

Mata Kuliah Coding & Machine Learning

# **Laporan Tugas Studi Kasus 2a: Supervised - Klasifikasi**

**Dosen Pengampu: Sri Wulandari, S.Kom., M.Cs.**



**Disusun oleh:**

**Lathif Ramadhan (5231811022)**

**PROGRAM STUDI SAINS DATA**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS TEKNOLOGI YOGYAKARTA**

**YOGYAKARTA**

**2025**



# 1. Algoritma Naïve Bayes

## a. Berikan penjelasan data yang saudara dapatkan

Dataset ini terdiri dari 8124 baris (sampel) dan 23 kolom (fitur). Setiap baris mewakili satu sampel jamur, dan setiap kolom mewakili atribut atau karakteristik dari jamur tersebut. Kolom terakhir adalah kolom target yang menunjukkan apakah jamur tersebut beracun (`p`) atau dapat dimakan (`e`).

Berikut adalah daftar kolom beserta penjelasannya:

1. class: Target variabel. Nilainya adalah:

- `e` untuk edible (dapat dimakan)
- `p` untuk poisonous (beracun)

2. cap-shape: Bentuk topi jamur. Nilainya adalah:

- `b` = bell (lonceng)
- `c` = conical (kerucut)
- `x` = convex (cembung)
- `f` = flat (datar)
- `k` = knobbed (bertonjolan)
- `s` = sunken (cekung)

3. cap-surface: Permukaan topi jamur. Nilainya adalah:

- `f` = fibrous (berserat)
- `g` = grooves (beralur)
- `y` = scaly (bersisik)
- `s` = smooth (halus)

4. cap-color: Warna topi jamur. Nilainya adalah:

- `n` = brown (coklat)
- `b` = buff (kuning muda)

- `c` = cinnamon (kayu manis)
- `g` = gray (abu-abu)
- `r` = green (hijau)
- `p` = pink (merah muda)
- `u` = purple (ungu)
- `e` = red (merah)
- `w` = white (putih)
- `y` = yellow (kuning)

5. bruises: Apakah jamur memiliki memar atau tidak. Nilainya adalah:

- `t` = true (ya)
- `f` = false (tidak)

6. odor: Bau jamur. Nilainya adalah:

- `a` = almond (almond)
- `l` = anise (adas)
- `c` = creosote (kreosot)
- `y` = fishy (amis)
- `f` = foul (busuk)
- `m` = musty (apak)
- `n` = none (tidak ada bau)
- `p` = pungent (menyengat)
- `s` = spicy (pedas)

7. gill-attachment: Cara insang menempel pada tangkai. Nilainya adalah:

- `a` = attached (menempel)
- `d` = descending (turun)
- `f` = free (bebas)
- `n` = notched (berlekuk)

8. gill-spacing: Jarak antara insang. Nilainya adalah:

- `c` = close (rapat)
- `w` = crowded (padat)
- `d` = distant (jauh)

9. gill-size: Ukuran insang. Nilainya adalah:

- `b` = broad (lebar)
- `n` = narrow (sempit)

10. gill-color: Warna insang. Nilainya sama dengan `cap-color`.

11. stalk-shape: Bentuk tangkai. Nilainya adalah:

- `e` = enlarging (melebar)
- `t` = tapering (menyempit)

12. stalk-root: Bentuk akar tangkai. Nilainya adalah:

- `b` = bulbous (berbonggol)
- `c` = club (berbentuk klub)
- `u` = cup (berbentuk cangkir)
- `e` = equal (sama)
- `z` = rhizomorphs (berbentuk akar)
- `r` = rooted (berakar)
- `?` = missing (hilang)

13. stalk-surface-above-ring: Permukaan tangkai di atas cincin. Nilainya sama dengan `cap-surface`.

14. stalk-surface-below-ring: Permukaan tangkai di bawah cincin. Nilainya sama dengan `cap-surface`.

15. stalk-color-above-ring: Warna tangkai di atas cincin. Nilainya sama dengan `cap-color`.

16. stalk-color-below-ring: Warna tangkai di bawah cincin. Nilainya sama dengan `cap-color`.

17. veil-type: Jenis selubung. Nilainya adalah:

- `p` = partial (sebagian)
- `u` = universal (universal)

18. veil-color: Warna selubung. Nilainya sama dengan `cap-color`.

19. ring-number: Jumlah cincin. Nilainya adalah:

- `n` = none (tidak ada)
- `o` = one (satu)
- `t` = two (dua)

20. ring-type: Jenis cincin. Nilainya adalah:

- `c` = cobwebby (seperti sarang laba-laba)
- `e` = evanescent (cepat hilang)
- `f` = flaring (melebar)
- `l` = large (besar)
- `n` = none (tidak ada)
- `p` = pendant (bergantung)
- `s` = sheathing (berpelindung)
- `z` = zone (zona)

21. spore-print-color: Warna spora. Nilainya sama dengan `cap-color`.

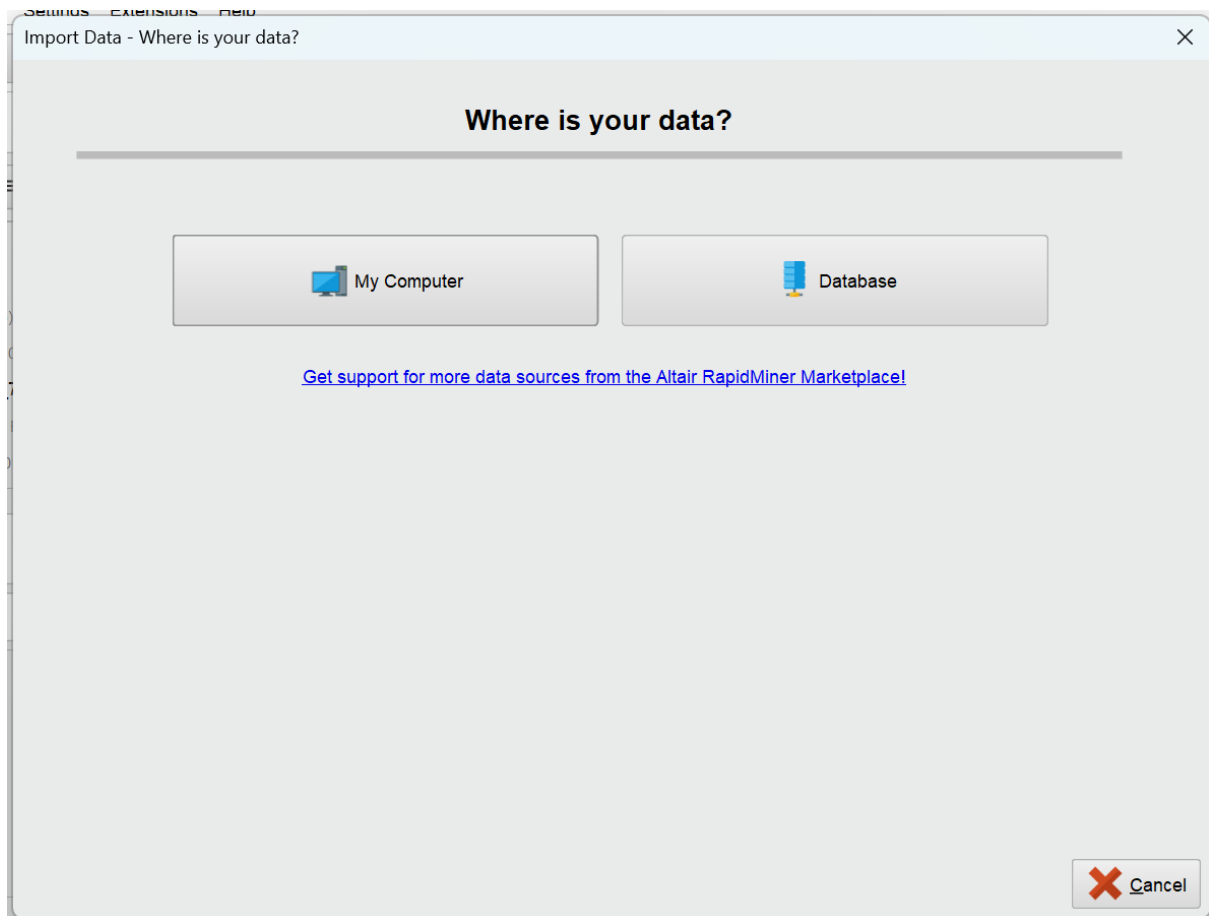
22. population: Populasi jamur. Nilainya adalah:

- `a` = abundant (melimpah)
- `c` = clustered (berkelompok)
- `n` = numerous (banyak)
- `s` = scattered (tersebar)
- `v` = several (beberapa)
- `y` = solitary (tunggal)

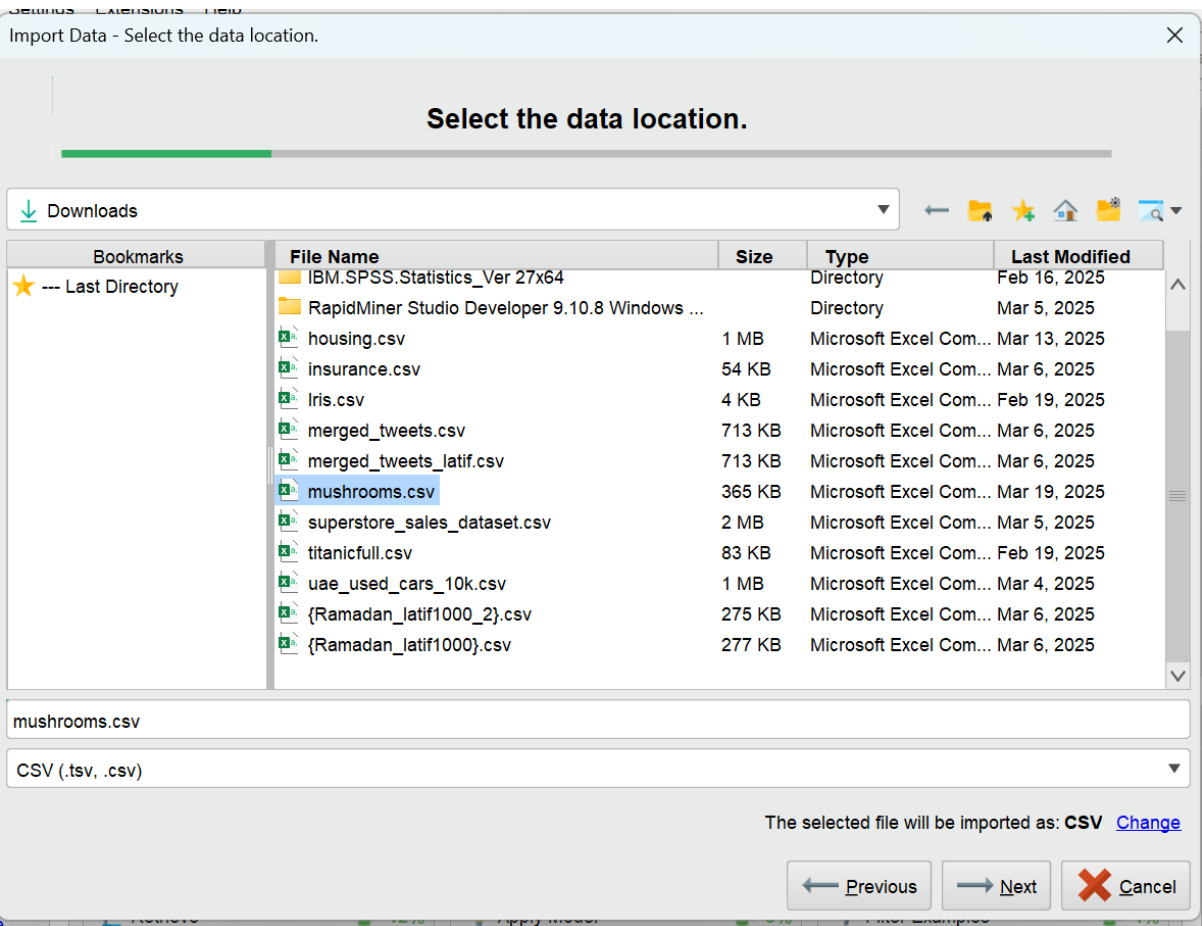
23. habitat: Habitat jamur. Nilainya adalah:

- `g` = grasses (rumput)
- `l` = leaves (daun)
- `m` = meadows (padang rumput)
- `p` = paths (jalan setapak)
- `u` = urban (perkotaan)
- `w` = waste (limbah)
- `d` = woods (hutan)

**b. Jelaskan langkah-langkah saudara dalam mengklasifikasi jamur tersebut mulai dari pengambilan data sampai terbangunnya model.**







Seabirds - Extensions - New

Import Data - Specify your data format

Specify your data format

☒ Header Row

1

Start Row

1

Column Separator

Comma ", "

File Encoding

windows-1252

Escape Character

\

Decimal Character

.

☒ Use Quotes

"

☒ Skip Comments

#

☐ Trim Lines

☐ Multiline Text

1	class	cap-sha...	cap-surf...	cap-color	bruises	odor	gill-attac...	gill-spac...	gill-size	gill-col
2	p	x	s	n	t	p	f	c	n	k
3	e	x	s	y	t	a	f	c	b	k
4	e	b	s	w	t	l	f	c	b	n
5	p	x	y	w	t	p	f	c	n	n
6	e	x	s	g	f	n	f	w	b	k
7	e	x	y	y	t	a	f	c	b	n
8	e	b	s	w	t	a	f	c	b	g
9	e	b	y	w	t	l	f	c	b	n
10	p	x	y	w	t	p	f	c	n	p
11										

☒ no problems.

Previous

Next

Cancel

Settings Extensions Help

Import Data - Format your columns.

### Format your columns.

Date format

☐ Replace errors with missing values ⓘ

	class <i>polynomial</i>	cap-shape <i>polynomial</i>	cap-surface <i>polynomial</i>	cap-color <i>polynomial</i>	bruises <i>polynomial</i>	odor <i>polynomial</i>
1	p	x	s	n	t	p
2	e	x	s	y	t	a
3	e	b	s	w	t	l
4	p	x	y	w	t	p
5	e	x	s	g	f	n
6	e	x	y	y	t	a
7	e	b	s	w	t	a
8	e	b	y	w	t	l
9	p	x	y	w	t	p
10	e	b	s	y	t	a
11	e	x	y	y	t	l

no problems.

Previous Next Cancel

ExampleSet (/Local Repository/processes/studi\_kasus\_2a\_modul\_hal71)

Open in Turbo Prep Auto Model Interactive Analysis

Filter (8,124 / 8,124 examples): all

Row No.	class	cap-shape	cap-surface	cap-color	bruises	odor	gill-attachm...	gill-spacing	gill-size	gill-color	stal
1	p	x	s	n	t	p	f	c	n	k	e
2	e	x	s	y	t	a	f	c	b	k	e
3	e	b	s	w	t	l	f	c	b	n	e
4	p	x	y	w	t	p	f	c	n	n	e
5	e	x	s	g	f	n	f	w	b	k	t
6	e	x	y	y	t	a	f	c	b	n	e
7	e	b	s	w	t	a	f	c	b	g	e
8	e	b	y	w	t	l	f	c	b	n	e
9	p	x	y	w	t	p	f	c	n	p	e
10	e	b	s	y	t	a	f	c	b	g	e
11	e	x	y	y	t	l	f	c	b	g	e
12	e	x	y	y	t	a	f	c	b	n	e
13	e	b	s	w	t	l	f	c	b	n	e

ExampleSet (8,124 examples,0 special attributes,23 regular attributes)

Berikut adalah penjelasan langkah-langkah yang saya lakukan untuk mengklasifikasikan jamur menggunakan model Decision Tree:

### 1. Retrieve Mushrooms (Pengambilan Data)

- Operator: `Retrieve` atau `Read CSV`
- Tujuan: Mengimpor dataset "Mushroom Classification" ke dalam RapidMiner.
- Detail: Dataset ini berisi 8124 sampel jamur dengan 23 atribut, termasuk kolom target `class` yang menunjukkan apakah jamur tersebut beracun (`p`) atau dapat dimakan (`e`).

### 2. Set Role (Menentukan Peran Kolom)

- Operator: `Set Role`
- Tujuan: Menentukan kolom mana yang akan digunakan sebagai label (target) untuk klasifikasi.
- Detail: Saya menetapkan kolom `class` sebagai label (target) yang akan diprediksi oleh model. Ini dilakukan dengan mengatur peran kolom `class` sebagai `label`.

### 3. Split Data (Pembagian Data)

- Operator: `Split Data`
- Tujuan: Membagi dataset menjadi dua bagian: data training dan data testing.
- Detail: Biasanya, dataset dibagi dengan rasio 70:30 atau 80:20. Data training digunakan untuk melatih model, sedangkan data testing digunakan untuk menguji performa model.

### 4. Naive Bayes (Membangun Model)

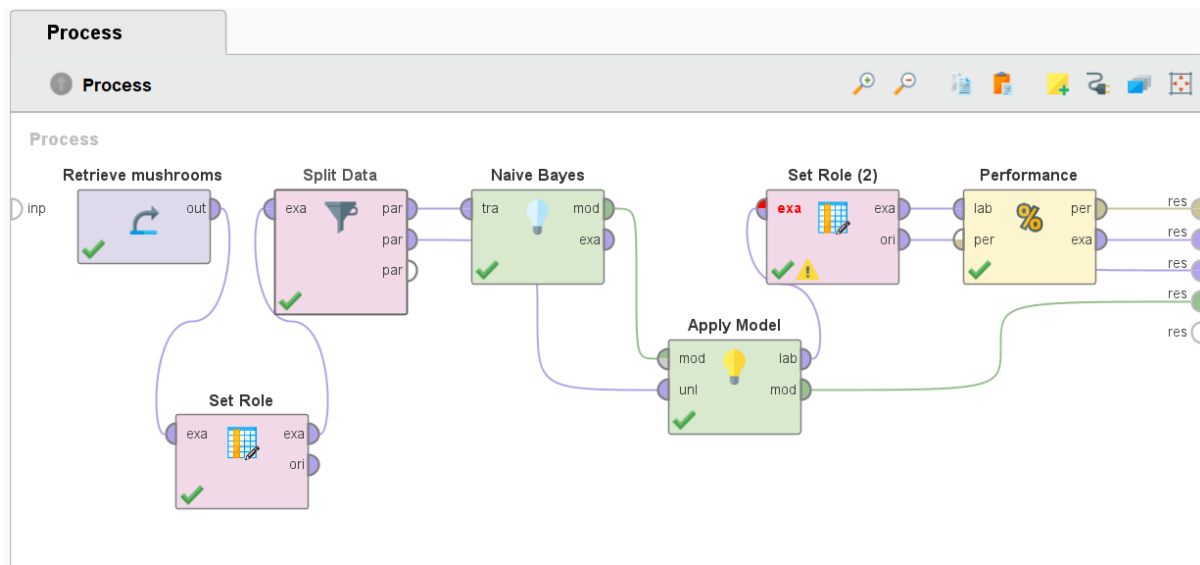
- Operator: `Naive Bayes`
- Tujuan: Membangun model klasifikasi menggunakan algoritma Naive Bayes.
- Detail: Model ini akan mempelajari pola dari data training berdasarkan atribut-atribut yang diberikan untuk memprediksi apakah suatu jamur beracun atau dapat dimakan.

### 5. Apply Model (Menerapkan Model)

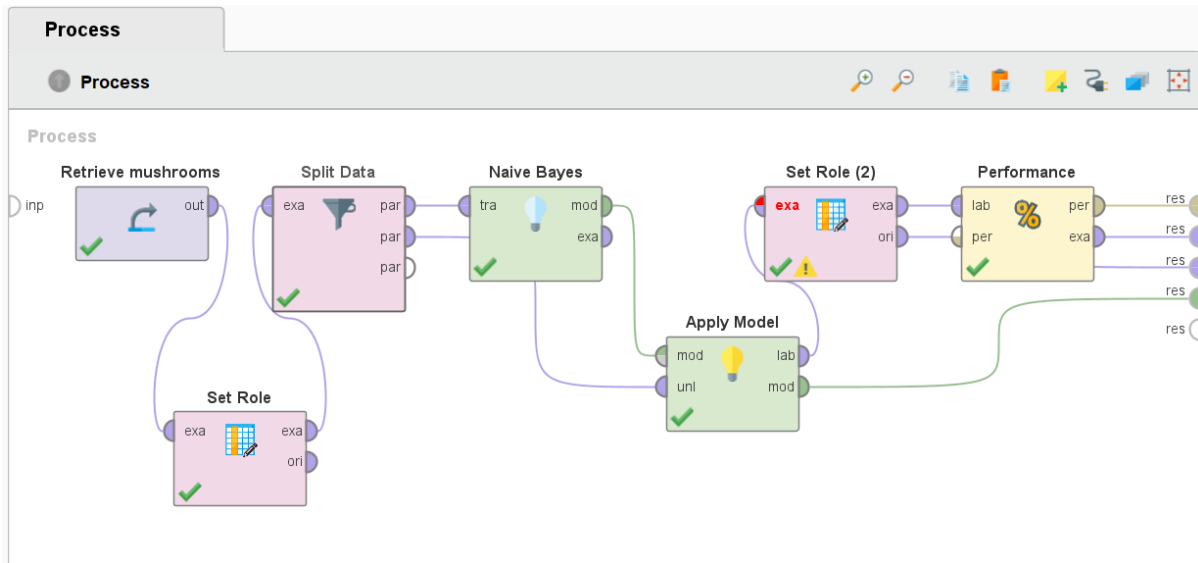
- Operator: `Apply Model`
- Tujuan: Menerapkan model yang telah dibangun (Naive Bayes) pada data testing.
- Detail: Model akan memprediksi label (`class`) untuk setiap sampel dalam data testing berdasarkan pola yang telah dipelajari selama proses training.

## 6. Performance (Evaluasi Model)

- Operator: `Performance`
- Tujuan: Mengevaluasi performa model dengan membandingkan prediksi model dengan label sebenarnya dari data testing.
- Detail: Beberapa metrik evaluasi yang umum digunakan adalah akurasi, precision, recall, dan F1-score. Metrik ini membantu saya memahami seberapa baik model Anda dalam mengklasifikasikan jamur.



