

MODUL 8

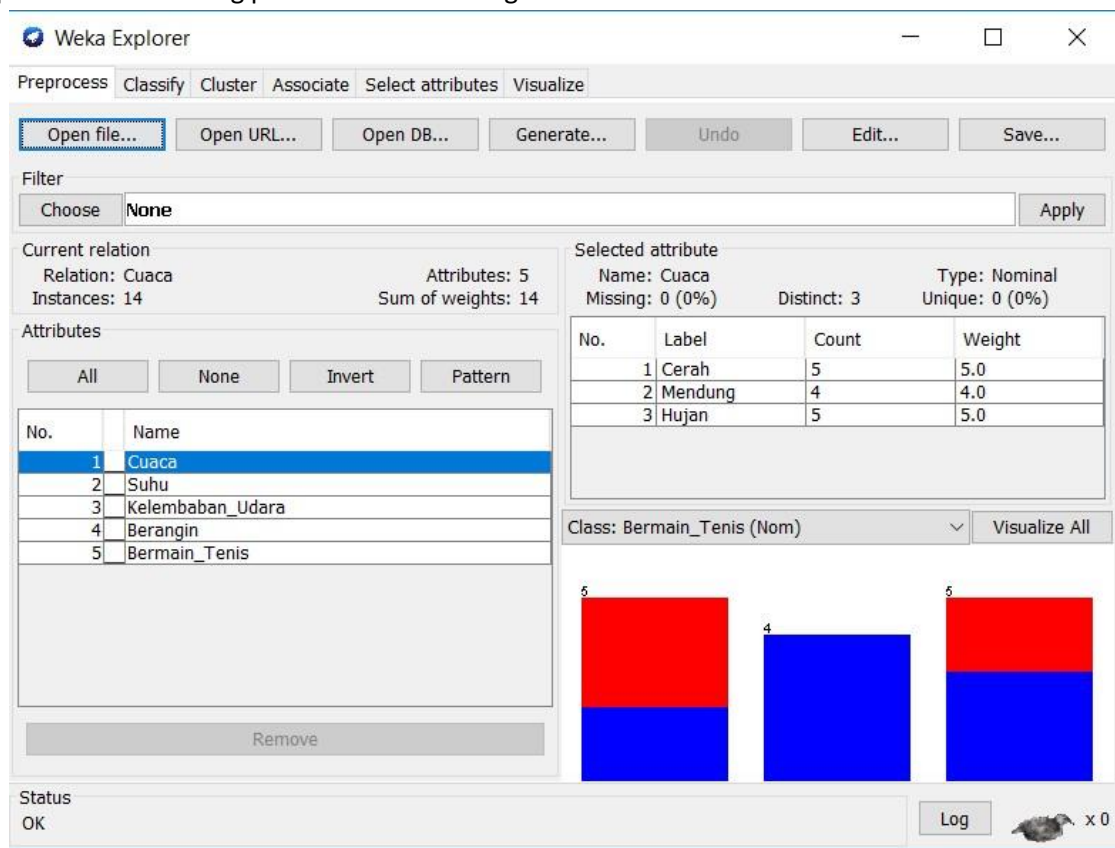
KLASIFIKASI : NAIVE BAYES

8.4 Langkah - langkah Praktikum

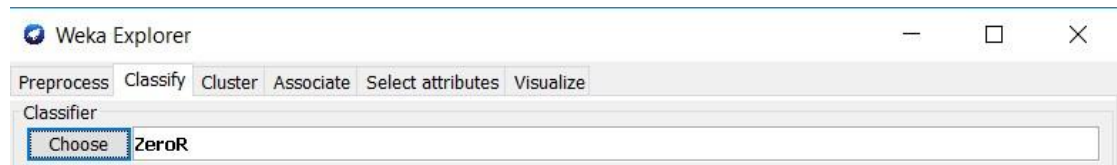
8.4.1 Implementasi Naive Bayes dengan Weka

Langkah - langkah menggunakan algoritma naive bayes dengan Weka sebagai berikut :

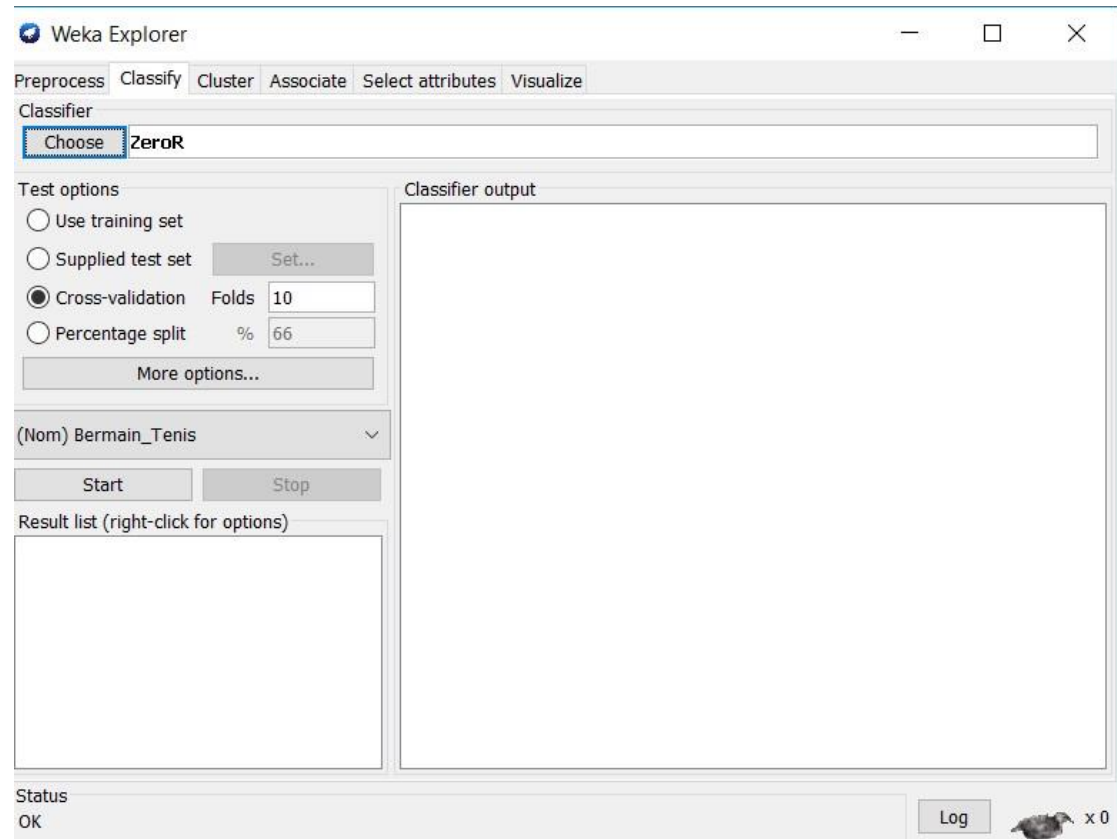
1. Persiapkan file Cuaca.arff dari hasil percobaan kegiatan 7.4.1 pada Modul 7. file ini akan kita gunakan sebagai data training.
2. Buatlah sebuah data testing dengan format ARFF dari tabel 8.1 sebagai data uji yang akan diprediksi dengan memiliki variabel- variabel independen dan variabel dependen yang sama. Dengan ketentuan variabel dependen diisi dengan tanda tanya (?). asumsi bahwa kita belum mengetahui nilai / kelas dari variabel tersebut. Nilai / kelas inilah yang akan kita prediksi dengan menggunakan algoritma naive bayes
3. Simpan dengan nama CuacaTesting.arff
4. Jika telah selesai membuat buka aplikasi weka, masuk dalam menu Weka Explorer
5. Buka kembali file Cuaca.arff dari hasil kegiatan 7.4.1 pada Modul 7 dengan menggunakan Weka Explorer. File ini akan kita gunakan sebagai data pelatihan untuk memprediksi data testing pada file CuacaTesting.arff

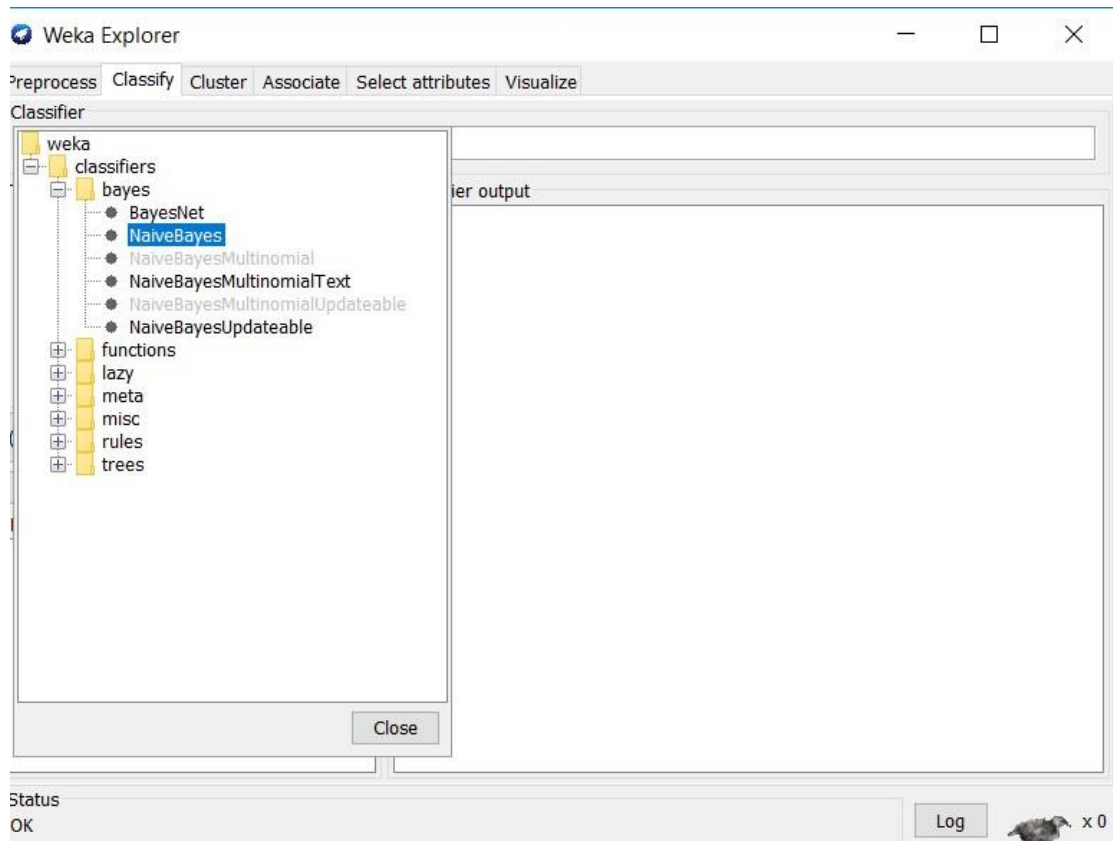


6. Masih pada jendela Weka Explorer, pilih tab Classify



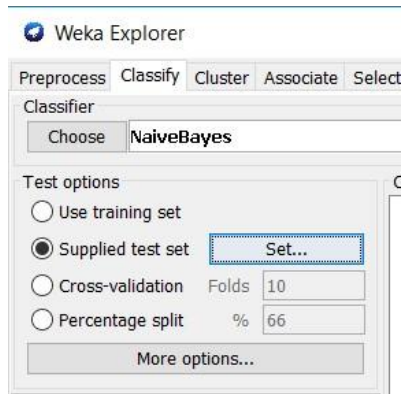
7. Sehingga akan muncul jendela Weka Explorer pada tab Classify. Pada kotak Classifier tombol Choose untuk memilih metode / algoritma Naive Bayes



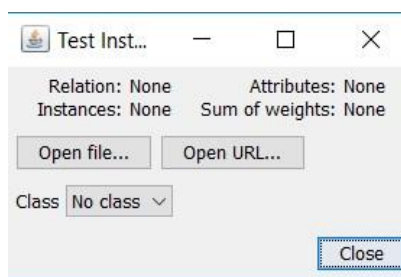


8. Selanjutnya adalah memnetentukan data testing sebagai data yang akan diprediksi variabel dependennya. File Cuacatesting.arff ditentukan sebagai data testing pada kegiatan ini.

