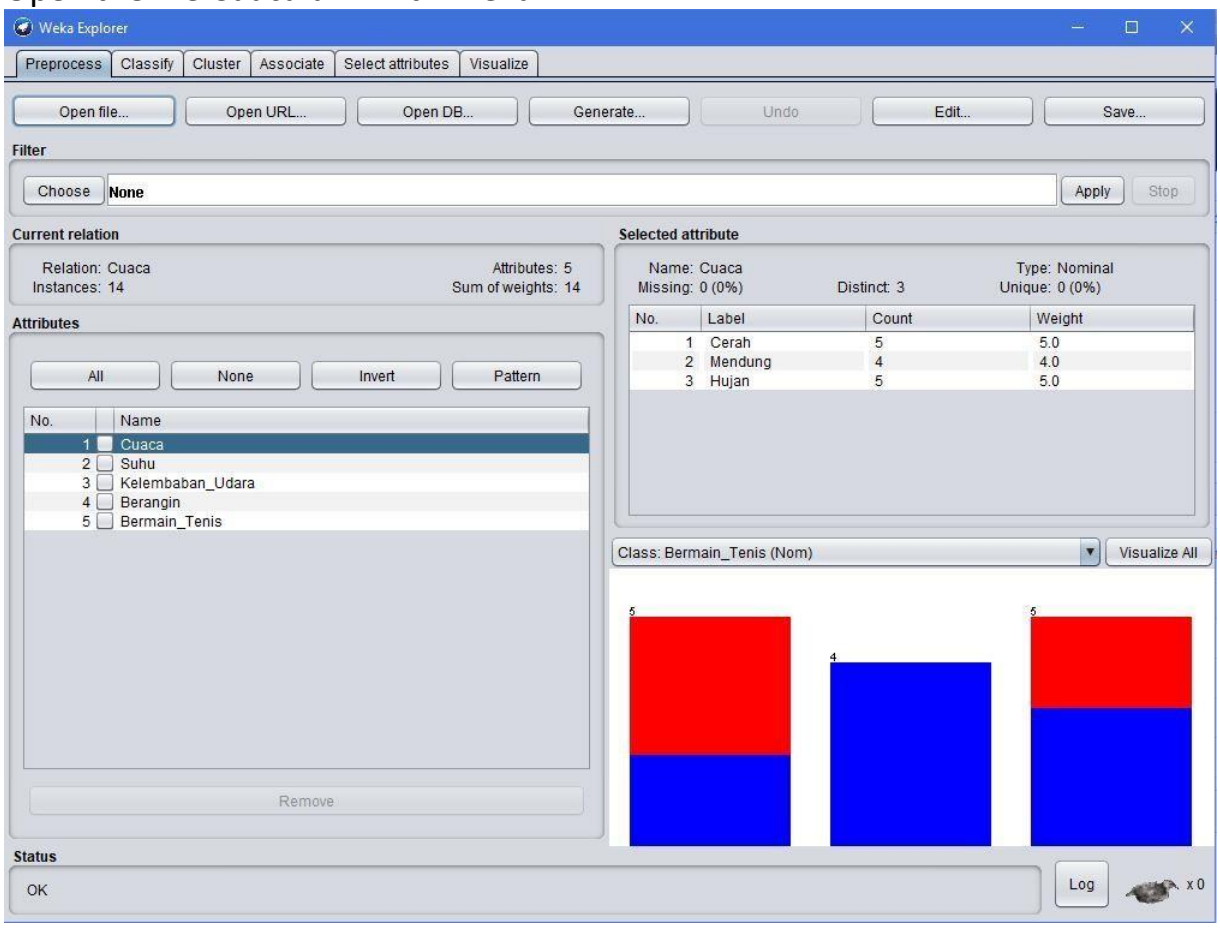
Name : Ridho Achmad F
NIM : L200174005
Class : X

