

BAB IV HASIL PEMBAHASAN

4.1 Implementasi Sistem

Pada bab ini akan dilakukan implementasi dari data siswa prestasi yang mana data tersebut ada dua dataset yaitu *Training* dan data *Testing* menggunakan Rapidminer 10.3. RapidMiner merupakan sebuah aplikasi untuk pengolahan data dengan menggunakan prinsip dan algoritma *Data Mining*.



Gambar 4.1 RapidMiner Studio 10.3

RapidMiner mampu mengstrak pola-pola dari dataset yang besar dengan mengkombinasikan metode statistika, kecerdasan buatan dan database sehingga sangat cocok di gunakan untuk metode data mining menggunakan Algoritma Naïve Bayes. Aplikasi ini dapat menyelesaikan masalah *Data Mining* diantaranya untuk menghasilkan keputusan.

4.2 Persiapan Data

Pada data ini bertujuan untuk mempersiapkan data dalam bentuk excel dan terdapat dua data set yaitu data *Training* dan data *Testing* agar kita bisa menguji akurasi menggunakan RapidMiner yang terdapat pada gambar 4.2 dan 4.3 setelah itu, data dapat di proses menggunakan RapidMiner.

Nama	Kelas	Nilai akhir	Tahun	Status
x1	X	<1300	2020	Tidak Dapat
x2	X	<1300	2020	Tidak Dapat
x3	XI	<1300	2020	Tidak Dapat
x4	X	>1300	2020	Dapat
x5	XII	>1300	2020	Dapat
x6	XI	<1300	2020	Tidak Dapat
x7	XII	<1300	2020	Tidak Dapat
x8	XII	>1300	2020	Dapat
x9	XI	>1300	2020	Dapat
x10	X	<1300	2020	Tidak Dapat
x11	XI	>1300	2020	Dapat
x12	X	<1300	2020	Tidak Dapat
x13	XII	>1300	2020	Dapat
x14	X	>1300	2020	Dapat
x15	XII	>1300	2020	Dapat
x16	XI	>1300	2020	Dapat
x17	X	<1300	2020	Tidak Dapat
x18	XI	>1300	2020	Dapat
x19	XI	<1300	2020	Tidak Dapat
x20	XII	<1300	2020	Tidak Dapat
x21	X	<1300	2020	Tidak Dapat

MENENTUKAN PROBABILITAS CLASS					
Dapat	0,53				
Tidak Dapat	0,47				
Jumlah	100%				

Atribut	Dapat	Tidak Dapat	Atribut	Dapat	Tidak Dapat
X	14	18	X	0,25	0,37
XI	22	23	XI	0,39	0,45
XII	20	9	XII	0,36	0,18
Jumlah	56	50	Jumlah	100%	100%

Atribut	Dapat	Tidak Dapat	Atribut	Dapat	Tidak Dapat
2020	11	13	2020	0,20	0,24
2021	11	6	2021	0,20	0,12
2022	9	14	2022	0,16	0,29
2023	14	9	2023	0,25	0,18
2024	11	8	2024	0,20	0,16
Jumlah	56	50	Jumlah	100%	100%

Gambar 4.2 Excel Data Training

data nb - Microsoft Excel (Product Activation Failed)

FileHomeInsertPage LayoutFormulasDataReviewViewNitro Pro 9Acrobat

PasteCutCopyFormat PainterClipboardFontAlignmentNumberConditional Formatting as Table StylesInsertDelete Format CellsAutosum Fill Clear Sort & Find & Filter Select

Calibri11Wrap TextGeneralCell

B I U A Merge & Center % + , # 3/4

Conditional Formatting as Table StylesInsertDelete Format CellsAutosum Fill Clear Sort & Find & Filter Select