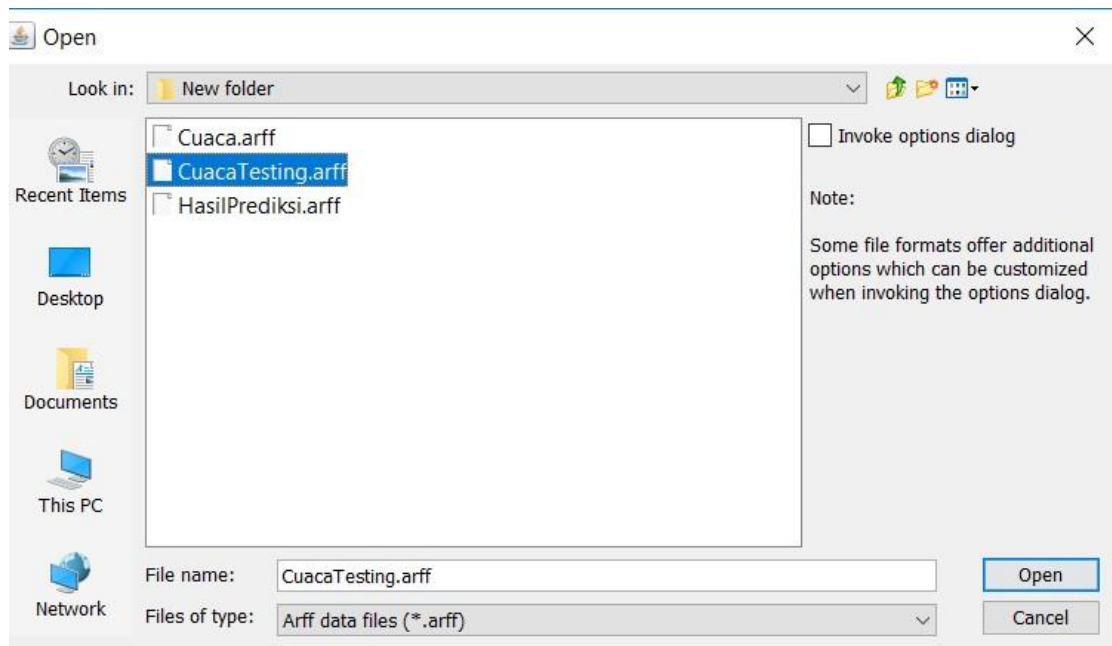
9. Pada percobaan kali ini, kita akan menggunakan pilihan Supplied test set Klik tombol Set untuk menentukana file ARFF sebagai data uji.



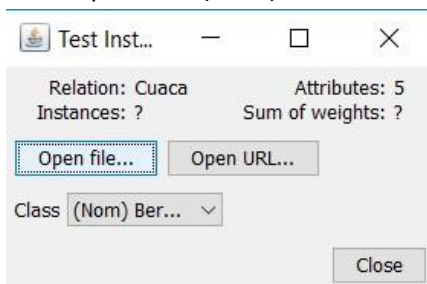
10. Sehingga akan muncul jendela test Intance. Klik Open file..



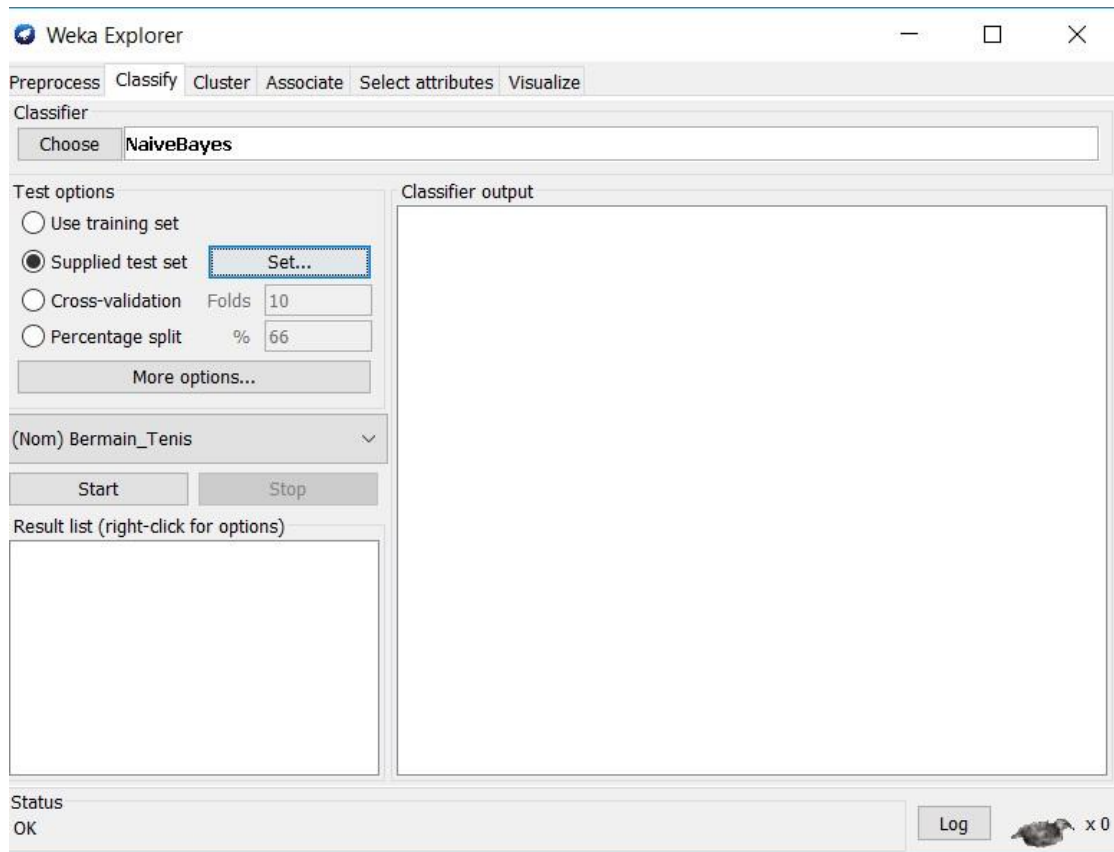
11. Pilih file Cuaca Testing.arff sebagai data uji. Klik Open



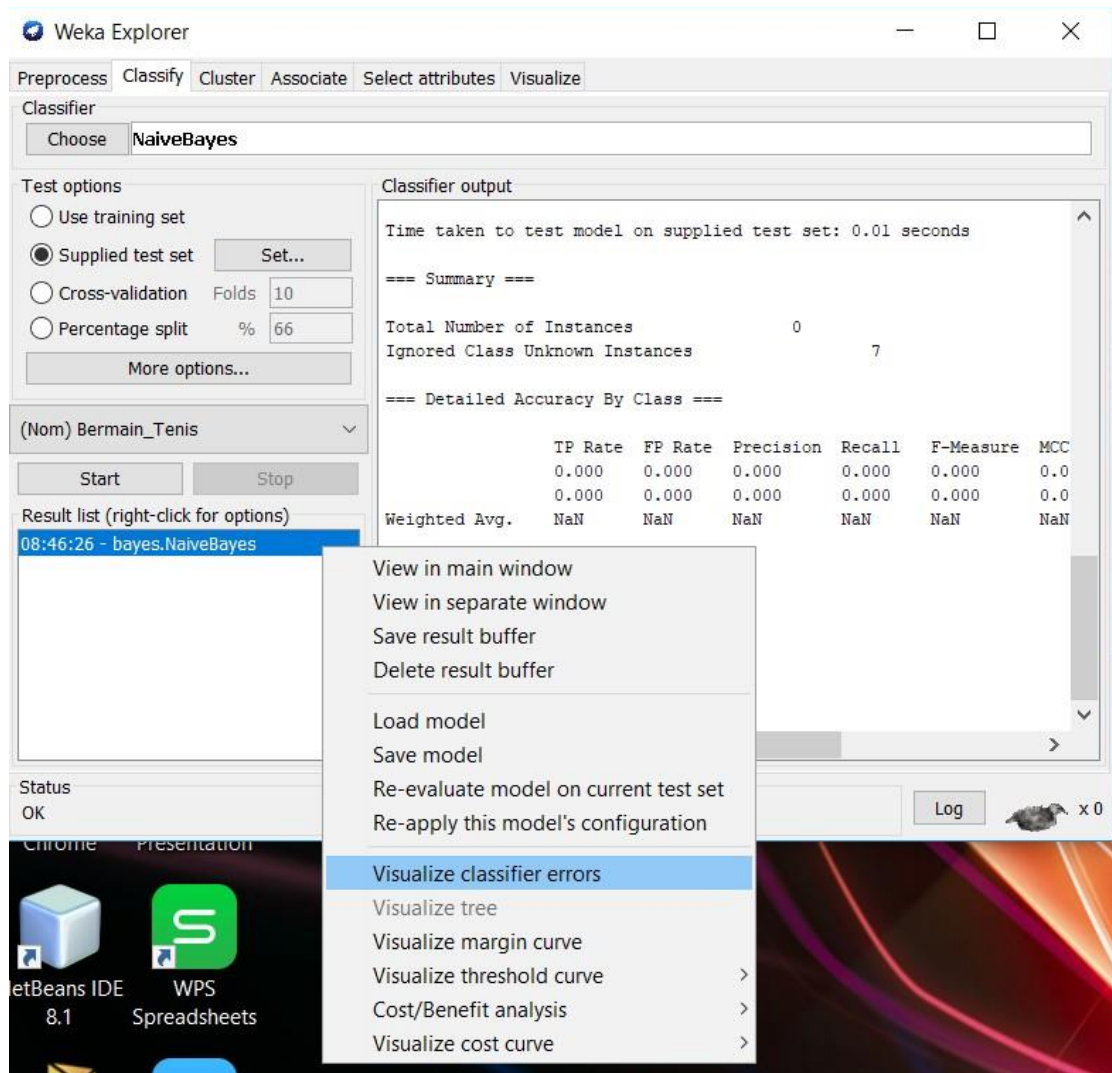
12. File CuacaTesting.arff akan diset sebagai data uji pada jendela test Instances dengan variabel predictor(Class) adalah Bermain_Tenis. Klik Close



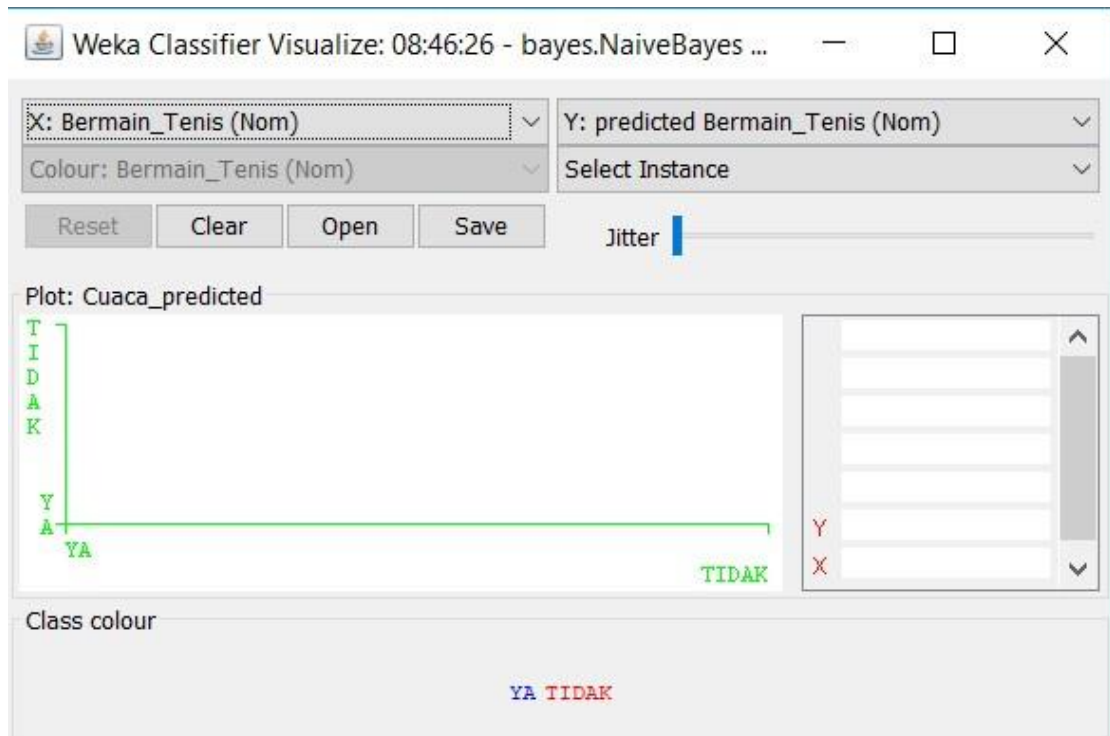
13. Klik Start untuk memulai proses naive bayes