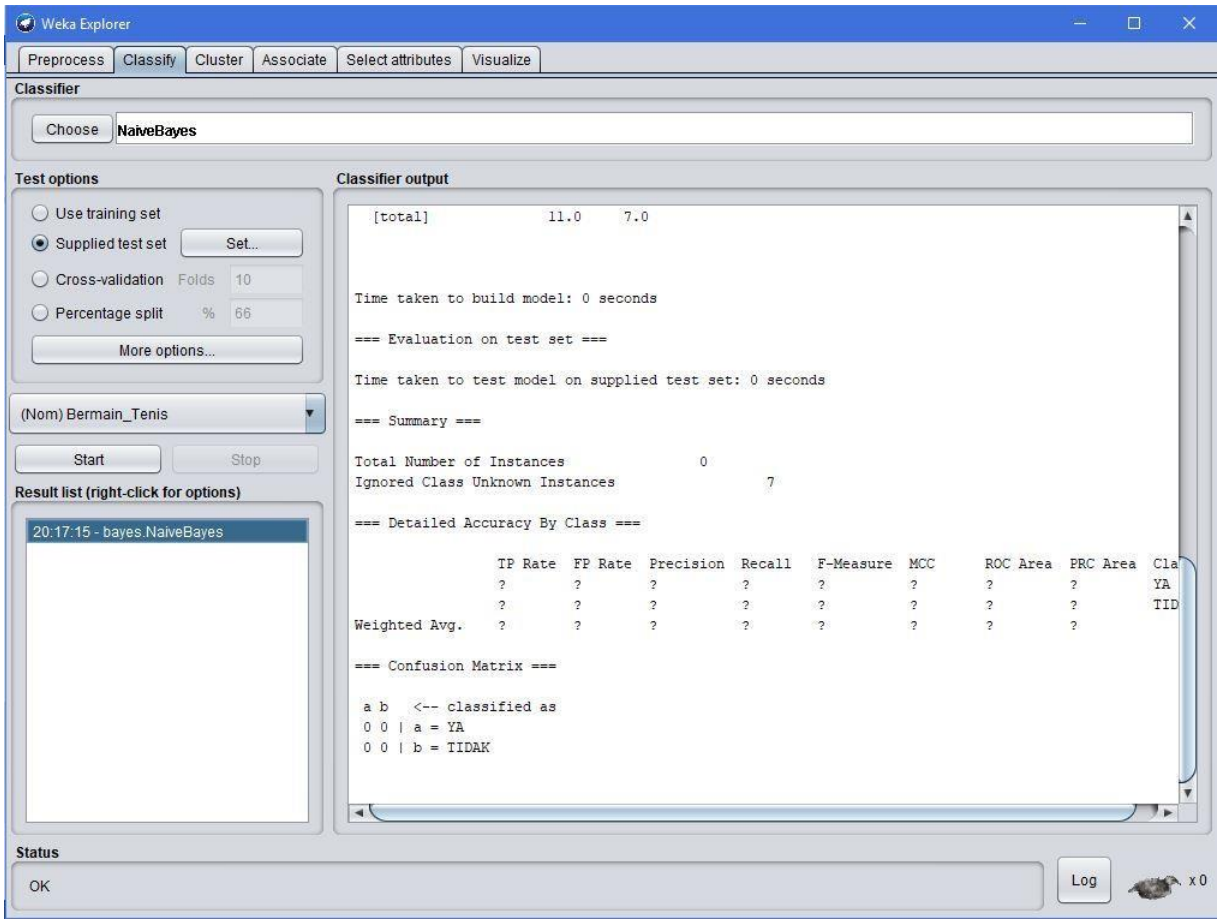
Practicum Report Module 8

Practicum

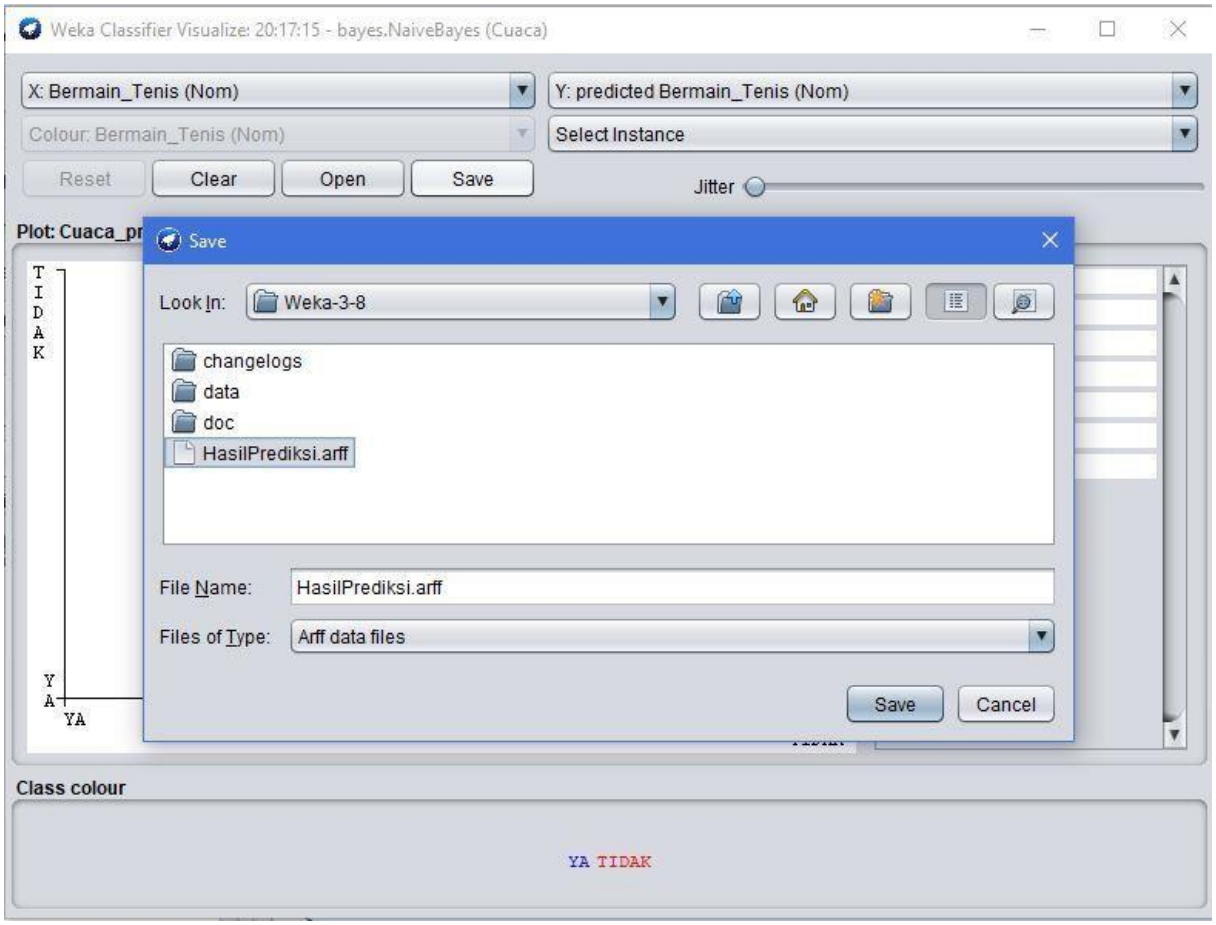
1. Implementation of Naive Bayes with Weka
   a. Open the file Cuaca.arff with Weka



