne 192 Cotto persuadan kitid per taun dan 193 Kenagi Cotta Pengunan Kital (A) Tang Kenawa 199 Kenagi Cotta Pengunan Kital (A) Tang Kenawa	500 12.43 40.43	Type Monset O Monset O	Nor Dec.	Leet Monthe Jul 31, 2009 Aug 10, 2009	Î	tanoretina atom.d.4000	
Branding (82) Crawneng (29) Hosteing (16) Braing (16)	Menap Ciele Penjuarian Bita Ad Files	n UD Trans Derivatidaria Januari April 2018 Assi	The o	orched for wi	i be impo	forf as Excel	*	Auto Cree	3

Figure 5 Import UD Seed Sales Data. Tiara Bersaudara

4. Select Data and Specify Id to Process

After the data import is carried out, the next process is to determine which cells will be used for testing the data that will be used in the Rapid Miner application. Please select the sheet, select INITIAL DATA, and Block cell range from (B6:G49) as seen in Figure 6 and Figure 7:

Edit Process View Conne	ctions Se	ettinos Extensions Hel	D						D	
) 📁 🔚 🔹 🗍	Import Da	ta - Select the cells to impo	ort.	-	Taxan State	-	-		fors_etc 🔑 All	Studio 🖣
pository ×			Sele	ct the c	ells to impo	ort.			*	
O Import Data = +	Obest	DATA ANNAL		86.040	Rolad All	Defes her	-			
Training Resources (connects	Sneet.	DATA AWAL	Cell range	B0:049	Select All	Denne nea	1		init	*
Samples		ITERASI 1	c .	n	F	r	6	н		
Community Samples (connec	111	ITERASI 2	Bibit Terong	5.000	2.000	3.000	1.000			
Local Repository (Local)	12	ITERASI 4	Bibit Terong	1.000	7.000	5.000	8.000			
DB (Legacy)	13	HASIL DAN KESIMPULAN	Bibit Terong	1.000	4.000	7.000	2.000			
	14	7.000	Bibit Terong	8.000	8.000	1.000	10.000			
	15	8.000	Bibit Jagun	79.000	123.000	98.000	116.000		-	
11 > J	16	9.000	Bibit Terong	4.000	2.000	9.000	1.000			
rators ×	17	10.000	Bibit Sawit	35.000	80.000	25.000	90.000		at parameters	
	18	11.000	Bibit Sawit	70.000	18.000	47.000	28.000		and the second second	
ch for Operatora	19	12.000	Bibit Pare R	9.000	3.000	2.000	0.000		http://doi.org/10.4000	
Data Access (53)	20	13.000	Bibit Pare P	5.000	9.000	1.000	9.000			
Blending (82)	21	14.000	Bibit Pare P	7.000	9.000	1.000	0.000			
Cleansing (29)	22	15.000	Bibit Onyon	9.000	8.000	11.000	2.000			
Modeling (166)	23	16.000	Bibit Jagun	178.000	156.000	149.000	188.000		Studio Core	
Scoring (14)	24	17.000	Bibit Menti	5.000	4.000	8.000	4.000			
Validation (30)	in the second	A CONTRACTOR OF A							·	
Set more operators from the						- Erevio	us $\rightarrow Ne$	t X Cancel	process.	ic.

Figure 6 Select Data to Process

De De Doom Jan Cond	Impor	t Date - Form	all year to	rivers.			-		-	-		-	-		i Mudio +
Repeatury =					Format	your	columns.	8						×	
O import Cata = *		Bardes	a kerker i	The rest of the local data	-			_							
W. Tunny Resnarces		Contraction of the local distance of the loc		and control or other										and	
Conversely Samples		No. imper	**	Bible # -	Januari potynominar	**	Fabruari polynominar	* •	Marut potyneminar	**	April polynomical	۰.	1		
Enter Reportery Lanat	1			Bird Jacane Port.	40		14		21		29		~		
DB Lourse	2	2		Bind Tornal Office.											
	3	3		Bine Tornat Sause.	3						6		E		
	4	4		Book Tarong Ung .	4		2.1		3						
				Brod Terrang Ung .			1.		8						
Operators X	4			Bolt Tarong Ung.	1		4		3		2			a supervision of	
and the second second		7		Biolt Tanong Ung .							18			and the second	
lawon the Cavination				Bill Japang Ref.	79		123		99		115			CONTRACTOR OF CONTRACTOR	
Data Accasa (53)				Biok Terospites.	4		2				1				
Bending (82)	10	10		Bild Santpoon.	35		40		25		90				
Deansing (28)	- 11	. 11		Pot Savt Bisco	20		18		47		28				
Modeling (198)	12	12		Bild Pare Radet.	8		3		2				5	their con ;	
Ecosing (14)											One	mether			
- Get none insection ton the									teres Diameters	-	tiet 🗙	Quero		which is the outer me	st

