

Nama : Angieta Putri Wahendra

NIM : L200170096

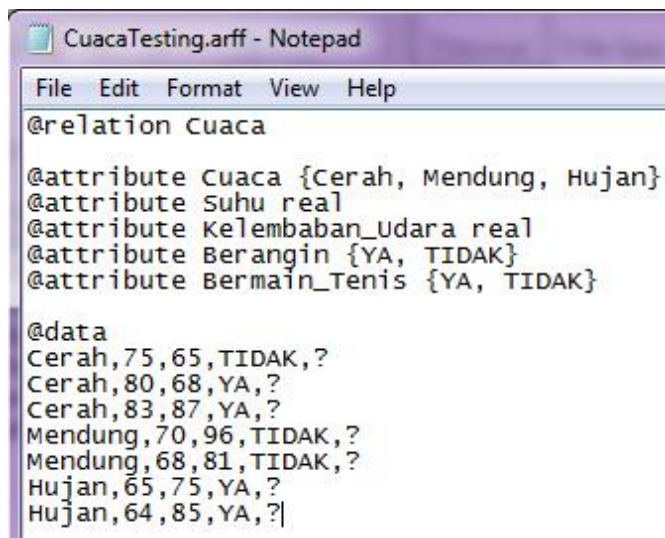
Kelas : D

## KEGIATAN

### 1. Persiapkan file Cuaca.arff :

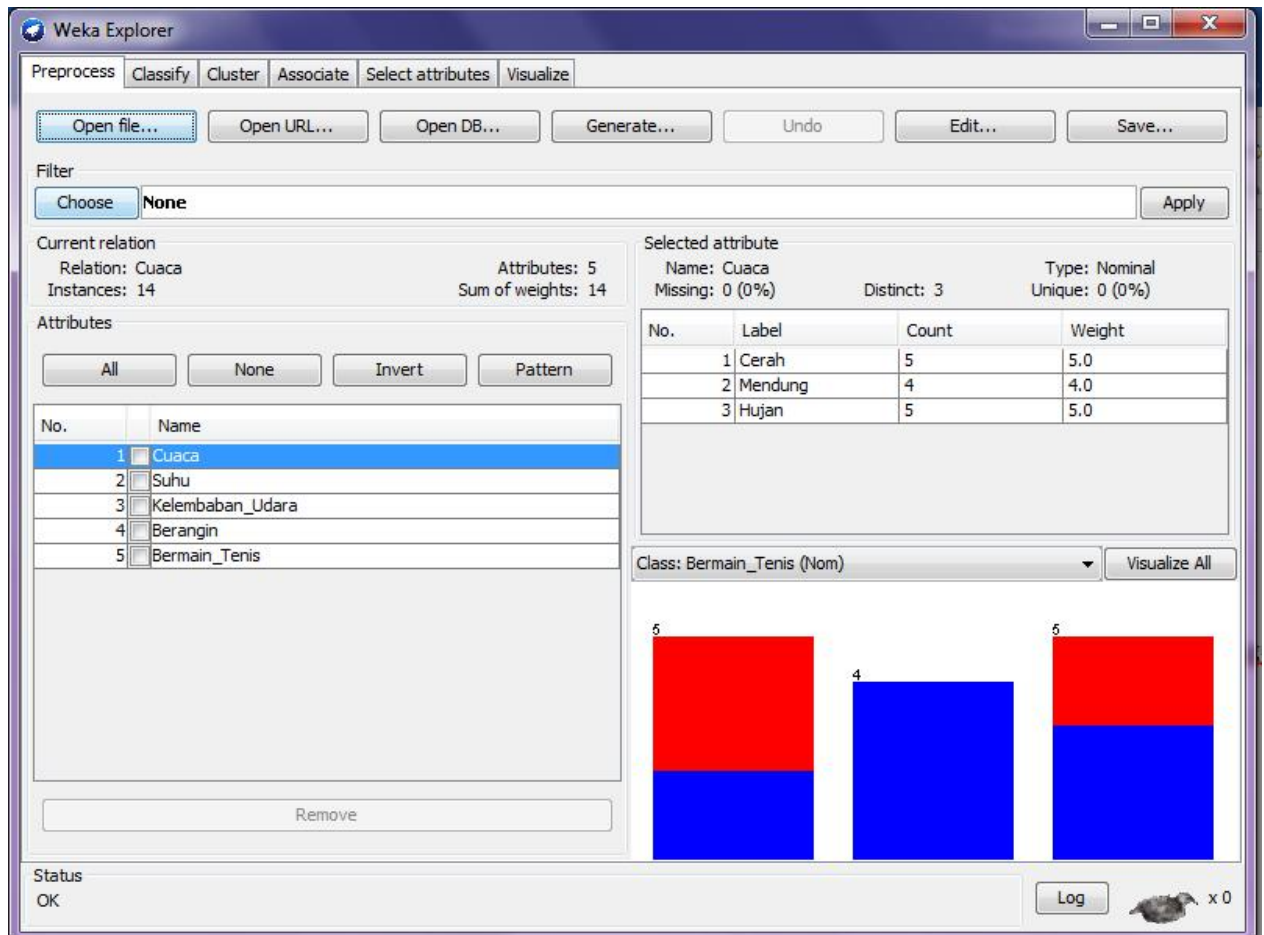
```
1 @relation Cuaca
2
3 @attribute Cuaca {Cerah, Mendung, Hujan}
4 @attribute Suhu real
5 @attribute Kelembaban_Udara real
6 @attribute Berangin {YA, TIDAK}
7 @attribute Bermain_Tenis {YA, TIDAK}
8
9 @data
10 Cerah,85,85,TIDAK,TIDAK
11 Cerah,80,90,YA,TIDAK
12 Mendung,83,86,TIDAK,YA
13 Hujan,70,96,TIDAK,YA
14 Hujan,68,80,TIDAK,YA
15 Hujan,65,70,YA,TIDAK
16 Mendung,64,65,YA,YA
17 Cerah,72,95,TIDAK,TIDAK
18 Cerah,69,70,TIDAK,YA
19 Hujan,75,80,TIDAK,YA
20 Cerah,75,70,YA,YA
21 Mendung,72,90,YA,YA
22 Mendung,81,75,TIDAK,YA
23 Hujan,71,91,YA,TIDAK
```