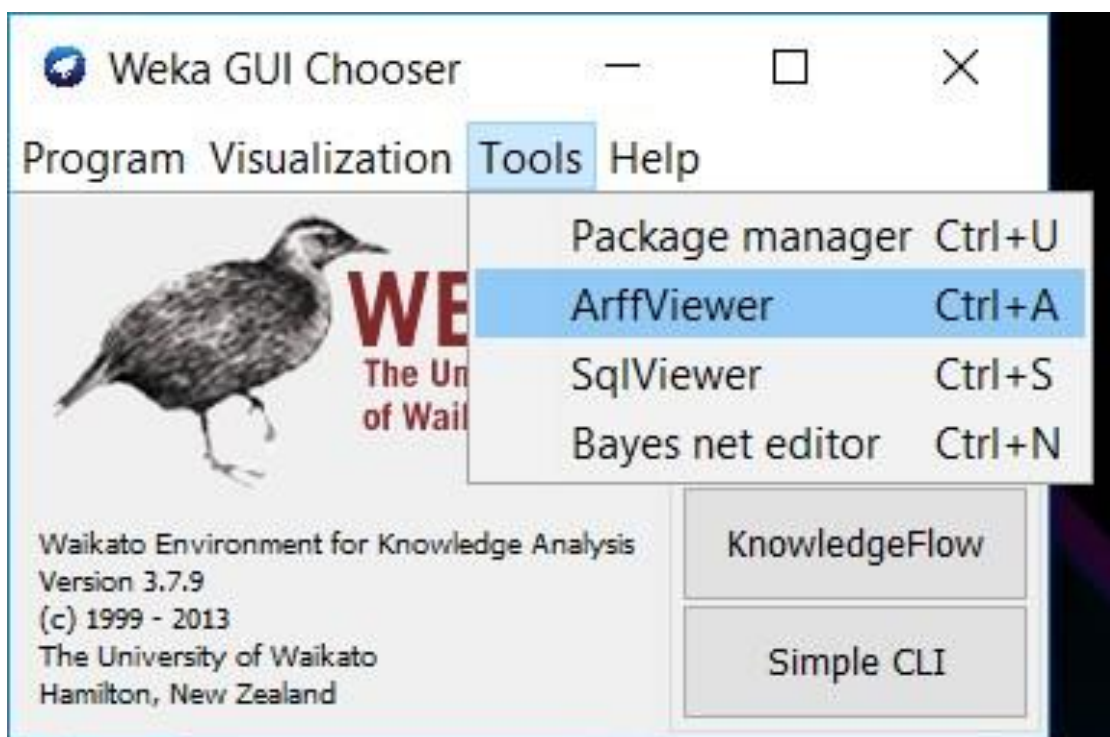
14. Jika muncul jendela pesan Classifier Panel, kita abaikan dengan mengklik Yes. Sehingga algoritma naive bayes akan diproses
15. Karena pada percobaan ini kita memproses data uji yang belum diketahui nilai kelas dari variabel dependen yang diajukan, maka kita abaikan nilai-nilai yang ditampilkan dalam jendela Classifier Output.
16. Untuk melihat hasil prediksi terhadap data uji, yang perlu kita lakukan berikutnya adalah dengan melihat nilai Classifie errors . klik kanan pada hasil proses dalam kotak result list. Pilih menu Visualize classifier errors.



17. Pada jendela Weka Classifie Visualize, abaikan hasil apapun yang disampaikan. Klik Save.
Simpan dengan nama file HasilPrediksi.arff



18. Tutup semua jendela termasuk Weka Explorer dan kembali ke Weka GUI Chooser. Pilih menu Tools-ArffViewer



19. Jendela ARFF-Viewer akan ditampilkan. Buka menu File-Open. Tunjukkan pada file HasilPrediksi.arff yang telah anda simpan pada langkah ke-18. Lihatlah, hasil prediksi telah diketahui pada kolom predicted Bermain_tenis Nominal

ARFF-Viewer - C:\Users\dessyna\Documents\Kuliah\Semester 5\Data Warehouse dan...

File Edit View

HasilPrediksi.arff

Relation: Cuaca_predicted

No.	1: Cuaca Nominal	2: Suhu Numeric	3: Kelembaban_Udara Numeric	4: Berangin Nominal	5: prediction margin Numeric	6: predicted Bermain_Tenis Nominal	7: Bermain_Tenis Nominal
1	Cerah	75.0	65.0	TIDAK	0.762765	YA	
2	Cerah	80.0	68.0	YA	0.087878	YA	
3	Cerah	83.0	87.0	YA	-0.676866	TIDAK	
4	Mendung	70.0	96.0	TIDAK	0.628523	YA	
5	Mendung	68.0	81.0	TIDAK	0.833996	YA	
6	Hujan	65.0	75.0	YA	0.253733	YA	
7	Hujan	64.0	85.0	YA	-0.160143	TIDAK	

8.4.2 Implementasi Naive Bayes dengan RapidMiner

1. Persiapkan file Tabel_Cuaca.xls yang terdiri dari 2 sheet
2. Sheet 1 digunakan sebagai data training, dan sheet 2 digunakan sebagai data uji
3. Masing – masing tabel memiliki attribute yang sama, yaitu :
 - a) Cuaca (X1)
 - b) Suhu (X2)
 - c) Kelembaban_udara(X3)
 - d) Bermain_tening(Y), sebagai variabel predictor.

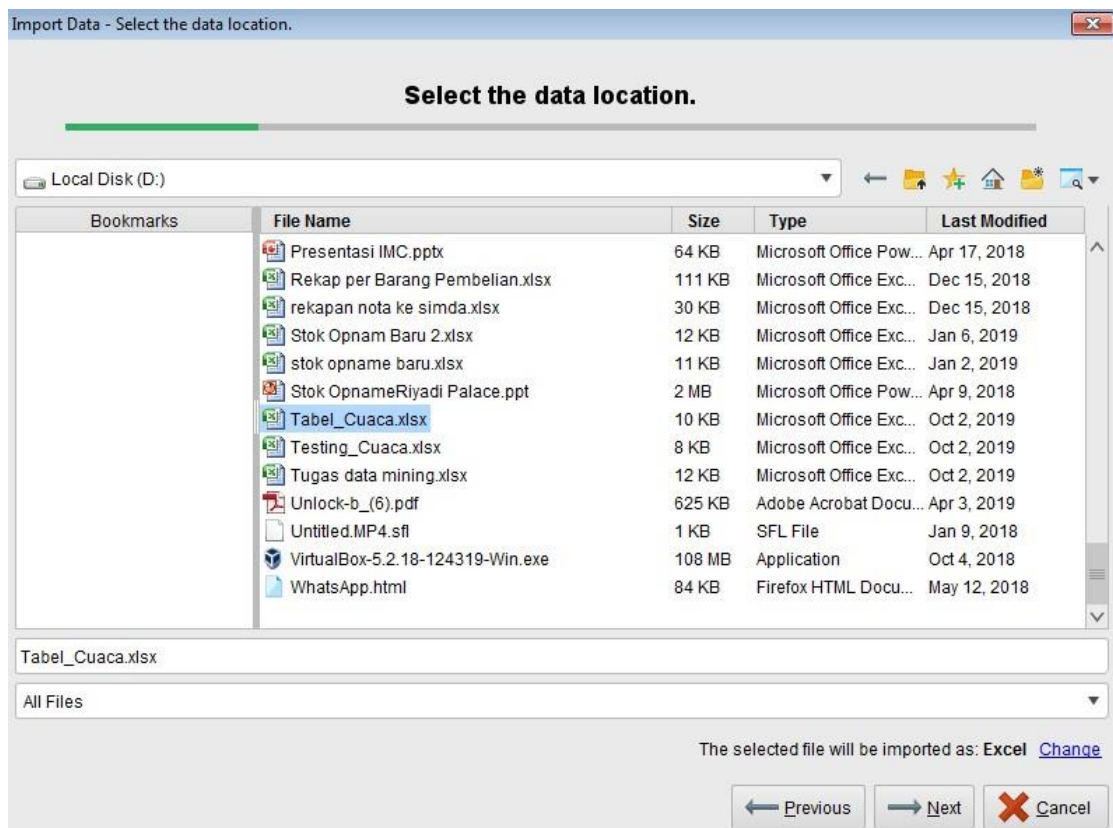
Tabel data tarining pada sheet1

F15					
	A	B	C	D	E
1	Cuaca	Suhu	Kelembapan_udara	Berangin	Bermain_Tenis
2	Cerah	85	85	TIDAK	TIDAK
3	Cerah	80	90	YA	TIDAK
4	Mendung	83	86	TIDAK	YA
5	Hujan	70	96	TIDAK	YA
6	Hujan	68	80	TIDAK	YA
7	Hujan	65	70	YA	TIDAK
8	Mendung	64	65	YA	YA
9	Cerah	72	95	TIDAK	TIDAK
10	Cerah	69	70	TIDAK	YA
11	Hujan	75	80	TIDAK	YA
12	Cerah	75	70	YA	YA
13	Mendung	72	90	YA	YA
14	Mendung	81	75	TIDAK	YA
15	Hujan	71	91	YA	TIDAK

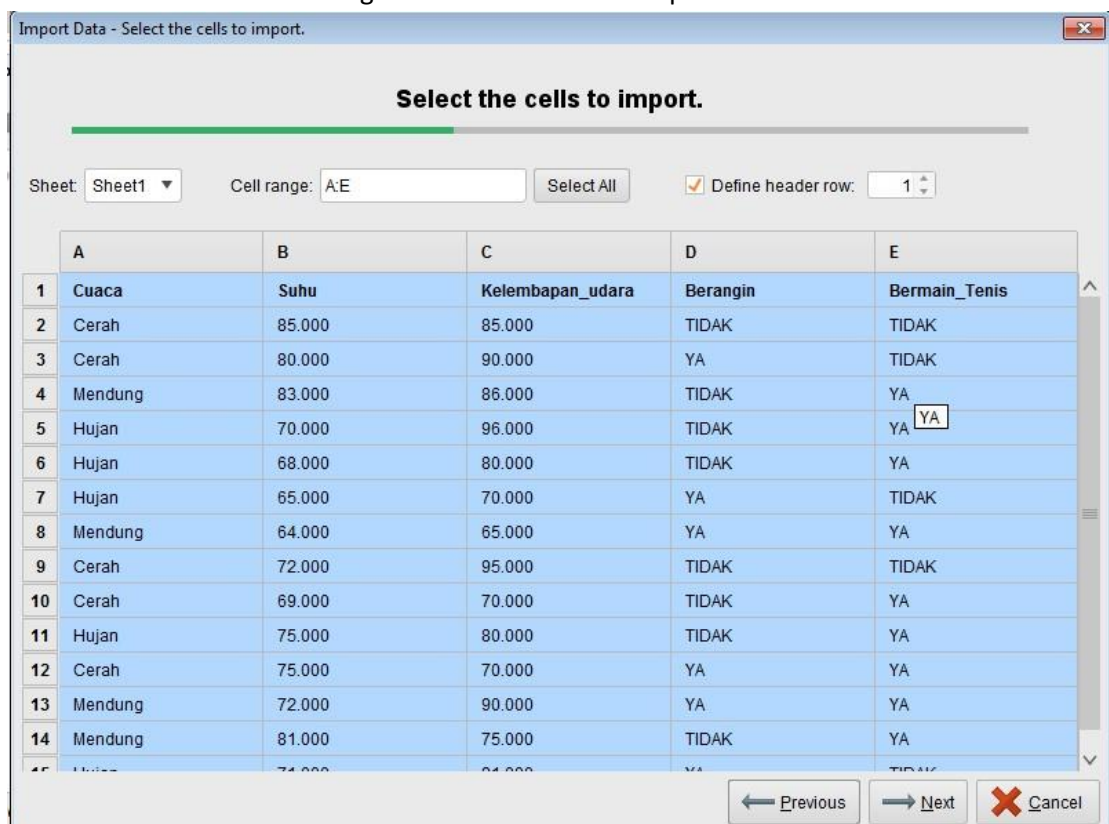
Tabel data uji pada sheet2 tanpa ada variabel Bermain_Tenis

