BAB IV IMPLEMENTASI DAN PEMBAHASAN

4.1 Hasil Implementasi RapidMiner

Pada bab ini akan membahas hasil dari implementasi rapidminer menggunakan algoritma K-Means *Clustering* dalam mengukur Tingkat kedisiplinan pegawai di Badan Pengelola Keuangan dan Aset Daerah Labuhanbatu. Untuk mendukung hasil perhitungan manual yang telah dibahas sebelumnya maka data tersebut diuji menggunakan software RapidMiner. Berikut tampilan awal ketika aplikasi rapidminer dijalankan.



Gambar 4. 1 Tampilan Awal RapidMiner

Setelah itu muncul tampilan awal yaitu membuat lembar kerja baru atau membuka file yang sudah ada. Klik blank process untuk memulai lembar kerja baru

Welcome to	Welcome to RapidMiner Studio!								
Start	Recent Learn								
Start	with								
	Blank Process Start a new process from scratch in the design view.	5	Turbo Prep Prepare your data interactively: transform, clean and combine data sets.	ı¶	Auto Model Build and optimize models using automated machine learning.				
Choo	se a template to start from								
(((⁴)))	Churn Modeling Predict which of your customers will churn and why with a decision tree.	2	Direct Marketing Predict response to campaigns and increase the conversion rate of your campaign.		Credit Risk Modeling Model credit default risk by training an optimized Support Vector Machine (SVM) model.				
	Market Basket Analysis Find products frequently purchased together and turn them into rules for recommendations.	₫.	Predictive Maintenance Model equipment failures to schedule maintenance pre-emptively	**	Price Risk Clustering Cluster price developments using X-Means to unveil price-risk-relationships.				
	Lift Chart Create a lift chart to visualize the improvement that a model provides compared to guessing.	¢	Operationalization Embed predictive models into business processes to trigger the right actions automatically.	\mathcal{P}	Outlier Detection Detect anomalies in data resulting from a chemical analysis of wines.				
	Geographic Distances	Î	Medical Fraud Detection		Web Analytics				

Gambar 4. 2 Memilih Lembar Kerja Awal

📊 <new process=""> – RapidMiner Studio Trial</new>	I 10.3.001 @ LAPTOP-90F3U	UP5							-		×
Eile Edit Process View Connections	Settings Extensions	<u>H</u> elp									
	Views	Design	Results	Turbo Prep	Auto Model	More 🔻	Find	data, operatorsetc	2	All Studio	, •
Repository ×	Process							Parameters	×		
🕒 Import Data 🛛 🖃 💌	Process				P P 🐚	🖪 🛃	i 🥏 🖸	Process			
Training Resources (connected)	Process							logverbosity	init	*	•
Samples	line							logfile			D
Community Samples (connected) Local Repository (Local)	Dub						ies (resultfile]@
DB (Legacy)											
								random seed	2001		٢
								send mail	never	•	· D
Operators ×					rst.			encoding	SYSTEM	•	1
Search for Operators			ta or o								
Data Access (59)											
Blending (81)											
Cleansing (28)											
Modeling (167)											
Sconing (13)											
validation (30)	Leverage the Wisdom of	Crowds to get operat	tor recommendati	ions based on you	ir process design!			Hide advance	d paramete	<u>trs</u>	
Get more operators from the Marketplace			Activate Wis	dom of Crowds				Change com	patibility (10	.3.001)	

Gambar 4. 3 Halaman Lembar Kerja RapidMiner

Selanjutnya proses yang dilakukan ialah penginputan data ke dalam lembar kerja. Maka dibutuhkan operator read excel kemudian *drag and drop* kedalam *view process* sehingga operator tersebut dapat tampil pada lembar kerja seperti Gambar 4.5. Lalu klik import configuration wizard pada parameters operator read excel.

