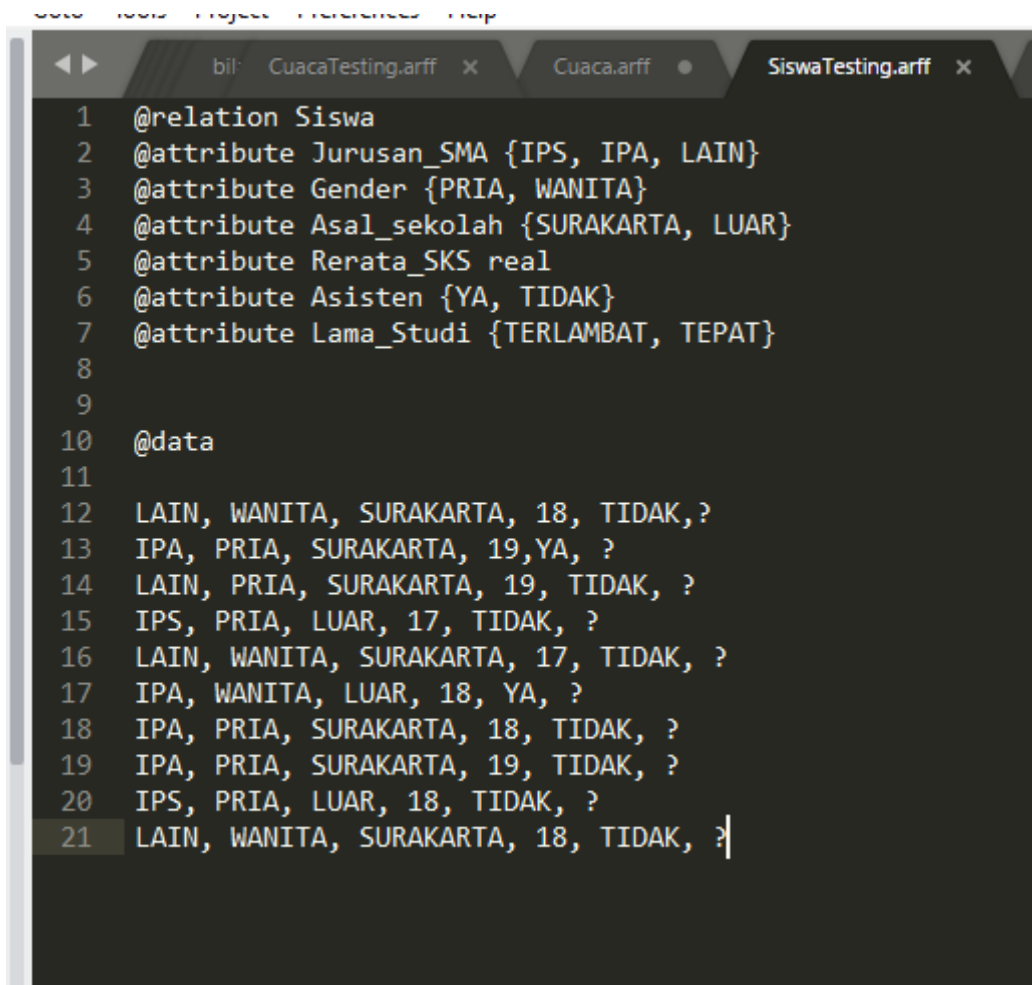


Nama : Intan Larasati
NIM : L200170091
Kelas : D

TUGAS

1. persiapkan file Siswa.arff
2. buatlah sebuah data testing dengan format arff sebagai data yang akan diprediksi dengan memiliki variabel variabel independen dan variabel dependen yang sama.
3. simpan dengan nama SiswaTesting.arff

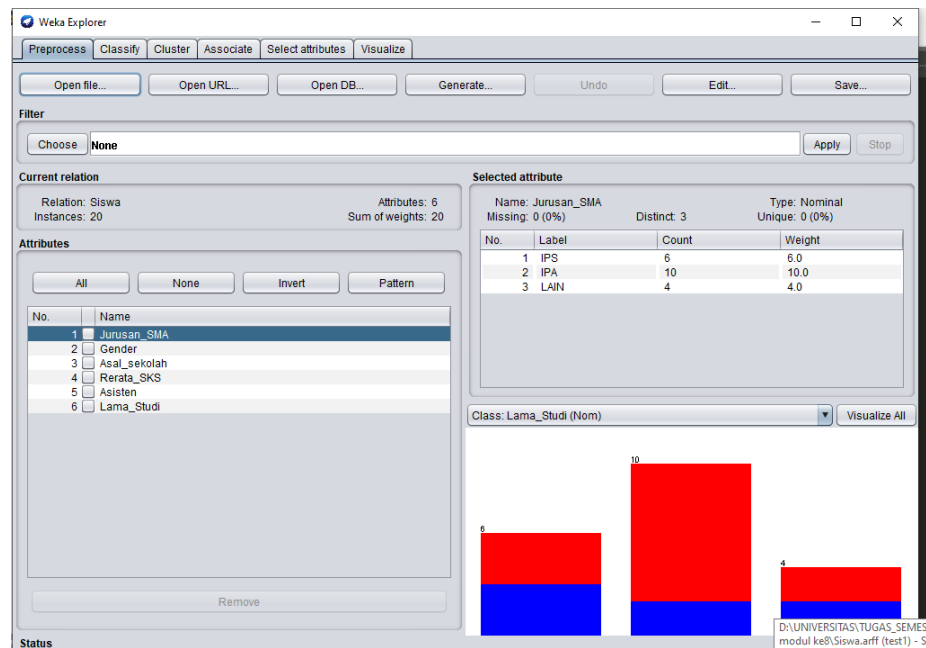
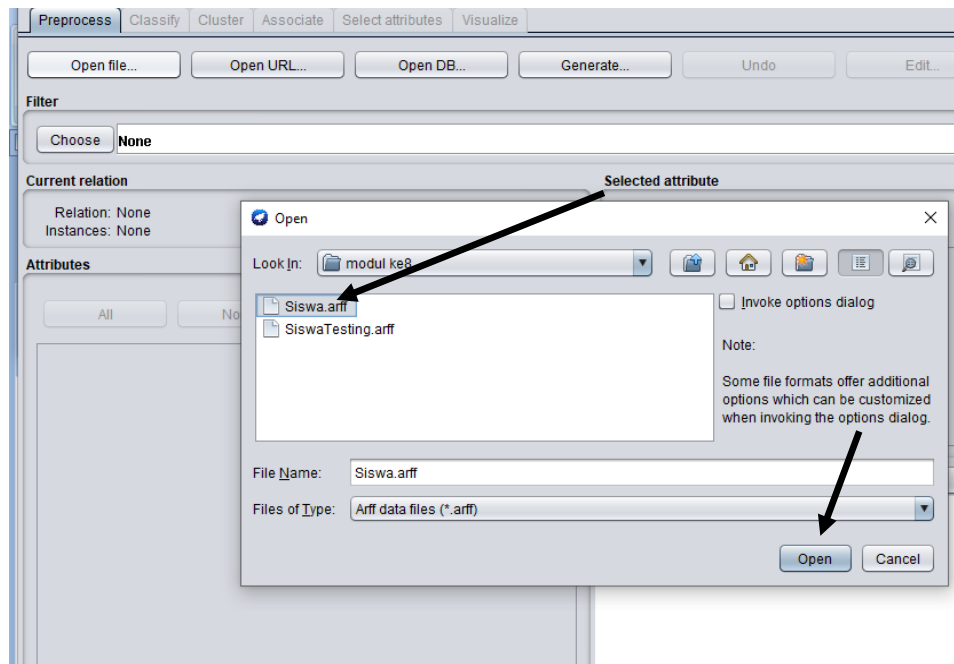


```
@relation Siswa
@attribute Jurusan_SMA {IPS, IPA, LAIN}
@attribute Gender {PRIA, WANITA}
@attribute Asal_sekolah {SURAKARTA, LUAR}
@attribute Rerata_SKS real
@attribute Asisten {YA, TIDAK}
@attribute Lama_Studi {TERLAMBAT, TEPAT}

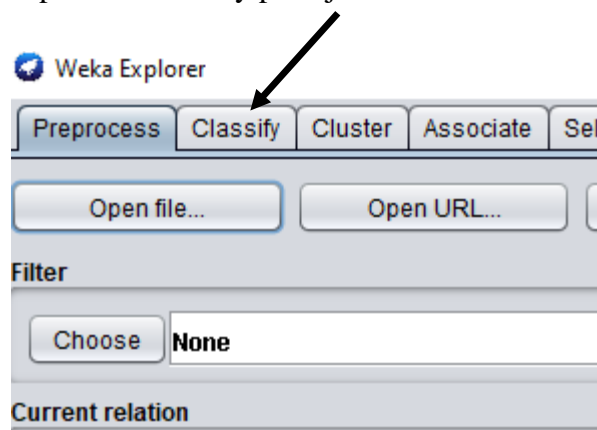
@data
12 LAIN, WANITA, SURAKARTA, 18, TIDAK, ?
13 IPA, PRIA, SURAKARTA, 19, YA, ?
14 LAIN, PRIA, SURAKARTA, 19, TIDAK, ?
15 IPS, PRIA, LUAR, 17, TIDAK, ?
16 LAIN, WANITA, SURAKARTA, 17, TIDAK, ?
17 IPA, WANITA, LUAR, 18, YA, ?
18 IPA, PRIA, SURAKARTA, 18, TIDAK, ?
19 IPA, PRIA, SURAKARTA, 19, TIDAK, ?
20 IPS, PRIA, LUAR, 18, TIDAK, ?
21 LAIN, WANITA, SURAKARTA, 18, TIDAK, ?
```