G15				
	A	B	C	D
1	Cuaca	Suhu	Kelembapan_udara	Berangin
2	Cerah	75	65	TIDAK
3	Cerah	80	68	YA
4	Cerah	83	87	YA
5	Mendung	70	96	TIDAK
6	Mendung	68	81	TIDAK
7	Hujan	65	75	YA
8	Hujan	64	85	YA
9				

4. Buka aplikasi RapidMiner. Menjalankan RapidMiner untuk pertama kali, terlebih dahulu membuat repositori baru. Repositori ini berfungsi sebagai lokasi penyimpanan terpusat untuk dapat dan proses analisa
5. Klik Import Data. Arahkan direktori tempat penyimpanan file pada langkah Select the data location, kemudian pilih file yang digunakan dan klik Next.



6. Pastikan sel excel sesuai di langkah Select the cells to import



7. Pada langkah Format your columns ubah kolom Bermain_Tenis dengan tipe data binomial karena hanya ada dua keputusan (YA dan TIDAK)

Import Data - Format your columns.

Format your columns.

☐ Replace errors with missing values ⓘ

	Cuaca <i>polynomial</i>	Suhu <i>integer</i>	Kelembapan_u... <i>integer</i>	Berangin <i>polynomial</i>	Bermain_Tenis <i>binomial</i>
1	Cerah	85	85	TIDAK	TIDAK
2	Cerah	80	90	YA	TIDAK
3	Mendung	83	86	TIDAK	YA
4	Hujan	70	96	TIDAK	YA
5	Hujan	68	80	TIDAK	YA
6	Hujan	65	70	YA	TIDAK
7	Mendung	64	65	YA	YA
8	Cerah	72	95	TIDAK	TIDAK
9	Cerah	69	70	TIDAK	YA
10	Hujan	75	80	TIDAK	YA
11	Cerah	75	70	YA	YA
12	Mendung	72	90	YA	YA
13	Mendung	81	75	TIDAK	YA

no problems.

Previous Next Cancel

8. Ubah pula sebagai label pada Change Role

Import Data - Format your columns.

Format your columns.

☐ Replace errors with missing values ⓘ

	Cuaca <i>polynomial</i>	Suhu <i>integer</i>	Kelembapan_u... <i>integer</i>	Berangin <i>polynomial</i>	Bermain_Tenis <i>binomial</i>
1	Cerah				
2	Cerah				
3	Mendung				
4	Hujan				
5	Hujan				
6	Hujan				
7	Mendung				
8	Cerah				
9	Cerah	69	70	TIDAK	YA
10	Hujan	75	80	TIDAK	YA
11	Cerah	75	70	YA	YA
12	Mendung	72	90	YA	YA
13	Mendung	81	75	TIDAK	YA

no problems.

Previous Next Cancel

Change role

Please enter the new role:

Enter value...

label
id
weight

Import Data - Format your columns.

Format your columns.

☐ Replace errors with missing values ⓘ

	Cuaca <i>polynomial</i>	Suhu <i>integer</i>	Kelembapan_u... <i>integer</i>	Berangin <i>polynomial</i>	Bermain_Tenis <i>binominal label</i>
1	Cerah	85	85	TIDAK	TIDAK
2	Cerah	80	90	YA	TIDAK
3	Mendung	83	86	TIDAK	YA
4	Hujan	70	96	TIDAK	TIDAK
5	Hujan	68	80	TIDAK	YA
6	Hujan	65	70	YA	TIDAK
7	Mendung	64	65	YA	YA
8	Cerah	72	95	TIDAK	TIDAK
9	Cerah	69	70	TIDAK	YA
10	Hujan	75	80	TIDAK	YA
11	Cerah	75	70	YA	YA
12	Mendung	72	90	YA	YA
13	Mendung	81	75	TIDAK	YA

no problems.

Previous Next Cancel

9. Simpan dengan nama DataCuaca_Training dilanjutkan klik tombol Finish

Import Data - Where to store the data?

Where to store the data?

- Local Repository (asus)
 - Connections (asus)
 - data (asus)
 - processes (asus)

Name DataCuaca_Training

Location //Local Repository/DataCuaca_Training

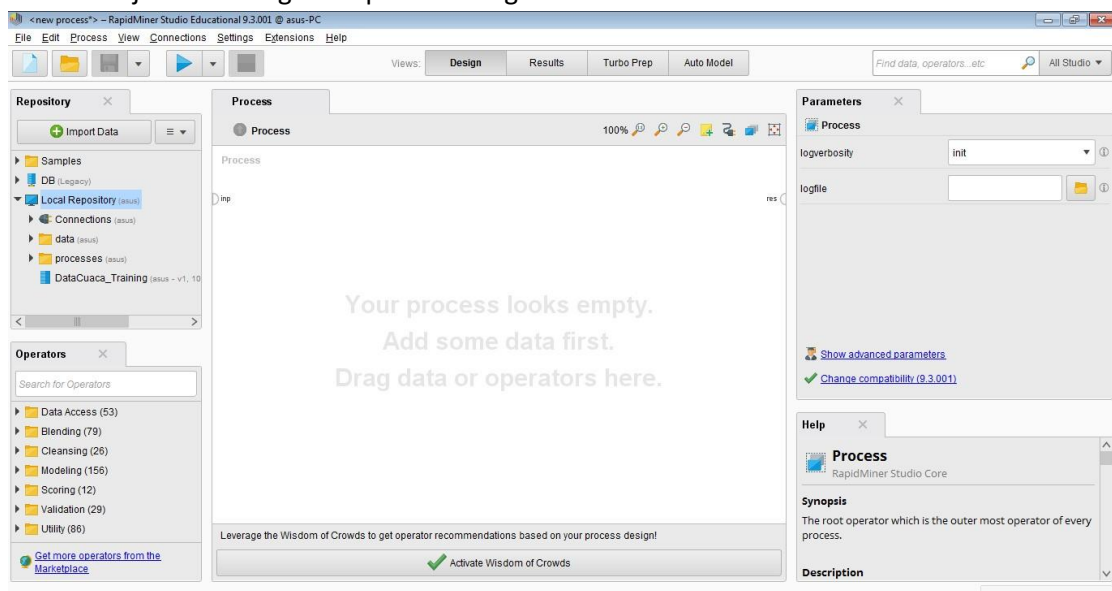
Previous Finish Cancel

10. Hasil import file Tabel_Cuaca.xls pada Sheet1 akan ditampilkan

ExampleSet (14 examples, 1 special attribute, 4 regular attributes)

Row No.	Bermain_Te...	Cuaca	Suhu	Kelembapan...	Berangin
1	TIDAK	Cerah	85	85	TIDAK
2	TIDAK	Cerah	80	90	YA
3	YA	Mendung	83	86	TIDAK
4	YA	Hujan	70	96	TIDAK
5	YA	Hujan	68	80	TIDAK
6	TIDAK	Hujan	65	70	YA
7	YA	Mendung	64	65	YA
8	TIDAK	Cerah	72	95	TIDAK
9	YA	Cerah	69	70	TIDAK
10	YA	Hujan	75	80	TIDAK
11	YA	Cerah	75	70	YA
12	YA	Mendung	72	90	YA
13	YA	Mendung	81	75	TIDAK
14	TIDAK	Hujan	71	91	YA

11. Kembali ke jendela Design Perspective dengan shortcut tombol F8



12. Lakukan hal yang sama untuk data testing yang diambil dari Tabel_Cuaca.xls pada Sheet2 (Testing) dengan mengulang langkah 5. Pastikan semua variabel data testing terpilih (ada 4 variabel), bedanya pada langkah ini tidak ada variabel yang diubah bertipe label seperti pada langkah 8.

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Sheet2 Cell range: A:D Select All ☒ Define header row: 1

	A	B	C	D
1	Cuaca	Suhu	Kelembapan_udara	Berangin
2	Cerah	75.000	65.000	TIDAK
3	Cerah	80.000	68.000	YA
4	Cerah	83.000	87.000	YA
5	Mendung	70.000	96.000	TIDAK
6	Mendung	68.000	81.000	TIDAK
7	Hujan	65.000	75.000	YA
8	Hujan	64.000	85.000	YA

Previous Next Cancel

<new process> - RapidMiner Studio Educational 9.3.001 © asus-PC

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Find data, operators, etc. All Studio

Result History ExampleSet (/Local Repository/DataCuaca_Testing) ExampleSet (/Local Repository/DataCuaca_Training)

Open in Turbo Prep Auto Model Filter (7 / 7 examples): all

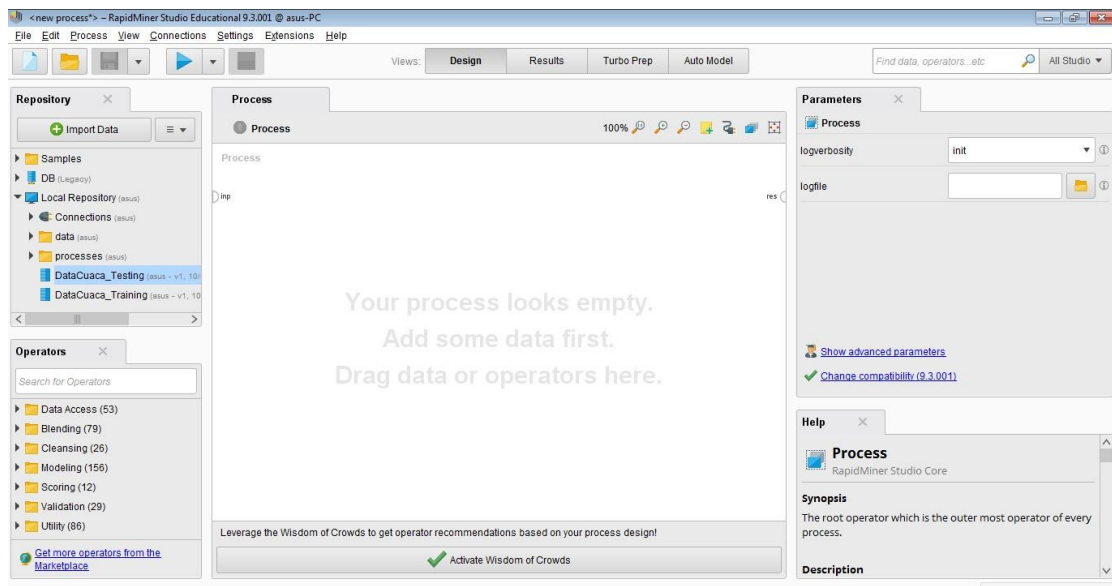
Row No.	Cuaca	Suhu	Kelembapan...	Berangin
1	Cerah	75	65	TIDAK
2	Cerah	80	68	YA
3	Cerah	83	87	YA
4	Mendung	70	96	TIDAK
5	Mendung	68	81	TIDAK
6	Hujan	65	75	YA
7	Hujan	64	85	YA

ExampleSet (7 examples, 0 special attributes, 4 regular attributes)

