

Nama : Oki Kus Mahesa

Nim : L200170064

Kelas : C

MODUL 8

KLASIFIKASI : NAÏVE BAYES

- TUGAS

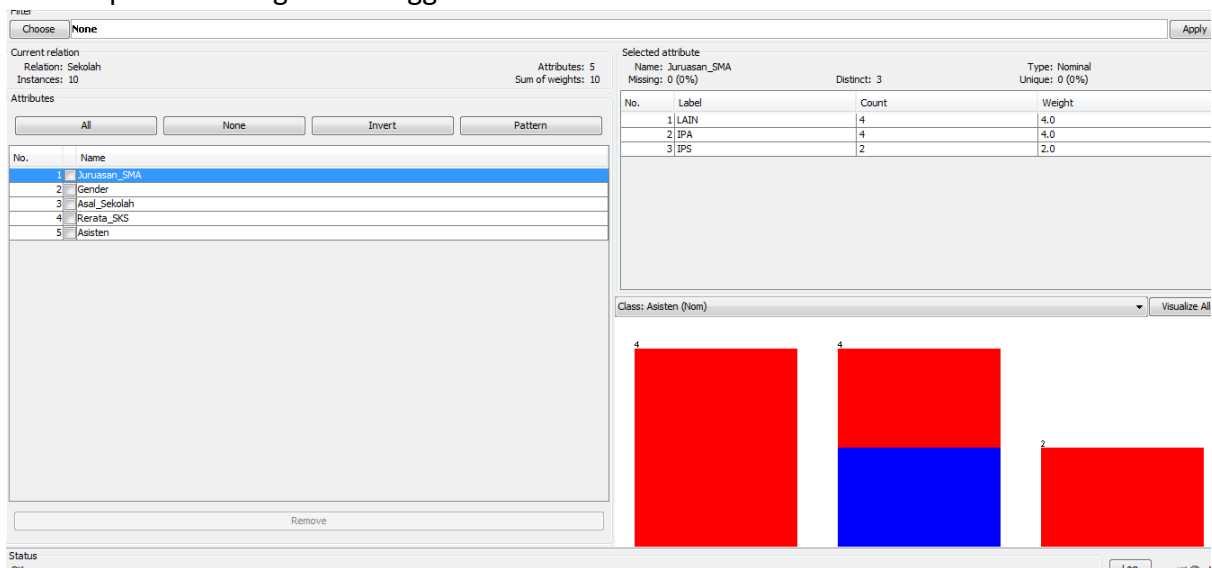
1. Berdasarkan tabel berikut, buatlah file dalam format.xls dan format.arff, Data ini akan di gunakan sebagai DataTesting

The image shows two side-by-side windows. The left window is Microsoft Excel, titled 'data testing - Excel', displaying a table with 4 columns: Jurusan_SMA, Gender, Asal_Sekolah, and Real. The right window is Visual Studio Code, titled 'Cuaca.arff - Visual Studio Code [A...]', showing the content of a file named 'Cuaca.arff'.