4. jika telah selesai membuat buka aplikasi weka, masuk dalam menu Weka explorer
5. buka kembali file Siswa.arff

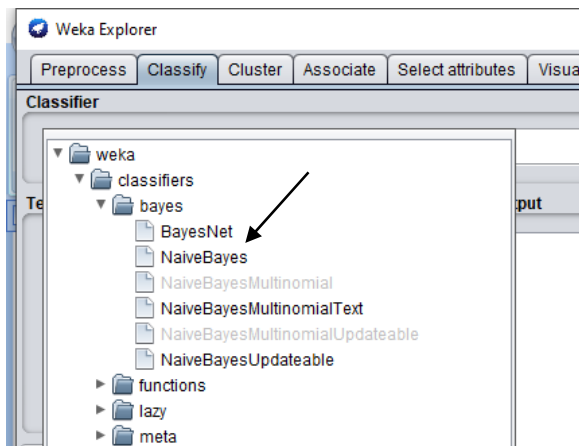
```
Siswa1.arff x  siswa.arff x  Siswa.arff x
1 @relation Siswa
2 @attribute Jurusan_SMA {IPS, IPA, LAIN}
3 @attribute Gender {PRIA, WANITA}
4 @attribute Asal_sekolah {SURAKARTA, LUAR}
5 @attribute Rerata_SKS real
6 @attribute Asisten {YA, TIDAK}
7 @attribute Lama_Studi {TERLAMBAT, TEPAT}
8
9 @data
10 IPS, WANITA, SURAKARTA, 18, TIDAK, TERLAMBAT
11 IPA, PRIA, SURAKARTA, 19, YA, TEPAT
12 LAIN, PRIA, SURAKARTA, 19, TIDAK, TERLAMBAT
13 IPA, PRIA, LUAR, 17, TIDAK, TERLAMBAT
14 IPA, WANITA, SURAKARTA, 17, TIDAK, TEPAT
15 IPA, WANITA, LUAR, 18, YA, TEPAT
16 IPA, PRIA, SURAKARTA, 18, TIDAK, TERLAMBAT
17 IPA, PRIA, SURAKARTA, 19, TIDAK, TEPAT
18 IPS, PRIA, LUAR, 18, TIDAK, TERLAMBAT
19 LAIN, WANITA, SURAKARTA, 18, TIDAK, TEPAT
20 IPA, WANITA, SURAKARTA, 19, TIDAK, TEPAT
21 IPS, PRIA, SURAKARTA, 20, TIDAK, TEPAT
22 IPS, PRIA, SURAKARTA, 19, TIDAK, TEPAT
23 IPA, PRIA, SURAKARTA, 19, TIDAK, TEPAT
24 IPA, PRIA, LUAR, 22, YA, TEPAT
25 LAIN, PRIA, SURAKARTA, 16, TIDAK, TERLAMBAT
26 IPS, PRIA, LUAR, 20, TIDAK, TEPAT
27 LAIN, PRIA, LUAR, 23, YA, TEPAT
28 IPA, PRIA, SURAKARTA, 21, YA, TEPAT
29 IPS, PRIA, SURAKARTA, 19, TIDAK, TERLAMBAT
```



6. pilih tab classify pada jendela weka



7. klik, lalu akan muncul tombol choose dan klik untuk memilih metode / algoritma naive bayes

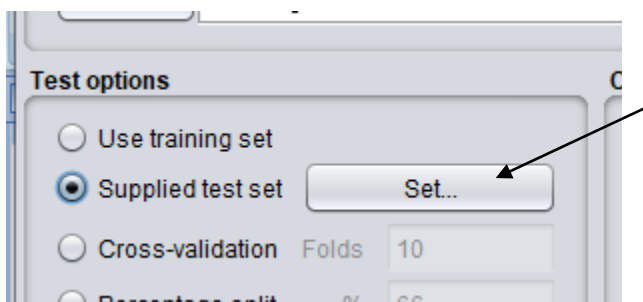


8. menentukan data testing sebagai data yang akan diprediksi variabel independennya. File SiswaTesting.

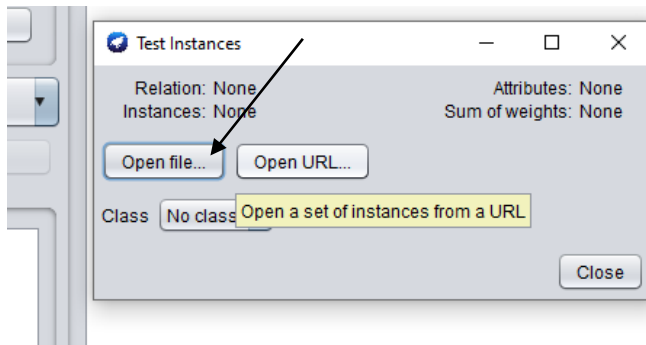
9. pada menu test options terdapat 4 pilihan pengujian yaitu :

- Use training set
- Supplied test set
- Cross-Validation
- Percentege split

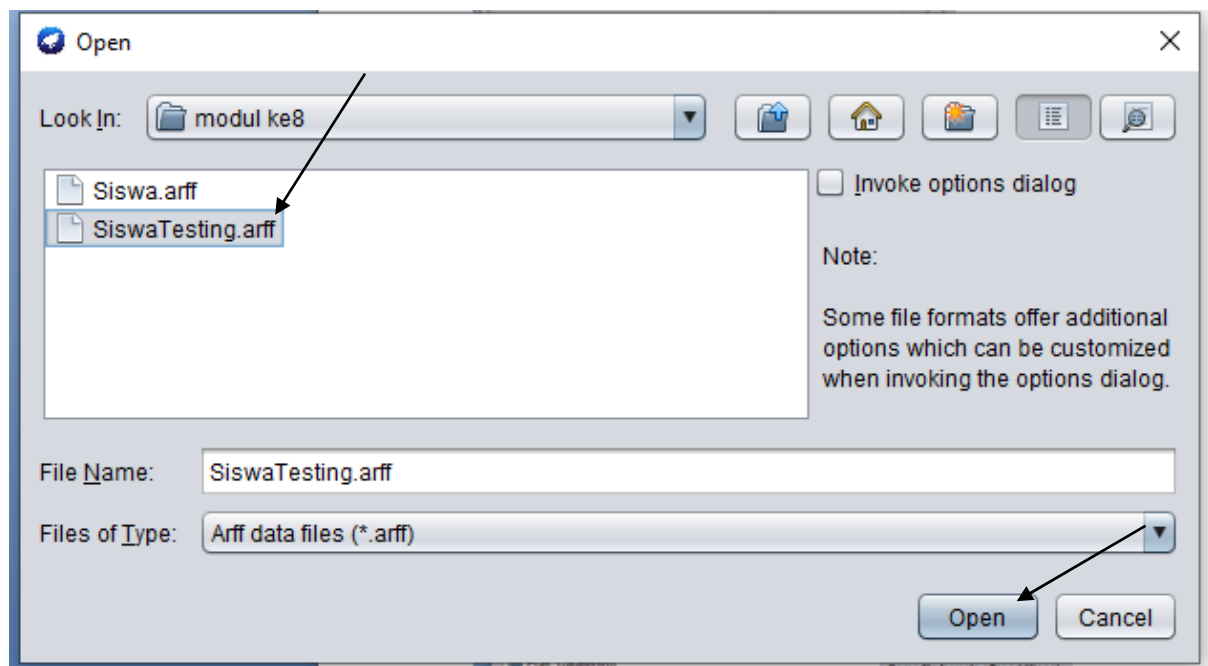
10. gunakan pilihan Supplied test set



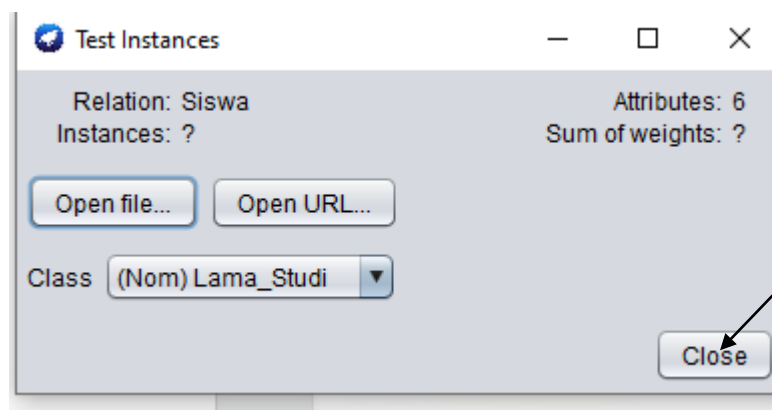
11. sehingga akan muncul jendela test instance Klik open file