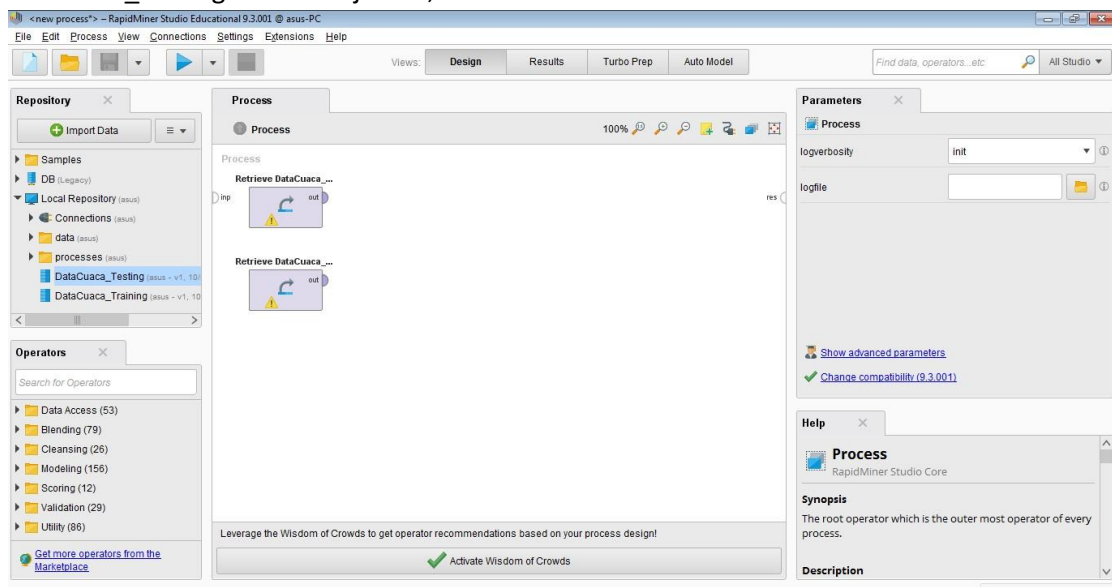
Repository

- Import Data
- Samples
- DB (Legacy)
- Local Repository (asus)
 - Connections (asus)
 - data (asus)
 - processes (asus)
 - DataCuaca_Testing (asus - v1, 10/5/19 9)
 - DataCuaca_Training (asus - v1, 10/5/19 9)

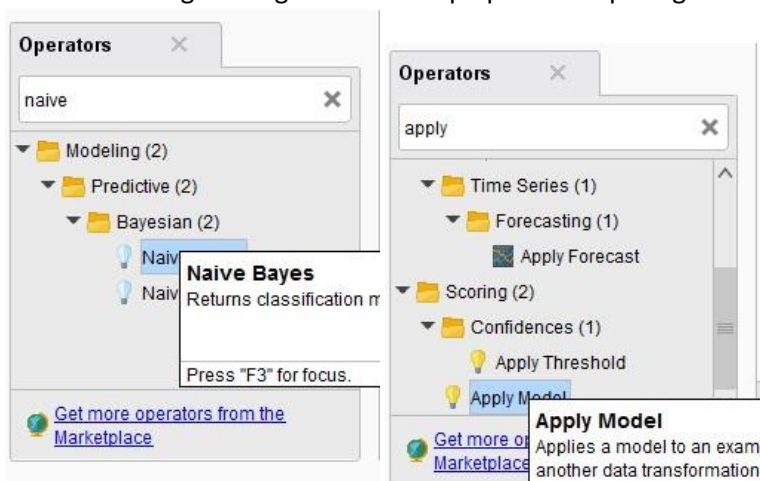
13. Simpan dengan nama DataCuaca_Testing

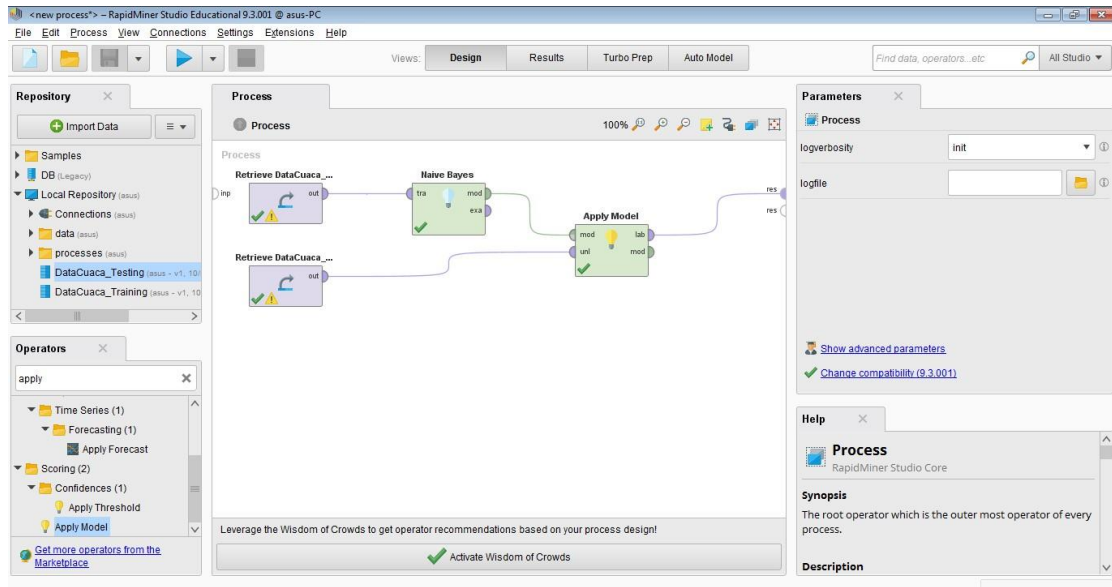


14. Langkah selanjutnya adalah membuat desain Naive Bayes. Drag DataCuaca_Training dan DataCuaca_Testing ke dalam jendela; a Process View



15. Masukkan juga operator Naive Bayes dan Apply Model ke dalam Process View. Hubungkan konektor masing-masing data terhadap operator seperti gambar





16. Jalankan proses naive bayes dengan menekan tombol Run (atau menekan tombol F11)

17. Perhatikan hasilproses klasifikasi naive bayes. Pada tab Data, dapat dilihat prediksi terhadap data testing serta tingkat confidence nilai kelas pada masing – masing data

The screenshot shows the 'Results' tab in RapidMiner Studio. The table displays the output of the 'Apply Model' operator, showing predictions and confidence scores for each example. The table has columns for Row No., prediction(B..., confidence(..., confidence(..., Cuaca, Suhu, Kelembapan..., and Berangin.

Row No.	prediction(B...	confidence(...	confidence(...	Cuaca	Suhu	Kelembapan...	Berangin
1	YA	0.154	0.846	Cerah	75	65	TIDAK
2	YA	0.498	0.502	Cerah	80	68	YA
3	TIDAK	0.856	0.144	Cerah	83	87	YA
4	YA	0.019	0.981	Mendung	70	96	TIDAK
5	YA	0.007	0.993	Mendung	68	81	TIDAK
6	YA	0.371	0.629	Hujan	65	75	YA
7	TIDAK	0.568	0.432	Hujan	64	85	YA

Pada tab Statistics, dapat dilihat bahwa distribusi nilai kelas pada variabel Y (Bermain_Tenis) rerata nilai confidence sebesar 0,353 untuk nilai TIDAK , dan 0,647 untuk nilai YA

ExampleSet (/Local Repository/DataCuaca_Testing) ExampleSet (/Local Repository/DataCuaca_Training)

Views: Design Results Turbo Prep Auto Model

Find data, operators, etc: All Studio

Result History

ExampleSet (Apply Model)

Name	Type	Missing	Statistics	Filter (7 / 7 attributes):	Search for Attributes
✓ Prediction(Bermain_Tenis)	Binominal	0	Least TIDAK (2) Most YA (5)	Values	YA (5), TIDAK (2)
✓ confidence(TIDAK)	Real	0	Min 0.007 Max 0.856	Average	0.353
✓ confidence_YA	Real	0	Min 0.144 Max 0.993	Average	0.647
✓ Cuaca	Polynomial	0	Least Mendung (2) Most Cerah (3)	Values	Cerah (3), Hujan (3), Mendung (2)
✓ Suhu	Integer	0	Min 64 Max 83	Average	72.143
✓ Kelembapan_udara	Integer	0	Min 65 Max 96	Average	79.571
✓ Berangin	Polynomial	0	Least TIDAK (3) Most YA (4)	Values	YA (4), TIDAK (3)

Showing attributes 1 - 7 Examples: 7 Special Attributes: 3 Regular Attributes: 4

Repository

- Import Data
- Samples
- DB (Legacy)
- Local Repository (asus)
 - data (asus)
 - processes (asus)
 - DataCuaca_Testing (asus - v1, 10/5/19 9)
 - DataCuaca_Training (asus - v1, 10/5/19 9)

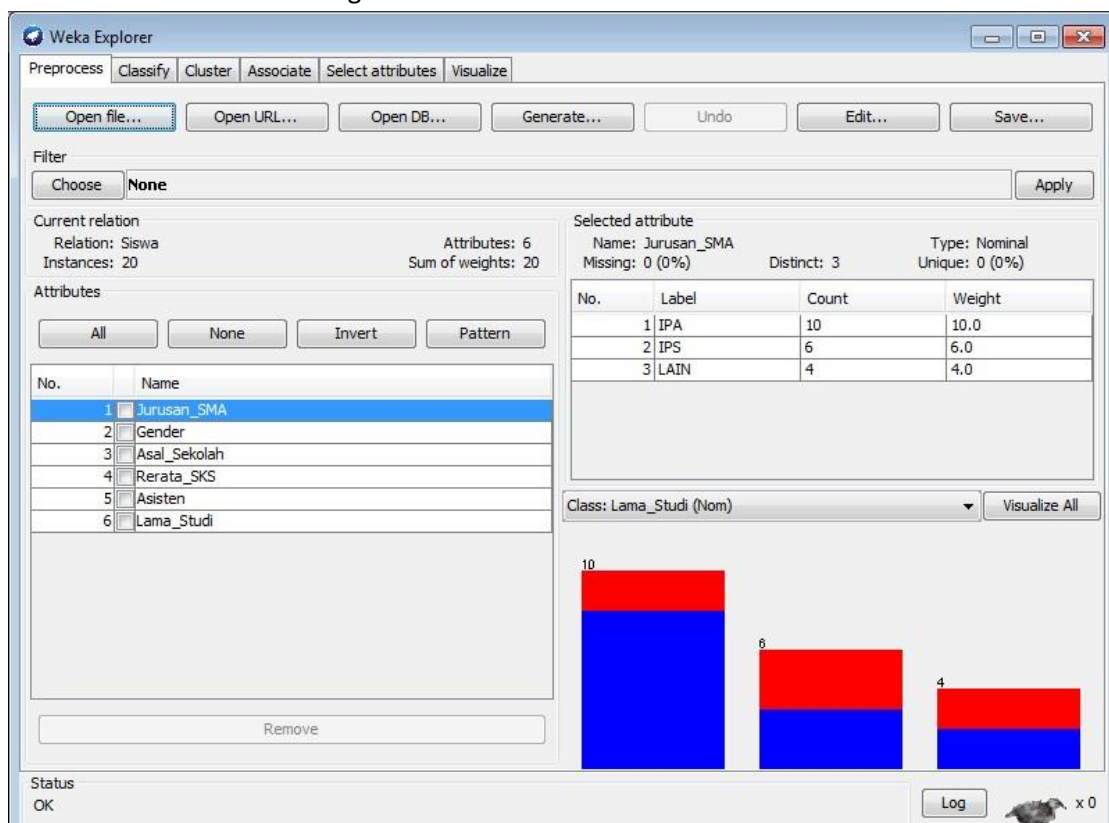
18. Bandingkan dengan hasil prediksi menggunakan WEKA. Dapat dilihat bahwa prediksi masing - masing aplikasi menunjukkan hasil yang sama