Figure 7 Specify ID Data

5. Save the Process to Repository

Save data that has been implemented can be saved to the repository. The data storage process can be seen in Figure 8:



Figure 8 Process Saving Data to Repository

3.2 Data Testing Process with Rapid Miner studio 9.7

After the data storage process is carried out, the next process is testing the data that has been stored using the Rapid Miner application

1. Testing Data

Testing data process is done by dragging and dropping data that has been stored in the repository to the process as shown in figure 9. Then on the operator interface select the operator K-Means then drag and drop it also to the process as seen in the picture 10.

Bit Bit Strates junc presented juncy Generative pres Bit Bit Strates junc presented in graph Generative pres Bit Bit Strates junc presented in graph Generative pres Bit Bit Strates junc presented in graph Generative pres Bit Bit Strates junc presented in graph Generative pres Bit Bit Strates junc presented in graph Generative press Bit Bit Strates junc press Bit Bit Strat	3 crew prices:"> - Republicer Sta	the Educational 9.7.002 (@ epu-PC					Local Alla good in
Speaklary Press I read finders Press	For For Loose Ann Con	actions getterge dig	Harabara (garg)	Results	fame from	Hufu Moder	Mon *	finishis security, at P Attacks
Outstand in a reserve (5) Outstand (2010) Outst	Repository =	Frecase						Paramatana =
Operative: Image: Control (Control (Contro) (Contro) (Control (Control (Control (Contro) (Control (Contro) (support Data suport Data suportData supportData suppo		Traine Contraction	Antonia and	₽ ₽ }~~[000		4 • R	Antonio Cala Pergenale Sal Antonio
Constants (see Constant	Operators =							C the associate contraction
Gold and a statistical data and a statis	 Statistics (20) Validation (20) 	Learspille Welling	en of Crowda to get speciale	r recommendations t	land on your process	is design!		Tege Louis Searce, Real Distance, Sourcains Eventure, Son Lable, Pennshira, Data Arcono
THE OWNERS AND ADDRESS OF ADDRESS	Gat men apprairie for the Manufation		~	A A CRUSH WITH BUT OF	Crowda			Byreagets This Operator can alcess shored information in

Figure 9 Drag and Drop data to Process

Crew process*a - RapidMiner In	dia Shuntonai 97,002 4	B ayu-PC		-				best@
ine Bat Lacess Ann Con	 edune Setings Edu • • • • • • • • • • • • • • • • • • •	views Beads	Reputs	Turbo Prep	Auto Mudei	51015 ×	Profidate operation, wh	p waas
Repetitory X	Process						Parameters ×	
O Import Data = +	@ Process +			PP		a 🖝 E3	C Ratifieve Data Perguatan B repository entry Data Pe	itat (Retreve) musion Bisk 🔠
Commando Services and Commando Services and Commando Services and Commando Services Commando Commando Services Commando	Antrono Dela Propue Constanta del Constanta Constanta del Constanta del	Bet Bet Role	Anthema		alan B	-	Z Ros adversa and second	
Budeling (5) Segmentation (5) Literate Advante Advante Advante							Hulp X	
B Alleans #	for focus.				a decord		Tage Unsupervised Outwith Services Services Eaching	s Segmentins Sea en Diverses Centre
		1	taute Western of Car	en or logs brook			Kontana Kontana	

Figure 10 Specify Operators

2. Specify Parameter and Doing the Testing Process

After the operator determination process is carried out, the next process is to determine the parameters in the process to be carried out, parameter determination can be carried out by Double clicking the clustering operator then in the parameters section fill in the K value with the number "2" because we will form 2 clusters. Under measure types, select Numerical Measures. In numerical measure select Euclidean Distance as shown in figure 11. To perform the data grouping process with K-Means, click the star as shown in figure 12:







Figure 12 Testing Process

3.3 System Results and Evaluation

A /Const Repository Tempetensolan Data Dergan V. Maan 1 Rep

The results of system testing using Rapid Miner software on previously prepared seed sales data can be seen in figure 13, figure 14, figure 15 and 16. The picture shows the results of grouping data that has been carried out in the Rapid Miner software. In figure 17 can also be seen Grampic and the data is included in cluster 1. In figure 18 can also be seen the Grampic and data included in cluster 2.

	and the second	Contraction of the local division of the loc								- Louis and												
famult History		Charles Martial (2	Destantigi -	I Example!	fat (Charlering	· · · ·	Parlomancelvie	(to) (Fudurmatica)	I	Repeatury =												
-		farts free	4 second				EBM (42)	the constraints		O ingest Data 8 +												
Dette	Box Ba		elected	-	JANEAR .	TERMAN	MARY	AP46.		* Contracting Researchers												
		. 4	autor_0	BORJARON -	88	10	.21	20.	10	* 🗸 Contracto Sangles												
12	1	12	chotec.2	Enter Tormation		4	3			Lord Reporting Long												
Patrice		A	shaller_3	Bid Turnel R.	8					* * 1000												
		4	shales_2	Bood Tarlong -			1	9.	- 81	a 🔤 processes												
			North C	and faring		3				Data Paripanan Elid (concertants and concertants)												
			denter_0	Biot Taning			7	2	- 81	C Parapetersponder Data Hamper A Monte												
		()).	(P)	(2)	(P)	19	2	(P)	(P)	2	(P)	(P)	(P)	(P)	status_3	Book Talong			1	36		• 1 DB
			manual, A	manue, t	manual, 1	mater, t	riterier, t	states, t	riteriar, 4	manue, t	mater, t	manue, t	manue, t	manue, t	BRA.NEUTE	78	943	**	110			
			ranne_0	david Turong	4																	
Contraction of the local division of the loc	10	(44.1)	theolog_0	and head pr.	86	**	20	-														
	**		Dame_P	Avertunera.	.19	74	47	24														
	12	10	skoler_5	BOX Fore Re.																		
	10		there, 5	BOLFING PS.																		
	14	.98	maner, i	-	(F)																	
	10		states, 8	Bank Drucking				1	101													



8.7	aformance/lector (Pr	domanco)	Erange	Set (A.o.a D	ecentry/Data Pa	sjular Ditti			Repeatory =
Stead 11	intery		Charles Model (Charloning)		I Examp	defect (Chasteria	4 ×		O transf Date 8 +
	Open in The	oo The A	an blocket		Filter (42)	42 econotect)	at	*	* 📴 Burligten * 🗃 Linca Reproduty
Deta		citoslar	DIGHT	JANDARI	reaminate	INVIET	APR.		• Catteriliana
0 = 0	34	storter_R	Bod Para Panast	¥.,				10	1 monthlese
12	40.	shader, D	Box Onung Pena Fi			44	1		Cultur President Bladt, Science and Science Burger
Detailes		shaller_3	Bood Jagung NR 7228	376	194	148	188		Pengelangahan Data Denger K. Meana
	47	Scotter, 2	Bibil Hertman Parjang Salara Fil		4				* B DB Laparty
	10	sheeter_P	Bild Workham Highs Califia F1					11	
Vasiatestors	18	shoter,2	Brok Water Public Hars P1		7		2		
	24	datas_2	Brod Stream, Brogga Ras P.7	F		18	22	. 11	
	21	states,0	Bibli Late Bulat Hasama FT	3	2		1		
	10	Maddar_P	Bild Holgsong Data Legen & Staur Le.	17	30				
	-83	theory, 0	Biok Kanphung Agterpona	-9					
	24	skeller_0	Brad Hostang Partyling Humania	8					
	25	shater_0	Biok Karang Papang Kanto Tao	3		4	6		
	29	and the second	Bild Jagang Planter 23	40	10	80	1.8		
	100	adate a	81111 M 10	4.14	100		144	M	