### 2. Buatlah seperti pada gambar dan save dengan nama CuacaTesting.arff :



```
CuacaTesting.arff - Notepad
File Edit Format View Help
@relation Cuaca
@attribute Cuaca {Cerah, Mendung, Hujan}
@attribute Suhu real
@attribute Kelembaban_Udara real
@attribute Berangin {YA, TIDAK}
@attribute Bermain_Tenis {YA, TIDAK}
@data
Cerah,75,65,TIDAK,?
Cerah,80,68,YA,?
Cerah,83,87,YA,?
Mendung,70,96,TIDAK,?
Mendung,68,81,TIDAK,?
Hujan,65,75,YA,?
Hujan,64,85,YA,?
```

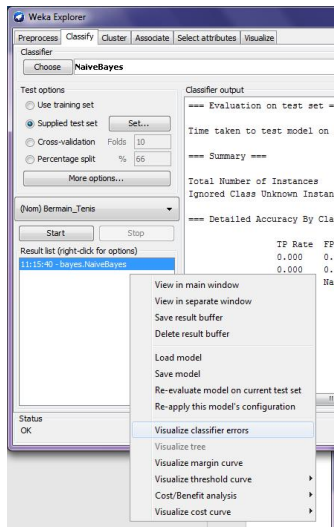
3. Buka kembali file Cuaca.arff dengan menggunakan Weka Explorer. Buka Weka Explorer – Klik Open File – Pilih Cuaca.arff