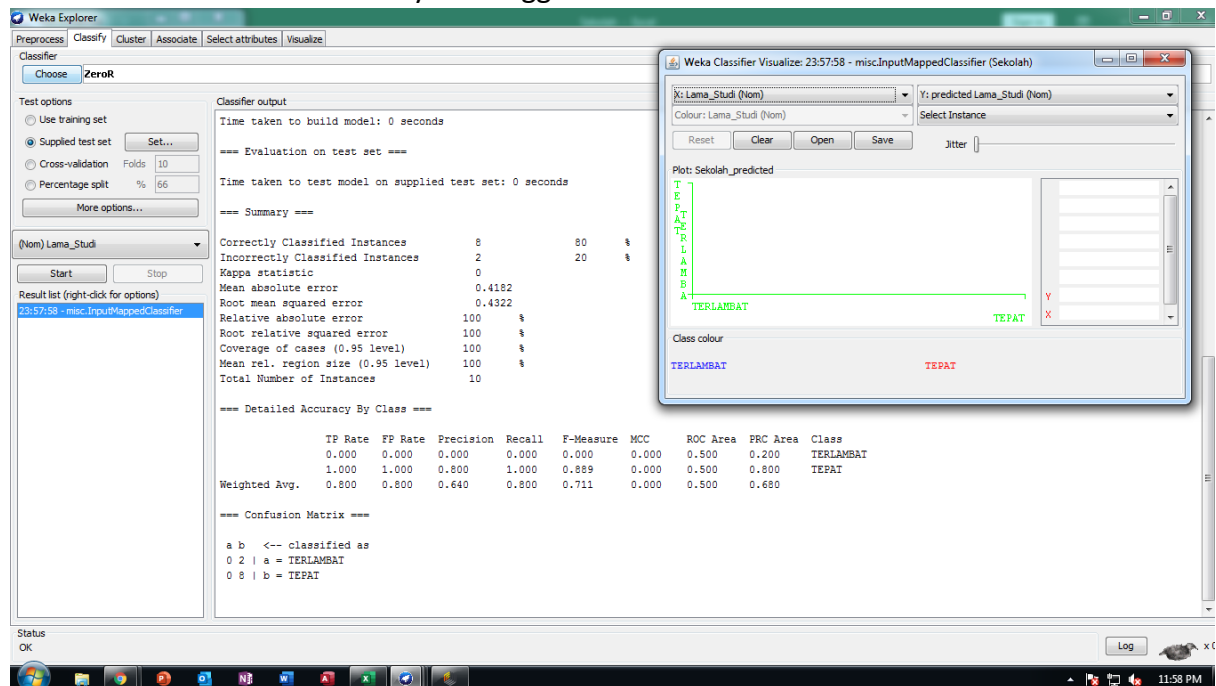
1	Jurusan_SMA	Gender	Asal_Sekolah	Real
2	LAIN	WANITA	SURAKARTA	
3	IPA	PRIA	SURAKARTA	
4	LAIN	PRIA	SURAKARTA	
5	IPS	PRIA	LUAR	
6	LAIN	WANITA	SURAKARTA	
7	IPA	WANITA	LUAR	
8	IPA	PRIA	SURAKARTA	
9	IPA	PRIA	SURAKARTA	
10	IPS	PRIA	LUAR	
11	LAIN	WANITA	SURAKARTA	
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				

```
@relation Sekolah
1
2
3 @attribute Jurusan_SMA {LAIN, IPA, IPS}
4 @attribute Gender {WANITA, PRIA}
5 @attribute Asal_Sekolah {SURAKARTA, LUAR}
6 @attribute Berata_SKS real
7 @attribute Asisten {YA, TIDAK}
8
9 @data
10 LAIN, WANITA, SURAKARTA, 18, TIDAK
11 IPA, PRIA, SURAKARTA, 19, YA
12 LAIN, PRIA, SURAKARTA, 19, TIDAK
13 IPS, PRIA, LUAR, 17, TIDAK
14 LAIN, WANITA, SURAKARTA, 17, TIDAK
15 IPA, WANITA, LUAR, 18, YA
16 IPA, PRIA, SURAKARTA, 18, TIDAK
17 IPA, PRIA, SURAKARTA, 19, TIDAK
18 IPS, PRIA, LUAR, 18, TIDAK
19 LAIN, WANITA, SURAKARTA, 18, TIDAK
```

- Gunakan file arff pada tugas no 1 dalam modul 7 sebagai DataTraining, lakukan prediksi terhadap DataTesting.arff menggunakan weka



Hasil Prediksi metode Naive Bayes menggunakan Weka



Hasil di prediction menggunakan weka

ARFF-Viewer - C:\Users\Person\Documents\modul 8\tugas\hasilprediksi.arff

File Edit View

hasilprediksi.arff

Relation: Sekolah_predicted

No.	1: Jurusan_SMA Nominal	2: Gender Nominal	3: Asal_Sekolah Nominal	4: Rerata_SKS Numeric	5: prediction margin Numeric	6: predicted Asisten Nominal	7: Asisten Nominal
1	LAIN	WANITA	SURAKARTA	18.0	0.785601	TIDAK	TIDAK
2	IPA	PRIA	SURAKARTA	19.0	-0.033585	TIDAK	YA
3	LAIN	PRIA	SURAKARTA	19.0	0.684916	TIDAK	TIDAK
4	IPS	PRIA	LUAR	17.0	0.93308	TIDAK	TIDAK
5	LAIN	WANITA	SURAKARTA	17.0	0.973646	TIDAK	TIDAK
6	IPA	WANITA	LUAR	18.0	0.166955	YA	YA
7	IPA	PRIA	SURAKARTA	18.0	0.42833	TIDAK	TIDAK
8	IPA	PRIA	SURAKARTA	19.0	0.033585	TIDAK	TIDAK
9	IPS	PRIA	LUAR	18.0	0.525209	TIDAK	TIDAK
10	LAIN	WANITA	SURAKARTA	18.0	0.785601	TIDAK	TIDAK

3. Gunakan file excel pad tugas no1 dalam modul 6 sebagai data training, lakukan prediksi terhadap, data testing (excel) di atas menggunakan RapidMiner!
DataTraining.xls dan DataTraining.arff.