1114fx

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
110		Nama	Kelas	Nilai	Tahun	Status Bantuan											
111	X1	XI	>1300	2022		Dapat											
112	X2	XII	<1300	2023		Tidak Dapat											
113	X3	XII	>1300	2024		Dapat											
114	X4	XI	>1300	2024		Dapat											
115	X5	XII	<1300	2024		Tidak Dapat											
116	X6	XI	>1300	2024		Dapat											
117	X7	X	>1300	2024		Dapat											
118	X8	X	<1300	2024		Tidak Dapat											
119	X9	XI	<1300	2024		Tidak Dapat											
120	X10	X	>1300	2024		Dapat											
121	X11	XI	<1300	2024		Tidak Dapat											
122	X12	XI	>1300	2024		Dapat											
123	X13	XII	>1300	2024		Dapat											
124	X14	XII	>1300	2024		Dapat											
125	X15	X	<1300	2024		Tidak Dapat											
126	X16	XI	>1300	2024		Dapat											
127	X17	XI	>1300	2024		Dapat											
128	X18	XI	<1300	2024		Tidak Dapat											
129	X19	X	<1300	2024		Tidak Dapat											
130	X20	XII	<1300	2024		Tidak Dapat											
131																	

CONFUSION TABLE

PREDIKSI

Dapat

Tidak Dapat

Dapat

10

1

Tidak Dapat

1

8

ACCURACY

90%

DATA TESTING NO 1

P(Dapat)

P(XII|Dapat)*P(>13000|Dapat)

P(2022|Dapat)*P(Kategori|Dapat)

0,36*1,0*0,16*0,56 =

0,03225600

0,030528

P(TidakDapat)

P(Laki-Laki|TidakDapat)*P(XII|TidakDapat)*P(>13000|TidakDapat)

P(2022|TidakDapat)*P(Kategori|TidakDapat)

Page: 1 of 1

Word: 0 French (BEL)

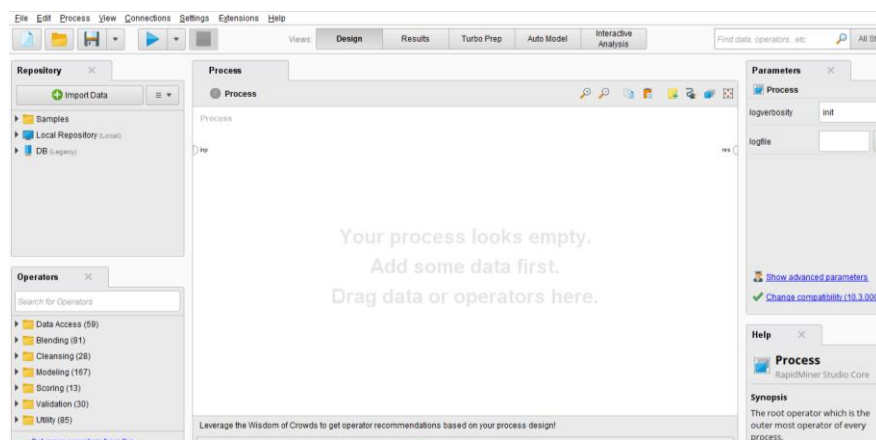
1046

10:27

Gambar 4.3 Excel Data Testing

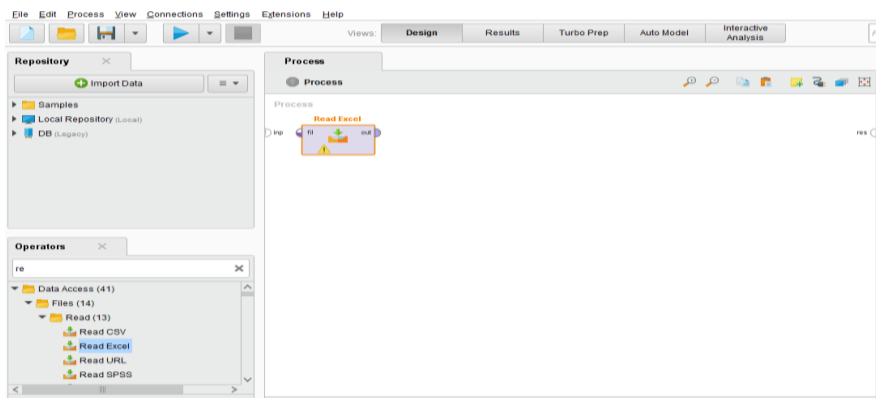
4.3 Pengujian Akurasi dengan RapidMiner Studio 10.3

Pengujian ini bertujuan untuk mengetahui akurasi dari metode Naïve Bayes untuk pengolahan data bantaun siswa prestasi dengan menggunakan data *training* dari angkatan 2020, 2021, 2022, 2023 dan 2024 sebanyak 106 data. Berikut ini proses pengujian data dan tampilan utama RapidMiner.