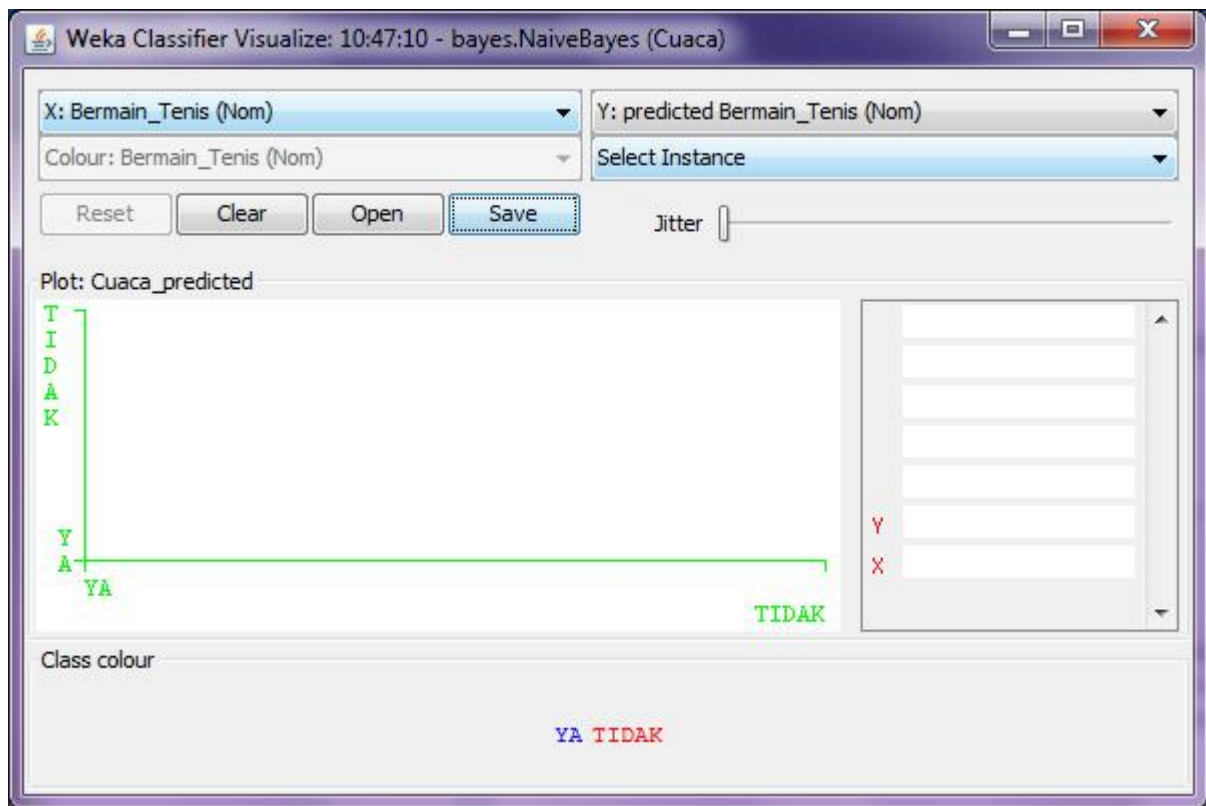
4. Masih pada Weka Explorer, pilih tab **Classify**
5. Sehingga akan muncul jendela Weka Explorer pada tab Classify, pada kotak **Classifier** klik tombol **Choose** untuk memilih metode/algorithm **Naïve Bayes**
6. Klik tombol Set untuk menentukan file ARFF sebagai data uji
7. Sehingga akan muncul jendela Test Instance, klik **Open file . . . .**



8. Pilih file **CuacaTesting.arff**
9. Kemudian klik **Close**
10. Klik **Start** untuk memulai proses naïve bayes
11. Klik kanan pada hasil proses dalam kotak **result list**. Pilih **Visual classifier errors**



12. Klik **Save**. Simpan dengan nama file **HasilPrediksi.arff**



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

14. Buka menu File-Open. Tunjukkan pada file HasilPrediksi.arff yang telah Anda simpan

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	

## Implementasi Naïve Bayes dengan RapidMiner

1. Persiapkan file **Tabel\_Cuaca.xls** yang terdiri dari 2 sheet.
2. **Tabel data training** pada Sheet1

	A	B	C	D	E
	<b>Cuaca</b>	<b>Suhu</b>	<b>Kelembapan Udara</b>	<b>Berangin</b>	<b>Bermain_Tenis</b>
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
14	Hujan	71	91	YA	TIDAK

Training Testing (+)

3. Tabel data uji pada Sheet2 tanpa ada variabel **Bermain\_Tenis**

	A	B	C	D
	<b>Cuaca</b>	<b>Suhu</b>	<b>Kelembapan_udara</b>	<b>Berangin</b>
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

Training Testing (+)

4. Buka aplikasi Rapid Miner
5. Klik **Import Data**. Arahkan direktori tempat penyimpanan file pada langkah **Select data location**, kemudian pilih file yang akan digunakan dan klik **Next**.
6. Pastikan sel Excel sesuai di langkah **Select the cells to import**.