Sekolah - Excel

File Home Insert Page Layout Formulas Data Review View Help Tell me Share

Paste Font Alignment Number Conditional Formatting Format as Table Cells Editing

H14

	A	B	C	D	E	F	G
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Lama_Studi	
2	IPS	WANITA	SURAKARTA	18	TIDAK	TERLAMBAT	
3	IPA	PRIA	SURAKARTA	19	YA	TEPAT	
4	LAIN	PRIA	SURAKARTA	19	TIDAK	TERLAMBAT	
5	IPA	PRIA	LUAR	17	TIDAK	TERLAMBAT	
6	IPA	WANITA	SURAKARTA	17	TIDAK	TEPAT	
7	IPA	WANITA	LUAR	18	YA	TEPAT	
8	IPA	PRIA	SURAKARTA	18	TIDAK	TERLAMBAT	
9	IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT	
10	IPS	PRIA	LUAR	18	TIDAK	TERLAMBAT	IP
11	LAIN	WANITA	SURAKARTA	18	TIDAK	TEPAT	
12	IPA	WANITA	SURAKARTA	19	TIDAK	TEPAT	
13	IPS	PRIA	SURAKARTA	20	TIDAK	TEPAT	
14	IPS	PRIA	SURAKARTA	19	TIDAK	TEPAT	
15	IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT	
16	IPA	PRIA	LUAR	22	YA	TEPAT	
17	LAIN	PRIA	SURAKARTA	16	TIDAK	TERLAMBAT	
18	IPS	PRIA	LUAR	20	TIDAK	TEPAT	
19	LAIN	PRIA	LUAR	23	YA	TEPAT	
20	IPA	PRIA	SURAKARTA	21	YA	TEPAT	
21	IPS	PRIA	SURAKARTA	19	TIDAK	TERLAMBAT	
22							
23							

Sheet1 Sheet2

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Open file... Open UR... Open DB... Generat... Undo Edit... Save...

Filter Choose None Apply

Current relation: Relation: Sekolah Instances: 20 Attributes: 6 Sum of weights: 20

Selected attribute: Name: Jurusan_SMA Type: Nominal Missing: 0 (0%) Distinct: 3 Unique: 0 (0%)

No.	Label	Count	Weight
1	IPS	6	6.0
2	IPA	10	10.0
3	LAIN	4	4.0

Attributes: All None Invert Pattern

No. Name

1	Jurusan_SMA
2	Gender
3	Asal_Sekolah
4	Rerata_SKS
5	Asisten
6	Lama_Studi

Class: Lama_Studi (Nom) Visualize All

Status OK Log x 0

11:54 PM

Prediksi menggunakan RapidMiner.

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>	Rearata_SKS <i>integer</i>	Asisten <i>binominal</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.

Where to store the data?

Local Repository (LABSI-07)

- Connections (LABSI-07)
- data (LABSI-07)
- processes (LABSI-07)
 - DataCuaca_Testing (LABSI-07 - v1, 10/9/19 11:25 PM - 359 bytes)
 - DataCuaca_Training (LABSI-07 - v1, 10/9/19 11:19 PM - 529 bytes)
 - DataSekolah_Training (LABSI-07 - v1, 10/10/19 12:06 AM - 675 bytes)

Name

Location //Local Repository/DataSekolah_Testing

Previous Finish Cancel

Activate Wisdom of Crowds

4. Hasil percobaan Tugas no 3, berapakah nilai rerata confidence untuk atribut Lama_Studi, dengan nilai TEPAT? Dan berapakah nilai rerata confidence untuk atribut lama_studi dengan nilai TERLAMBAT?