<new process*=""> – RapidMiner Studio Trial 10.3.001</new>	@ LAPTOP-90F3UUP5	– o ×
Eile Edit Process View Connections Setting	s Extensions Help	
	Views: Design Results Turbo Prep Auto Model Interactive Analysis	Find data, operatorsetc 🔎 All Studio 🔻
Repository ×	Process	Parameters ×
🕒 Import Data 🛛 🗉 💌	🕘 Process 90% 🔑 🔎 🖄 📑 📑 🥁 🐲 🖾	Read Excel
Local Repository (Local)	Process	Market Configuration Wizard
Connections	hp Rt	excel file
data processes	Read Excel	sheet selection sh ▼ ①
Uji Data Set Absensi Pegawai (2/13/25	✓ —	sheet number 1
Uji Data Set Absensi Pegawai (TA) (2/		imported cell range
Operators ×		encoding SY 🔻 🛈
Search for Operators		✓ use header row ①
Tota Access (59)		header row 1
 Files (19) Read (13) 		date format
Read CSV		time zone
Read Excel		
Aread SPSS	Leverage the Wisdom of Crowds to get operator recommendations based on your process design!	Iocale En ▼ ① ✓
Get more operators from the Marketplace	Activate Wisdom of Crowds	Change compatibility (10.3.001)

Gambar 4. 4 Import Configuration Wizard

Setelah mengklik import configuration wizard maka akan muncul form data yang akan dipilih untuk diimport pada lembar kerja.

Import Data - Select the data	a location.				\times
	Select the data loc	ation.			
TUGAS AKHIR			▼ ← 🧧	🖡 🛧 🏤 🐸	
Bookmarks	File Name	Size	Туре	Last Modified	
★ Last Directory	Artikel Artikel Berkas TA Grace Selvi Monica Zebua Keuangan Serttifikat Data Set Absensi Pegawai xisx Data setxisx	10 KB 44 KB 19 KB	File Folder File Folder File Folder File Folder File Folder File Folder Microsoft Excel Wor. Microsoft Excel Wor.	Dec 5, 2024 Nov 29, 2024 Oct 28, 2024 Oct 28, 2024 Feb 16, 2025 Oct 28, 2024 Feb 4, 2025 Feb 9, 2025 Jan 20, 2025 Feb 16, 2025	
Data Set Absensi Pegawai.	xlsx				
Excel (.xlsx, .xls)					-
			- Previous	▶ <u>N</u> ext 🔀 ⊆	ancel

Gambar 4. 5 Alur Proses Import Data Tahap 1

Pada Gambar 4.5, Tahap 1 memilih lokasi file data yang akan digunakan. Pilih lokasi penyimpanan data, pilih data yang akan digunakan kemudian klik next. Maka akan muncul form tahap 2 yaitu select the cells to import seperti pada Gambar 4.6 lalu klik next.

<	.3.001 <u>S</u> etting	i⊚ L⁄ gs I	APTOP-90F3UUP5							-	(5	×
		Imp	ort Data - Select the	cells to import.			Inte	ractive	×	Find data, operatorsetc 🛛 🖇	A	JI Stud	Jio 🔻
Repository ×					Select the	cells to impo	ort.			Parameters X			
									Import Configuration Wizard		•		
Connections		Sne	eet Sheeti *	Cell range: A:E		Select All	 Define header row: 	1.*					
data			۵	в	с		D	F		xcei me			J.
processes		1	Nama		т .		DA .	Cuti	~	heet selection	sh		Ð
📌 Tugas Akhir (2/13/25 3:11 PM – 9 kB)		2	Penawai 1	64.000	0.000		0.000	0.000					
Uji Data Set Absensi Pegawai (2/13/2		3	Pegawai 2	63.000	1.000		1.000	0.000	=	heet number 1	1		Ð
📕 Uji Data Set Absensi Pegawai (TA) (2		4	Pegawai 3	55.000	9.000		9.000	0.000					
DB (Legacy)	\sim	5	Pegawai 4	60.000	4.000		6.000	0.000		nported cell range	A1) @
<		6	Pegawai 5	58.000	6.000		6.000	0.000		ncodina	S		œ
Operators ×		7	Pegawai 6	64.000	0.000		3.000	0.000					
·		8	Pegawai 7	59.000	5.000		7.000	0.000		✓ use header row			Ð
Search for Operators		9	Pegawai 8	55.000	9.000		12.000	0.000					
▼ 🛅 Data Access (59)	^	10	Pegawai 9	64.000	0.000		0.000	0.000		eader row	1		(D)
🕶 📇 Files (19)	-	11	Pegawai 10	61.000	3.000		7.000	0.000			D	-	
🔻 📇 Read (13)		12	Pegawai 11	58.000	6.000		7.000	0.000		ate tormat	U	12	W
tead CSV		13	Pegawai 12	60.000	4.000		7.000	0.000		me zone	SY		Ð
📩 Read Excel		14	Pegawai 13	60.000	4.000		6.000	0.000					
📩 Read URL										cale	En	I T	• v
📩 Read SPSS							- Previous	→ Next X Cance	əl	R Hide advanced parameters			
Get more operators from the Marketplace	V				Activate	Nisdom of Crowds				Change compatibility (10.3.001	1		