**Select the cells to import.**

---

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

	A	B	C	D	E	F
1	Cuaca	Suhu	Kelembaban_Udara	Berangin	Bermain_Tenis	
2	Cerah	85.000	85.000	TIDAK	TIDAK	
3	Cerah	80.000	90.000	YA	TIDAK	
4	Mendung	83.000	86.000	TIDAK	YA	
5	Hujan	70.000	96.000	TIDAK	YA	
6	Hujan	68.000	80.000	TIDAK	YA	
7	Hujan	65.000	70.000	YA	TIDAK	
8	Mendung	64.000	65.000	YA	YA	
9	Cerah	72.000	95.000	TIDAK	TIDAK	
10	Cerah	69.000	70.000	TIDAK	YA	
11	Hujan	75.000	80.000	TIDAK	YA	
12	Cerah	75.000	70.000	YA	YA	
13	Mendung	72.000	90.000	YA	YA	
14	Mendung	81.000	75.000	TIDAK	YA	
15	Hujan	71.000	91.000	YA	TIDAK	

Previous Next Cancel

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

**Format your columns.**

---

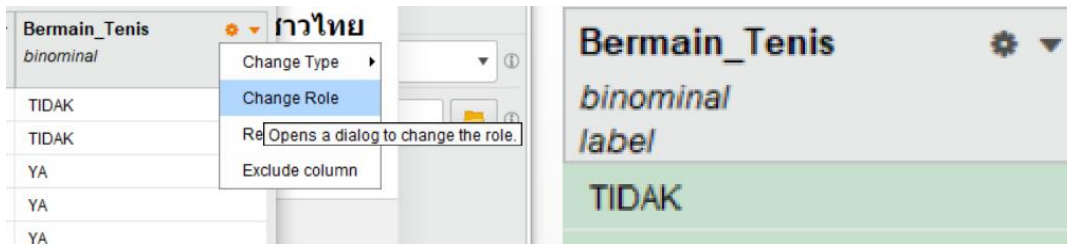
☐ Replace errors with missing values

	Cuaca <i>polynomial</i>	Suhu <i>integer</i>	Kelembaban_Udara <i>integer</i>	Berangin <i>polynomial</i>	Bermain_Tenis <i>polynomial</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

Change Type ☒ polynomial  
 Change Role ☒ binomial  
 Rename column  
 Exclude column  
 real  
 integer  
 date\_time  
 date  
 time



8. Ubah pula sebagai **label** pada **Change Role**



9. Simpan dengan nama **DataCuaca\_Training** dilanjutkan klik tombol **Finish**

10. Hasil import file **Tabel\_Cuaca.xls** pada Sheet1 akan di tampilkan.

ExampleSet (//Local Repository/DataCuaca\_Training)

Open in Turbo Prep Auto Model Filter (14 / 14 examples): all

Row No.	Bermain_Tenis	Cuaca	Suhu	Kelembaban_Udara	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 dari langkah 5

**Select the cells to import.**



---

Sheet: **Testing** Cell range: **A:D**  ☒ Define header row: **1**

	A	B	C	D
1	<b>Cuaca</b>	<b>Suhu</b>	<b>Kelembaban_Udara</b>	<b>Berangin</b>
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