**c. Berikan penjelasan model yang saudara dapatkan.**



Berikut penjelasan tentang model yang saya buat:

1. **Ambil Data:** Saya mengambil dataset jamur yang berisi informasi tentang berbagai jenis jamur dan apakah mereka beracun atau bisa dimakan.
2. **Bagi Data:** Dataset dibagi menjadi dua bagian: satu untuk melatih model (training data) dan satu lagi untuk menguji model (testing data).
3. **Set Role:** Saya menentukan kolom "class" sebagai target yang ingin diprediksi, yaitu apakah jamur beracun atau tidak.
4. **Naive Bayes:** Saya menggunakan algoritma Naive Bayes untuk melatih model. Algoritma ini belajar dari data training untuk memprediksi kategori jamur berdasarkan fitur-fiturnya.
5. **Apply Model:** Model yang sudah dilatih kemudian digunakan untuk memprediksi kategori jamur pada data testing.
6. **Evaluasi:** Saya mengevaluasi performa model dengan membandingkan prediksi model dengan kategori sebenarnya dari data testing. Ini membantu saya melihat seberapa akurat model saya.

Jadi, model saya adalah sebuah sistem yang belajar dari data jamur untuk memprediksi apakah suatu jamur beracun atau bisa dimakan, dan saya menguji seberapa baik sistem ini bekerja.

accuracy: 99.55%

	true p	true e	class precision
pred. p	1169	5	99.57%
pred. e	6	1257	99.52%
class recall	99.49%	99.60%	

Row No.	class	prediction(c...	confidence(p)	confidence(e)	cap-shape	cap-surface	cap-color	bruises	odor	gill-attachm...	gill-
1	p	p	0.999	0.001	x	s	n	t	p	f	c
2	e	e	0	1	x	y	y	t	a	f	c
3	e	e	0	1	b	s	w	t	a	f	c
4	p	p	1.000	0.000	x	y	w	t	p	f	c
5	e	e	0.000	1.000	x	y	y	t	a	f	c
6	e	e	0.000	1.000	b	s	y	t	a	f	c
7	e	e	0.000	1.000	f	f	w	f	n	f	w
8	p	p	0.993	0.007	x	s	n	t	p	f	c
9	p	p	0.998	0.002	x	y	w	t	p	f	c
10	p	p	0.999	0.001	x	s	n	t	p	f	c
11	e	e	0.000	1.000	b	s	w	t	a	f	c

ExampleSet (2,437 examples,4 special attributes,22 regular attributes)

# SimpleDistribution

Distribution model for label attribute class

Class p (0.482)  
22 distributions

Class e (0.518)  
22 distributions





**d. Cobalah model tersebut untuk mengklasifikasi jamur, silahkan saudara gunakan data yang saudara kembangkan sendiri**

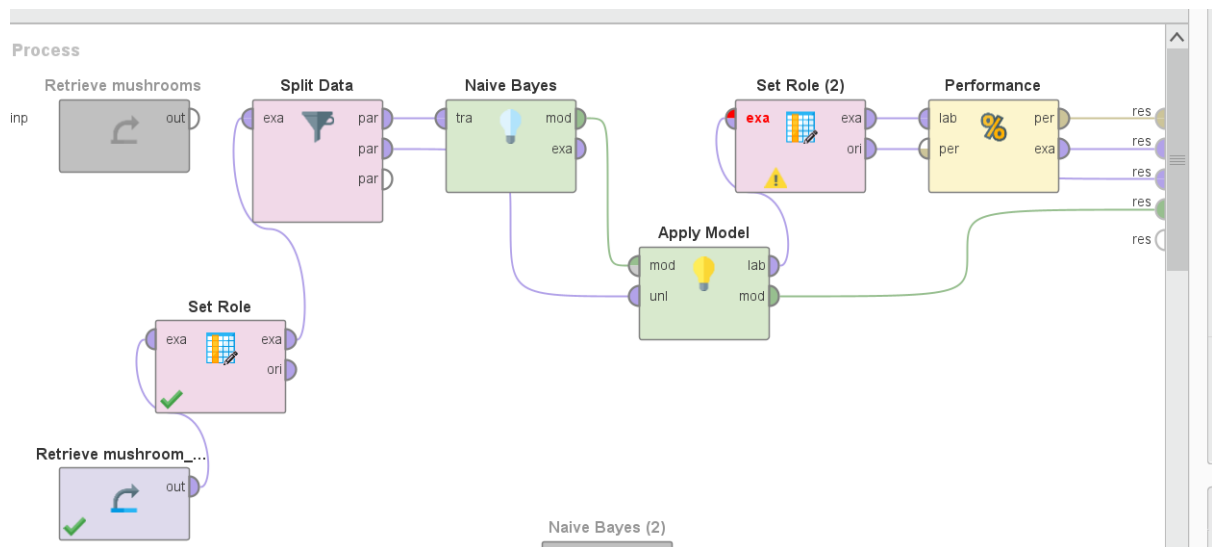
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
1	class	cap-shape	cap-surface	cap-color	bruises	odor	gill-attach	gill-spacing	gill-size	gill-color	stalk-shape	stalk-root	stalk-surf	stalk-surf	stalk-color	stalk-color	veil-type	veil-color	ring-num	ring
2	p	f	s	e	f	s	f	c	n	b	t	?	k	s	w	p	p	w	o	e
3	p	f	s	n	f	y	f	c	b	b	t	?	s	k	w	w	p	w	o	e
4	p	f	y	e	f	y	f	c	n	b	t	?	k	s	w	w	p	w	o	e
5	p	x	y	e	f	s	f	c	n	g	t	?	k	s	p	p	p	w	o	e
6	p	k	s	e	f	f	f	c	n	b	e	?	s	s	w	w	p	w	o	e
7	p	f	y	n	f	s	f	c	n	b	t	?	s	s	w	w	p	w	o	e
8	p	f	s	n	f	y	f	c	n	b	t	?	s	k	w	p	p	w	o	e
9	p	x	s	e	f	y	f	c	n	b	t	?	k	k	p	p	p	w	o	e
10	p	f	s	n	f	f	f	c	n	b	t	?	s	k	w	p	p	w	o	e
11	p	f	s	e	f	s	f	c	n	b	t	?	s	k	w	p	p	w	o	e
12	p	f	y	e	f	y	f	c	n	b	t	?	s	s	p	w	p	w	o	e
13	e	b	f	w	f	n	f	w	b	b	e	?	s	s	w	w	p	w	t	p
14	p	x	y	e	f	s	f	c	n	g	t	?	s	s	w	p	p	w	o	e
15	p	x	y	e	f	f	f	c	b	b	e	?	k	k	p	p	p	w	o	e
16	p	f	y	e	f	s	f	w	n	b	t	?	k	s	w	w	p	w	o	e
17	p	f	y	e	f	f	f	c	n	b	t	?	s	k	w	p	p	w	o	e
18	e	k	f	w	f	s	f	w	b	g	e	?	k	s	w	w	p	w	t	p
19	p	x	s	w	f	s	f	c	n	b	t	?	k	k	p	w	p	w	o	e
20	p	f	y	n	f	s	f	c	n	b	t	?	k	s	p	p	p	w	o	e
21	p	f	s	n	f	s	f	c	n	b	t	?	s	s	p	p	p	w	o	e

Open in [Turbo Prep](#) [Auto Model](#) [Interactive Analysis](#) Filter (20 / 20 examples):

Row No.	class	cap-shape	cap-surface	cap-color	bruises	odor	gill-attachm...	gill-spacing	gill-size	gill-color
1	p	f	s	e	f	s	f	c	n	b
2	p	f	s	n	f	y	f	c	b	b
3	p	f	y	e	f	y	f	c	n	b
4	p	x	y	e	f	s	f	c	n	g
5	p	k	s	e	f	f	f	c	n	b
6	p	f	y	n	f	s	f	c	n	b
7	p	f	s	n	f	y	f	c	n	b
8	p	x	s	e	f	y	f	c	n	b
9	p	f	s	n	f	f	f	c	n	b
10	p	f	s	e	f	s	f	c	n	b
11	p	f	y	e	f	y	f	c	n	b
12	e	b	f	w	f	n	f	w	b	b

ExampleSet (20 examples, 0 special attributes, 23 regular attributes)

Name	Type	Missing	Statistics	Filter (23 / 23 attributes): <input type="text" value="Search for Attributes"/>
<input checked="" type="checkbox"/> class	Nominal	0	Least e (2)      Most p (18)	Values p (18), e (2)
<input checked="" type="checkbox"/> cap-shape	Nominal	0	Least b (1)      Most f (12)	Values f (12), x (5), ...[2 more]
<input checked="" type="checkbox"/> cap-surface	Nominal	0	Least f (2)      Most s (9)	Values s (9), y (9), ...[1 more]
<input checked="" type="checkbox"/> cap-color	Nominal	0	Least w (3)      Most e (11)	Values e (11), n (6), ...[1 more]
<input checked="" type="checkbox"/> bruises	Nominal	0	Least f (20)      Most f (20)	Values f (20)
<input checked="" type="checkbox"/> odor	Nominal	0	Least n (1)      Most s (10)	Values s (10), y (5), ...[2 more]
Showing attributes 1 - 23				
Examples: 20    Special Attributes: 0    Regular Attributes: 2				