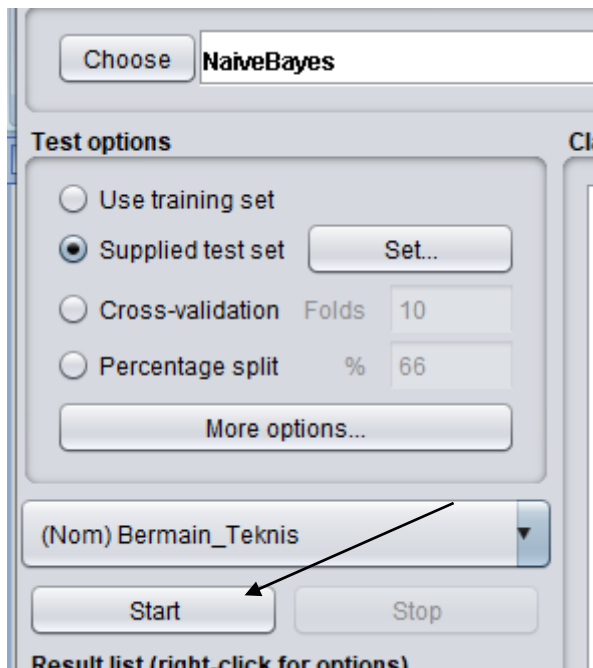
12. pilih file SiswaTesting.arff sebagai data uji. klik open



13. Lalu klik close



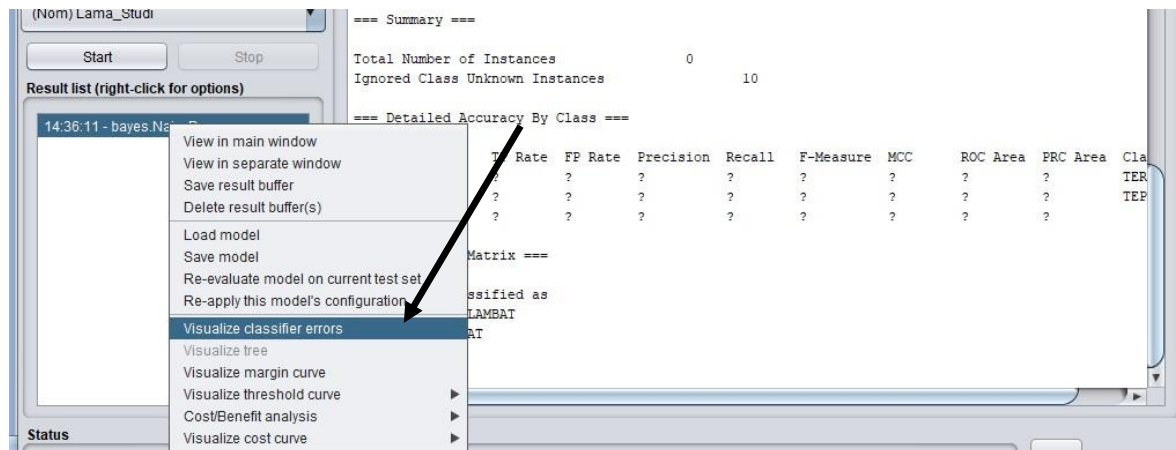
14. klik start untuk memulai proses naive bayes



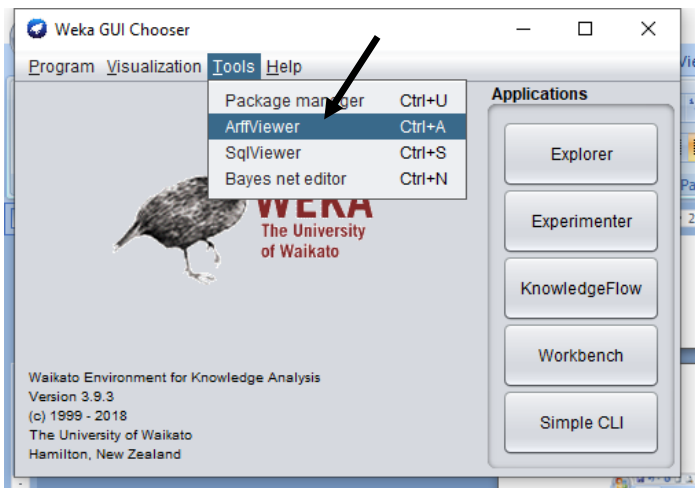
15. jika muncul jendela pesan classifier panel. Kita abaikan dengan mengklik yes. Sehingga algoritma naive bayes akan diproses

16. kita abaikan nilai nilai yang ditampilkan dalam jendela classifier output

17. untuk melihat hasil prediksi terhadap data uji, lihat nilai classifier errors. Klik kanan pada hasil proses dalam kotak result list. Pilih menu visual classifier errors



18. pada jendela weka classifier visualize, abaikan hasil apapun yang ditampilkan. Klik save . simpan dengan nama file HasilPrediksi.arff



20. jendela arff-viewers akan ditampilkan. Buka menu file-open. Tunjukkan pada file HasilPrediksi.arff jendela arff-viewers akan ditampilkan. Buka menu file-open. Tunjukkan pada file HasilPrediksi.arff yang telah anda simpan. Lihatlah hasil prediksi

ARFF-Viewer - D:\UNIVERSITAS\TUGAS_SEMESTER5\Prak_DataMining\modul ke8\HasilPrediksi.arff

File Edit View

HasilPrediksi.arff

Relation: Siswa_predicted

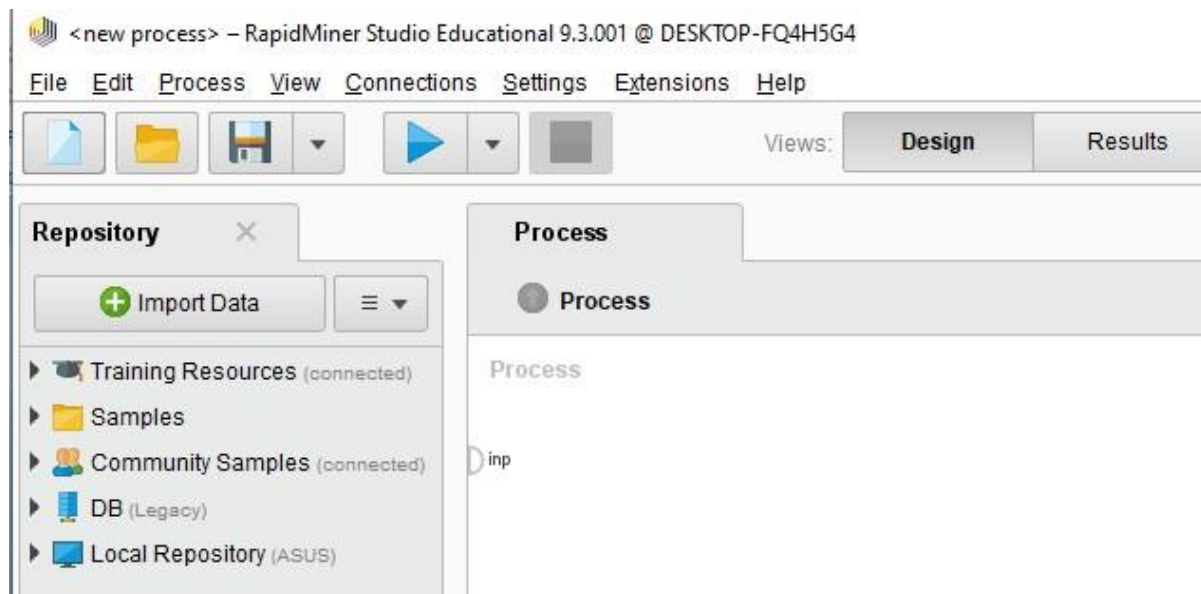
No.	1: Jurusan_SMA	2: Gender	3: Asal_sekolah	4: Rerata_SKS	5: Asisten	6: prediction margin	7: predicted Lama_Studi	8: Lama_Studi
	Nominal	Nominal	Nominal	Numeric	Nominal	Numeric	Nominal	Nominal
1	LAIN	WANITA	SURAKARTA	18.0	TIDAK	0.375862	TERLAMBAT	
2	IPA	PRIA	SURAKARTA	19.0	YA	-0.836469	TEPAT	
3	LAIN	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	