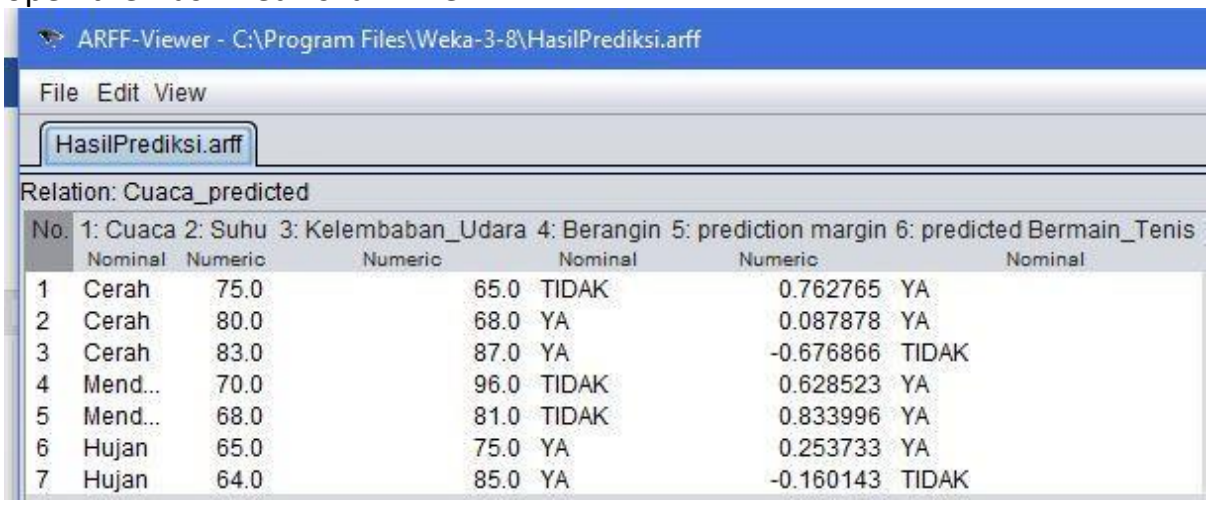
   b. Choose the Naive Bayes method and open the CuacaTesting.arff file.



c. After that save the data with the name HasilPrediksi.arff.



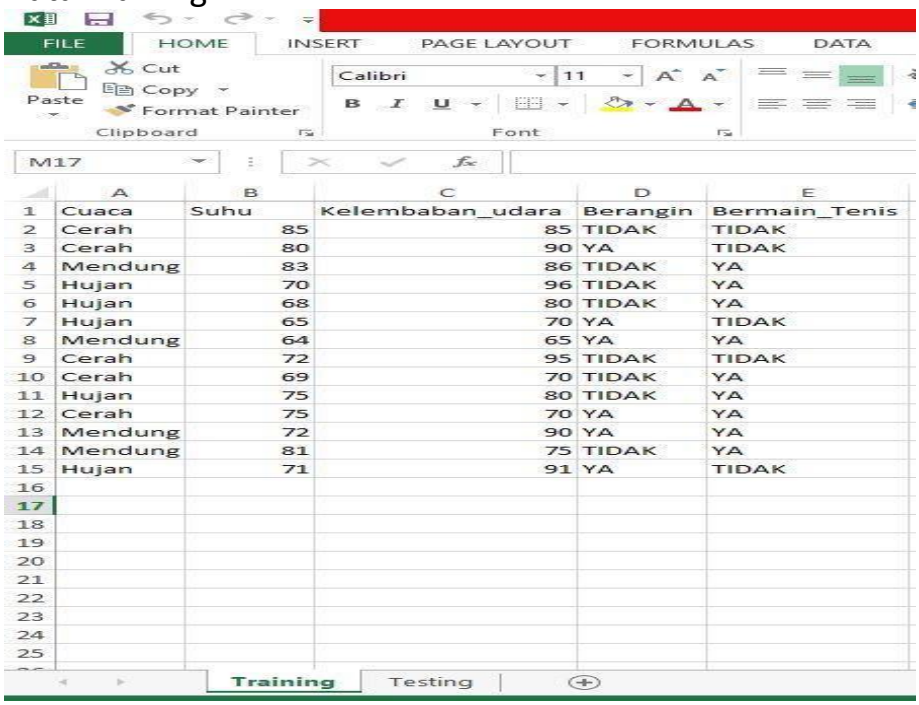
d. Back in the Weka GUI Chooser select the Tools menu - ArffViewer and open the HasilPrediksi.arff file.



No.	1: Cuaca	2: Suhu	3: Kelembaban_Udara	4: Berangin	5: prediction margin	6: predicted Bermain_Tenis
	Nominal	Numeric	Numeric	Nominal	Numeric	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	Mend...	70.0	96.0	TIDAK	0.628523	YA
5	Mend...	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

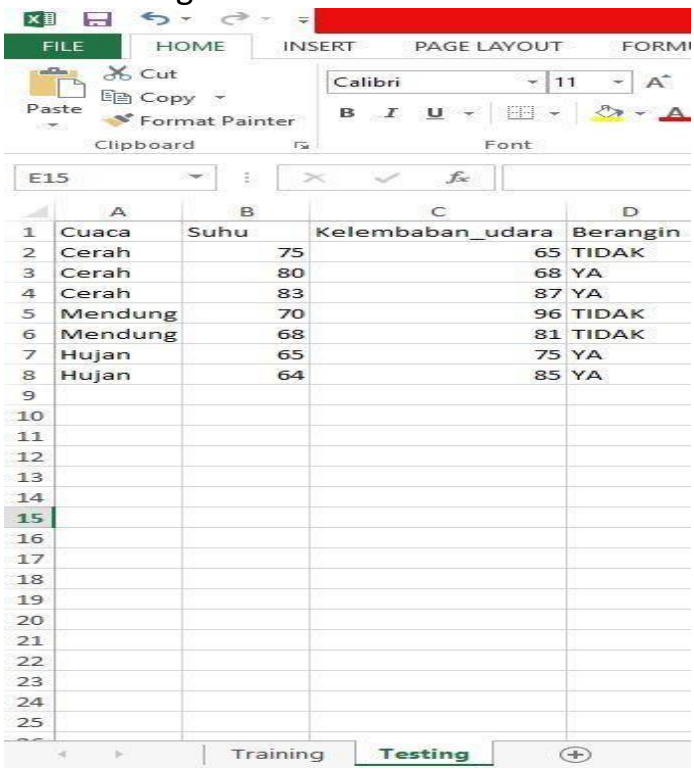
2. Implementation of Naive Bayes with RapidMiner
   a. Prepare the file Table_Cuaca.xls which contains 2 sheets, namely for data training and data testing.