# SimpleDistribution

Distribution model for label attribute class

Class p (0.929)  
22 distributions

Class e (0.071)  
22 distributions

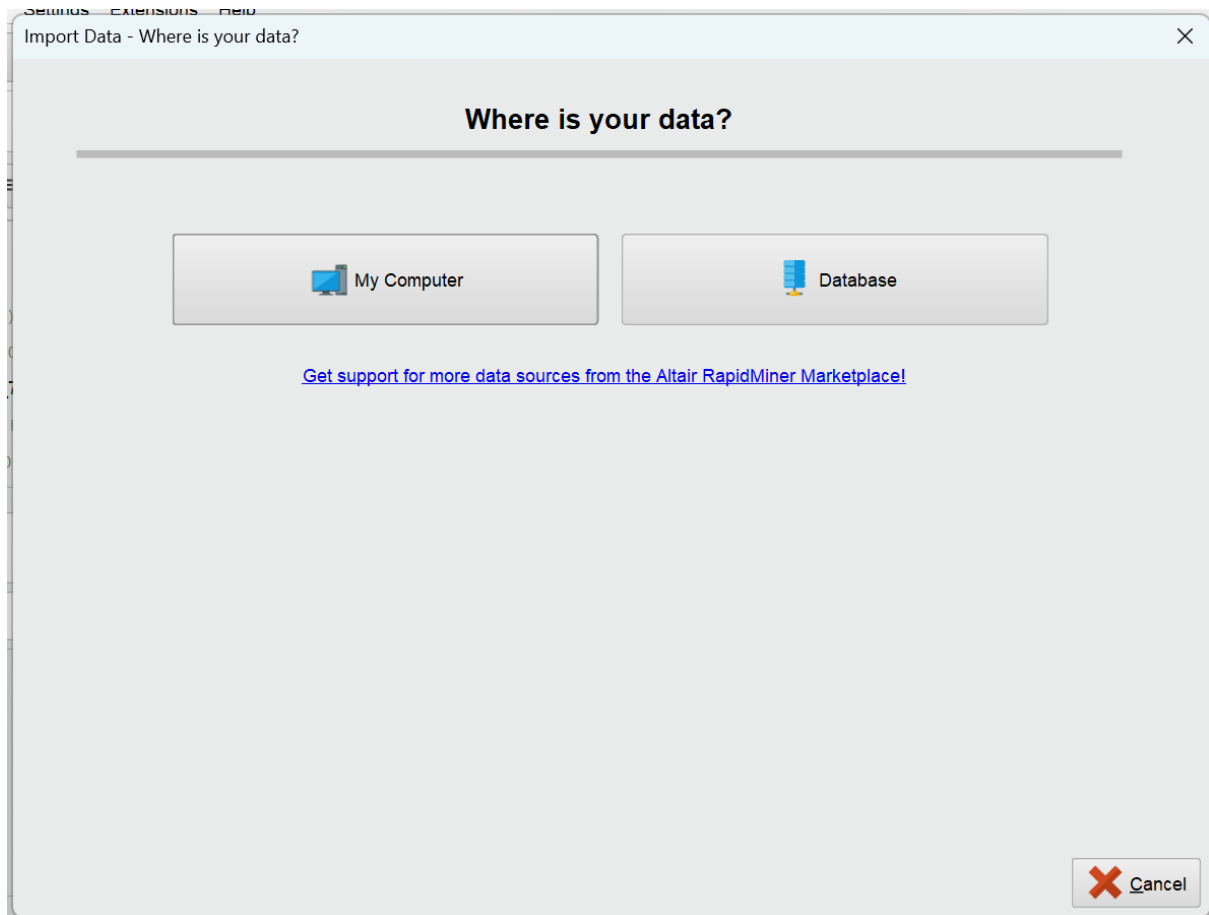
Row No.	class	prediction(c...	confidence(p)	confidence(e)	cap-shape	cap-surface	cap-color	bruises	odor	gill-attachm...	gill-sp
1	p	p	1.000	0.000	x	y	e	f	s	f	c
2	p	p	1	0	f	s	n	f	y	f	c
3	p	p	1	0	f	s	e	f	s	f	c
4	e	e	0.000	1.000	b	f	w	f	n	f	w
5	p	p	1.000	0.000	x	y	e	f	f	f	c
6	p	p	1.000	0.000	f	y	e	f	s	f	w

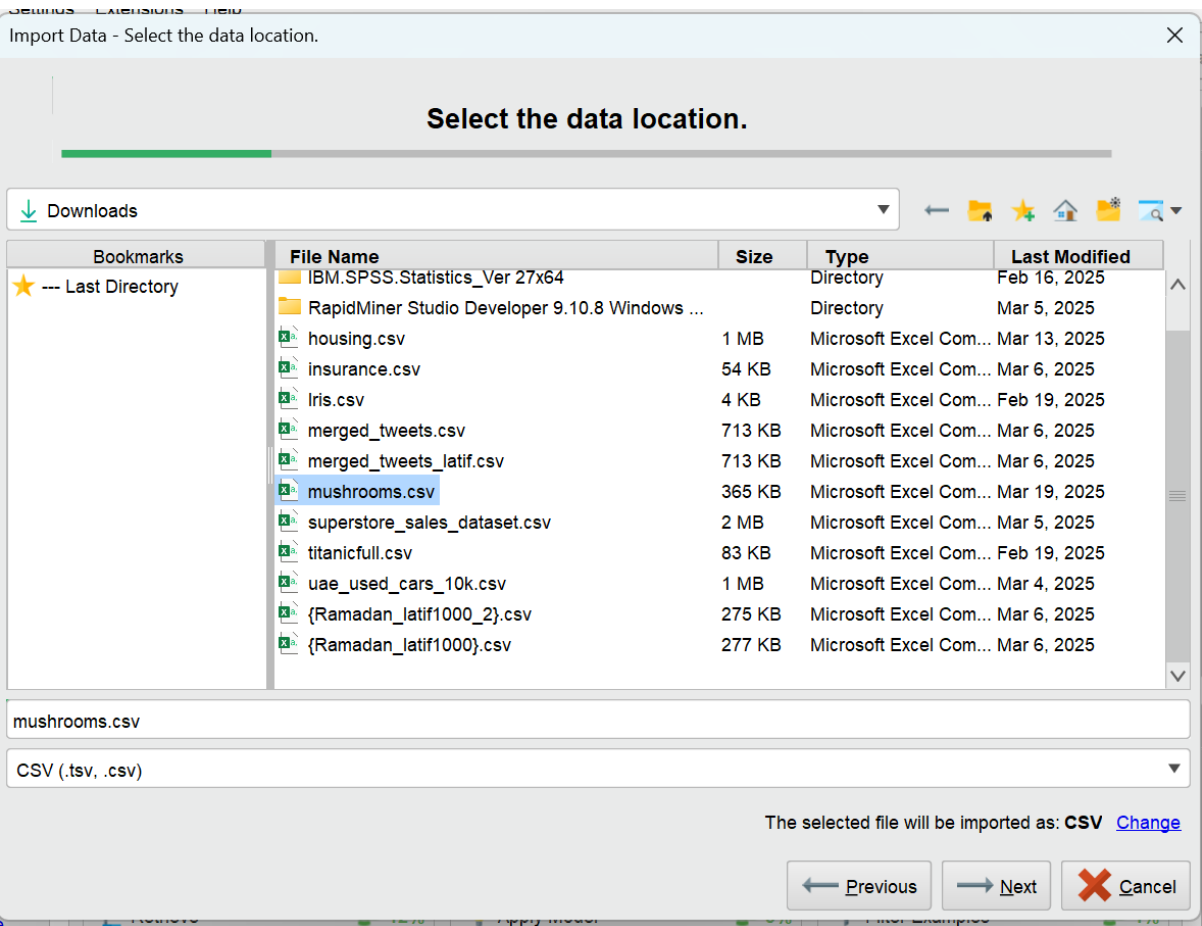
accuracy: 100.00%

	true p	true e	class precision
pred. p	5	0	100.00%
pred. e	0	1	100.00%
class recall	100.00%	100.00%	

## 2. Decision Tree

- a. Jelaskan langkah-langkah saudara dalam mengklasifikasi jamur tersebut mulai dari pengambilan data sampai terbangunnya model.





Seabirds - Extensions - New

Import Data - Specify your data format

Specify your data format

☒ Header Row

1

Start Row

1

Column Separator

Comma ", "

File Encoding

windows-1252

Escape Character

\

Decimal Character

.

☒ Use Quotes

"

☒ Skip Comments

#

☐ Trim Lines

☐ Multiline Text

1	class	cap-sha...	cap-surf...	cap-color	bruises	odor	gill-attac...	gill-spac...	gill-size	gill-col
2	p	x	s	n	t	p	f	c	n	k
3	e	x	s	y	t	a	f	c	b	k
4	e	b	s	w	t	l	f	c	b	n
5	p	x	y	w	t	p	f	c	n	n
6	e	x	s	g	f	n	f	w	b	k
7	e	x	y	y	t	a	f	c	b	n
8	e	b	s	w	t	a	f	c	b	g
9	e	b	y	w	t	l	f	c	b	n
10	p	x	y	w	t	p	f	c	n	p
11										

☒ no problems.

Previous

Next

Cancel

Settings Extensions Help

Import Data - Format your columns.

### Format your columns.

Date format

☐ Replace errors with missing values ⓘ

	class <i>polynomial</i>	cap-shape <i>polynomial</i>	cap-surface <i>polynomial</i>	cap-color <i>polynomial</i>	bruises <i>polynomial</i>	odor <i>polynomial</i>
1	p	x	s	n	t	p
2	e	x	s	y	t	a
3	e	b	s	w	t	l
4	p	x	y	w	t	p
5	e	x	s	g	f	n
6	e	x	y	y	t	a
7	e	b	s	w	t	a
8	e	b	y	w	t	l
9	p	x	y	w	t	p
10	e	b	s	y	t	a
11	e	x	y	y	t	l

no problems.

Previous Next Cancel

ExampleSet (//Local Repository/processes/studi\_kasus\_2a\_modul\_hal71)

Open in Turbo Prep Auto Model Interactive Analysis

Filter (8,124 / 8,124 examples): all

Row No.	class	cap-shape	cap-surface	cap-color	bruises	odor	gill-attachm...	gill-spacing	gill-size	gill-color	stal
1	p	x	s	n	t	p	f	c	n	k	e
2	e	x	s	y	t	a	f	c	b	k	e
3	e	b	s	w	t	l	f	c	b	n	e
4	p	x	y	w	t	p	f	c	n	n	e
5	e	x	s	g	f	n	f	w	b	k	t
6	e	x	y	y	t	a	f	c	b	n	e
7	e	b	s	w	t	a	f	c	b	g	e
8	e	b	y	w	t	l	f	c	b	n	e
9	p	x	y	w	t	p	f	c	n	p	e
10	e	b	s	y	t	a	f	c	b	g	e
11	e	x	y	y	t	l	f	c	b	g	e
12	e	x	y	y	t	a	f	c	b	n	e
13	e	b	s	w	t	l	f	c	b	w	e

ExampleSet (8,124 examples,0 special attributes,23 regular attributes)

Berikut adalah penjelasan langkah-langkah yang saya lakukan untuk mengklasifikasikan jamur menggunakan model Decision Tree:

### 1. Retrieve Mushrooms (Pengambilan Data)

- Operator: `Retrieve` atau `Read CSV`
- Tujuan: Mengimpor dataset "Mushroom Classification" ke dalam RapidMiner.
- Detail: Dataset ini berisi 8124 sampel jamur dengan 23 atribut, termasuk kolom target `class` yang menunjukkan apakah jamur tersebut beracun (`p`) atau dapat dimakan (`e`).

### 2. Set Role (Menentukan Peran Kolom)

- Operator: `Set Role`
- Tujuan: Menentukan kolom mana yang akan digunakan sebagai label (target) untuk klasifikasi.
- Detail: Anda menetapkan kolom `class` sebagai label (target) yang akan diprediksi oleh model. Ini dilakukan dengan mengatur peran kolom `class` sebagai `label`.

### 3. Split Data (Pembagian Data)

- Operator: `Split Data`
- Tujuan: Membagi dataset menjadi dua bagian: data training dan data testing.
- Detail: Biasanya, dataset dibagi dengan rasio 70:30 atau 80:20. Data training digunakan untuk melatih model, sedangkan data testing digunakan untuk menguji performa model.

### 4. Decision Tree (Membangun Model)

- Operator: Decision Tree
- Tujuan: Membangun model klasifikasi menggunakan algoritma Decision Tree.
- Detail: Model ini akan mempelajari pola dari data training dengan membangun struktur pohon keputusan berdasarkan atribut-atribut yang diberikan. Setiap cabang pohon mewakili keputusan berdasarkan nilai atribut tertentu, dan daun pohon mewakili prediksi apakah jamur tersebut beracun atau dapat dimakan. Decision Tree memilih atribut terbaik untuk memisahkan data pada setiap langkah, sehingga membentuk aturan-aturan yang mudah dipahami untuk klasifikasi.