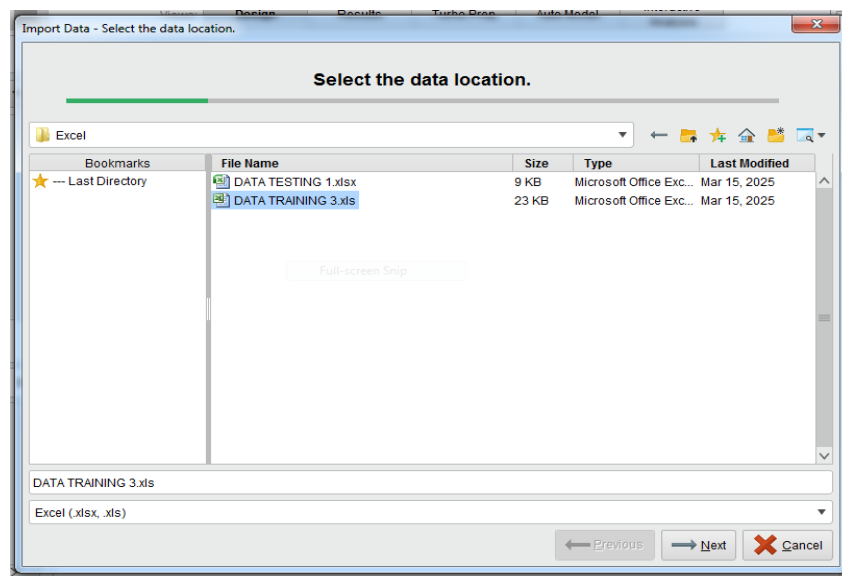
Gambar 4.4 Tampilan Lembaran Kerja

Buka aplikasi RapidMiner, setelah itu akan muncul gambar seperti pada gambar 4.1 maka akan muncul tampilan *start with* kemudian pilih *blank* proses untuk membuka lembaran baru seperti pada Gambar 4.4.



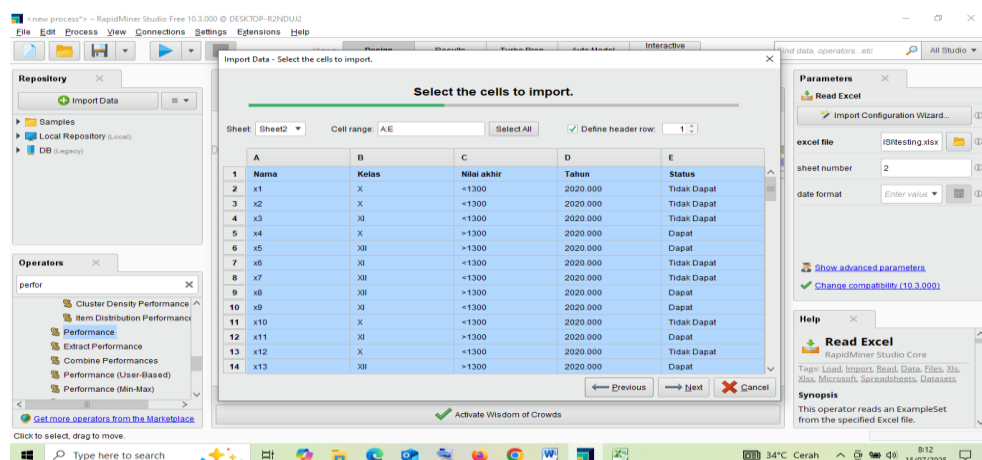
Gambar 4.5 Operator Read Excel

Selanjutnya proses *input* data membutuhkan *operator read excel* ke dalam lembar kerja seperti Gambar 4.5. Proses pengambilan data dengan *double click* pada *operatorread excel* lalu pilih file data lalu klik *Next* seperti Gambar 4.6.