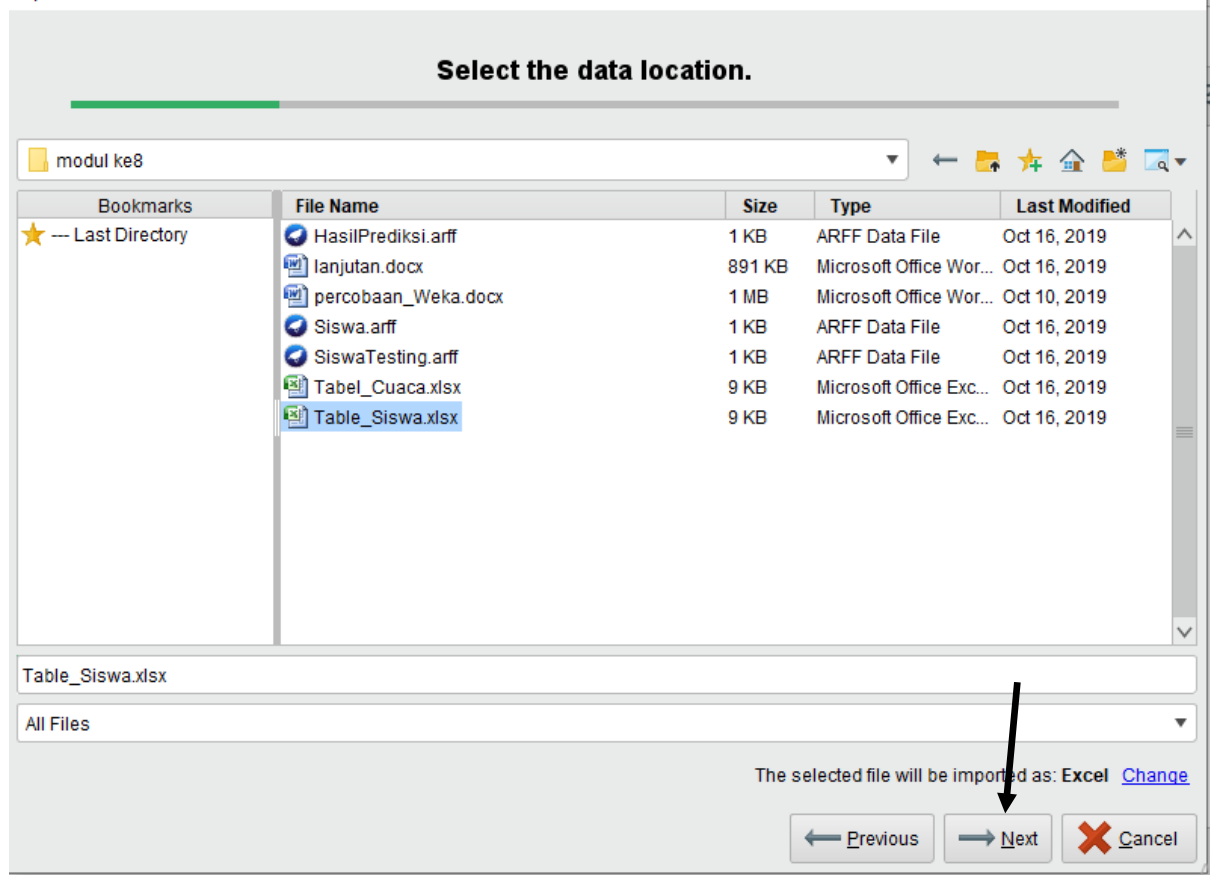
Menggunakan Naive Bayes dengan RapidMiner

1. siapkan Table_Siswa.xlsx

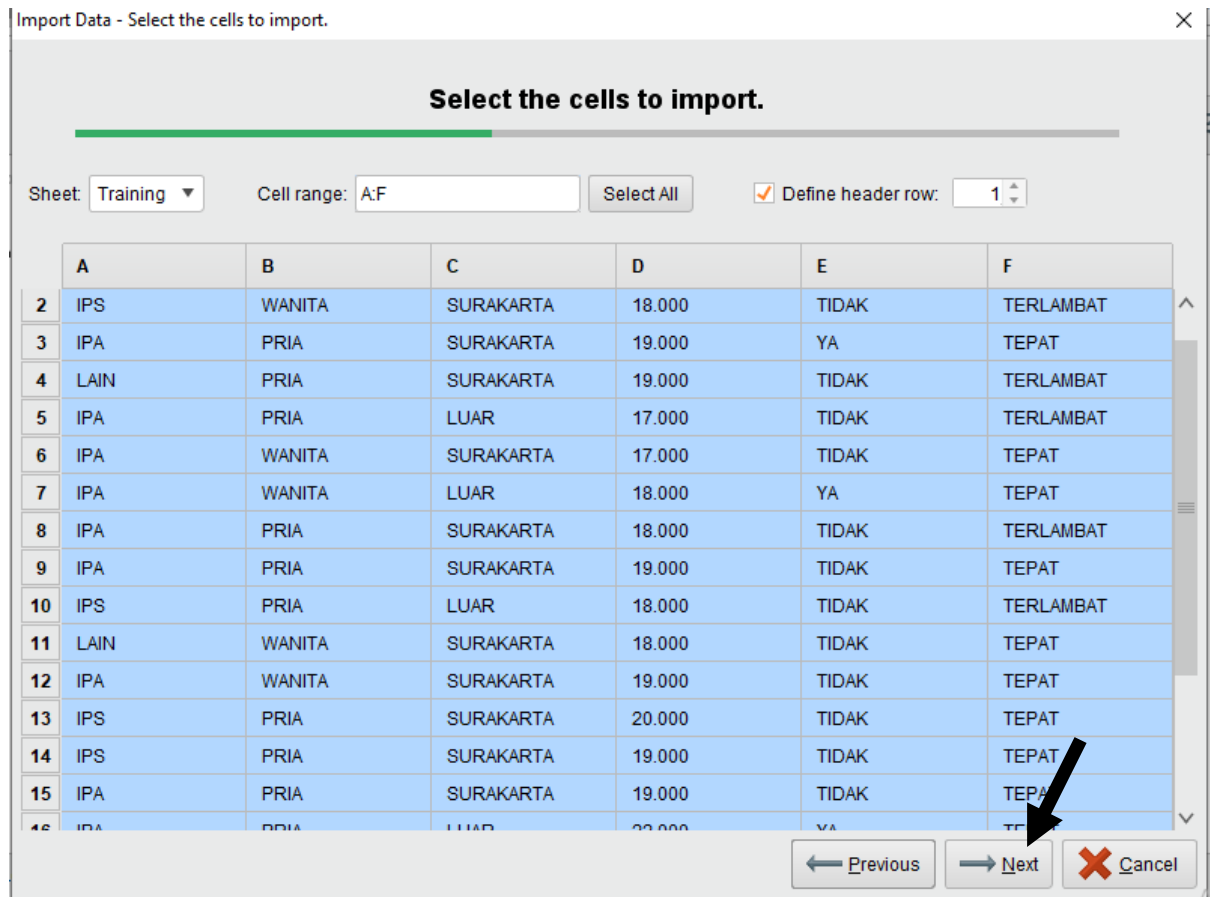
	A	B	C	D	E	
1	Jurusan_S	Gender	Asal_Sekolah	Rerata_SK	Asisten	
2	LAIN	WANITA	SURAKARTA	18	TIDAK	
3	IPA	PRIA	SURAKARTA	19	YA	
4	LAIN	PRIA	SURAKARTA	19	TIDAK	
5	IPS	PRIA	LUAR	17	TIDAK	
6	LAIN	WANITA	SURAKARTA	17	TIDAK	
7	IPA	WANITA	LUAR	18	YA	
8	IPA	PRIA	SURAKARTA	18	TIDAK	
9	IPA	PRIA	SURAKARTA	19	TIDAK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

2. buka aplikasi rapid miner, lalu klik import. Lalu arahkan direktori tempat penyimpanan file pada langkah select the data location. Lalu klik next