Gambar 4. 6 Alur Proses Import Data Tahap 2

Kemudian muncul tahap 3 yaitu format your columns, pada tahap ini data polynomial kita ubah role menjadi id seperti pada Gambar 4.7 kemudian klik finish.

Elle Edit Process View Connections S	ettings	Extensions Help							0	
	ſ	mport Data - Format y	our colur	nns.		Int	eractive ×	Find data, operatorsetc	All Stud	tio 🔻
Repository ×				For	mat your column	s.		Parameters ×		
Local Repository (Local)	^	Replace erro	ors with i	missing values				🌮 Import Configuration Wiza	rd	•
 Connections data processes 	i.	Nama polynominal Id	¢ •	H & v integer	TL & 🕈 🔻 integer	PA 🔹 👻	Cuti 🔹 👻 integer	xcel file	د 🚞 sh 🔻) (D) (D
🖑 Tugas Akhir (2/13/25 3:11 PM – 9 kB) 📕 Uji Data Set Absensi Pegawai (2/13/25		1 Pegawai 1 2 Pegawai 2		64 63	0	0	0	heet number	1]@
Uji Data Set Absensi Pegawai (TA) (2/	~	3 Pegawai 3 4 Pegawai 4		55 60	9	9	0	nported cell range	A1	0
Operators ×		5 Pegawar5 6 Pegawar6 7 Pegawar7		58 64 59	6 0 5	6 3 7	0	ncoding	SY *	0
Search for Operators		8 Pegawai 8 9 Pegawai 9		55 64	9 0	12 0	0	eader row	1	0
 Files (19) Eead (13) 		10 Pegawai 10 11 Pegawai 11		61 58	3 6	7 7	0	ate format	•	Ð
Aread CSV	Ļ	12 Pegawai 12		60	4	7	0 🔗 no problems.	me zone	SY ▼	0
Read SPSS	~		_			- Previous	Einish Cancel	Cale	En ▼	•
Get more operators from the Marketplace				× -	Activate Wisdom of Crowds			 Change compatibility (10.3.001) 		

Gambar 4. 7 Alur Proses Import Data Tahap 3

Setelah tahapan import data selesai selanjutnya *drag and drop* operator multiply, berfungsi untuk untuk membuat salinan objek. Operator multiply mengambil objek dari port *input* dan mengirimkan salinannya ke port *output* dimana setiap port yang terhubung membuat salinan yang tidak terikat. Sehingga jika ingin mengubah satu salinan tidak akan mempengaruhi salinan yang lain. Lalu hubungkan operator read

<new process*=""> – RapidMiner Studio Trial 10.3.001 <u>Elle Edit Process View Connections Setting</u></new>	© LAPTOP-90F3UUP5 3 Extensions <u>H</u> elp		-	o ×
	Views: Design Results Turbo Prep Auto Model Interactive Analysis	Find data, operato	rsetc 🔎 .	All Studio 🔻
Repository ×	Process	Parameters	×	
🕒 Import Data 🛛 = 👻	🕘 Process 90% 🔑 🔑 🍺 📑 🛃 🥁 📽 🖾	Process		
Local Repository (Local)	Process	logverbosity	init	•
Connections) up mr.(logfile		0
data				
Tugas Akhir (2/13/25 3:11 PM – 9 kB)		resulune		
Uji Data Set Absensi Pegawai (2/13/25	Read Excel Multiply	random seed	2001	D
Uji Data Set Absensi Pegawai (TA) (2/		rood mail	never	• @
DB (Legacy)		Senaman		
		encoding	SYSTEM	•
Operators ×				
performance ×				
1 Performance (Support Vector Co				
% Performance (Attribute Count)				
Segmentation (4)				
S Cluster Count Performance				
S Cluster Distance Performance				
S Cluster Density Performance				
S Item Distribution Performance	Leverage the Wisdom of Crowds to get operator recommendations based on your process design!	R Hide advance	d parameters	
<		1.01		
No results were found.	Activate Wisdom of Crowds	✓ <u>Unange comp</u>	aubility (10.3.001)	