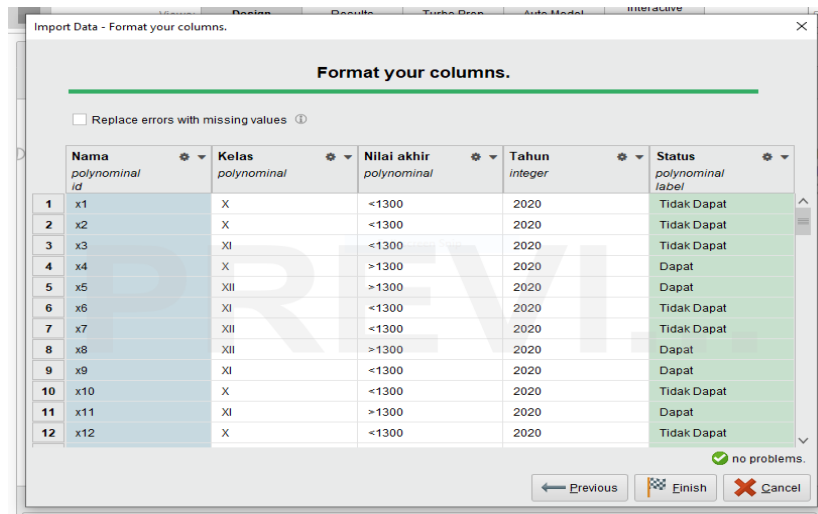


Gambar 4.6 Import Data Training

Selanjutnya pada data yang dipilih seperti Gambar 4.7 setelah itu tetukan tipe data lalu pada kelas data di beri atribut label dengan pilihan *role* dan klik *finish* maka tampilannya seperti Gambar 4.8.

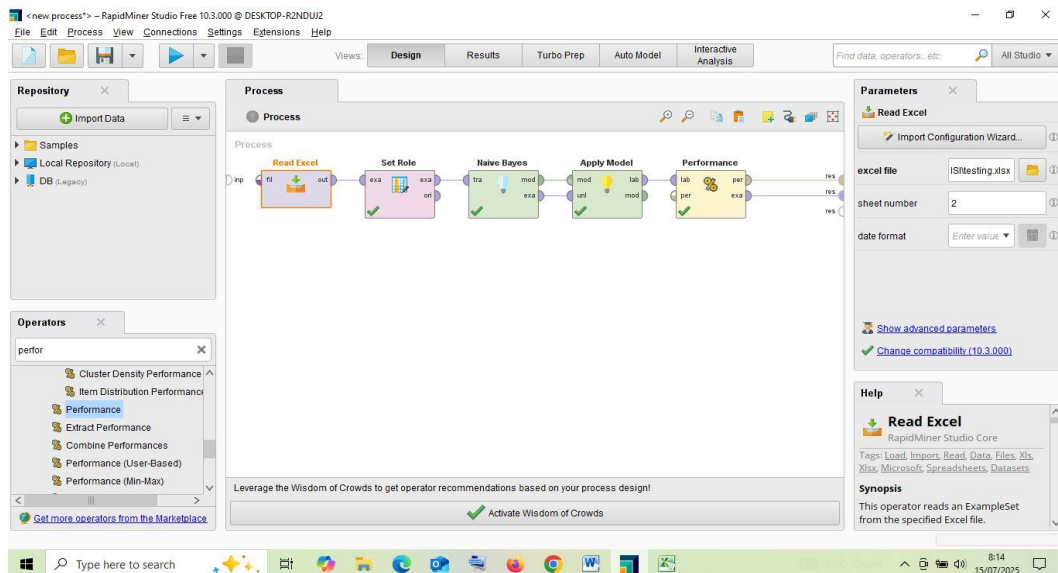


Gambar 4.7 Proses Import Data



Gambar 4.8 Penentuan Tipe Data

Data yang telah selesai diimportasikan di cari akurasi. Kemudian *drag and drop* operator Naïve Bayes, *Set Role*, *Apply Model* dan *Performance* dan hubungkan dengan menarik garis seperti Gambar 4.9.



Gambar 4.9 Proses Akurasi

Setelah proses pada Gambar 4.9 selanjutnya klik ikon *run* untuk menampilkan hasilnya seperti Gambar 4.10.