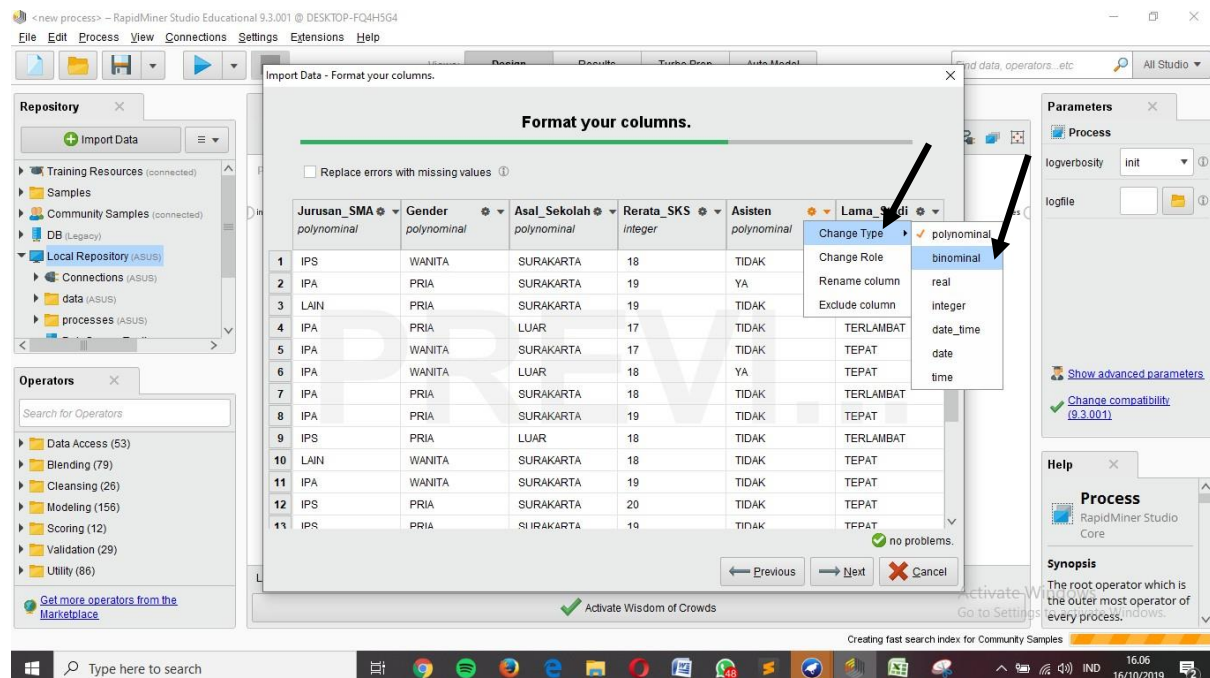


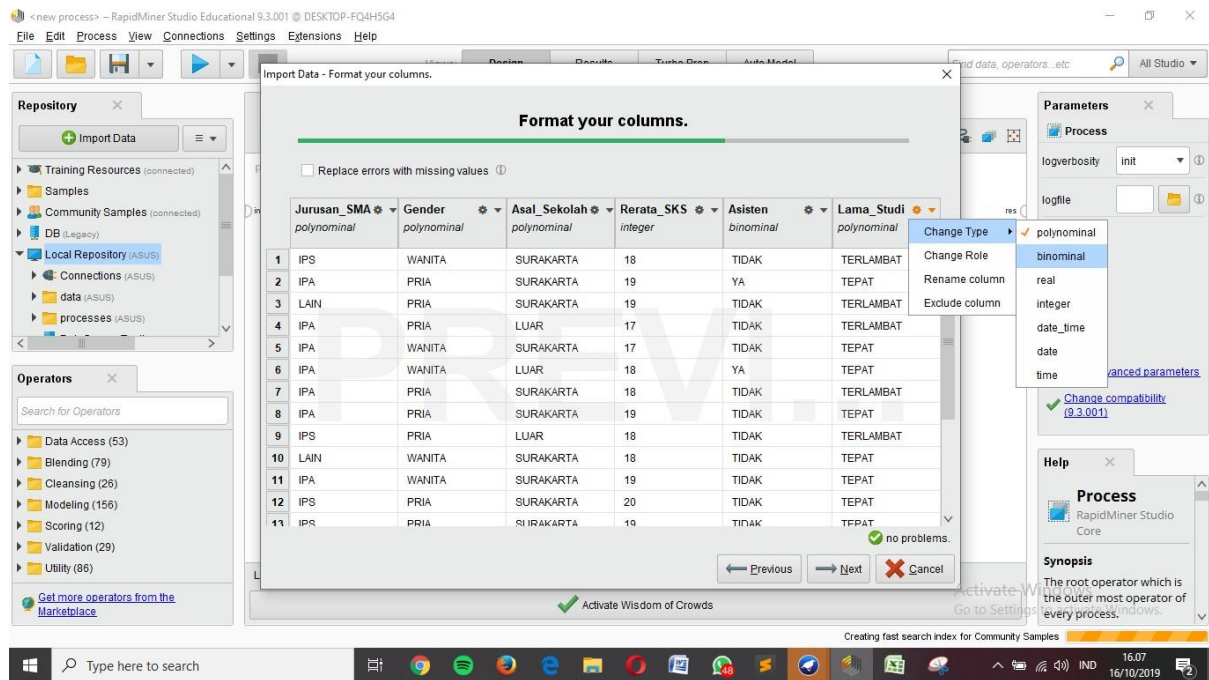


3. pastikan sel excel sesuai di langkah select the cells to import

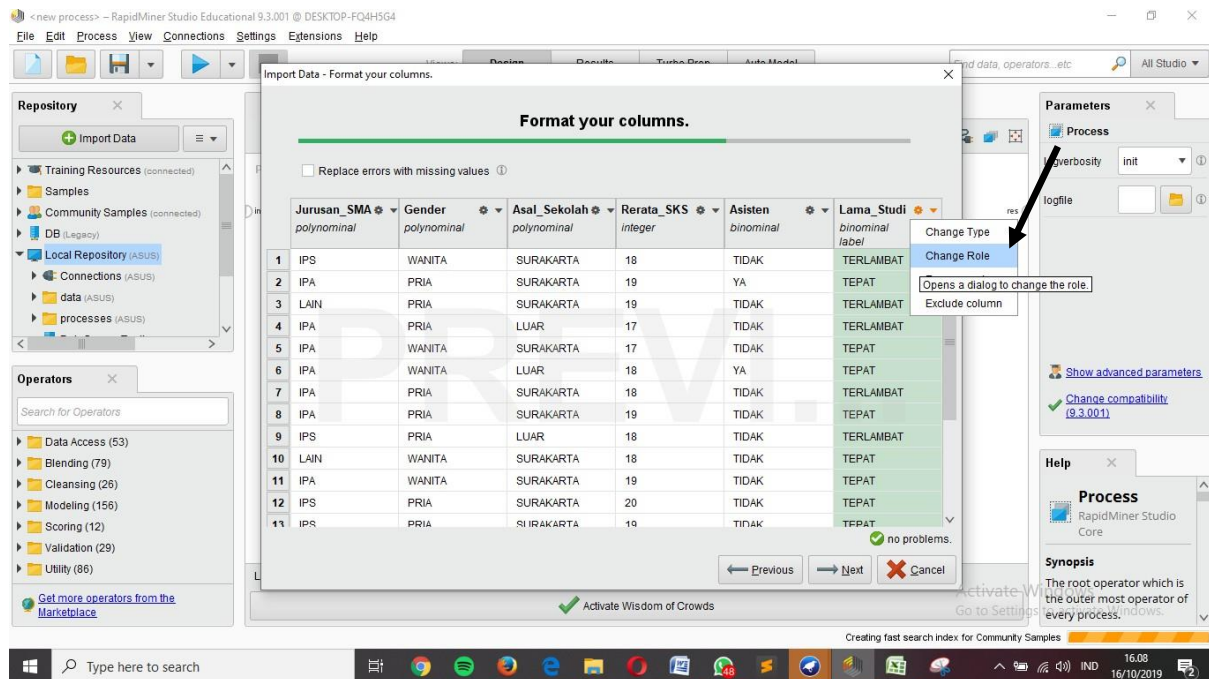


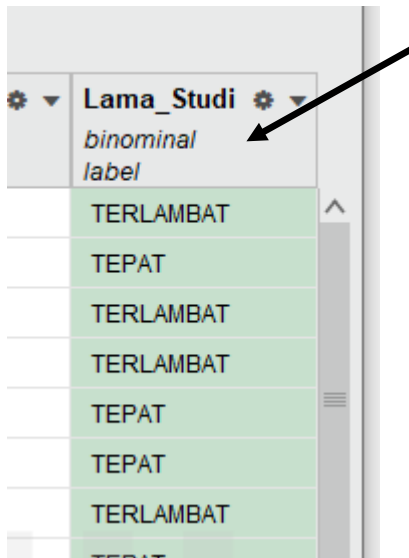
4. pada langkah format your columns ubah kolom Lama_Studi dan Asisten dengan tipe data binomial karena hanya ada dua keputusan YA/TIDAK