### 5. Apply Model (Menerapkan Model)

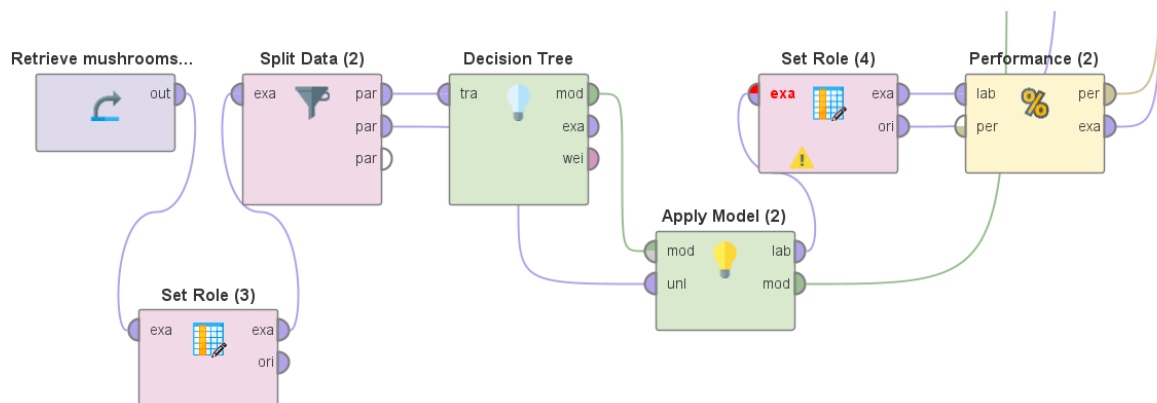
- Operator: `Apply Model`



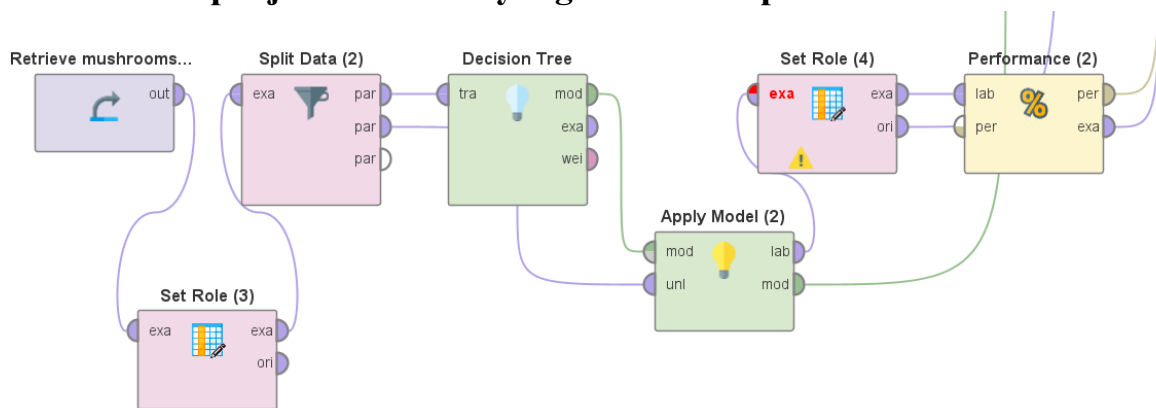
- Tujuan: Menerapkan model yang telah dibangun (Naive Bayes) pada data testing.
- Detail: Model akan memprediksi label (`class`) untuk setiap sampel dalam data testing berdasarkan pola yang telah dipelajari selama proses training.

## 6. Performance (Evaluasi Model)

- Operator: `Performance`
- Tujuan: Mengevaluasi performa model dengan membandingkan prediksi model dengan label sebenarnya dari data testing.
- Detail: Beberapa metrik evaluasi yang umum digunakan adalah akurasi, precision, recall, dan F1-score. Metrik ini membantu Anda memahami seberapa baik model Anda dalam mengklasifikasikan jamur.



**b. Berikan penjelasan model yang saudara dapatkan.**



Berikut penjelasan singkat tentang model yang dibuat:

1. **Ambil Data:** Dataset jamur diambil untuk dianalisis. Dataset ini berisi informasi tentang berbagai jenis jamur dan apakah mereka beracun atau bisa dimakan.
2. **Bagi Data:** Dataset dibagi menjadi dua bagian, yaitu data training untuk melatih model dan data testing untuk menguji model.
3. **Set Role:** Kolom "class" ditetapkan sebagai target yang ingin diprediksi, yaitu kategori jamur (beracun atau bisa dimakan).
4. **Decision Tree:** Model dibangun menggunakan algoritma Decision Tree. Algoritma ini membuat pohon keputusan berdasarkan atribut-atribut dalam data untuk memprediksi kategori jamur. Setiap cabang pohon mewakili keputusan berdasarkan nilai atribut tertentu.
5. **Apply Model:** Model yang sudah dilatih digunakan untuk memprediksi kategori jamur pada data testing.
6. **Evaluasi:** Performa model diukur dengan membandingkan prediksi model dengan kategori sebenarnya dari data testing. Ini membantu melihat seberapa akurat model dalam mengklasifikasikan jamur.

accuracy: 99.88%

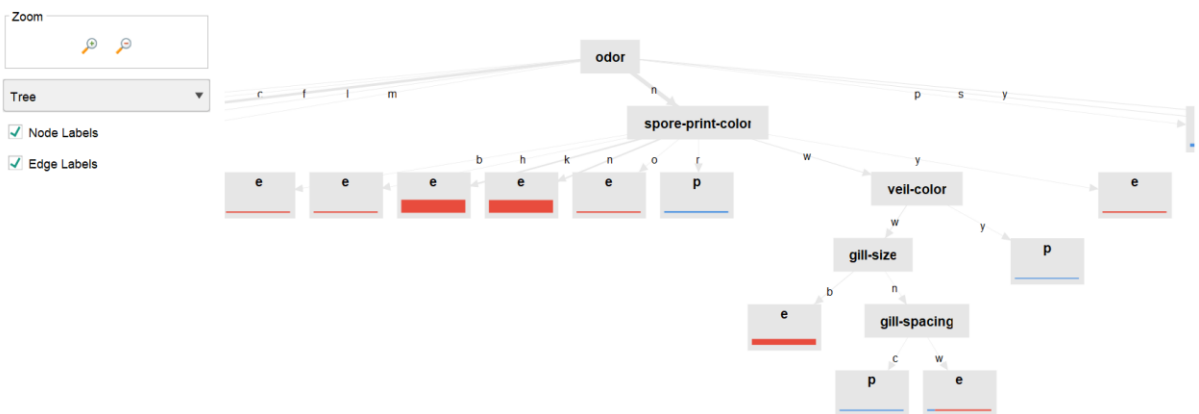
	true p	true e	class precision
pred. p	1172	0	100.00%
pred. e	3	1262	99.76%
class recall	99.74%	100.00%	

Open in  Turbo Prep  Auto Model  Interactive Analysis Filter (2,437 / 2,437 examples):

Row No.	class	prediction(c...	confidence(p)	confidence(e)	cap-shape	cap-surface	cap-color	bruises	odor	gill-attachm...	gill-
1	p	p	1	0	x	s	n	t	p	f	c
2	e	e	0	1	x	y	y	t	a	f	c
3	e	e	0	1	b	s	w	t	a	f	c
4	p	p	1	0	x	y	w	t	p	f	c
5	e	e	0	1	x	y	y	t	a	f	c
6	e	e	0	1	b	s	y	t	a	f	c
7	e	e	0	1	f	f	w	f	n	f	w
8	p	p	1	0	x	s	n	t	p	f	c
9	p	p	1	0	x	y	w	t	p	f	c
10	p	p	1	0	x	s	n	t	p	f	c
11	e	e	0	1	b	s	w	t	a	f	c

ExampleSet (2,437 examples,4 special attributes,22 regular attributes)

ory Tree (Decision Tree) ExampleSet (Set Role (4)) ExampleSet (Apply Model (2))



c. Cobalah model tersebut untuk mengklasifikasi jamur, silahkan saudara gunakan data yang saudara kembangkan sendiri

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
1	class	cap-shape	cap-surface	cap-color	bruises	odor	gill-attach	gill-spacing	gill-size	gill-color	stalk-shape	stalk-root	stalk-surface	stalk-surface	stalk-color	stalk-color	veil-type	veil-color	ring-number	ring-color
2	p	f	s	e	f	s	f	c	n	b	t	?	k	s	w	p	p	w	o	e
3	p	f	s	n	f	y	f	c	b	b	t	?	s	k	w	w	p	w	o	e
4	p	f	y	e	f	y	f	c	n	b	t	?	k	s	w	w	p	w	o	e
5	p	x	y	e	f	s	f	c	n	g	t	?	k	s	p	p	p	w	o	e
6	p	k	s	e	f	f	f	c	n	b	e	?	s	s	w	w	p	w	o	e
7	p	f	y	n	f	s	f	c	n	b	t	?	s	s	w	w	p	w	o	e
8	p	f	s	n	f	y	f	c	n	b	t	?	s	k	w	p	p	w	o	e
9	p	x	s	e	f	y	f	c	n	b	t	?	k	k	p	p	p	w	o	e
10	p	f	s	n	f	f	f	c	n	b	t	?	s	k	w	p	p	w	o	e
11	p	f	s	e	f	s	f	c	n	b	t	?	s	k	w	p	p	w	o	e
12	p	f	y	e	f	y	f	c	n	b	t	?	s	s	p	w	p	w	o	e
13	e	b	f	w	f	n	f	w	b	b	e	?	s	s	w	w	p	w	t	p
14	p	x	y	e	f	s	f	c	n	g	t	?	s	s	w	p	p	w	o	e
15	p	x	y	e	f	f	f	c	b	b	e	?	k	k	p	p	p	w	o	e
16	p	f	y	e	f	s	f	w	n	b	t	?	k	s	w	w	p	w	o	e
17	p	f	y	e	f	f	f	c	n	b	t	?	s	k	w	p	p	w	o	e
18	e	k	f	w	f	s	f	w	b	g	e	?	k	s	w	w	p	w	t	p
19	p	x	s	w	f	s	f	c	n	b	t	?	k	k	p	w	p	w	o	e
20	p	f	y	n	f	s	f	c	n	b	t	?	k	s	p	p	p	w	o	e
21	p	f	s	n	f	s	f	c	n	b	t	?	s	s	p	p	p	w	o	e