Row No.	Nama	Status	prediction(S...	confidence(...	confidence(...	Kelas	Nilai akhir	Tahun
1	x1	Tidak Dapat	Tidak Dapat	0.978	0.022	X	<1300	2020
2	x2	Tidak Dapat	Tidak Dapat	0.978	0.022	X	<1300	2020
3	x3	Tidak Dapat	Tidak Dapat	0.970	0.030	XI	<1300	2020
4	x4	Dapat	Dapat	0.000	1.000	X	>1300	2020
5	x5	Dapat	Dapat	0.000	1.000	XII	>1300	2020
6	x6	Tidak Dapat	Tidak Dapat	0.970	0.030	XI	<1300	2020
7	x7	Tidak Dapat	Tidak Dapat	0.936	0.064	XII	<1300	2020
8	x8	Dapat	Dapat	0.000	1.000	XII	>1300	2020
9	x9	Dapat	Tidak Dapat	0.970	0.030	XI	<1300	2020
10	x10	Tidak Dapat	Tidak Dapat	0.978	0.022	X	<1300	2020
11	x11	Dapat	Dapat	0.000	1.000	XI	>1300	2020
12	x12	Tidak Dapat	Tidak Dapat	0.978	0.022	X	<1300	2020
13	x13	Dapat	Dapat	0.000	1.000	XII	>1300	2020
14	x14	Dapat	Dapat	0.000	1.000	X	>1300	2020
15	x15	Dapat	Dapat	0.000	1.000	XII	>1300	2020

Gambar 4.10 Hasil Uji Data

PerformanceVector

PerformanceVector:
accuracy: 98.11%
ConfusionMatrix:
True: Tidak Dapat Dapat
Tidak Dapat: 48 2
Dapat: 0 56
precision: 100.00% (positive class: Dapat)
ConfusionMatrix:
True: Tidak Dapat Dapat
Tidak Dapat: 48 2
Dapat: 0 56
recall: 96.85% (positive class: Dapat)
ConfusionMatrix:
True: Tidak Dapat Dapat
Tidak Dapat: 48 2
Dapat: 0 56
AUC (optimistic): 0.995 (positive class: Dapat)
AUC: 0.994 (positive class: Dapat)
AUC (pessimistic): 0.994 (positive class: Dapat)

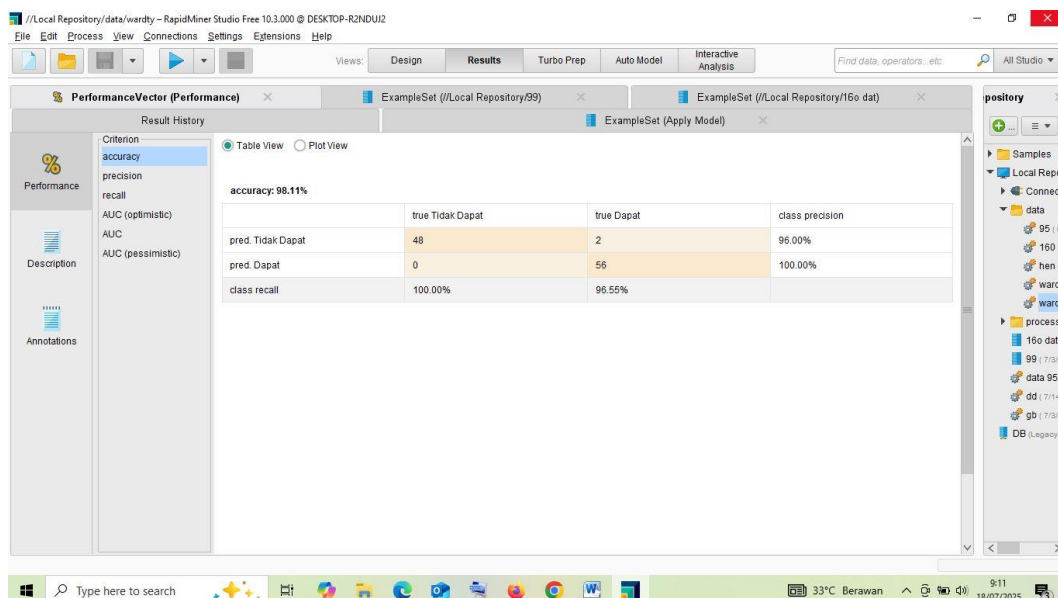
Gambar 4.11 Performance Vector

Pada Gambar 4.11 tersebut menunjukkan hasil evaluasi model di RapidMiner, yang menampilkan metrik kinerja model klasifikasi. Berikut adalah penjelasannya:

- Akurasi (*accuracy*): 98,11%, menunjukkan bahwa nilai model mengklasifikasikan semua data cukup baik.

- b. Confusion Matrix: Menampilkan jumlah prediksi benar dan salah untuk masing-masing kelas (Tidak Dapat dan Dapat). Model memprediksi semua sampel cukup baik.
- c. Precision: 100%, menunjukkan bahwa semua prediksi positif (Dapat) benar.
- d. Recall: 96,55%, berarti model berhasil mengidentifikasi data pada kasus positif Cukup baik.
- b. AUC (Area Under Curve): 0.995 untuk semua skenario (*optimistic*, *standard*, dan *pessimistic*), menunjukkan bahwa model memiliki performa baik dalam membedakan kelas.

Secara keseluruhan, hasil ini mengindikasikan bahwa model bekerja sangat baik dalam klasifikasi.

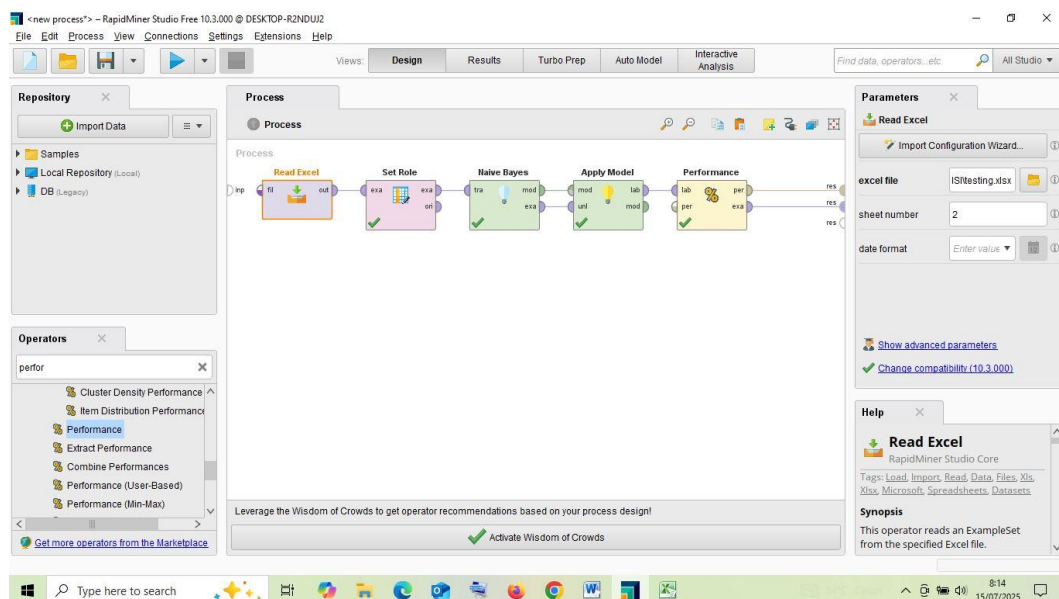


Gambar 4.12 Hasil Performance

Hasil akurasi dari *performance* vektor mendapatkan nilai akurasi sebesar 96,55%. dengan jumlah 106 data *Training*, sebanyak 50 diprediksi tidak dapat, 56 siswa dinyatakan dapat bantuan.