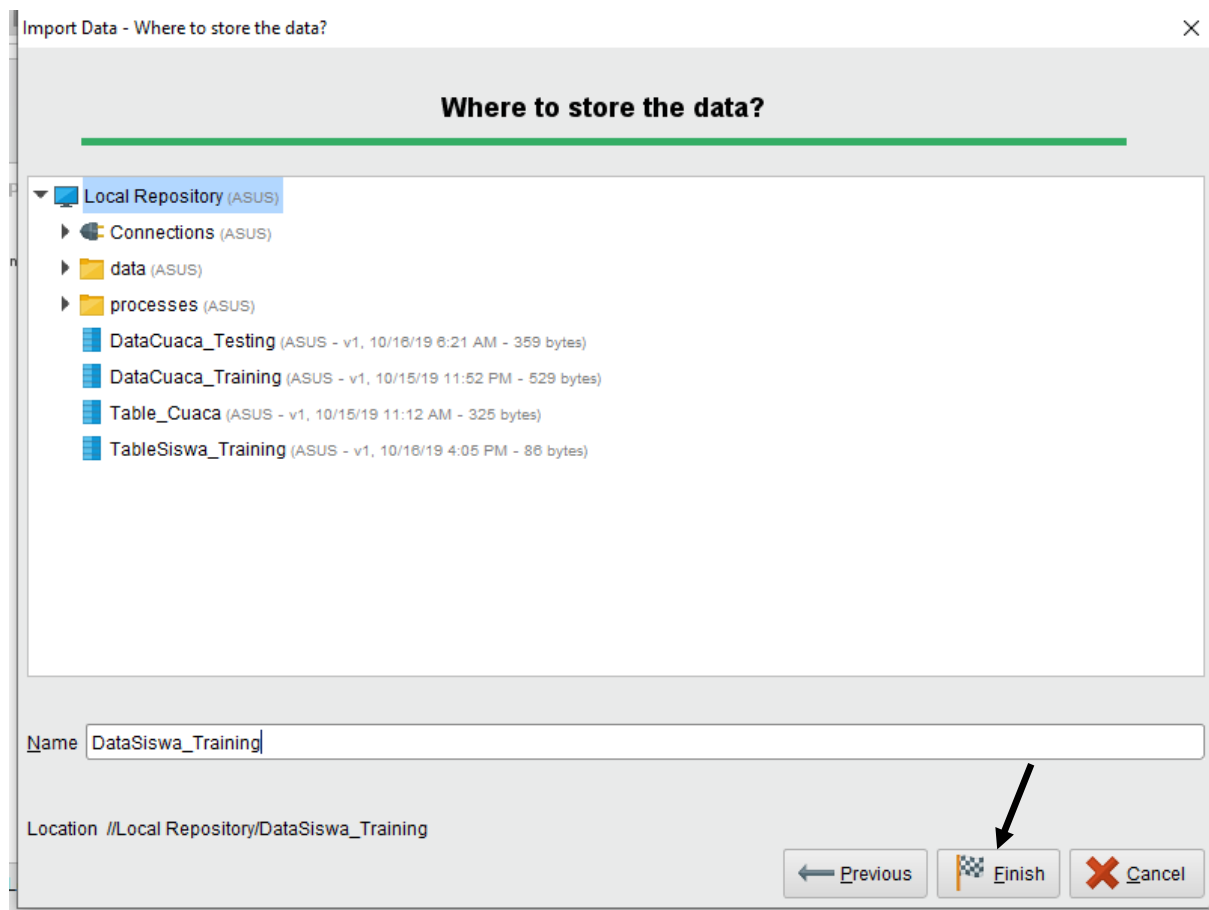


5. ubah juga label pada change Role





6. simpan dengan nama DataSiswa_Training dilanjutkan klik tombol finish



7. hasil import file Tabel_Siswa.xls akan ditampilkan

ExampleSet (/Local Repository/TableSiswa_Training)

Result History

ExampleSet (/Local Repository/DataSiswa_Training)

Open in

Turbo Prep

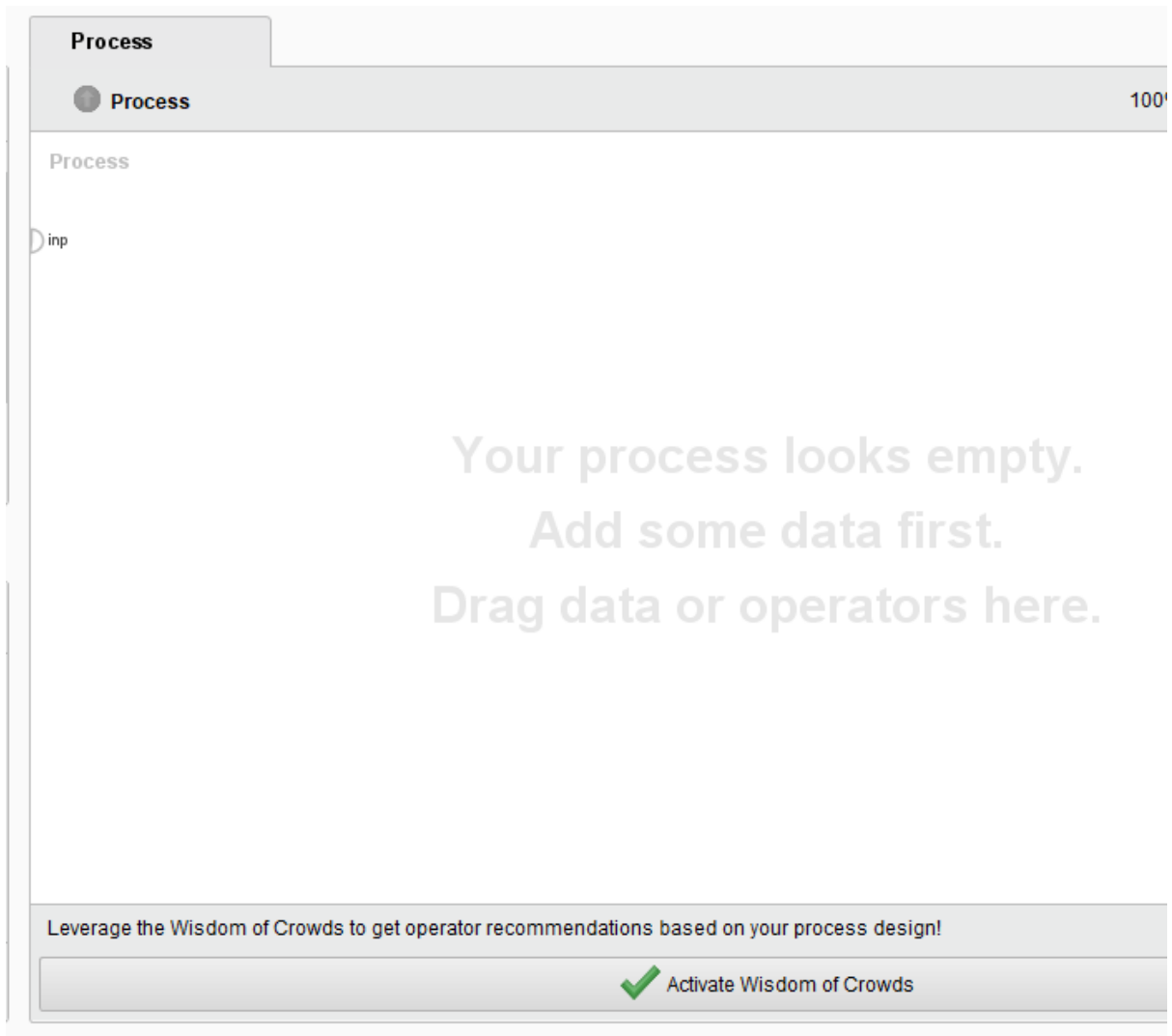
Auto Model

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)

8. kembali ke jendela design perspective dengan shorcut tombol F8



9. lakukan hal yg sama untuk data testing yg diambil dari Table_Cuaca.xls pda sheet 2. Dengan mengulang langkah 2. Pastikan semua variable data testing terpilih 4 variabel, beda pada langkah ini tidak ada variable yg diubah tipe label

Import Data - Select the cells to import.