Data Training:



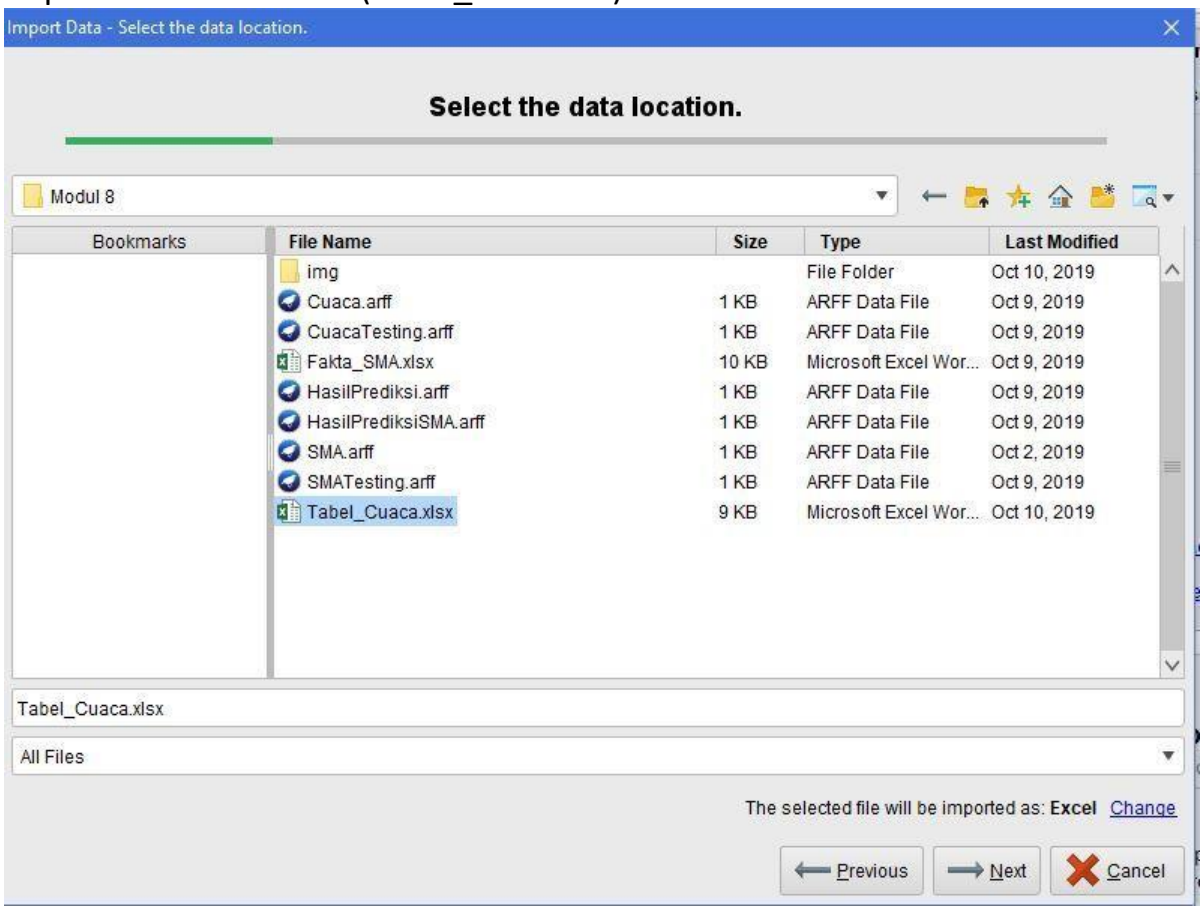
	A	B	C	D	E
1	Cuaca	Suhu	Kelembaban_udara	Berangin	Bermain_Tenis
2	Cerah	85	85	TIDAK	TIDAK
3	Cerah	80	90	YA	TIDAK
4	Mendung	83	86	TIDAK	YA
5	Hujan	70	96	TIDAK	YA
6	Hujan	68	80	TIDAK	YA
7	Hujan	65	70	YA	TIDAK
8	Mendung	64	65	YA	YA
9	Cerah	72	95	TIDAK	TIDAK
10	Cerah	69	70	TIDAK	YA
11	Hujan	75	80	TIDAK	YA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

Data Testing:

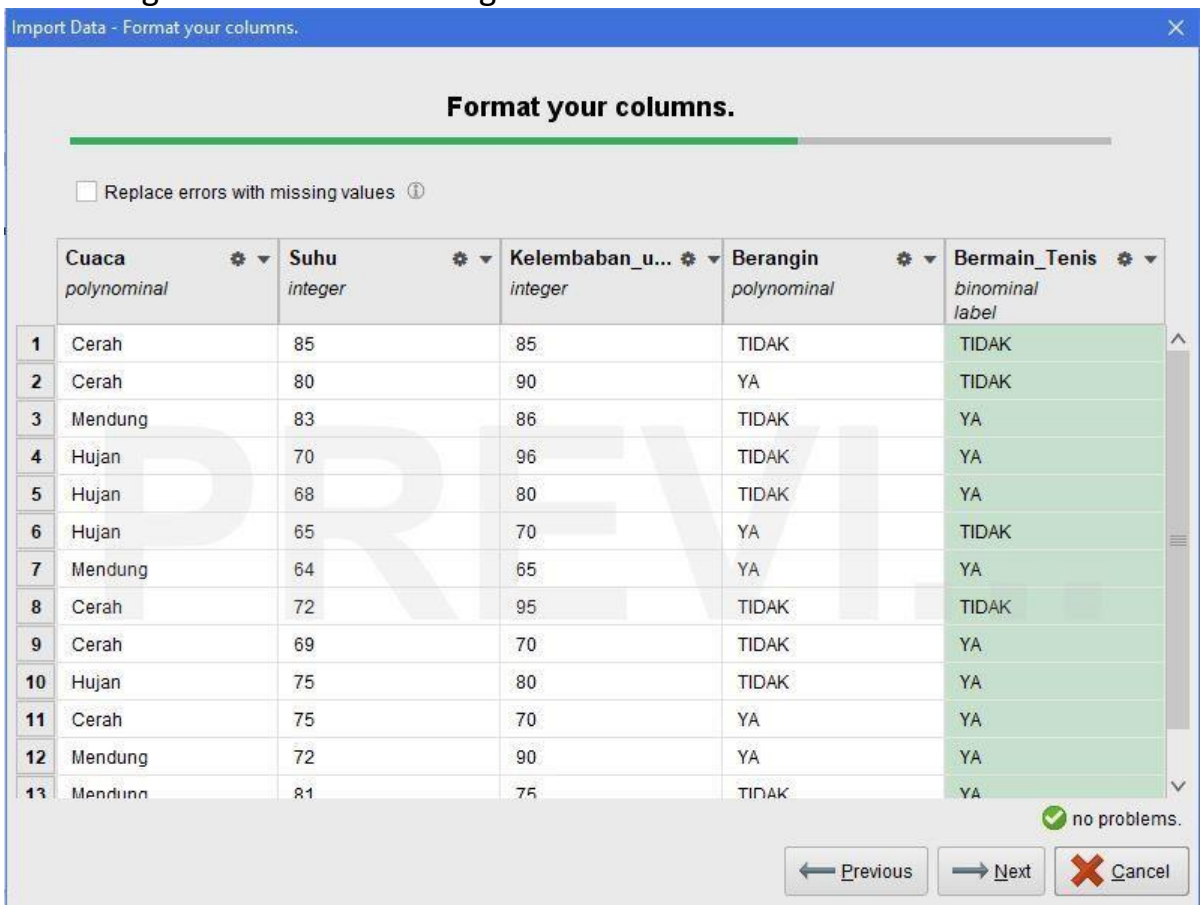


	A	B	C	D
1	Cuaca	Suhu	Kelembaban_udara	Berangin
2	Cerah	75	65	TIDAK
3	Cerah	80	68	YA
4	Cerah	83	87	YA
5	Mendung	70	96	TIDAK
6	Mendung	68	81	TIDAK
7	Hujan	65	75	YA
8	Hujan	64	85	YA

b. Import data to be used (Table_Cuaca.xls).