Tugas

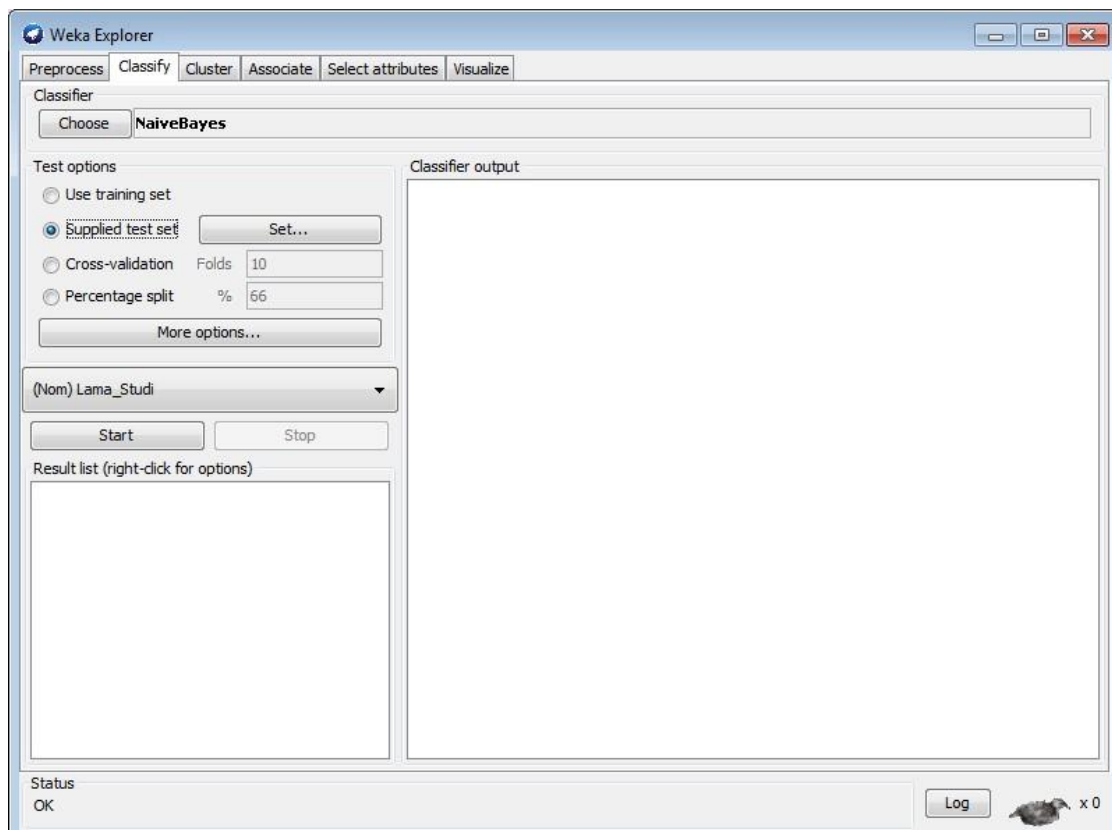
1. Berdasarkan tabel berikut, buatlah file dalam format Excel(.xls) dan format ARFF(.arff) ! Data ini akan digunakan sebagai data testing

	A	B	C	D	E	F
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	
2	LAIN	WANITA	SURAKARTA	18	TIDAK	
3	IPA	PRIA	SURAKARTA	19	YA	
4	LAIN	PRIA	SURAKARTA	19	TIDAK	
5	IPA	PRIA	LUAR	17	TIDAK	
6	IPA	WANITA	SURAKARTA	17	TIDAK	
7	IPA	WANITA	LUAR	18	YA	
8	IPA	PRIA	SURAKARTA	18	TIDAK	
9	IPA	PRIA	SURAKARTA	19	TIDAK	
10	IPS	PRIA	LUAR	18	TIDAK	
11	LAIN	WANITA	SURAKARTA	18	TIDAK	
12						

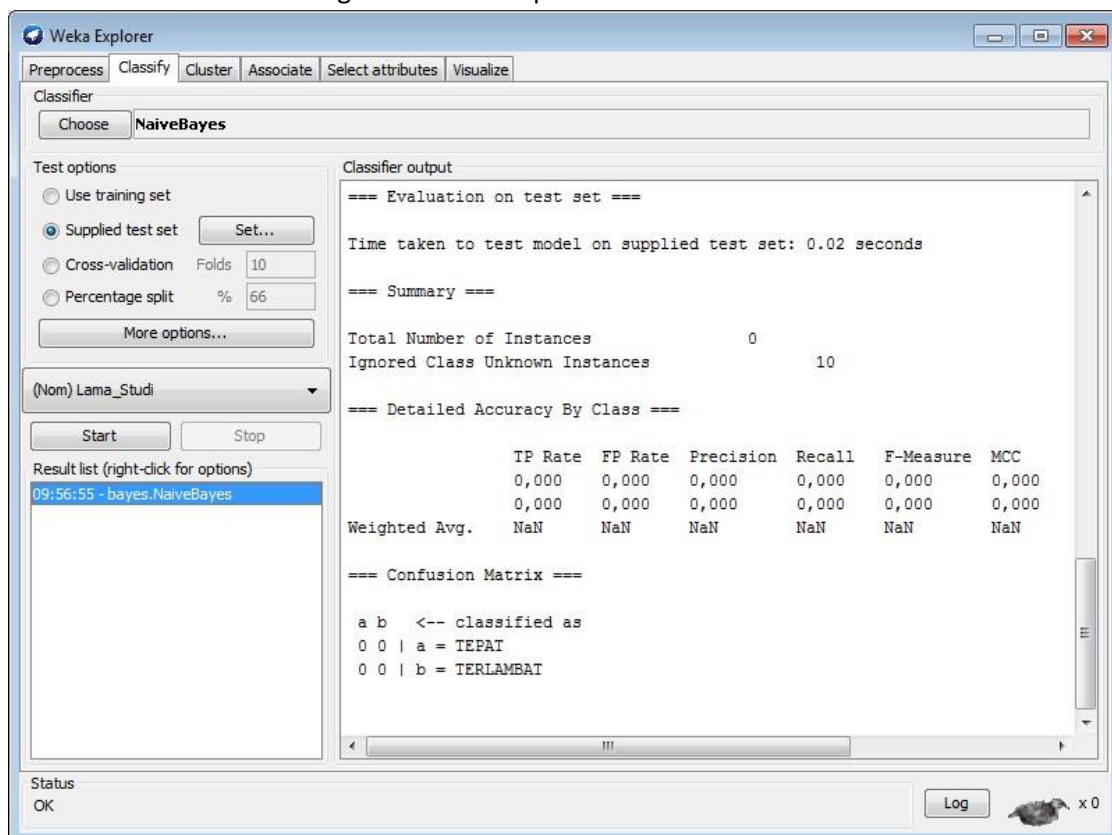
2. Gunakan file ARFF yang dikerjakan pada Tugas nomor 1 dalam Modul 7 sebagai data training. Lakukan prediksi terhadap data testing (ARFF) di atas menggunakan WEKA!
 - a. Membuka file Training



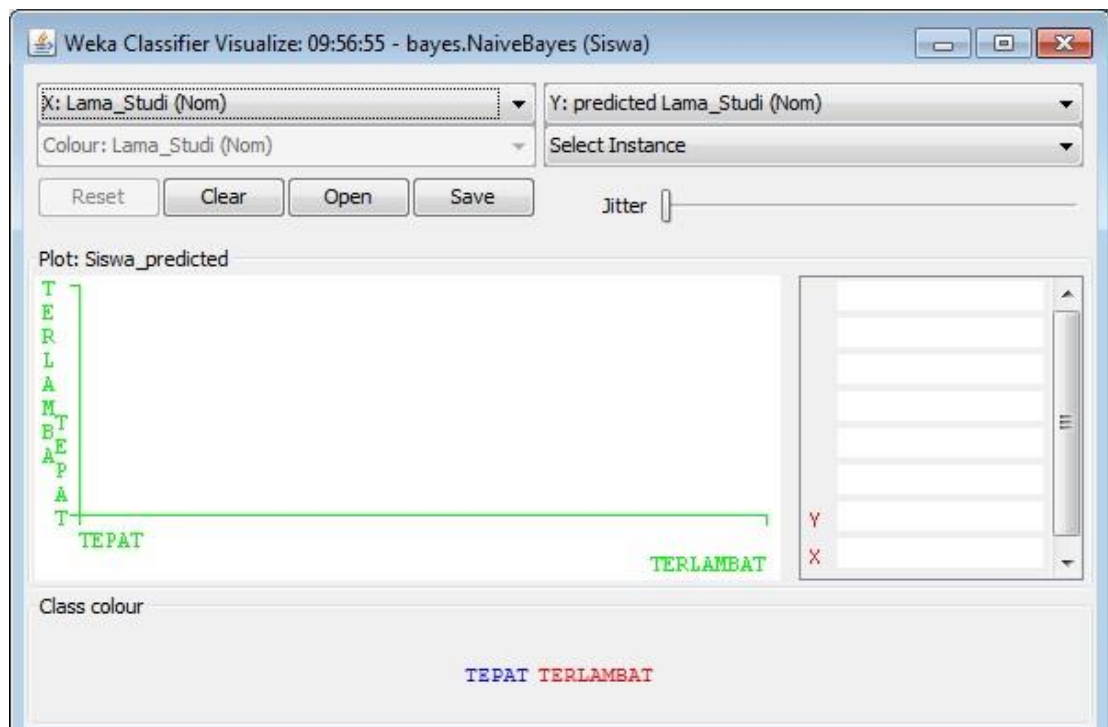
- b. Memilih metode Naive Bayes



c. Membuka file Testing lalu klik kanan pilih Visualize classifier errors



d. Menyimpan file hasil



e.

Hasil Prediksi

ARFF-Viewer - D:\dwdm\modul 8\tugas\HasilTugas.arff

File Edit View

HasilTugas.arff

Relation: Siswa_predicted

No.	1: Jurusan_SMA Nominal	2: Gender Nominal	3: Asal_Sekolah Nominal	4: Rerata_SKS Numeric	5: Asisten Nominal	6: prediction margin Numeric	7: predicted Lama_Studi Nominal	8: Lama_Studi Nominal
1	LAIN	WANITA	SURAKARTA	18.0	TIDAK	-0.375862	TERLAMBAT	
2	IPA	PRIA	SURAKARTA	19.0	YA	0.836469	TEPAT	
3	LAIN	PRIA	SURAKARTA	19.0	TIDAK	-0.175169	TERLAMBAT	
4	IPS	PRIA	LUAR	17.0	TIDAK	-0.713206	TERLAMBAT	
5	LAIN	WANITA	SURAKARTA	17.0	TIDAK	-0.546846	TERLAMBAT	
6	IPA	WANITA	LUAR	18.0	YA	0.757815	TEPAT	
7	IPA	PRIA	SURAKARTA	18.0	TIDAK	-0.125076	TERLAMBAT	
8	IPA	PRIA	SURAKARTA	19.0	TIDAK	0.356012	TEPAT	
9	IPS	PRIA	LUAR	18.0	TIDAK	-0.588286	TERLAMBAT	
10	LAIN	WANITA	SURAKARTA	18.0	TIDAK	-0.375862	TERLAMBAT	

- Gunakan file Excel yang dikerjakan pada Tugas nomor 1 dalam Modul 6 sebagai data training. Laukan prediksi terhadap data testing (Excel) di atas menggunakan RapidMiner! a. Data Training

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Sheet1
Cell range: A:F
Select All
☒ Define header row: 1

	A	B	C	D	E	F
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Lama_Studi
2	IPS	WANITA	SURAKARTA	18.000	TIDAK	TERLAMBAT
3	IPA	PRIA	SURAKARTA	19.000	YA	TEPAT
4	LAIN	PRIA	SURAKARTA	19.000	TIDAK	TERLAMBAT
5	IPA	PRIA	LUAR	17.000	TIDAK	TERLAMBAT
6	IPA	WANITA	SURAKARTA	17.000	TIDAK	TEPAT
7	IPA	WANITA	LUAR	18.000	YA	TEPAT
8	IPA	PRIA	SURAKARTA	18.000	TIDAK	TERLAMBAT
9	IPA	PRIA	SURAKARTA	19.000	TIDAK	TEPAT
10	IPS	PRIA	LUAR	18.000	TIDAK	TERLAMBAT
11	LAIN	WANITA	SURAKARTA	18.000	TIDAK	TEPAT
12	IPA	WANITA	SURAKARTA	19.000	TIDAK	TEPAT
13	IPS	PRIA	SURAKARTA	20.000	TIDAK	TEPAT
14	IPS	PRIA	SURAKARTA	19.000	TIDAK	TEPAT

Previous
Next
Cancel

Import Data - Format your columns.