Open in

Turbo Prep

Auto Model

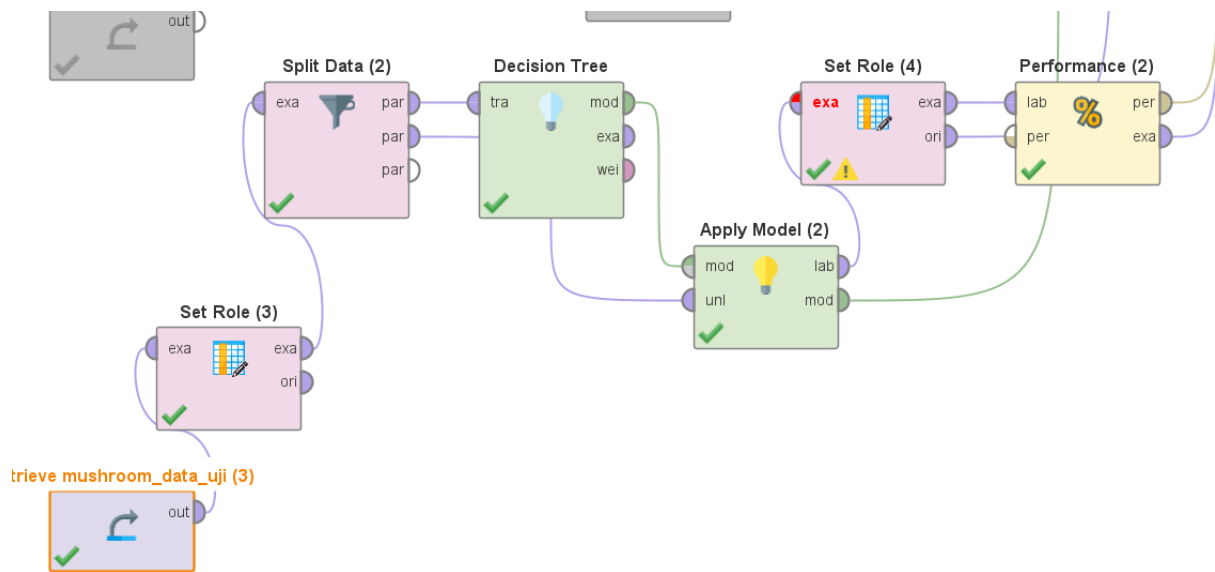
Interactive Analysis

Filter (20 / 20 examples): all

Row No.	class	cap-shape	cap-surface	cap-color	bruises	odor	gill-attachm...	gill-spacing	gill-size	gill-color
1	p	f	s	e	f	s	f	c	n	b
2	p	f	s	n	f	y	f	c	b	b
3	p	f	y	e	f	y	f	c	n	b
4	p	x	y	e	f	s	f	c	n	g
5	p	k	s	e	f	f	f	c	n	b
6	p	f	y	n	f	s	f	c	n	b
7	p	f	s	n	f	y	f	c	n	b
8	p	x	s	e	f	y	f	c	n	b
9	p	f	s	n	f	f	f	c	n	b
10	p	f	s	e	f	s	f	c	n	b
11	p	f	y	e	f	y	f	c	n	b
12	e	b	f	w	f	n	f	w	b	g

ExampleSet (20 examples,0 special attributes,23 regular attributes)





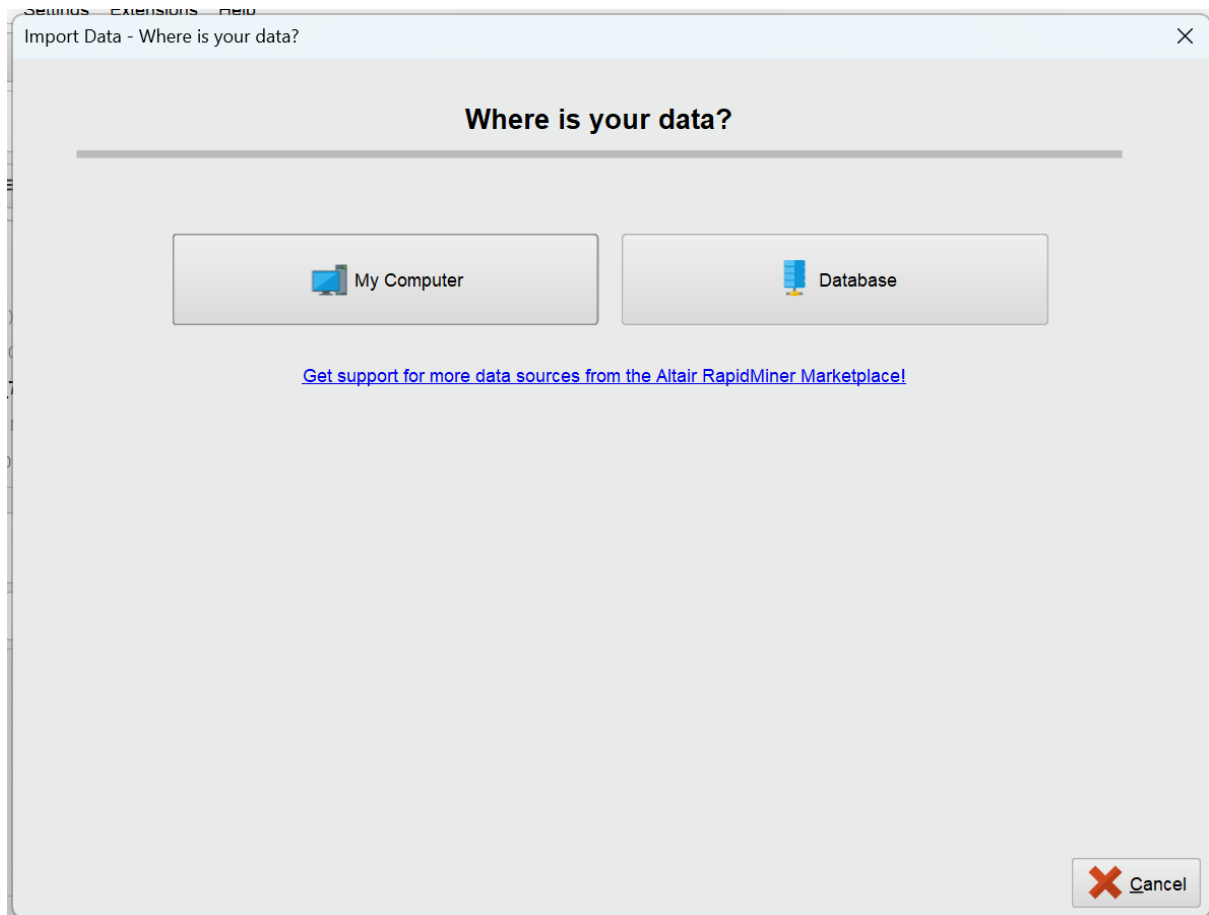
☒ Table View ☐ Plot View

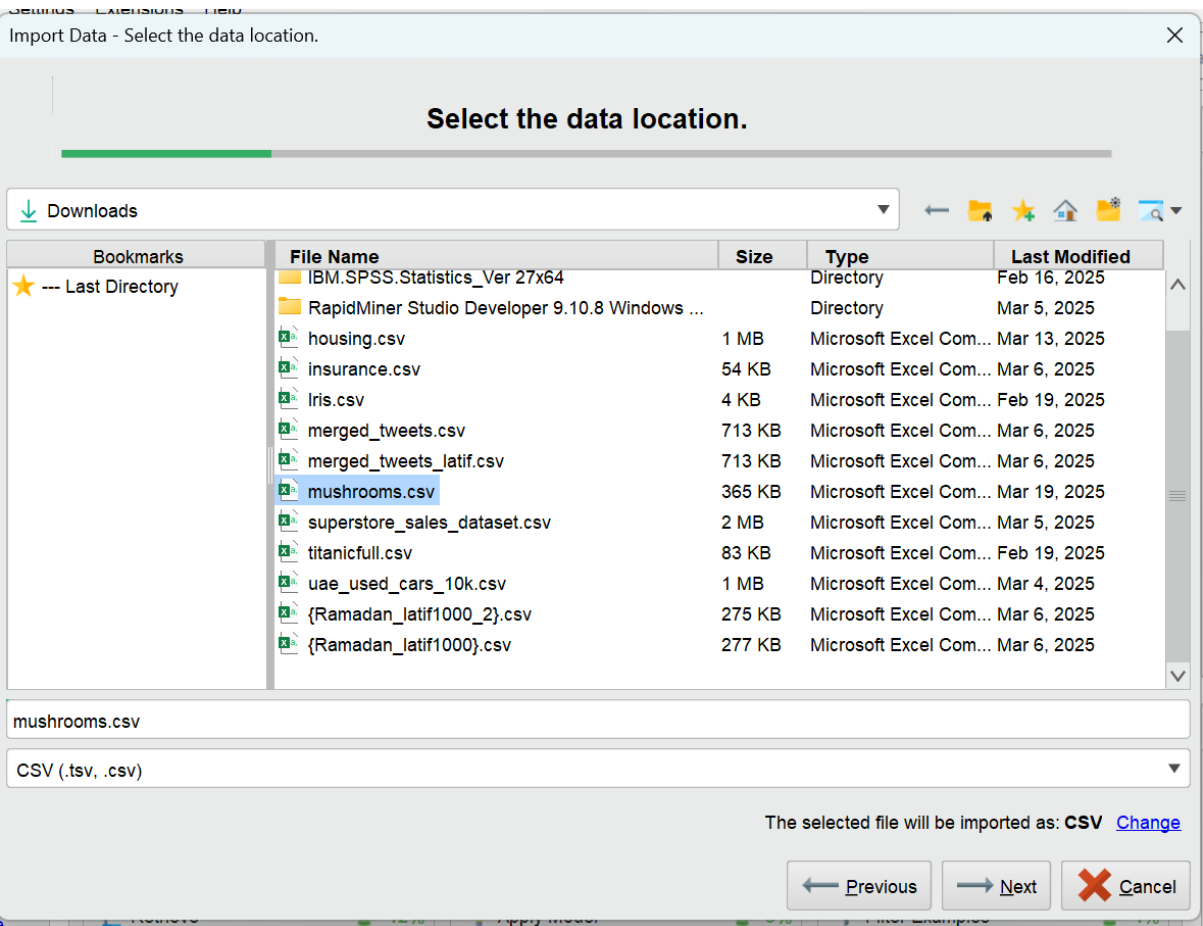
accuracy: 83.33%

	true p	true e	class precision
pred. p	5	1	83.33%
pred. e	0	0	0.00%
class recall	100.00%	0.00%	

### 3. Neural Nets

- a. Jelaskan langkah-langkah saudara dalam mengklasifikasi jamur tersebut mulai dari pengambilan data sampai terbangunnya model.







Seabirds - Extensions - New

Import Data - Specify your data format

Specify your data format

☒ Header Row

1

Start Row

1

Column Separator

Comma ", "

File Encoding

windows-1252

Escape Character

\

Decimal Character

.

☒ Use Quotes

"

☒ Skip Comments

#

☐ Trim Lines

☐ Multiline Text

no problems.

Previous

Next

Cancel

Settings Extensions Help

Import Data - Format your columns.

**Format your columns.**

Date format

☐ Replace errors with missing values ⓘ

	class <i>polynomial</i>	cap-shape <i>polynomial</i>	cap-surface <i>polynomial</i>	cap-color <i>polynomial</i>	bruises <i>polynomial</i>	odor <i>polynomial</i>
1	p	x	s	n	t	p
2	e	x	s	y	t	a
3	e	b	s	w	t	l
4	p	x	y	w	t	p
5	e	x	s	g	f	n
6	e	x	y	y	t	a
7	e	b	s	w	t	a
8	e	b	y	w	t	l
9	p	x	y	w	t	p
10	e	b	s	y	t	a
11	e	x	y	y	t	l

no problems.