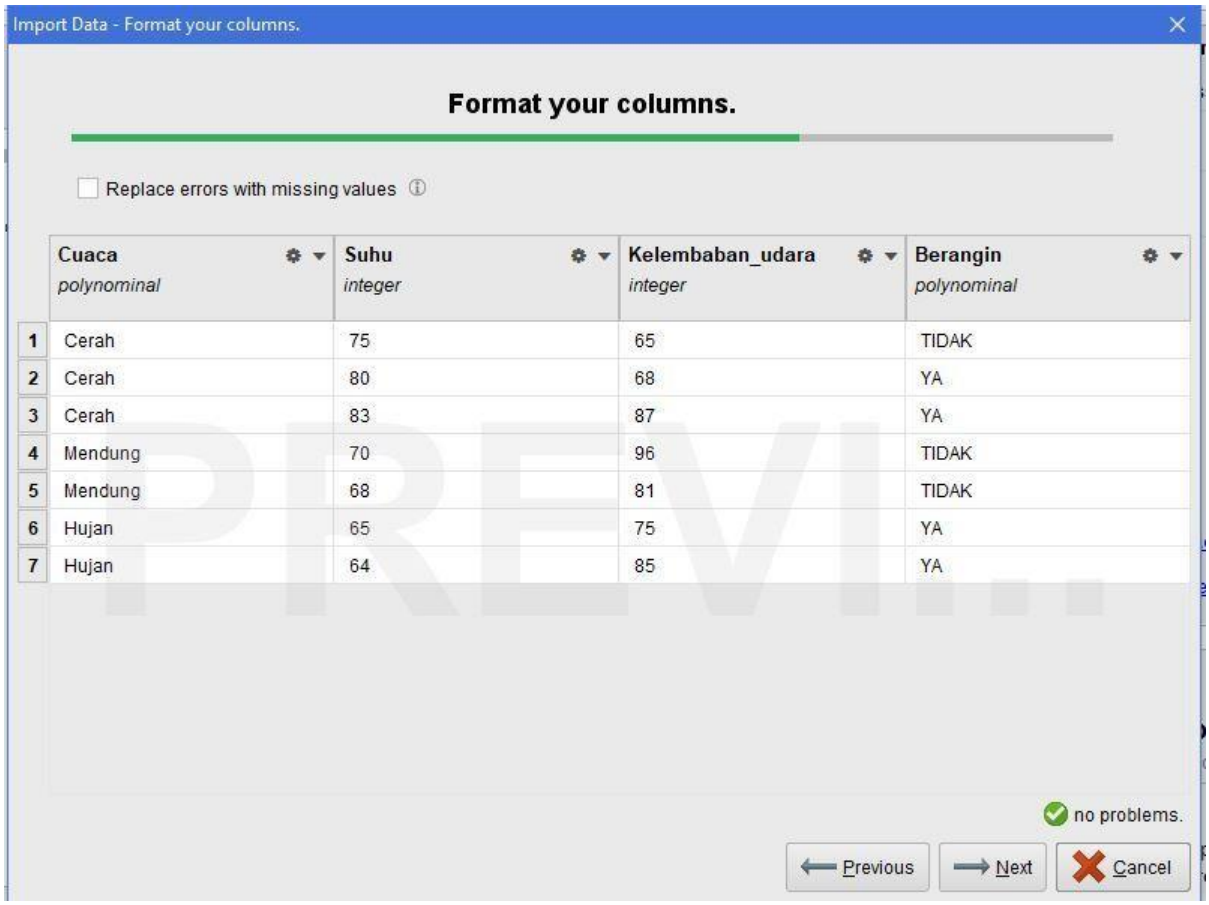


c. In the training data, change the Bermain_Tenis column to binomial type and change it as a label in Change Role.



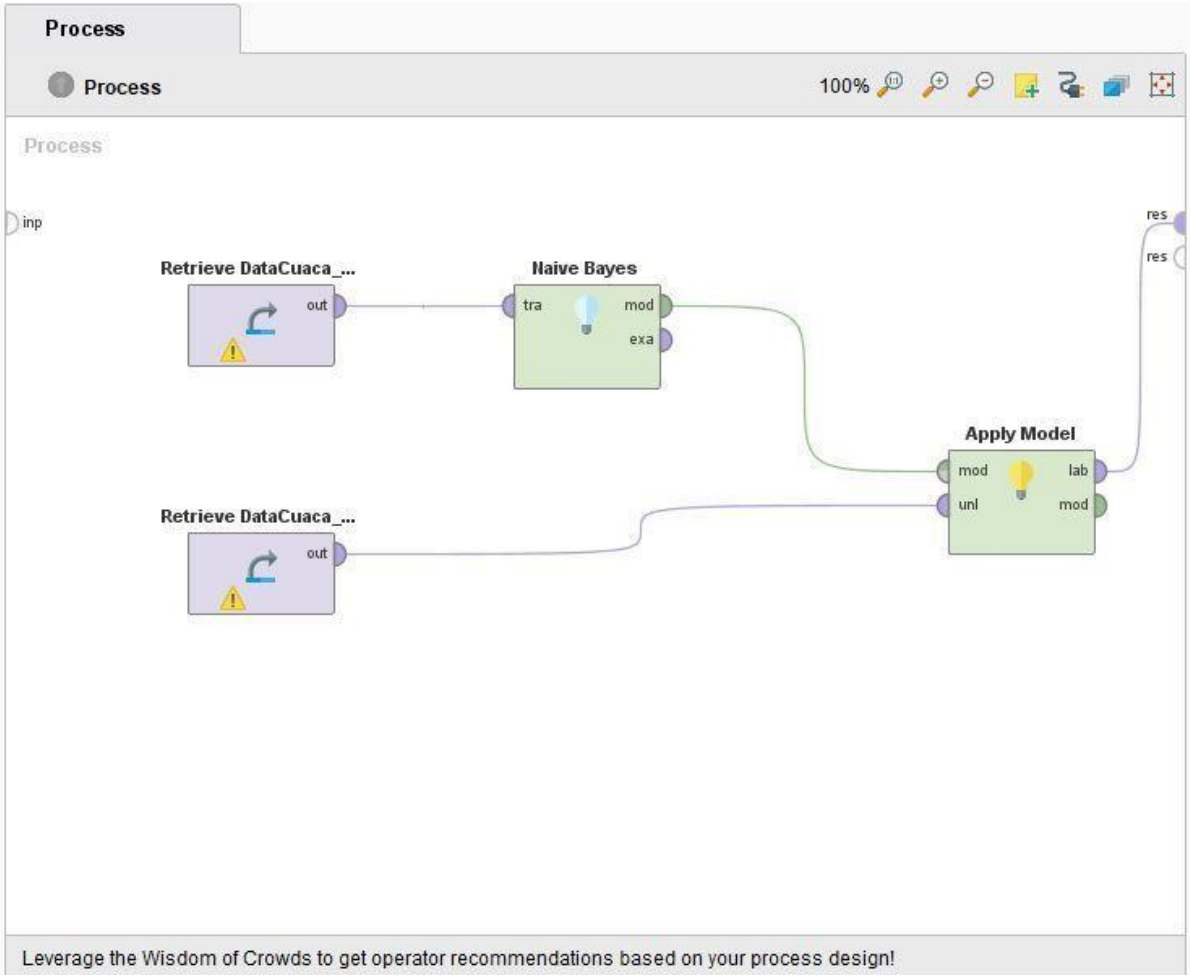
	Cuaca *polynominal*	Suhu *integer*	Kelembaban_u... *integer*	Berangin *polynominal*	Bermain_Tenis *binominal label*
1	Cerah	85	85	TIDAK	TIDAK
2	Cerah	80	90	YA	TIDAK
3	Mendung	83	86	TIDAK	YA
4	Hujan	70	96	TIDAK	YA
5	Hujan	68	80	TIDAK	YA
6	Hujan	65	70	YA	TIDAK
7	Mendung	64	65	YA	YA
8	Cerah	72	95	TIDAK	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

d. Save the data with the name DataCuaca_Training after that, do the same with the data testing but do not make changes to the type and label.



e. After that, save it with the file name DataCuaca_Testing.

f. Next create a Naive Bayes design using Naive Bayes and Apply Models into the process view and connect each data connector to the operator.



g. Run the naive bayes process and we can see the results of the data and statistics.