Select the cells to import.

Sheet: **Testing** Cell range: **A1:A1** **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

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>binomial</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

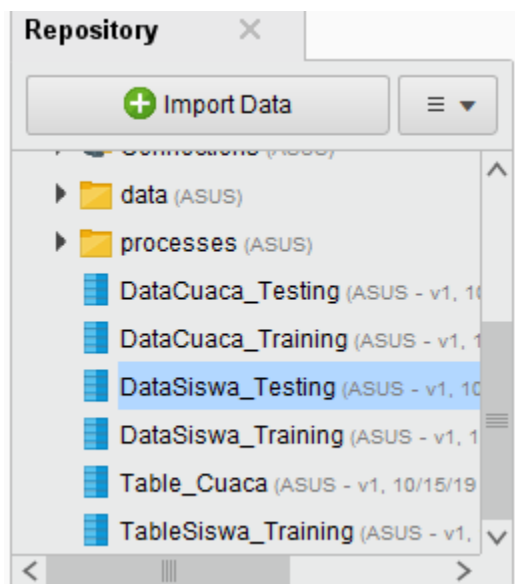
no problems.

ExampleSet (//Local Repository/DataSiswa_Testing)

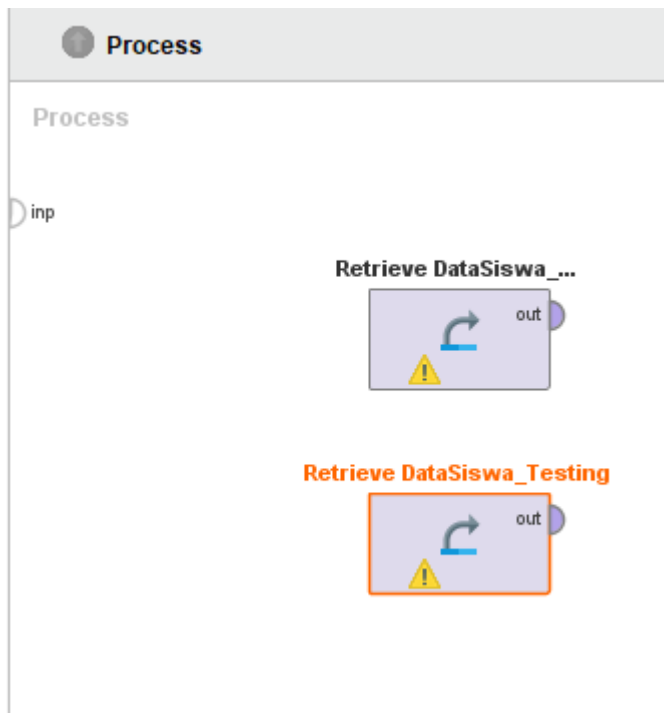
Open in Turbo Prep Auto Model Filter (10 / 10 examples):

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

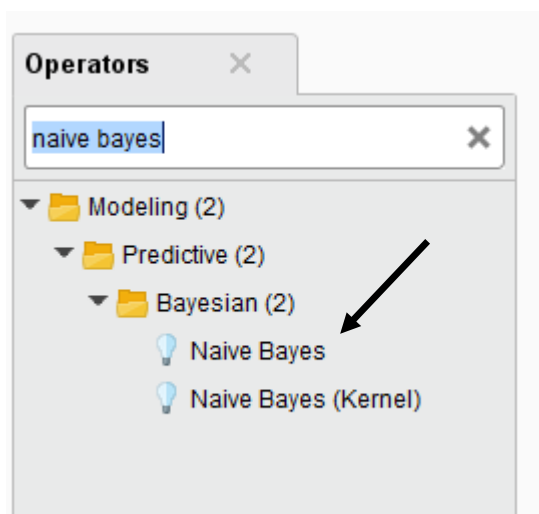
10. simpan dengan nama DataSiswa_Testing

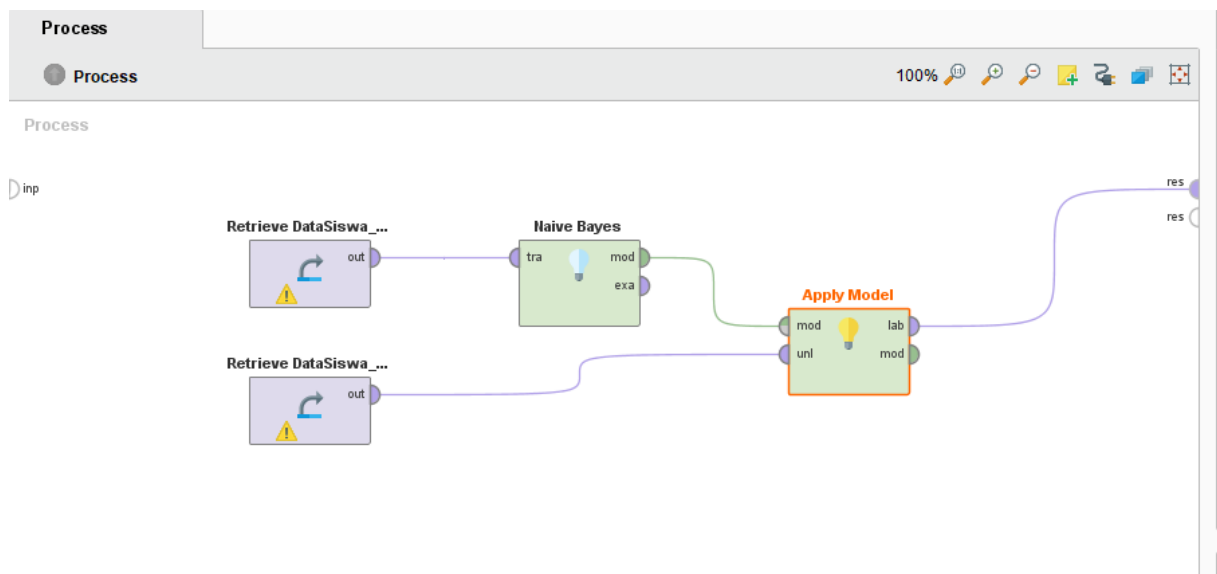
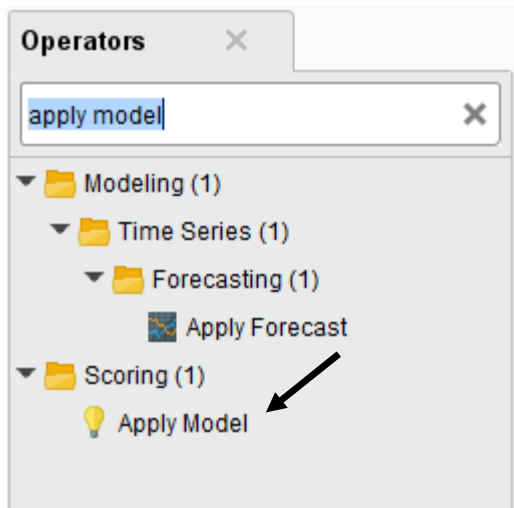


11. lalu buat design naive bayes. Drag DataSiswa_Training dan DataSiswa_Testing ke dalam jendela proses view



12. Masukkan juga operator naive bayes dan apply model ke dalam process view. Hubungkan konektor masing masing data terhadap operator seperti gambar





13. jalankan proses naive bayes dengan menekan tombol Run

14. perhatikan hasil proses klasifikasi naive bayes. Pada tab Data, dapat kita lihat hasil prediksi terhadap data testing serta tingkat confidence nilai kelas pada masing masing data

ExampleSet (Apply Model)		ExampleSet (//Local Repository/DataSiswa_Testing)						
Open in		Turbo Prep	Auto Model	Filter (10 / 10 examples): all				
No.	prediction(L...	confidence(...	confidence(...	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
	TERLAMBAT	0.648	0.352	LAIN	WANITA	SURAKARTA	18	TIDAK
	TEPAT	0.005	0.995	IPA	PRIA	SURAKARTA	19	YA
	TERLAMBAT	0.650	0.350	LAIN	PRIA	SURAKARTA	19	TIDAK
	TERLAMBAT	0.868	0.132	IPS	PRIA	LUAR	17	TIDAK
	TERLAMBAT	0.738	0.262	LAIN	WANITA	SURAKARTA	17	TIDAK
	TEPAT	0.005	0.995	IPA	WANITA	LUAR	18	YA
	TERLAMBAT	0.547	0.453	IPA	PRIA	SURAKARTA	18	TIDAK
	TEPAT	0.321	0.679	IPA	PRIA	SURAKARTA	19	TIDAK
	TERLAMBAT	0.811	0.189	IPS	PRIA	LUAR	18	TIDAK
	TERLAMBAT	0.648	0.352	LAIN	WANITA	SURAKARTA	18	TIDAK