Figure 14 Next Data testing results

8.0	elamanor/vector ph	(fermanin)	- Kental	stict (D. and R.	equilibrium Data Pre	granal Black			Repository -	
Restort 100	atory		Cuari Moni Stateng		E Enamp	in Set (Chestories			Orquittas	1.44
1	Open m	in the A	An Marine		Piller 1421	42 menaness (a	•		Earlyten	
Dans		sherter	8887		remain	MANT	Arrest.		• Correctiona	
	87	iduited_7	West Jugang Phanes 32	218	186	254	100	10.1	* TO entrement	
12	28	chater_0	Blod Mechanish Pander Standard 71	4					Duto Personan Ditti	
Rebellers	29	shaster, t	Stat Jacong NE 6572	179	49	**	-54		🖉 Pergelangakan Data Danga	e killband
	34	steater_0	Elled Tarring Waruppin Private FT	4	14	**	R	117	Ell (apro)	
	38	(Ayettar_B	Rold Japang Harm Danama	0		18	31			
tranklabora	.10	stantar_0	Rood Jospeng Materia Dig Publishers Loose.	4		1				
	- 10	United_B	Bird Japang Marin 91 Stated	98	18	-11	11			
-	24	shatler, if	BILL LINDA FORMED WIN CHI PLAN	1	19	-16	19			
And a local division of the local division o	36	realer_0	Bith Cabe Rawl Parks F1	4	22		1			
	28	shater_0	Bitt Cabe Rout Tatura	11		30	28			
	.37	steater_0	Reed Coston Reariting Name P1				Ŧ.			
	34	ekenini_3	Rold Caller Handing Cardio		1	1	1			
		states_P	Biol Cale tarking Bast 71	4			1			
	4	and the second second		4				of 11		

Figure 15 Next Data testing results

	100 · ·	• (100)	Under Davigs B	instra 1	atis Page - Au	Au bitudai - Disp	animents .		Part and Approximate or P At Shalls
	adamatic Vector (F	witemance)	Elampi	allel (Rocal Ro	postory Data Par	puter Billy			Repeatury =
Read In	latory.		Curu Mohl (During)		Eramp	deSet (Chastaring	6 - K		O Input Data = +
	Open In 📑 Ta	nine ğir	do juydai		Fiber (42)	42 animpted y			* 📜 Damptes * 记 Local Reportery
Data	и.	cleater	807		FEBRUAR	NARET	APRIL		• Connections
-		duning, 6	Biol Tarong Manggis Proeta-F1		14	10		1	* To processes
12	38	charles_0	Bill Juping Haris Donards	0	4	18	21		Colla Pressular Dist. Inde to state and an
nanto.	3.0	challer_0	Bitst Jugung Mante Bill Publi New Loren.			2			Progetompetan Cala Denger K Keane
-	38	challer_8	Bild Jugong Maria 01 Sweet	16	18	36	11		
2	34	chuller_0	Biod Cabe Rawll Sala Of Public	F	90	18	43		
natabes	34	those,0	BOECIDA Raall Palls P1	10	22		8.1		
	.34	chulter_0	Biot Cape Reell Tanna	16	7	30	28		
3	37	chune_0	Biol Calls Karling Klys F1	10		:18	T		
enstations	24	Outer_0	Bolt Care Karling Carlin		2		ŧ		
	38	cluster_8	BORCups torting Baja/1	,	4	4.	3		
	40	chalter_6	Bild Depart Matal Mile		3		4		
	41	shater_6	Bild Stays In Higher Manadhe	19	32	10	43		
	42	thurse, if	Bild Decars How Katha					. 81	

Figure 16 Next Data testing results



Figure 17 Graph and Data Views Included in Cluster 1

	H · • H	treat Design Results	Turtse Prop Auto Model Degloyments		Perg dala, spergitre, etc.	P Al Dubi •
8 P	aftimation/Vector (Performation)	🕺 🚺 ExampleBat (M.a	cal Repository Data Perjudian Bibli	8.	pository 🔀	
Result Ho	atory	Contar Madel (Contaring)	ExampleSet (Clustering)		O Import Cata	
Description	Zoon PP		to 20 30 40 50		Samples Local Repository (Local Connections Connections	
Folder New	🗭 Noce Labers 🗩 Edge Labers	root set	1 00 78 00 100 100 100 100 100 100 100 100 100		Data Pergustan Data /	lengar Kideana i s
A			130 140 150 170			
			190 200 21.5			
Talia			0 220 230 240			
Put			21.0 21.0 20.0 30.0 31.5			
-			32.0	4		