13. Simpan dengan nama **DataCuaca\_Testing**.

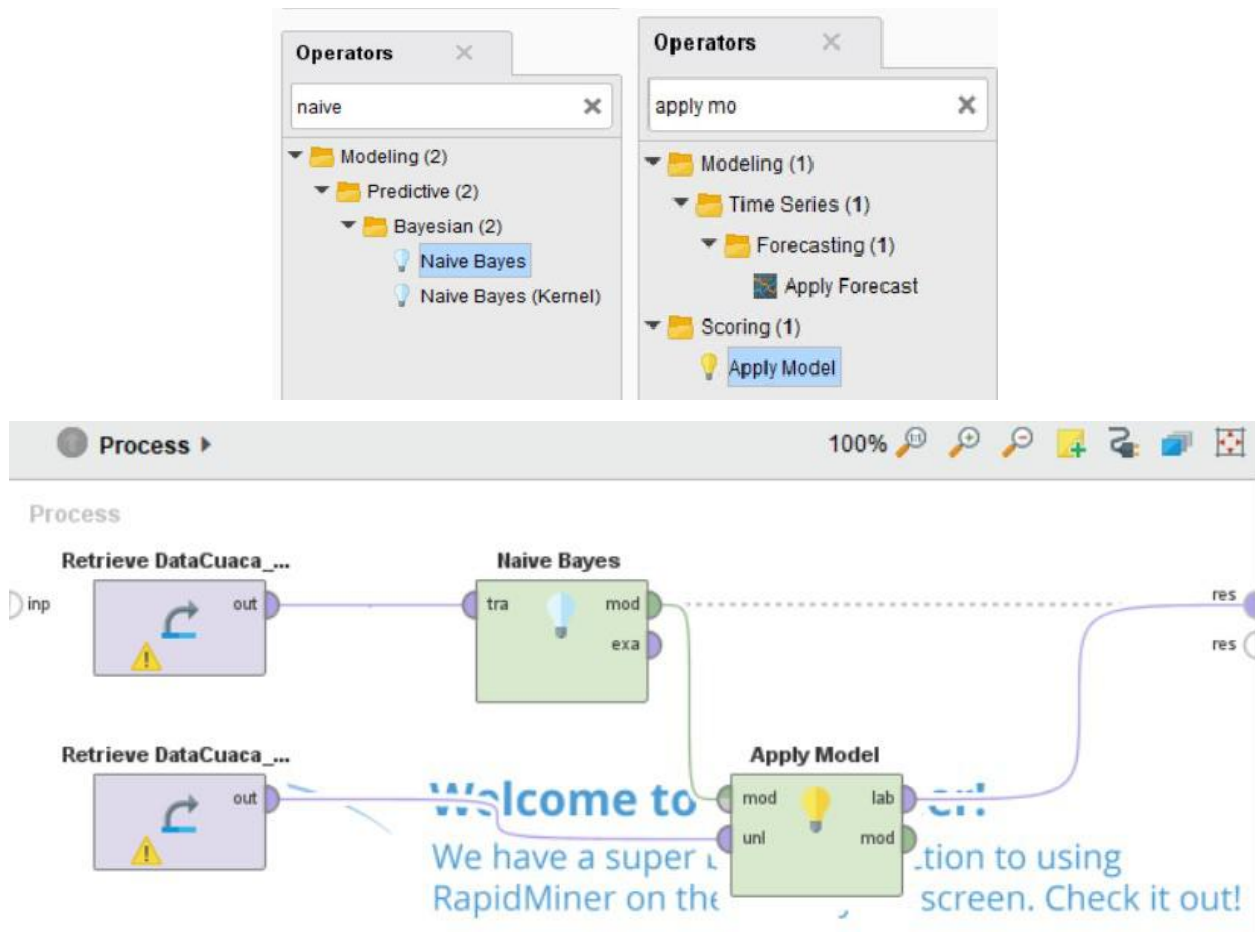
tory **ExampleSet (//Local Repository/DataCuaca\_Testing)**

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

Row No.	Cuaca	Suhu	Kelembaban_Udara	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

14. Drag **DataCuaca\_Training** dan **DataCuaca\_Testing** kedalam jendela Process View

15. Masukkan juga operator **Naive Bayes** dan **Apply Model** ke dalam Proses 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 hasil proses klasifikasi naive bayes. Pada tab **Data**, dapat dilihat hasil prediksi serta tingkat confidence nilai kelas pada masing-masing data.

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



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

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