Result History		ExampleSet (Apply Model)							
Open in		Turbo Prep		Auto Model		Filter (10 / 10 examples): all			
Ro...	prediction(L...	confidence(TERLAMBAT)	confidence(TEPAT)	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)

- Nilai Rerata Confidence untuk attribute Lama_Studi dengan nilai TEPAT:

Confidence_TEPAT	Real	0	Min	Max	Average
confidence(TEPAT)	Real	0	0.132	0.995	0.476

- Nilai Rerata Confidence untuk attribute Lama_Studi dengan nilai TERLAMBAT:

Confidence_TERLAMBAT	Real	0	Min	Max	Average
confidence(TERLAMBAT)	Real	0	0.005	0.868	0.524

5. Dari hasil percobaan tugas no 3 berapa jumlah banyaknya orang yang akan lulus TEPAT, dan berapa jumlah banyaknya orang yang akan lulus TERLAMBAT

- Jumlah banyaknya Orang yang akan lulus TEPAT :

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

- Jumlah banyaknya Orang yang akan lulus TERLAMBAT :

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

6. Prediksikan 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

- Memilih Nama dewi untuk di prediksi

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Testing Cell range: A1:F2 Select All Define header row: 1

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

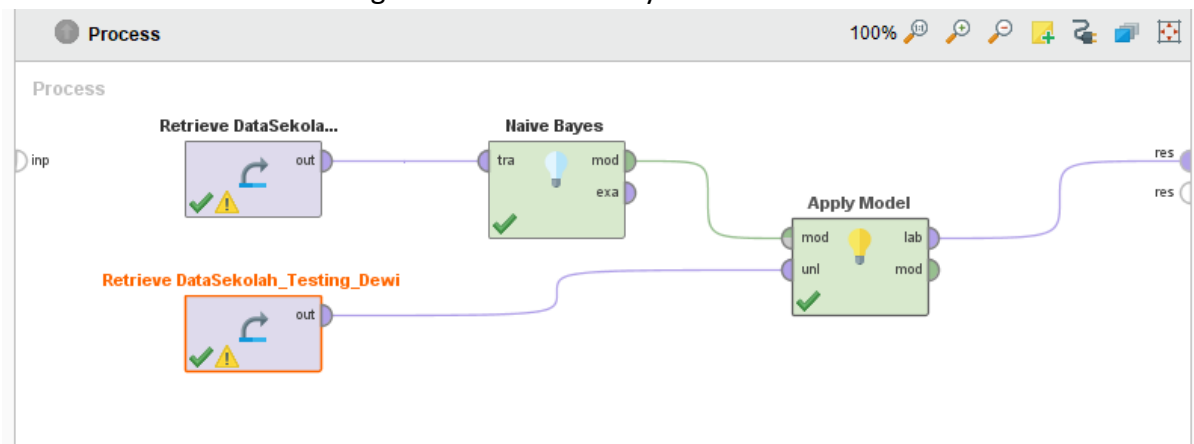
- Tidak ada yang di ubah dan di kasih label karena ini data testing

Format your columns.

☐ Replace errors with missing values ⓘ

	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Nama
	polynomial	polynomial	polynomial	integer	polynomial	polynomial
1	IPA	WANITA	LUAR	18	TIDAK	Dewi

- Melakukan model dengan metode naive bayes



- Hasil ketepatan Lama_Studi Dewi adalah (TEPAT) tidak TERLAMBAT

Row No.	prediction(Lama_Studi)	confidence(TERLAMBAT)	confidence(TEPAT)	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Nama
1	TEPAT	0.298	0.702	IPA	WANITA	LUAR	18	TIDAK	Dewi

- Prediksikan ketepatan studi si Jono, jika jono adalah PRIA jurusan LAIN pada SMA, asal sekolah dari SURAKARTA, mengambil sks rata rata sebanyak 17 SKS, dan pernah menjadi Asisten

- Memilih nama jono untuk di prediksi

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Testing Cell range: A1:F2 Select All ☒ Define header row: 1

	A	B	C	D	E	F
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Nama
2	Lain	PRIA	SURAKARTA	17.000	YA	JONO
3		PRIA	SURAKARTA	19.000	YA	
4	LAIN	PRIA	SURAKARTA	19.000	TIDAK	
5	IPS	PRIA	LUAR	17.000	TIDAK	

- Tidak ada yang di ubah dan di kasih label, karena ini data testing

Import Data - Format your columns.

Format your columns.

☐ Replace errors with missing values ⓘ

	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Nama
	<i>polynomial</i>	<i>polynomial</i>	<i>polynomial</i>	<i>integer</i>	<i>polynomial</i>	<i>polynomial</i>
1	Lain	PRIA	SURAKARTA	17	YA	JONO

Views: Design Results Turbo Prep Auto Model Deployments

Import Data - Where to store the data?

Where to store the data?