excel dengan operator multiply

Gambar 4. 8 Input Operator Multiply

Selanjutnya pilih model algoritma yang akan diuji, dalam penelitian ini menggunakan algoritma k-means. Kemudian setting parameter dari masing-masing model clustering yang akan diuji, model ke 1 dengan jumlah *cluster* sebanyak 3 dan model ke 2 dengan jumlah *cluster* sebanyak 4 dan measure types yang digunakan ialah mixed measures. Lalu hubungkan operator multiply dengan operator clustering



Gambar 4. 9 Input Operator Algoritma K-Means

Selanjutnya untuk menilai performance dari hasil *clustering* data yaitu menggunakan operator cluster distance performance, *drag* dan *drop* operator kedalam proses. Kemudian *setting* parameter pada operator tersebut ceklis *normalize* dan *maximize*, dan ubah *main criterion* menjadi davies bouldin untuk mengevaluasi kualitas hasil *clustering* dengan mengukur rasio antar jarak antar klaster dan jarak dalam klaster, jika nilai DBI lebih rendah maka mengindikasikan bahwa clustering tersebut lebih baik dan lebih terstruktur.



Gambar 4. 10 Susunan Operator Read Excel,Multiply, K-Means dan Performance Setelah di *drag* dan *drop* kemudian hubungkan operator clustering dengan operator performance dan hubungkan performance dengan result. Kemudian klik tombol run untuk mengeksekusi data dan menampilkan hasil.



Gambar 4. 11 Ikon Tombol Run

Maka akan muncul hasil performance vector dari kedua clustering yang dijalankan, sehingga dapat dilihat bahwasannya nilai DBI yang lebih rendah adalah nilai dari clustering dengan jumlah klaster sebanyak 3 dengan nilai 0.160, maka yang akan digunakan adalah clustering dengan jumlah klaster sebanyak 3.

Result History % PerformanceVector (Performance (K=3)) ×		Result History		🐒 PerformanceVector (Performance (K=4)) 🛛 🛛		
%	Per	formanceVector	%	Per	formanceVector	
Performance	Perfor Avg. W	rmanceVector: within centroid distance: 4.409	Performance	Perfo Avg. v	rmanceVector: within centroid distance: 2.824	
Description	Avg. v Avg. v Avg. v Davies	<pre>within centroid distance_cluster_0: 3.880 within centroid distance_cluster_1: 11.428 within centroid distance_cluster_2: 1.396 s Bouldin: 0.160</pre>	Description	Avg. n Avg. n Avg. n Avg. n Davies	within centroid distance_cluster_0: 1.254 within centroid distance_cluster_1: 2.671 within centroid distance_cluster_2: 10.594 within centroid distance_cluster_3: 4.182 s Bouldin: 0.168	
Annotations			Annotations			

Gambar 4. 12 Evaluasi Kualitas Hasil Clustering Menggunakan DBI

Sehingga hasil dari klasterisasi data kedisiplinan pegawai yang diperoleh sebanyak 3 klaster yaitu cluster 0 dengan jumlah data sebanyak 30 item, cluster 1 dengan jumlah data sebanyak 16 item dan cluster 2 dengan jumlah data sebanyak 32 item dengan total keseluruhan data ialah 78 item.