4.3.1 Prediksi dengan Rapidminer Studio 10.3

Prediksi bantaun siswa prestasi dengan data *testing* berjumlah 20 data dan data *training* berjumlah 106 data. Proses prediksi dimulai dengan tambahkan dua operator read excel untuk data training dan data testing, lalu tambahkan operator Naïve Bayes, *Apply Model* dan *Performance* kemudian hubungkan dengan cara menarik garis data training ke operator Naïve Bayes, lalu hubungkan operator *Set Role* Naïve Bayes dan data testing ke *Apply Model*, selanjutnya hubungkan *Apply Model* ke *Performance* dan selanjutnya hubungkan *Performance* ke result disisi kanan seperti Gambar 4.13. Setelah itu jalankan dengan klik ikon *run* maka akan menampilkan hasil prediksi seperti Gambar 4.14.



Gambar 4.13 Tampilan Model Pengujian Data

Row No.	Nama	Status Bant...	prediction(S...	confidence(...	confidence(...	Kelas	Nilai	Tahun
1	X1	Dapat	Dapat	1.000	0.000	XI	>1300	2022
2	X2	Tidak Dapat	Tidak Dapat	0.453	0.547	XII	<1300	2023
3	X3	Dapat	Dapat	0.815	0.185	XII	>1300	2024
4	X4	Dapat	Dapat	0.897	0.103	XI	>1300	2024
5	X5	Tidak Dapat	Tidak Dapat	0.057	0.943	XII	<1300	2024
6	X6	Dapat	Dapat	0.897	0.103	XI	>1300	2024
7	X7	Dapat	Dapat	0.747	0.253	X	>1300	2024
8	X8	Tidak Dapat	Tidak Dapat	0.039	0.961	X	<1300	2024
9	X9	Tidak Dapat	Tidak Dapat	0.106	0.894	XI	<1300	2024
10	X10	Dapat	Dapat	0.747	0.253	X	>1300	2024
11	X11	Tidak Dapat	Tidak Dapat	0.106	0.894	XI	<1300	2024
12	X12	Dapat	Dapat	0.897	0.103	XI	>1300	2024
13	X13	Dapat	Dapat	0.815	0.185	XII	>1300	2024
14	X14	Dapat	Dapat	0.815	0.185	XII	>1300	2024
15	X15	Tidak Dapat	Tidak Dapat	0.039	0.961	X	<1300	2024

Gambar 4.14 Hasil Prediksi Pengujian Data

PerformanceVector

PerformanceVector:
accuracy: 90.00%
ConfusionMatrix:
True: Dapat Tidak Dapat
Dapat: 10 1
Tidak Dapat: 1 8
precision: 88.89% (positive class: Tidak Dapat)
ConfusionMatrix:
True: Dapat Tidak Dapat
Dapat: 10 1
Tidak Dapat: 1 8
recall: 88.89% (positive class: Tidak Dapat)
ConfusionMatrix:
True: Dapat Tidak Dapat
Dapat: 10 1
Tidak Dapat: 1 8
AUC (optimistic): 0.929 (positive class: Tidak Dapat)
AUC: 0.899 (positive class: Tidak Dapat)
AUC (pessimistic): 0.869 (positive class: Tidak Dapat)