Local Repository (OKI KUS MAHESA)

data (OKI KUS MAHESA)

processes (OKI KUS MAHESA)

DataSekolah_Testing (OKI KUS MAHESA - v1, 10/11/19 9:45 PM - 485 bytes)

DataSekolah_Testing_Dewi (OKI KUS MAHESA - v1, 10/11/19 10:11 PM - 415 bytes)

DataSekolah_Training (OKI KUS MAHESA - v1, 10/11/19 9:44 PM - 675 bytes)

Name

DataSekolah_Testing_Jono

Location

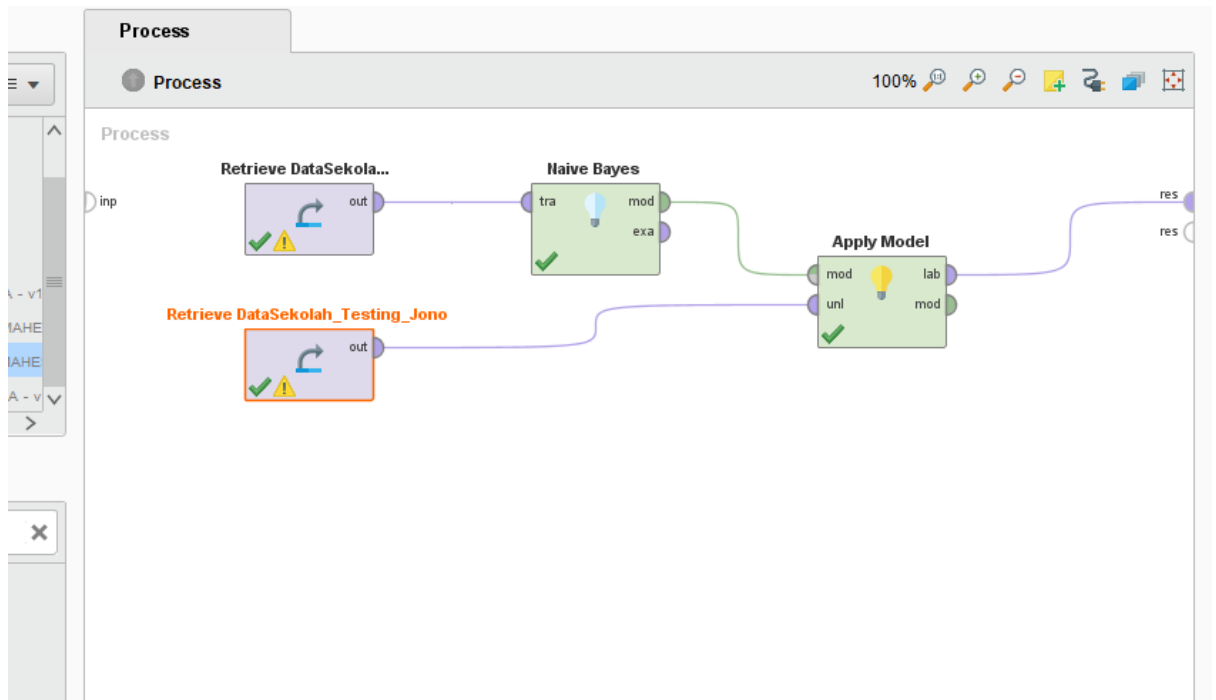
//Local Repository/DataSekolah_Testing_Jono

Previous

Finish

Cancel

- Melakukan model dengan metode naive bayes



- Hasil ketepatan Lama_Studi Jono adalah (TEPAT) tidak TERLAMBAT

ExampleSet (/Local Repository/DataSekolah_Testing)

ExampleSet (/Local Repository/DataSekolah_Testing)

ExampleSet (/Local Repository/DataSekolah_Testing_MilikDewi)

ExampleSet (/Local Repository/DataSekolah_Testing_Dewi)

ExampleSet (/Local Repository/DataSekolah_Training)

ExampleSet (/Local Repository/DataSekolah_Training)

ExampleSet (/Local Repository/DataSekolah_Testing_Dewi)

ExampleSet (/Local Repository/DataSekolah_Testing_Jono)

Result History

ExampleSet (Apply Model)

ExampleSet (/Local Repository/DataSekolah_Testing_Jono)

Open in

Turbo Prep

Auto Model

Filter (1 / 1 examples): all

Row No.	prediction(Lama_Studi)	confidence(TERLAMBAT)	confidence(TEPAT)	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Nama
1	TEPAT	0.076	0.924	Lain	PRIA	SURAKARTA	17	YA	JONO