Pada tab statistic, dapat dilihat bahwa rerata nilai confidence sebesar 0,524 untuk nilai TERLAMBAT, dan 0,476 untuk nilai TEPAT

ExampleSet (Apply Model)		ExampleSet (//Local Repository/DataSiswa_Testing)						
Open in		Turbo Prep	Auto Model	Filter (10 / 10 examples): all				
No.	prediction(L...	confidence(...	confidence(...	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
	TERLAMBAT	0.648	0.352	LAIN	WANITA	SURAKARTA	18	TIDAK
	TEPAT	0.005	0.995	IPA	PRIA	SURAKARTA	19	YA
	TERLAMBAT	0.650	0.350	LAIN	PRIA	SURAKARTA	19	TIDAK
	TERLAMBAT	0.868	0.132	IPS	PRIA	LUAR	17	TIDAK
	TERLAMBAT	0.738	0.262	LAIN	WANITA	SURAKARTA	17	TIDAK
	TEPAT	0.005	0.995	IPA	WANITA	LUAR	18	YA
	TERLAMBAT	0.547	0.453	IPA	PRIA	SURAKARTA	18	TIDAK
	TEPAT	0.321	0.679	IPA	PRIA	SURAKARTA	19	TIDAK
	TERLAMBAT	0.811	0.189	IPS	PRIA	LUAR	18	TIDAK
	TERLAMBAT	0.648	0.352	LAIN	WANITA	SURAKARTA	18	TIDAK

Name	Type	Missing	Filter (8 / 8 attributes): Search for Attributes
Binominal	0	Least TEPAT (3)	Most TERLAMBAT (7)
Real	0	Min 0.005	Max 0.868
Real	0	Min 0.132	Max 0.995
Polynomial	0	Least IPS (2)	Most IPA (4)
Polynomial	0	Least WANITA (4)	Most PRIA (6)
Polynomial	0	Least LUAR (3)	Most SURAKARTA (7)
Integer	0	Min 17	Max 19

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

15. tambahkan data testing berupa nama dewi dan jono

	A	B	C	D	E
1	Jurusan_S	Gender	Asal_Sekolah	Rerata_SK	Asisten
2	IPA	WANITA	LUAR	18	TIDAK
3	LAIN	PRIA	SURAKARTA	17	YA
4					
5					

16. lakukan import data

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	18.000	TIDAK
3	LAIN	PRIA	SURAKARTA	17.000	YA

17. ubah Asisten menjadi Binomial

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>binomial</i>
1	IPA	WANITA	LUAR	18	TIDAK
2	LAIN	PRIA	SURAKARTA	17	YA

18. simpan dengan nama Tugas_Testing

Where to store the data?

Local Repository (ASUS)

- ▶ Connections (ASUS)
- ▶ data (ASUS)
- ▶ processes (ASUS)
 - DataCuaca_Testing (ASUS - v1, 10/16/19 6:21 AM - 359 bytes)
 - DataCuaca_Training (ASUS - v1, 10/15/19 11:52 PM - 529 bytes)
 - DataSiswa_Testing (ASUS - v1, 10/16/19 4:16 PM - 483 bytes)
 - DataSiswa_Training (ASUS - v1, 10/16/19 4:10 PM - 673 bytes)
 - Table_Cuaca (ASUS - v1, 10/15/19 11:12 AM - 325 bytes)
 - TableSiswa_Training (ASUS - v1, 10/16/19 4:05 PM - 86 bytes)

Name

Location //Local Repository/Tugas_Testing

← Previous
 Finish
 Cancel

19. lihat hasil dari penyimpanan

reset (/Local Repository/DataSiswa_Testing)

ExampleSet (/Local Repository/Tugas_Testing) ✕

ExampleSet (/Local Repository/DataSiswa_Training)

ExampleSet (/Local Repository/Tugas_Testing) ✕

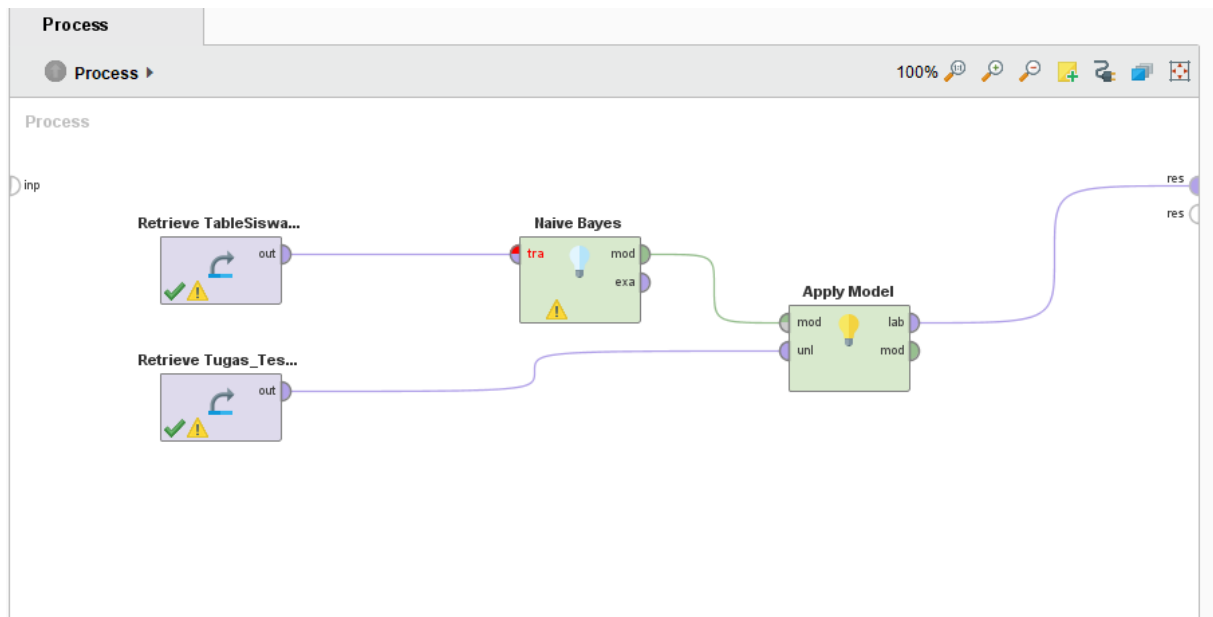
ExampleSet (Apply M

Open in Turbo Prep Auto Model

Filter (2 / 2 examples): all

Row No.	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	IPA	WANITA	LUAR	18	TIDAK
2	LAIN	PRIA	SURAKARTA	17	YA

20. lalu drag TableSiswa_Training dan Tugas_Testing



21. lalu run, dan liat hasilnya

ExampleSet (apply model)								
ExampleSet (Local repository: siswa11m11g)								
Open in		Turbo Prep	Auto Model	Filter (2 / 2 examples):				
Row No.	prediction(L...	confidence(...	confidence(...	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asist
1	TEPAT	0.298	0.702	IPA	WANITA	LUAR	18	TIDAK
2	TEPAT	0.076	0.924	LAIN	PRIA	SURAKARTA	17	YA

22. Pada tab statistic, dapat dilihat bahwa rerata nilai confidence sebesar 0,187 untuk nilai TERLAMBAT dan 0,813 untuk nilai TEPAT

ExampleSet (/Local Repository/DataSiswa_Training)

ExampleSet (/Local Repository/TableSiswa_Training)

ExampleSet (/Local Repository/Tugas_Testing)

ExampleSet (/Local Repository/DataSiswa_Testing)

Result History

ExampleSet (Apply Model)

ExampleSet (/Local Repository/siswatraining)

Data	Name	Type	Missing	Filter (8 / 8 attributes):	Search for Attributes
	Binominal	0	Least TERLAMBAT (0)	Most TEPAT (2)	Values TEPAT (2), TERLAMBAT (0)
Statistics	Real	0	Min 0.076	Max 0.298	Average 0.187
	Real	0	Min 0.702	Max 0.924	Average 0.813
Visualizations	Polynominal	0	Least LAIN (1)	Most IPA (1)	Values IPA (1), LAIN (1)
	Polynominal	0	Least WANITA (1)	Most PRIA (1)	Values PRIA (1), WANITA (1)
Annotations	Polynominal	0	Least SURAKARTA (1)	Most LUAR (1)	Values LUAR (1), SURAKARTA (1)
	Integer	0	Min 17	Max 18	Average 17.500

Showing attributes 1 - 8

Examples: 2 Special Attributes: 3 Regular Attributes: 5