Previous Next Cancel

ExampleSet (/Local Repository/processes/studi\_kasus\_2a\_modul\_hal71)

Open in Turbo Prep Auto Model Interactive Analysis

Filter (8,124 / 8,124 examples): all

Row No.	class	cap-shape	cap-surface	cap-color	bruises	odor	gill-attachm...	gill-spacing	gill-size	gill-color	stal
1	p	x	s	n	t	p	f	c	n	k	e
2	e	x	s	y	t	a	f	c	b	k	e
3	e	b	s	w	t	l	f	c	b	n	e
4	p	x	y	w	t	p	f	c	n	n	e
5	e	x	s	g	f	n	f	w	b	k	t
6	e	x	y	y	t	a	f	c	b	n	e
7	e	b	s	w	t	a	f	c	b	g	e
8	e	b	y	w	t	l	f	c	b	n	e
9	p	x	y	w	t	p	f	c	n	p	e
10	e	b	s	y	t	a	f	c	b	g	e
11	e	x	y	y	t	l	f	c	b	g	e
12	e	x	y	y	t	a	f	c	b	n	e
13	e	b	s	w	t	l	f	c	b	w	e

ExampleSet (8,124 examples,0 special attributes,23 regular attributes)

## 1. Retrieve Mushrooms (Pengambilan Data)

- **Operator:** Retrieve atau Read CSV

- **Tujuan:** Mengimpor dataset jamur ke dalam RapidMiner. Dataset ini berisi informasi tentang berbagai jenis jamur, termasuk atribut-atribut seperti bentuk topi, warna, bau, dan lainnya, serta kolom target (class) yang menunjukkan apakah jamur tersebut beracun (p) atau dapat dimakan (e).
  - **Detail:** Dataset yang digunakan adalah "Mushroom Classification" dari Kaggle, yang terdiri dari 8124 sampel dan 23 atribut.
- 

## 2. One-Hot Encoding (Transformasi Data)

- **Operator:** One-Hot Encoding
  - **Tujuan:** Mengubah data kategorikal (seperti warna, bentuk, atau bau) menjadi format numerik yang bisa diproses oleh model machine learning.
  - **Detail:** Karena algoritma Decision Tree (atau model lainnya) membutuhkan input numerik, atribut kategorikal seperti cap-shape, cap-color, atau odor diubah menjadi kolom-kolom biner (0 atau 1) untuk setiap nilai yang mungkin. Misalnya, jika cap-color memiliki nilai red, blue, dan green, maka akan dibuat tiga kolom baru: cap-color\_red, cap-color\_blue, dan cap-color\_green, dengan nilai 1 atau 0 tergantung pada nilai aslinya.
- 

## 3. Set Role (Menentukan Peran Kolom)

- **Operator:** Set Role
  - **Tujuan:** Menentukan kolom mana yang akan digunakan sebagai label (target) untuk klasifikasi.
  - **Detail:** Kolom class ditetapkan sebagai label yang akan diprediksi oleh model. Ini dilakukan dengan mengatur peran kolom class sebagai label. Kolom lainnya tetap sebagai atribut (features) yang digunakan untuk memprediksi label.
- 

## 4. Split Data (Pembagian Data)

- **Operator:** Split Data
  - **Tujuan:** Membagi dataset menjadi dua bagian: data training dan data testing.
  - **Detail:** Dataset dibagi dengan rasio tertentu, misalnya 70:30 atau 80:20. Data training digunakan untuk melatih model, sedangkan data testing digunakan untuk menguji performa model. Misalnya, jika dataset memiliki 1000 sampel, 700 sampel digunakan untuk training dan 300 sampel untuk testing.
- 

## 5. Decision Tree (Membangun Model)

- **Operator:** Decision Tree
- **Tujuan:** Membangun model klasifikasi menggunakan algoritma Decision Tree.
- **Detail:** Model ini mempelajari pola dari data training dengan membangun struktur pohon keputusan. Setiap cabang pohon mewakili keputusan berdasarkan nilai atribut tertentu (misalnya, "apakah bau jamur menyengat?"), dan daun pohon mewakili prediksi (beracun atau dapat dimakan). Decision Tree memilih atribut terbaik untuk memisahkan data pada setiap langkah, sehingga membentuk aturan-aturan yang mudah dipahami.

---

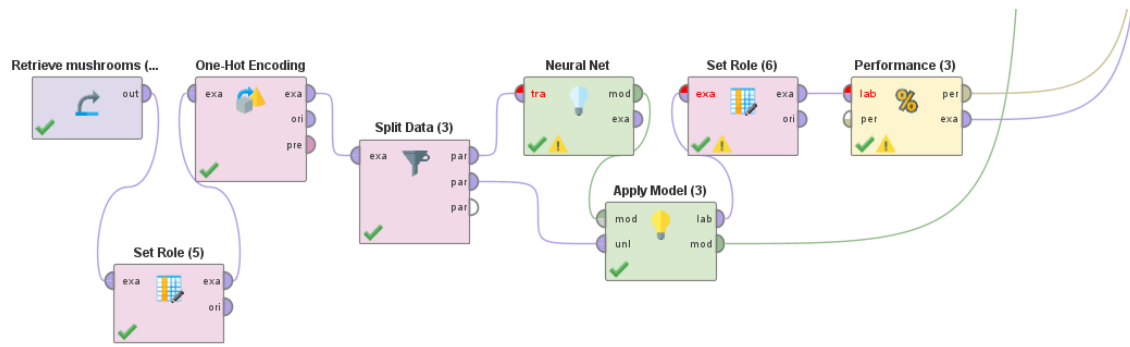
## 6. Apply Model (Menerapkan Model)

- **Operator:** Apply Model
- **Tujuan:** Menerapkan model yang telah dibangun (Decision Tree) pada data testing.
- **Detail:** Model menggunakan aturan-aturan yang telah dipelajari selama training untuk memprediksi label (class) pada data testing. Hasilnya adalah prediksi apakah setiap jamur dalam data testing beracun atau dapat dimakan.

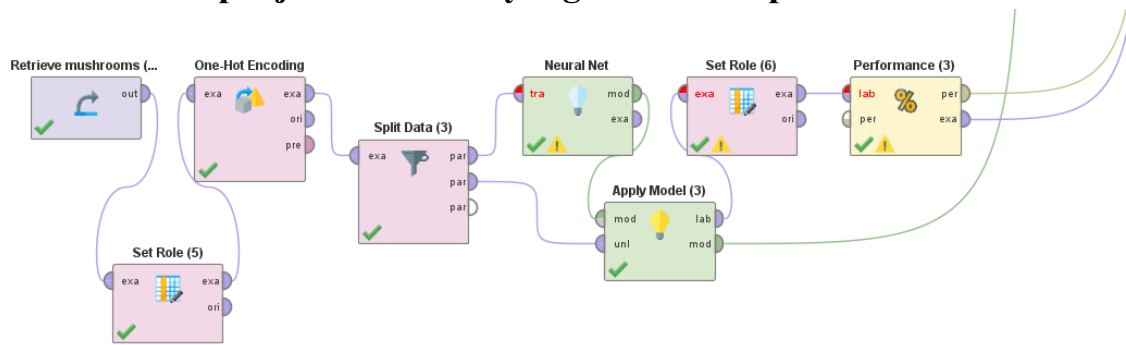
---

## 7. Performance (Evaluasi Model)

- **Operator:** Performance
- **Tujuan:** Mengevaluasi performa model dengan membandingkan prediksi model dengan label sebenarnya dari data testing.
- **Detail:** Beberapa metrik evaluasi yang digunakan adalah:
  - **Akurasi:** Seberapa sering model benar dalam memprediksi.
  - **Precision:** Seberapa akurat prediksi positif (misalnya, prediksi "beracun" yang benar).
  - **Recall:** Seberapa banyak kasus positif yang berhasil diidentifikasi.
  - **F1-Score:** Gabungan dari precision dan recall untuk mengukur keseimbangan keduanya.



**b. Berikan penjelasan model yang saudara dapatkan.**

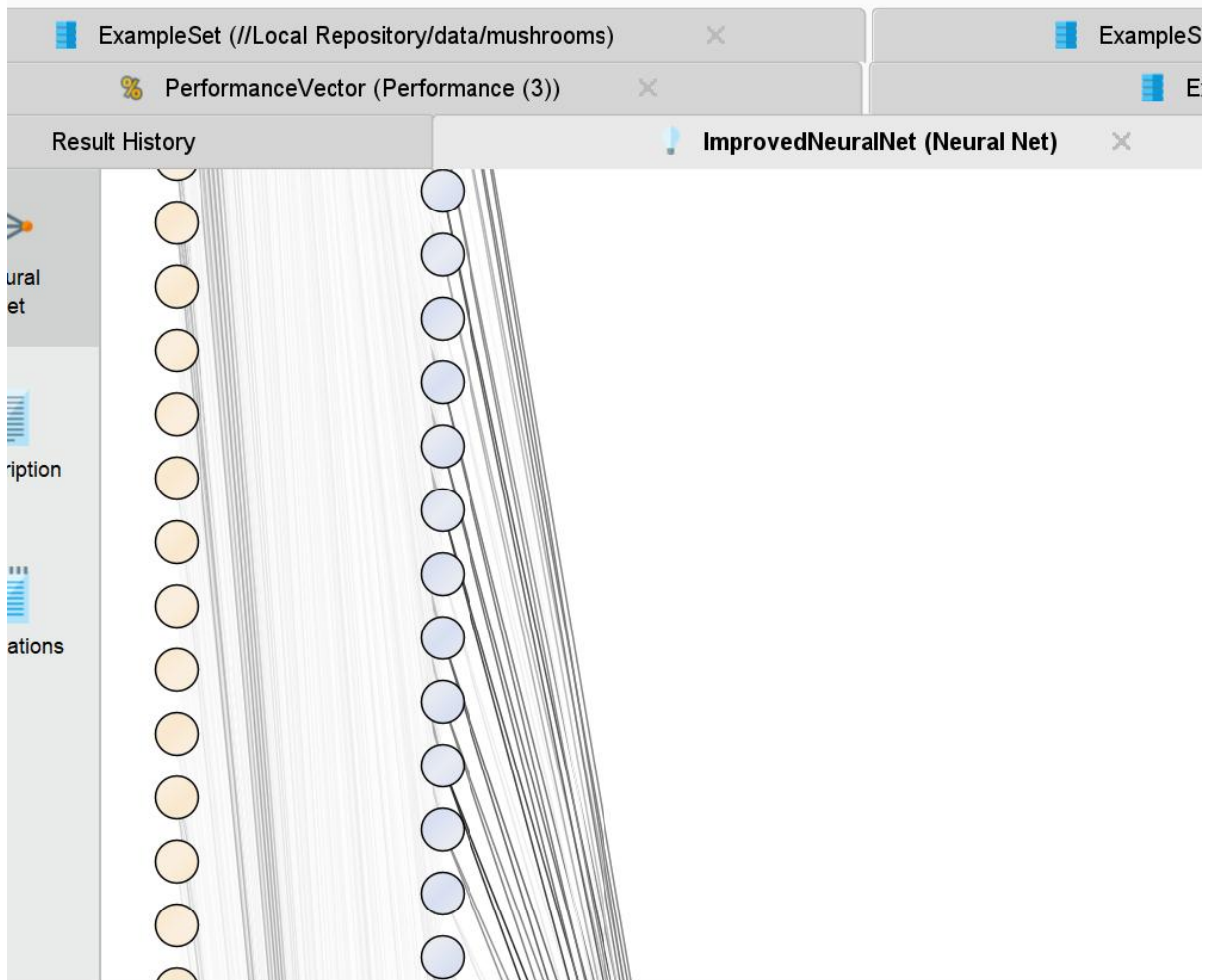


Berikut penjelasan singkat tentang model yang dibuat:

1. **Ambil Data:** Dataset skor musik diambil untuk dianalisis. Dataset ini berisi informasi tentang berbagai atribut musik yang relevan.
2. **One-Hot Encoding:** Data kategorikal diubah menjadi format numerik menggunakan teknik one-hot encoding. Ini dilakukan karena model machine learning membutuhkan input numerik. Misalnya, jika ada kategori seperti genre musik, setiap genre akan diubah menjadi kolom terpisah dengan nilai 0 atau 1.
3. **Set Role:** Kolom target ditetapkan sebagai label yang ingin diprediksi. Ini membantu model memahami mana kolom yang menjadi fokus prediksi.
4. **Split Data:** Dataset dibagi menjadi dua bagian, yaitu data training untuk melatih model dan data testing untuk menguji model.
5. **Apply Model:** Model yang sudah dilatih digunakan untuk memprediksi hasil pada data testing. Ini dilakukan untuk melihat seberapa baik model bekerja pada data yang belum pernah dilihat sebelumnya.
6. **Performance:** Performa model diukur dengan membandingkan prediksi model dengan hasil sebenarnya dari data testing. Metrik seperti akurasi, precision, recall, dan F1-score digunakan untuk mengevaluasi seberapa baik model ini bekerja.

accuracy: 100.00%

	true p	true e	class precision
pred. p	1175	0	100.00%
pred. e	0	1262	100.00%
class recall	100.00%	100.00%	



Open in [Turbo Prep](#) [Auto Model](#) [Interactive Analysis](#)

Filter (2,437 / 2,437 examples): [all](#)

Row No.	class	prediction(c...	confidence(p)	confidence(e)	cap-shape =...	cap-shape =...	cap-shape =...	cap-shape = f	cap-shape =...	cap-surface...	cap
1	p	p	0.990	0.010	1	0	0	0	0	1	0
2	e	e	0.001	0.999	1	0	0	0	0	0	1
3	e	e	0.006	0.994	0	1	0	0	0	1	0
4	p	p	0.997	0.003	1	0	0	0	0	0	1
5	e	e	0.002	0.998	1	0	0	0	0	0	1
6	e	e	0.006	0.994	0	1	0	0	0	1	0
7	e	e	0.000	1.000	0	0	0	1	0	0	0
8	p	p	0.987	0.013	1	0	0	0	0	1	0
9	p	p	0.998	0.002	1	0	0	0	0	0	1
10	p	p	0.991	0.009	1	0	0	0	0	1	0
11	e	e	0.014	0.986	0	1	0	0	0	1	0

ExampleSet (2,437 examples, 4 special attributes, 95 regular attributes)

Recommendations Update

c. Cobalah model tersebut untuk mengklasifikasi jamur, silahkan saudara gunakan data yang saudara kembangkan sendiri

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
1	class	cap-shape	cap-surface	cap-color	bruises	odor	gill-attach	gill-spacing	gill-size	gill-color	stalk-shape	stalk-root	stalk-surface	stalk-surface	stalk-color	stalk-color	veil-type	veil-color	ring-number	ring-color
2	p	f	s	e	f	s	f	c	n	b	t	?	k	s	w	p	p	w	o	e
3	p	f	s	n	f	y	f	c	b	b	t	?	s	k	w	w	p	w	o	e
4	p	f	y	e	f	y	f	c	n	b	t	?	k	s	w	w	p	w	o	e
5	p	x	y	e	f	s	f	c	n	g	t	?	k	s	p	p	p	w	o	e
6	p	k	s	e	f	f	f	c	n	b	e	?	s	s	w	w	p	w	o	e
7	p	f	y	n	f	s	f	c	n	b	t	?	s	s	w	w	p	w	o	e
8	p	f	s	n	f	y	f	c	n	b	t	?	s	k	w	p	p	w	o	e
9	p	x	s	e	f	y	f	c	n	b	t	?	k	k	p	p	p	w	o	e
10	p	f	s	n	f	f	f	c	n	b	t	?	s	k	w	p	p	w	o	e
11	p	f	s	e	f	s	f	c	n	b	t	?	s	k	w	p	p	w	o	e
12	p	f	y	e	f	y	f	c	n	b	t	?	s	s	p	w	p	w	o	e
13	e	b	f	w	f	n	f	w	b	b	e	?	s	s	w	w	p	w	t	p
14	p	x	y	e	f	s	f	c	n	g	t	?	s	s	w	p	p	w	o	e
15	p	x	y	e	f	f	f	c	b	b	e	?	k	k	p	p	p	w	o	e
16	p	f	y	e	f	s	f	w	n	b	t	?	k	s	w	w	p	w	o	e
17	p	f	y	e	f	f	f	c	n	b	t	?	s	k	w	p	p	w	o	e
18	e	k	f	w	f	s	f	w	b	g	e	?	k	s	w	w	p	w	t	p
19	p	x	s	w	f	s	f	c	n	b	t	?	k	k	p	w	p	w	o	e
20	p	f	y	n	f	s	f	c	n	b	t	?	k	s	p	p	p	w	o	e
21	p	f	s	n	f	s	f	c	n	b	t	?	s	s	p	p	p	w	o	e

Open in

Turbo Prep

Auto Model

Interactive Analysis

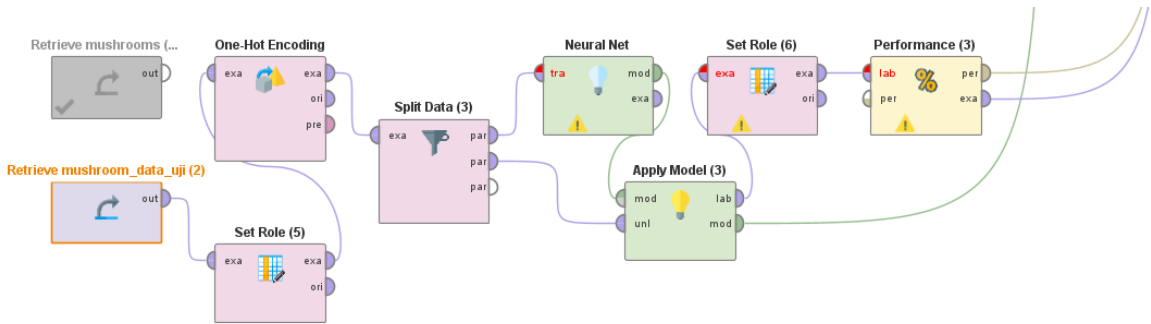
Filter (20 / 20 examples): all

Row No.	class	cap-shape	cap-surface	cap-color	bruises	odor	gill-attachm...	gill-spacing	gill-size	gill-color
1	p	f	s	e	f	s	f	c	n	b
2	p	f	s	n	f	y	f	c	b	b
3	p	f	y	e	f	y	f	c	n	b
4	p	x	y	e	f	s	f	c	n	g
5	p	k	s	e	f	f	f	c	n	b
6	p	f	y	n	f	s	f	c	n	b
7	p	f	s	n	f	y	f	c	n	b
8	p	x	s	e	f	y	f	c	n	b
9	p	f	s	n	f	f	f	c	n	b
10	p	f	s	e	f	s	f	c	n	b
11	p	f	y	e	f	y	f	c	n	b
12	e	b	f	w	f	n	f	w	b	g

ExampleSet (20 examples,0 special attributes,23 regular attributes)



Name	Type	Missing	Statistics			Filter (23 / 23 attributes): <input type="text" value="Search for Attributes"/>
class	Nominal	0	Least e (2)	Most p (18)	Values p (18), e (2)	
cap-shape	Nominal	0	Least b (1)	Most f (12)	Values f (12), x (5), ...[2 more]	
cap-surface	Nominal	0	Least f (2)	Most s (9)	Values s (9), y (9), ...[1 more]	
cap-color	Nominal	0	Least w (3)	Most e (11)	Values e (11), n (6), ...[1 more]	
bruises	Nominal	0	Least f (20)	Most f (20)	Values f (20)	
odor	Nominal	0	Least n (1)	Most s (10)	Values s (10), y (5), ...[2 more]	
			Least	Most	Values	
Showing attributes 1 - 23						Examples: 20 Special Attributes: 0 Regular Attributes: 2



Open in	Turbo Prep	Auto Model	Interactive Analysis	Filter (6 / 6 examples): <input type="text" value="all"/>							
Row No.	class	prediction(c...	confidence(p)	confidence(e)	cap-shape = f	cap-shape =...	cap-shape =...	cap-surface...	cap-surface...	cap-color = e	cap-co
1	p	p	0.957	0.043	0	1	0	0	1	1	0
2	p	p	0.982	0.018	1	0	0	1	0	0	1
3	p	p	0.980	0.020	1	0	0	1	0	1	0
4	e	e	0.449	0.551	0	0	0	0	0	0	0
5	p	p	0.967	0.033	0	1	0	0	1	1	0
6	p	p	0.939	0.061	1	0	0	0	1	1	0

# Kesimpulan

Berikut kesimpulan dari hasil yang diperoleh:

1. **Naive Bayes:** Model ini memberikan akurasi sebesar **99,55%**. Naive Bayes bekerja dengan mengasumsikan bahwa setiap atribut independen satu sama lain, dan meskipun sederhana, model ini sangat efektif untuk dataset jamur ini.
2. **Decision Tree:** Model ini memberikan akurasi lebih tinggi, yaitu **99,88%**. Decision Tree membangun pohon keputusan dengan memilih atribut terbaik untuk memisahkan data, sehingga menghasilkan aturan yang mudah dipahami dan sangat akurat.
3. **Neural Network:** Model ini mencapai akurasi sempurna, yaitu **100%**. Neural Network meniru cara kerja otak manusia dengan lapisan-lapisan neuron yang kompleks, sehingga mampu menangkap pola-pola yang lebih rumit dalam data.

**Kesimpulan:** Dari ketiga model yang diuji, Neural Network memberikan hasil terbaik dengan akurasi 100%, diikuti oleh Decision Tree (99,88%) dan Naive Bayes (99,55%). Meskipun Neural Network paling akurat, Decision Tree juga sangat kompetitif dan lebih mudah diinterpretasi. Naive Bayes, meskipun paling sederhana, tetap menunjukkan performa yang sangat baik. Pemilihan model dapat disesuaikan dengan kebutuhan, apakah mengutamakan akurasi tertinggi atau kemudahan interpretasi.