Format your columns.

☐ Replace errors with missing values ⓘ

	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Lama_Studi
	<i>polynominal</i>	<i>polynominal</i>	<i>polynominal</i>	<i>integer</i>	<i>binominal</i>	<i>binominal label</i>
1	IPS	WANITA	SURAKARTA	18	TIDAK	TERLAMBAT
2	IPA	PRIA	SURAKARTA	19	YA	TEPAT
3	LAIN	PRIA	SURAKARTA	19	TIDAK	TERLAMBAT
4	IPA	PRIA	LUAR	17	TIDAK	TERLAMBAT
5	IPA	WANITA	SURAKARTA	17	TIDAK	TEPAT
6	IPA	WANITA	LUAR	18	YA	TEPAT
7	IPA	PRIA	SURAKARTA	18	TIDAK	TERLAMBAT
8	IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT
9	IPS	PRIA	LUAR	18	TIDAK	TERLAMBAT
10	LAIN	WANITA	SURAKARTA	18	TIDAK	TEPAT
11	IPA	WANITA	SURAKARTA	19	TIDAK	TEPAT
12	IPS	PRIA	SURAKARTA	20	TIDAK	TEPAT
13	IPS	PRIA	SURAKARTA	19	TIDAK	TEPAT

☒ no problems.
Previous
Next
Cancel

<new process> - RapidMiner Studio Educational 9.3.001 @ asus-PC

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model

Find data, operators, etc. All Studio

ExampleSet (/Local Repository/DataCuaca_Testing) ExampleSet (/Local Repository/DataCuaca_Training)

Result History ExampleSet (/Local Repository/DataSekolah_Training) ExampleSet (Apply Model)

Open in Turbo Prep Auto Model

Filter (20 / 20 examples): all

Row No.	Lama_Studi	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TERLAMBAT	IPS	WANITA	SURAKARTA	18	TIDAK
2	TEPAT	IPA	PRIA	SURAKARTA	19	YA
3	TERLAMBAT	LAIN	PRIA	SURAKARTA	19	TIDAK
4	TERLAMBAT	IPA	PRIA	LUAR	17	TIDAK
5	TEPAT	IPA	WANITA	SURAKARTA	17	TIDAK
6	TEPAT	IPA	WANITA	LUAR	18	YA
7	TERLAMBAT	IPA	PRIA	SURAKARTA	18	TIDAK
8	TEPAT	IPA	PRIA	SURAKARTA	19	TIDAK
9	TERLAMBAT	IPS	PRIA	LUAR	18	TIDAK
10	TEPAT	LAIN	WANITA	SURAKARTA	18	TIDAK
11	TEPAT	IPA	WANITA	SURAKARTA	19	TIDAK
12	TEPAT	IPS	PRIA	SURAKARTA	20	TIDAK
13	TEPAT	IPS	PRIA	SURAKARTA	19	TIDAK
14	TEPAT	IPA	PRIA	SURAKARTA	19	TIDAK

ExampleSet (20 examples, 1 special attribute, 5 regular attributes)

Repository

- Import Data
- Samples
- DB (Legacy)
- Local Repository (asus)
 - Connections (asus)
 - data (asus)
 - processes (asus)
 - DataCuaca_Testing (asus - v1, 10/5/19 9:00 AM)
 - DataCuaca_Training (asus - v1, 10/5/19 9:00 AM)
 - DataSekolah_Training (asus - v1, 10/5/19 9:00 AM)
 - HasilCuaca (asus - v1, 10/5/19 10:00 AM)

b. Data Testing

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Sheet2 Cell range: A:E Select All ☒ Define header row: 1

	A	B	C	D	E
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
2	LAIN	WANITA	SURAKARTA	18.000	TIDAK
3	IPA	PRIA	SURAKARTA	19.000	YA
4	LAIN	PRIA	SURAKARTA	19.000	TIDAK
5	IPS	PRIA	LUAR	17.000	TIDAK
6	LAIN	WANITA	SURAKARTA	17.000	TIDAK
7	IPA	WANITA	LUAR	18.000	YA
8	IPA	PRIA	SURAKARTA	18.000	TIDAK
9	IPA	PRIA	SURAKARTA	19.000	TIDAK
10	IPS	PRIA	LUAR	18.000	TIDAK
11	LAIN	WANITA	SURAKARTA	18.000	TIDAK

Previous Next Cancel

Import Data - Format your columns.

Format your columns.

☐ Replace errors with missing values ⓘ

	Jurusan_SMA <i>polynomial</i>	Gender <i>polynomial</i>	Asal_Sekolah <i>polynomial</i>	Rerata_SKS <i>integer</i>	Asisten <i>polynomial</i>
1	LAIN	WANITA	SURAKARTA	18	TIDAK
2	IPA	PRIA	SURAKARTA	19	YA
3	LAIN	PRIA	SURAKARTA	19	TIDAK
4	IPS	PRIA	LUAR	17	TIDAK
5	LAIN	WANITA	SURAKARTA	17	TIDAK
6	IPA	WANITA	LUAR	18	YA
7	IPA	PRIA	SURAKARTA	18	TIDAK
8	IPA	PRIA	SURAKARTA	19	TIDAK
9	IPS	PRIA	LUAR	18	TIDAK
10	LAIN	WANITA	SURAKARTA	18	TIDAK

no problems.

Previous Next Cancel

<new process> - RapidMiner Studio Educational 9.3.001 © asus-PC

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model

Find data, operators, etc. All Studio

ExampleSet (/Local Repository/DataCuaca_Testing) ExampleSet (/Local Repository/DataCuaca_Training)

ExampleSet (/Local Repository/DataSekolah_Testing) ExampleSet (/Local Repository/DataSekolah_Training)

Result History ExampleSet (/Local Repository/DataSekolah_Testing)

Open in Turbo Prep Auto Model

Filter (10 / 10 examples): all

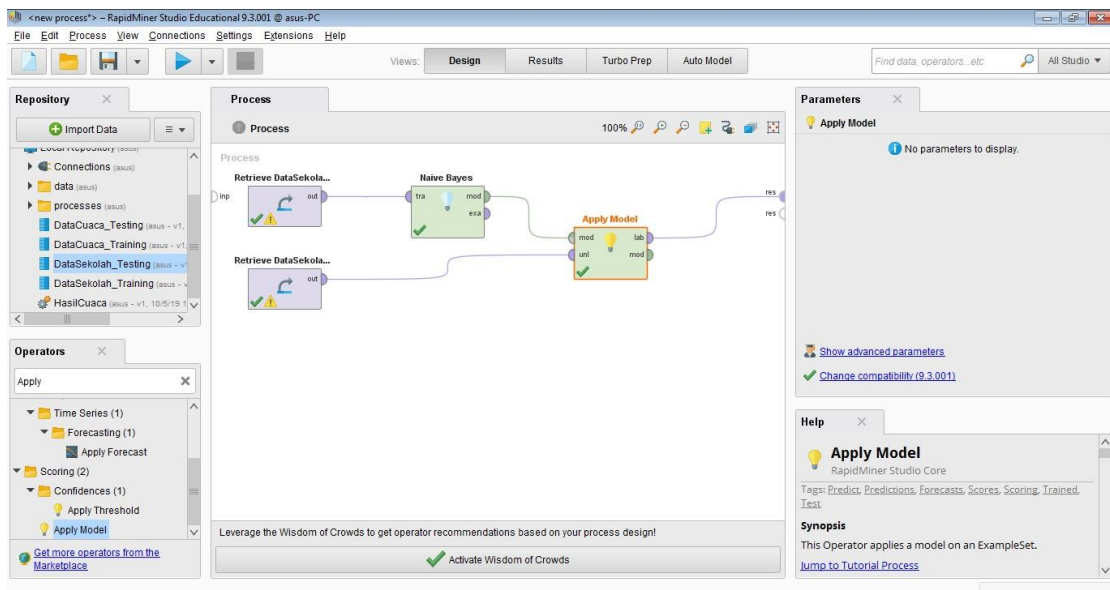
Row No.	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	LAIN	WANITA	SURAKARTA	18	TIDAK
2	IPA	PRIA	SURAKARTA	19	YA
3	LAIN	PRIA	SURAKARTA	19	TIDAK
4	IPS	PRIA	LUAR	17	TIDAK
5	LAIN	WANITA	SURAKARTA	17	TIDAK
6	IPA	WANITA	LUAR	18	YA
7	IPA	PRIA	SURAKARTA	18	TIDAK
8	IPA	PRIA	SURAKARTA	19	TIDAK
9	IPS	PRIA	LUAR	18	TIDAK
10	LAIN	WANITA	SURAKARTA	18	TIDAK

ExampleSet (10 examples, 0 special attributes, 5 regular attributes)

Repository

- Import Data
- Samples
- DB (Legacy)
- Local Repository (asus)
 - Connections (asus)
 - data (asus)
 - processes (asus)
 - DataCuaca_Testing (asus - v1, 10/5/19 9:10)
 - DataCuaca_Training (asus - v1, 10/5/19 9:10)
 - DataSekolah_Testing (asus - v1, 10/5/19 10:00)
 - DataSekolah_Training (asus - v1, 10/5/19 10:00)
 - HasilCuaca (asus - v1, 10/5/19 10:00 AM)

c. Implementasi Naive Bayes



d. Hasil

Result History

ExampleSet (Apply Model)

Filter (10 / 10 examples): all

Row No.	prediction(L...	confidence(L...	confidence(L...	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TERLAMBAT	0.648	0.352	LAIN	WANITA	SURAKARTA	18	TIDAK
2	TEPAT	0.005	0.995	IPA	PRIA	SURAKARTA	19	YA
3	TERLAMBAT	0.650	0.350	LAIN	PRIA	SURAKARTA	19	TIDAK
4	TERLAMBAT	0.868	0.132	IPS	PRIA	LUAR	17	TIDAK
5	TERLAMBAT	0.738	0.262	LAIN	WANITA	SURAKARTA	17	TIDAK
6	TEPAT	0.005	0.995	IPA	WANITA	LUAR	18	YA
7	TERLAMBAT	0.547	0.453	IPA	PRIA	SURAKARTA	18	TIDAK
8	TEPAT	0.321	0.679	IPA	PRIA	SURAKARTA	19	TIDAK
9	TERLAMBAT	0.811	0.189	IPS	PRIA	LUAR	18	TIDAK
10	TERLAMBAT	0.648	0.352	LAIN	WANITA	SURAKARTA	18	TIDAK

ExampleSet (10 examples, 3 special attributes, 5 regular attributes)

Result History

ExampleSet (Apply Model)

Filter (8 / 8 attributes): Search for Attributes

Name	Type	Missing	Statistics
prediction(Lama_Studi)	Binominal	0	Least TEPAT (3) Most TERLAMBAT (7) Values TERLAMBAT (7), TEPAT (3)
confidence(TERLAMBAT)	Real	0	Min 0.005 Max 0.868 Average 0.524
confidence(TEPAT)	Real	0	Min 0.132 Max 0.995 Average 0.476
Jurusan_SMA	Polynomial	0	Least IPS (2) Most IPA (4) Values IPA (4), LAIN (4), ...[1 more]
Gender	Polynomial	0	Least WANITA (4) Most PRIA (6) Values PRIA (6), WANITA (4)
Asal_Sekolah	Polynomial	0	Least LUAR (3) Most SURAKARTA (7) Values SURAKARTA (7), LUAR (3)

Annotations	▼ Rerata_SKS	Integer	0	Min 17	Max 19	Average 18.100
	▼ Asisten	Polynomial	0	Least YA (2)	Most TIDAK (8)	Values TIDAK (8), YA (2)

Showing attributes 1 - 8 Examples: 10 Special Attributes: 3 Regular Attributes: 5

4. Dari hasil percobaan Tugas nomor 3 di atas, berapakah nilai rerata confidence untuk atribut Lama_Studi dengan nilai TEPAT? Berapakah nilai rerata confidence untuk atribut Lama_Studi dengan nilai TERLAMBAT?