Gambar 4. 13 Hasil Klasterisasi Data



Gambar 4. 14 Hasil Cluster 0 (Pegawai yang kurang disiplin)

	▼ 📩 cluster_1
	📄 Pegawai 15
	📄 Pegawai 34
Folder	📄 Pegawai 40
View	📄 Pegawai 41
	📄 Pegawai 42
⊷,	📄 Pegawai 43
Δ	📄 Pegawai 46
Graph	📄 Pegawai 47
	📄 Pegawai 49
	📄 Pegawai 51
	📄 Pegawai 55
Centroid Table	📄 Pegawai 58
Table	📄 Pegawai 61
	📄 Pegawai 62
2	📄 Pegawai 65
Plot	🗋 Pegawai 73

Gambar 4. 15 Hasil Cluster 1 (Pegawai yang tidak disiplin)



Gambar 4. 16 Hasil Cluster 2 (Pegawai yang disiplin)

Dapat dilihat hasil dari data pada gambar 4.17 menggambarkan persebaran klaster dalam bentuk scatter bubble yang dihasilkan dari proses pengelompokan kedisiplinan pegawai. Scatter berwarna biru melambangkan cluster 2, scatter berwarna hijau melambangkan cluster 0, dan scatter berwarna jingga melambangkan cluster 1 dengan nilai klaster ditunjukkan pada sumbu x dan sumbu y. Kelompok pada cluster 0 dikelompokkan sebagai pegawai kurang disiplin (tingkat disiplin sedang), cluster 1 dapat dikategorikan sebagai pegawai tidak disiplin (tingkat disiplin rendah), dan cluster 2 dikelompokkan sebagai pegawai



Gambar 4. 17 Visualisasi Hasil Pemodelan Clustering K-means

Attribute	cluster_0	cluster_1	cluster_2
Н	57.367	49.250	62.844
TL	6.467	14.062	1.125
PA	7.633	15	1.531
Cuti	0.167	0.688	0.031

Gambar 4. 18 Hasil Rata-Rata Centroid Dari Tiap Cluster

4.2 Pembahasan

4.2.1 Perbandingan Hasil Perhitungan Manual dan RapidMiner

Untuk memastikan keakuratan hasil klasterisasi maka dilakukan perbandingan antara perhitungan manual menggunakan Microsoft Excel dan perhitungan otomatis menggunakan RapidMiner. Dari hasil perhitungan dari kedua metode ini menunjukkan adanya konsistensi dalam pengelompokan pegawai ke dalam tiga *cluster* utama, yaitu pegawai disiplin, kurang disiplin, dan tidak disiplin. Perbandingan dalam perhitungan manual (Microsoft Excel):

- a. Perhitungan dilakukan dengan menggunakan metode Euclidean Distance untuk menentukan jarak setiap pegawai ke pusat cluster.
- b. Setelah beberapa iterasi, nilai centroid diperbarui hingga mencapai konvergensi.
- c. Excel menghasilkan tiga kelompok dengan hasil yang sebanding dengan RapidMiner, meskipun membutuhkan waktu lebih lama untuk perhitungan manual.

Perbandingan dalam perhitungan dengan RapidMiner :

- a. RapidMiner mengotomatiskan proses klasterisasi dengan hasil yang lebih cepat dan akurat dibandingkan metode manual.
- b. Evaluasi hasil menggunakan Davies-Bouldin Index menunjukkan bahwa hasil klasterisasi memiliki tingkat pemisahan yang cukup baik.
- c. RapidMiner juga menghasilkan visualisasi klaster yang mempermudah interpretasi pola kedisiplinan pegawai.

4.2.2 Interpretasi Hasil Perhitungan

Hasil perhitungan manual dan RapidMiner menunjukkan bahwa :

- Pegawai dalam Cluster 1 memiliki tingkat kedisiplinan yang tinggi, dengan tingkat kehadiran yang hampir sempurna dan keterlambatan minimal.
- Pegawai dalam Cluster 2 memiliki tingkat kedisiplinan sedang, dengan keterlambatan yang lebih tinggi dibandingkan Cluster 1 dan beberapa pelanggaran aturan.

c. Pegawai dalam Cluster 3 memiliki kedisiplinan yang rendah, dengan jumlah keterlambatan dan pelanggaran aturan yang tinggi, serta beberapa pegawai dengan absensi yang tidak memenuhi standar.

Perhitungan menggunakan RapidMiner lebih efisien dibandingkan dengan metode manual di Excel, terutama dalam menangani dataset yang lebih besar.