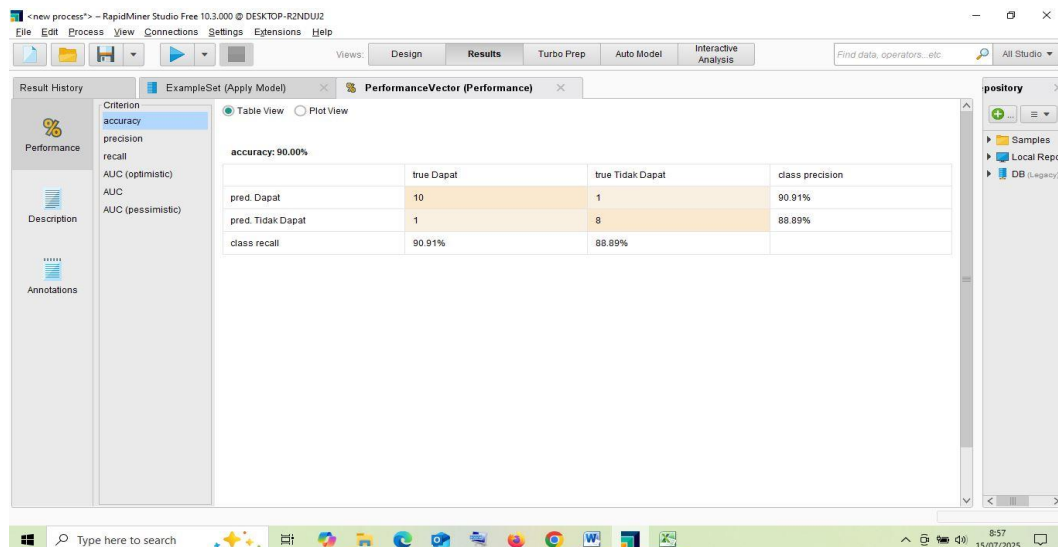
Gambar 4.15 Performance Vector

Pada Gambar 4.15 ini menampilkan hasil evaluasi model klasifikasi dalam RapidMiner melalui metrik *Performance Vector*. Model yang diuji menunjukkan performa optimal dengan hasil sebagai berikut:

- Akurasi (*accuracy*): 90,00%, yang menunjukkan bahwa seluruh data uji diklasifikasikan cukup baik.

- b. Confusion Matrix: Model berhasil memprediksi seluruh sampel dengan benar, di mana 10 data Dapat dan 8 data Tidak Dapat diklasifikasikan secara akurat tanpa kesalahan.
- c. Precision: 88,89% yang berarti semua prediksi positif (dapat) cukup baik.
- d. Recall: 88,89% yang menunjukkan bahwa model mampu mengidentifikasi seluruh kelas positif (dapat) dengan baik.
- e. AUC (Area Under Curve): 0,929 dalam semua skenario (*optimistic, standard, dan pessimistic*), yang mengindikasikan bahwa model memiliki kemampuan maksimal dalam membedakan kelas.

Berdasarkan hasil evaluasi ini, model yang digunakan memiliki tingkat akurasi yang baik dalam menentukan status klasifikasi.



Gambar 4.16 Performance

Pada Gambar 4.16 Hasil akurasi dari *Performance* mendapatkan nilai akurasi sebesar 90,00%. dengan jumlah 20 data *Testing*, sebanyak 8 diprediksi tidak dapat, 10 siswa dinyatakan Dapat Bantuan.

4.4 Pembahasan

Penelitian ini menerapkan metode Naive Bayes. Data yang didapatkan berasal dari data bantaun siswa prestasi SMA Swasta Has Sepakat Bilah Hilir. Data yang diperoleh dibersihkan dan dikelompokkan berdasarkan atribut yang digunakan. Adapun atribut yang digunakan yaitu jenis kelamin, kelas, nilai akhir, tahun angkatan, dan status bantuan. Data yang diambil dibagi dua kelompok yaitu sebagai data *Training* dengan jumlah 106 data dan data sebagai data *Testing* berjumlah 20 data. Pada penelitian ini memiliki dua kelas data yaitu (Dapat) dan (Tidak Dapat). Selanjutnya pengujian dengan RapidMiner Studio 10.3 sebagai alat bantu untuk memperoleh tingkat akurasi dari metode Naïve Bayes. Berikut adalah perhitungan untuk mendapatkan nilai akurasi dengan persamaan berikut: **Akurasi** = $\frac{\text{Jumlah Prediksi Benar}}{\text{Total Data}} \times 100\%$

Dari tabel yang diberikan, prediksi benar untuk kelas "Dapat" = 10, prediksi benar untuk kelas "Tidak Dapat" = 8, Total data uji = 10 + 8 = 18

Sehingga:

$$\text{akurasi} = \frac{10+8}{20} \times 100\% = \frac{18}{20} \times 100\% = 90\%$$

Hasil ini sesuai dengan nilai akurasi yang ditampilkan dalam RapidMiner, yaitu 90%. Nilai 90% membuktikan metode Naïve Bayes dapat digunakan untuk melakukan pengolahan data yang baik. Nilai akurasi ini sangat baik untuk melakukan prediksi pada *data testing*. Nilai akurasi 90% Prediksi dilakukan dengan 20 *data Testing* dengan menggunakan metode Naïve Bayes mendapatkan hasil sebanyak 10 diprediksikan (Dapat), 8 diprediksikan (Tidak Dapat).