Figure 18 Graph and Data Views Included in Cluster 2

Dis Doca	te fren Connections Segnet	a Epistenia (pri	Internet Cassign Results	Turba Prog AutoModel Depterments	And the growthen at (2) be there a
Pedenamolystar Pedenanco Inampelar (A.aca Repeating State Personan Bint)					Republicly
Read N	ulary	E Claster N	indel (Cheloring)	Evanylatiat (Chataning -	O ImportDate II +
	Attribute JPELAR PEDERLAR NORT NYEL		chuine_8 11285 11727 11474 8789	ckense_1 594.258 595.250 595.250 595.250	Stangies Quero Recentor scene A Consentance Sector Recentor scene Sector Recentor scene Sector Recentor Scene Recentor Rec

Figure 19 Center Cluster for the Last Iteration

Conclusion from the results of tests that have been carried out using RapidMining software, a type of seed that is in great demand by farmers at UD. The Tiara Brothers are NK 212 Corn Seeds, NK 7328 Corn Seeds, Pioneer 32 Corn Seeds, NK 617232 Corn Seeds

4. CONCLUSION

The conclusions that can be generated from research on the best seed clustering using Rapid Miner in increasing sales are as follows: With the application of the K-Means Algorithm to Rapid Miner, it can be applied in clustering and increasing seed sales based on seed sales data for 4 months, from January to April 2022. Grouping seed sales data by dividing seed sales data into 2 clusters. From the results of clustering, 4 types of seeds are in great demand by farmers, namely in cluster 1. NK 212 Corn Seeds, NK 7328 Corn Seeds, Pioneer 32 Corn Seeds, NK 617232 Corn Seeds are types of seeds that are in great demand by farmers. Cluster 2 shows seedlings that farmers are less interested in. With the application of the K-Means algorithm in Rapid Miner, the availability of seed stock can be managed properly, namely by maintaining the availability of seed stock that is most requested by farmers so that it can increase sales.

5. REFERENCES

- Ali, A. (2019). Klasterisasi Data Rekam Medis Pasien Menggunakan Metode K-Means Clustering di Rumah Sakit Anwar Medika Balong Bendo Sidoarjo. MATRIK: Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer, 19(1), 186–195. https://doi.org/10.30812/matrik.v19i1.529
- Alkhairi, P., & Windarto, A. P. (2019). Penerapan K-Means Cluster pada Daerah Potensi Pertanian Karet Produktif di Sumatera Utara. Seminar Nasional Teknologi Komputer & Sains, 762–767.
- Bhargava, A. (2019). Grouping of Medicinal Drugs Used for Similar Symptoms by Mining Clusters from Drug Benefits Reviews. SSRN Electronic Journal, 1049–1058. https://doi.org/10.2139/ssrn.3356314
- Liu, R.-Q., Lee, Y.-C., & Mu, H.-L. (2018). Customer Classification and Market Basket Analysis Using K-Means Clustering and Association Rules: Evidence from Distribution Big Data of Korean Retailing Company. *Knowledge Management Review*, 19(4), 59–76. https://doi.org/10.15813/kmr.2018.19.4.004
- S, I. T. J. (2019). Survey of Data Mining Algorithms for Intelligent Computing System. Journal of Trends in Computer Science and Smart Technology, 01(01), 14–23. https://doi.org/10.36548/jtcsst.2019.1.002
- Siregar, M. H. (2018). Data Mining Klasterisasi Penjualan Alat-Alat Bangunan Menggunakan Metode K-Means (Studi Kasus Di Toko Adi Bangunan). Jurnal Teknologi Dan Open Source, 1(2), 83–91. https://doi.org/10.36378/jtos.v1i2.24
- Yunita, F. (2018). Penerapan Data Mining Menggunkan Algoritma K-Means Clustring Pada Penerimaan Mahasiswa Baru. Sistemasi, 7(3), 238. https://doi.org/10.32520/stmsi.v7i3.388