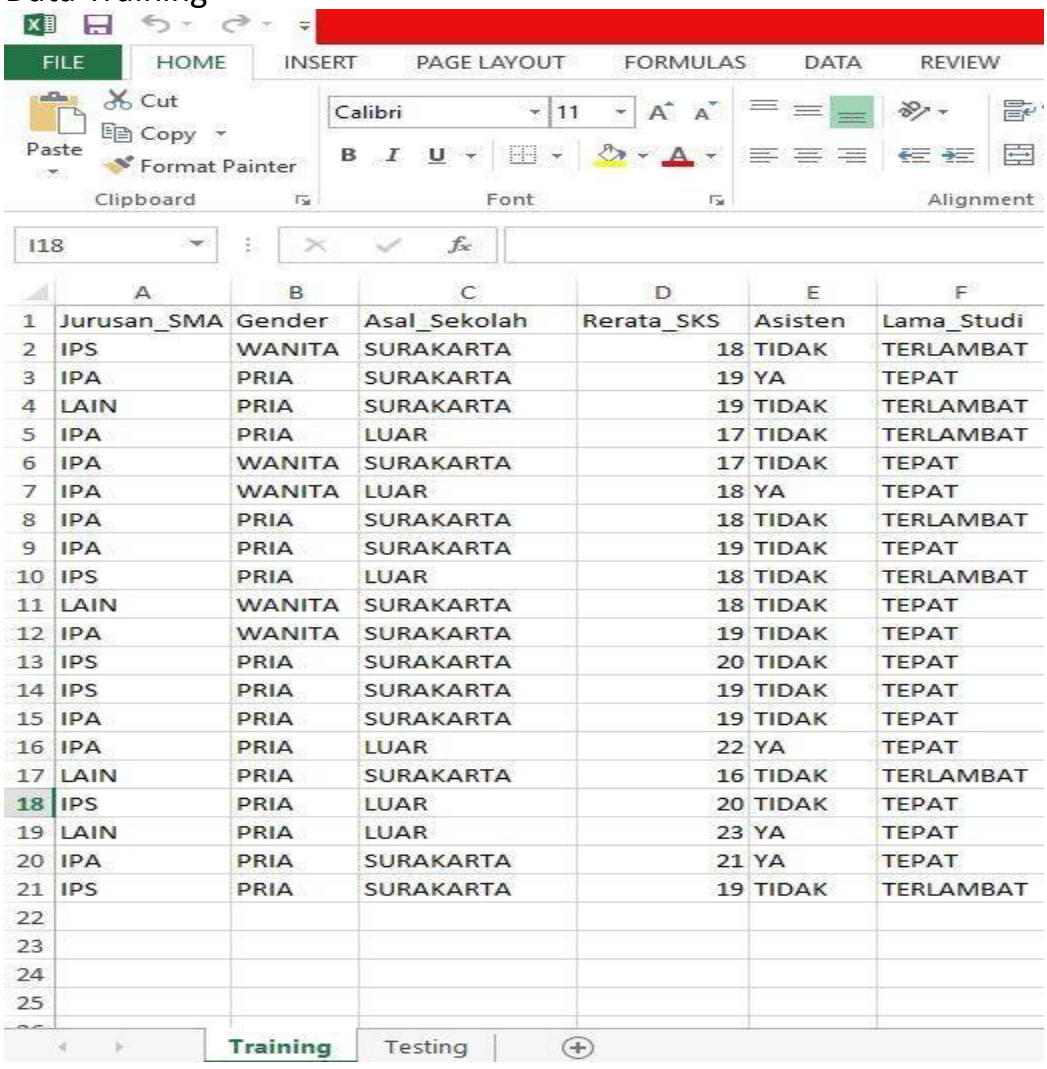
Open in Turbo Prep Auto Model Filter (7 / 7 examples): all

Row No.	prediction(B...	confidence(...	confidence(...	Cuaca	Suhu	Kelembaban...	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

Name	Type	Missing	Statistics		Filter (7 / 7 attributes): Search for Attributes
Prediction **prediction(Bermain_Tenis)**	Binominal	0	Least TIDAK (2)	Most YA (5)	Values YA (5), TIDAK (2)
Confidence_TIDAK **confidence(TIDAK)**	Real	0	Min 0.007	Max 0.856	Average 0.353
Confidence_YA **confidence(YA)**	Real	0	Min 0.144	Max 0.993	Average 0.647
Cuaca	Polynominal	0	Least Mendung (2)	Most Cerah (3)	Values Cerah (3), Hujan (2), ...[1 more]
Suhu	Integer	0	Min 64	Max 83	Average 72.143
Kelembaban_udara	Integer	0	Min 65	Max 96	Average 79.571
Berangin	Polynominal	0	Least TIDAK (3)	Most YA (4)	Values YA (4), TIDAK (3)