▼ Confidence_TERLAMBAT confidence(TERLAMBAT)	Real	0	Min 0.005	Max 0.868	Average 0.524
▼ Confidence_TEPAT confidence(TEPAT)	Real	0	Min 0.132	Max 0.995	Average 0.476

Jadi nilai rerata confidence untuk atribut Lama_Studi dengan nilai TEPAT adalah **0,524** dan nilai rerata confidence untuk atribut Lama_Studi dengan nilai TERLAMBAT adalah **0,476**

5. Dari hasil percobaan Tugas nomor 3 di atas, berapa orang yang akan lulus TEPAT, dan berapa orang yang akan lulus TERLAMBAT?

▼ Prediction prediction(Lama_Studi)	Binominal	0	Least TEPAT (3)	Most TERLAMBAT (7)	Values TERLAMBAT (7), TEPAT (3)
--	-----------	---	--------------------	-----------------------	------------------------------------

Jadi yang lulus dengan tepat sebanyak **3** orang sedangkan yang lulus terlambat sebanyak **7** orang

6. Prekdisikan ketepatan lama studi si Dewi, jika Dewi adalah seorang WANITA yang berasal dari jurusan IPA pada saat SMA, asal sekolah dari LUAR SURAKARTA, mengambil SKS dengan rata-rata sebanyak 18 SKS tiap semester, dan tidak pernah menjadi Asisten selama kuliah.

a. Mengimport data Dewi

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Sheet3 Cell range: A:E Select All ☒ Define header row: 1

	A	B	C	D	E
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
2	IPA	WANITA	LUAR SURAKARTA	18.000	TIDAK

← Previous → Next X Cancel

Import Data - Format your columns.

Format your columns.

☐ Replace errors with missing values ⓘ

	Jurusan_SMA <i>polynomial</i>	Gender <i>polynomial</i>	Asal_Sekolah <i>polynomial</i>	Rerata_SKS <i>integer</i>	Asisten <i>polynomial</i>
1	IPA	WANITA	LUAR SURAKARTA	18	TIDAK

PREVI...

no problems.

Previous Next Cancel

<new process> - RapidMiner Studio Educational 9.3.001 © asus-PC

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model

Find data, operators, etc. All Studio

ExampleSet (/Local Repository/DataCuaca_Testing) ExampleSet (/Local Repository/DataCuaca_Training)

ExampleSet (/Local Repository/DataSekolah_Testing) ExampleSet (/Local Repository/DataSekolah_Training)

ExampleSet (Apply Model) ExampleSet (/Local Repository/DataSekolah_Testing)

Result History ExampleSet (/Local Repository/DataDewi_Testing)

Open in Turbo Prep Auto Model

Filter (1 / 1 examples): all

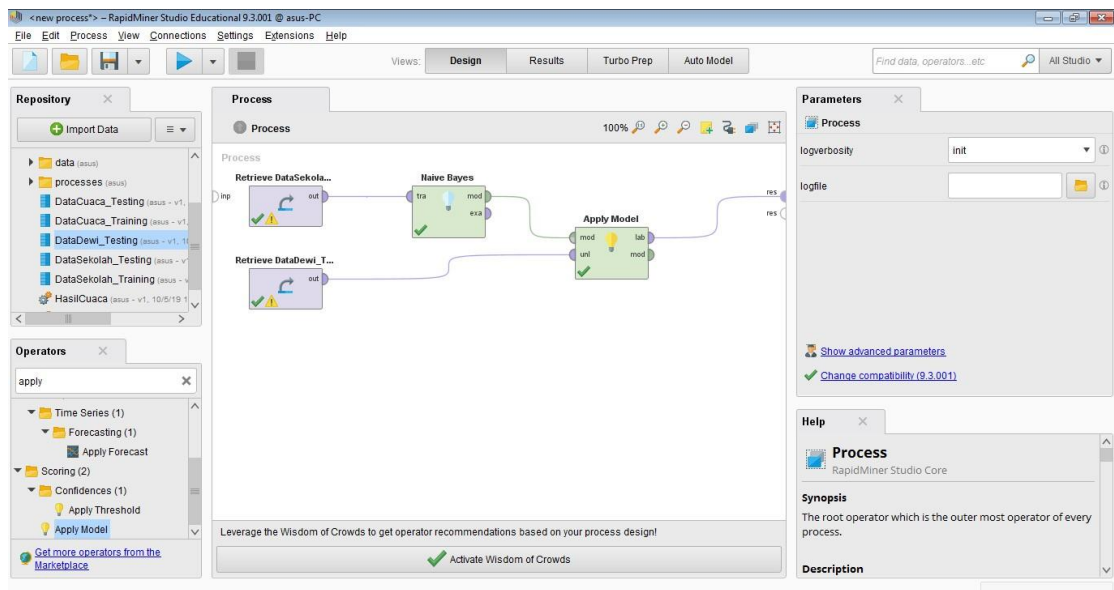
Row No.	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	IPA	WANITA	LUAR SURA...	18	TIDAK

ExampleSet (1 example, 0 special attributes, 5 regular attributes)

Repository

- Import Data
- Samples
- DB (Legacy)
- Local Repository (asus)
 - Connections (asus)
 - data (asus)
 - processes (asus)
 - DataCuaca_Testing
 - DataCuaca_Training
 - DataDewi_Testing
 - DataSekolah_Testin
 - DataSekolah_Trainin
 - HasilCuaca (asus - v1)
 - HasilSekolah (asus -

b. Mengimplementasikan Naive Bayes



C.

Hasil

Open in Turbo Prep Auto Model Filter (1 / 1 examples): all

Row No.	prediction(L...	confidence(...	confidence(...	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TEPAT	0.457	0.543	IPA	WANITA	LUAR SURA...	18	TIDAK

ExampleSet (Apply Model)

Name	Type	Missing	Statistics	Filter (8 / 8 attributes)
Prediction (Lama_Studi)	Binominal	0	Least: TERLAMBAT (0) Most: TEPAT (1)	Values: TEPAT (1), TERLAMBAT (0)
Confidence_TERLAMBAT	Real	0	Min: 0.457 Max: 0.457	Average: 0.457
Confidence_TEPAT	Real	0	Min: 0.543 Max: 0.543	Average: 0.543
Jurusan_SMA	Polynomial	0	Least: IPA (1) Most: IPA (1)	Values: IPA (1)
Gender	Polynomial	0	Least: WANITA (1) Most: WANITA (1)	Values: WANITA (1)
Asal_Sekolah	Polynomial	0	Least: LUAR SURAKARTA (1) Most: LUAR SURAKARTA (1)	Values: LUAR SURAKARTA (1)
Rerata_SKS	Integer	0	Min: 18 Max: 18	Average: 18
Asisten	Polynomial	0	Least: TIDAK (1) Most: TIDAK (1)	Values: TIDAK (1)

Showing attributes 1 - 8 Examples: 1 Special Attributes: 3 Regular Attributes: 5

Berdasarkan hasil di atas maka prediksi untuk Dewi dia akan lulus dengan tepat

7. Prekdisikan ketepatan lama studi si Jono, jika Jono adalah seorang PRIA yang berasal dari jurusan IPA pada saat SMA, asal sekolah dari SURAKARTA, mengambil SKS dengan rata-rata sebanyak 17 SKS tiap semester, dan pernah menjadi Asisten selama kuliah.
 - a. Mengimport Data Jono

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Sheet4
Cell range: A:E
Select All
☒ Define header row: 1

	A	B	C	D	E
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
2	LAIN	PRIA	SURAKARTA	17.000	YA

Previous
Next
Cancel

Import Data - Format your columns.

Format your columns.

☐ Replace errors with missing values ⓘ

	Jurusan_SMA <div>polynomial</div>	Gender <div>polynomial</div>	Asal_Sekolah <div>polynomial</div>	Rerata_SKS <div>integer</div>	Asisten <div>polynomial</div>
1	LAIN	PRIA	SURAKARTA	17	YA

PREVIEW

no problems.

Previous
Next
Cancel

ExampleSet (/Local Repository/DataCuaca_Testing) ExampleSet (/Local Repository/DataCuaca_Training) ExampleSet (/Local Repository/DataSekolah_Testing) ExampleSet (/Local Repository/DataSekolah_Training) ExampleSet (/Local Repository/DataDewi_Testing) ExampleSet (/Local Repository/DataSekolah_Testing) ExampleSet (/Local Repository/DataJono_Testing) ExampleSet (Apply Model)

Result History

Open in Turbo Prep Auto Model Filter (1 / 1 examples): all

Row No.	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	LAIN	PRIA	SURAKARTA	17	YA

ExampleSet (1 example, 0 special attributes, 5 regular attributes)

b. Mengimplementasikan Naive Bayes

Repository

Process

Parameters

Help

Process

Retrieve DataSekola... Naive Bayes Apply Model

Retrieve DataJono_T...

Log verbosity: init

Logfile:

Show advanced parameters

Change compatibility (9.3.001)

Process

RapidMiner Studio Core

Synopsis

The root operator which is the outer most operator of every process.

Description

c. Hasil

Open in Turbo Prep Auto Model Filter (1 / 1 examples): all

Row No.	prediction(L...	confidence(...	confidence(...	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TEPAT	0.076	0.924	LAIN	PRIA	SURAKARTA	17	YA

ExampleSet (/Local Repository/DataCuaca_Testing) ExampleSet (/Local Repository/DataCuaca_Training)
ExampleSet (/Local Repository/DataSekolah_Testing) ExampleSet (/Local Repository/DataSekolah_Training)
ExampleSet (/Local Repository/DataDewi_Testing) ExampleSet (/Local Repository/DataJono_Testing)

Result History ExampleSet (Apply Model) ExampleSet (/Local Repository/DataJono_Testing)

	Name	Type	Missing	Statistics		
Data	Prediction prediction(Lama_Studi)	Binominal	0	Least TERLAMBAT (0)	Most TEPAT (1)	Values TEPAT (1), TERLAMBAT (0)
Statistics	Confidence_TERLAMBAT confidence(TERLAMBAT)	Real	0	Min 0.076	Max 0.076	Average 0.076
	Confidence_TEPAT confidence(TEPAT)	Real	0	Min 0.924	Max 0.924	Average 0.924
Visualizations	Jurusan_SMA	Polynomial	0	Least LAIN (1)	Most LAIN (1)	Values LAIN (1)
	Gender	Polynomial	0	Least PRIA (1)	Most PRIA (1)	Values PRIA (1)
	Asal_Sekolah	Polynomial	0	Least SURAKARTA (1)	Most SURAKARTA (1)	Values SURAKARTA (1)
Annotations	Rerata_SKS	Integer	0	Min 17	Max 17	Average 17
	Asisten	Polynomial	0	Least YA (1)	Most YA (1)	Values YA (1)

Showing attributes 1 - 8 Examples: 1 Special Attributes: 3 Regular Attributes: 5

Repository

- Import Data
- Samples
- DB (Legacy)
- Local Repository (asus)
 - Connections (asus)
 - data (asus)
 - processes (asus)
 - DataCuaca_Testing
 - DataCuaca_Training
 - DataDewi_Testing
 - DataJono_Testing
 - DataSekolah_Testing
 - DataSekolah_Training
 - HasilCuaca (asus - v1)
 - HasilDewi (asus - v1)
 - HasilSekolah (asus - v1)

Berdasarkan hasil di atas maka prediksi untuk Jono dia akan lulus dengan tepat