Assignment

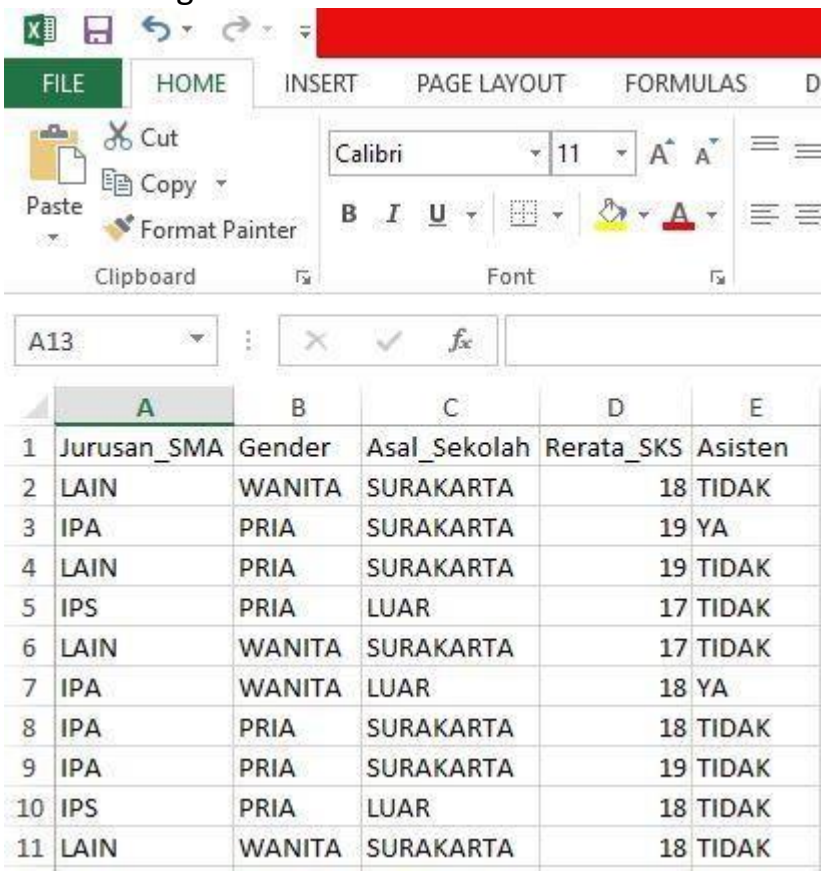
1. Data Training



	A	B	C	D	E	F
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
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						
24						
25						

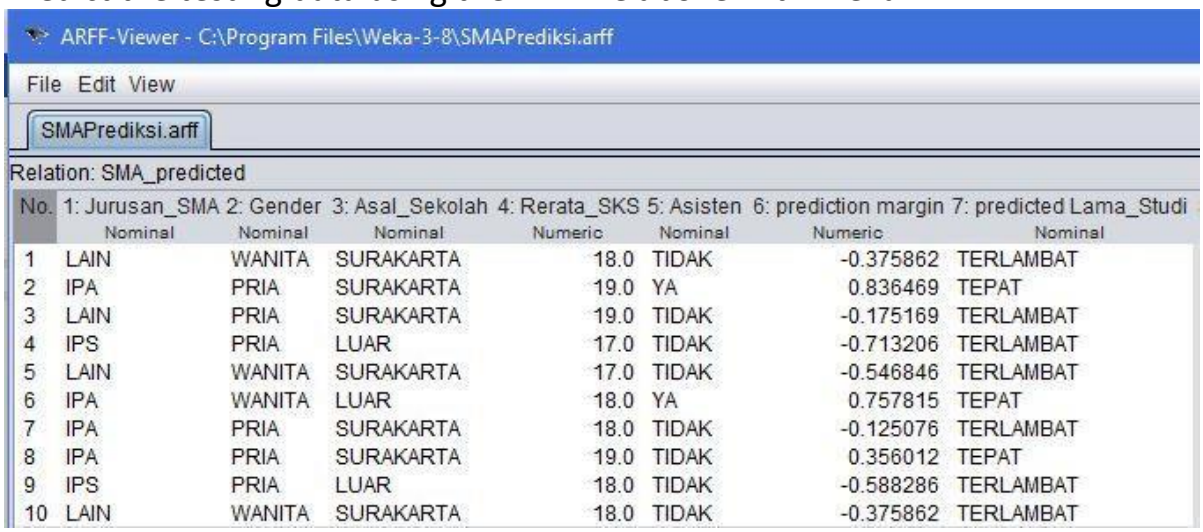
Training Testing

Data Testing



	A	B	C	D	E
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	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

2. Predict the testing data using the Arff file above with Weka



ARFF-Viewer - C:\Program Files\Weka-3-8\SMAPrediksi.arff

File Edit View

SMAPrediksi.arff

Relation: SMA_predicted

No.	1: Jurusan_SMA	2: Gender	3: Asal_Sekolah	4: Rerata_SKS	5: Asisten	6: prediction margin	7: predicted Lama_Studi
	Nominal	Nominal	Nominal	Numeric	Nominal	Numeric	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

3. Predict the testing data using the Excel file above with Rapidminer.

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

Row No.	prediction(L...	confidence(...	confidence(...	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

Name | Type | Missing | Statistics | Filter (8 / 8 attributes): Search for Attributes

	Type	Missing	Statistics		
Prediction **prediction(Lama_Studi)**	Binominal	0	Least TEPAT (3)	Most TERLAMBAT (7)	Values TERLAMBAT (7), TEPAT (3)
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
Jurusan_SMA	Polynominal	0	Least IPS (2)	Most IPA (4)	Values IPA (4), LAIN (4), ...[1 more]
Gender	Polynominal	0	Least WANITA (4)	Most PRIA (6)	Values PRIA (6), WANITA (4)
Asal_Sekolah	Polynominal	0	Least LUAR (3)	Most SURAKARTA (7)	Values SURAKARTA (7), LUAR (3)
Rerata_SKS	Integer	0	Min 17	Max 19	Average 18.100
Asisten	Polynominal	0	Least YA (2)	Most TIDAK (8)	Values TIDAK (8), YA (2)

4. The average confidence value for lama_studi with the TEPAT value is 0,476

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

Average confidence value for lama_studi with TERLAMBAT value is 0,524

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

5. The people who will graduate TEPAT are 3 while those who will graduate TERLAMBAT are 7.

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

6. Data prediction of the Dewi and Jono.

11	TEPAT	0.298	0.702	IPA	WANITA	LUAR	18	TIDAK
12	TEPAT	0.076	0.924	LAIN	PRIA	SURAKARTA	17	YA