

Mushroom

Group 6

408850443 林侑霆

408850112 劉姿妤

408850435 盧佳馨

408850393 葉柏甫

408850070 林廷毅

Content

Dataset

Research question

ML software

Decision tree model

Rules

Results

PCA

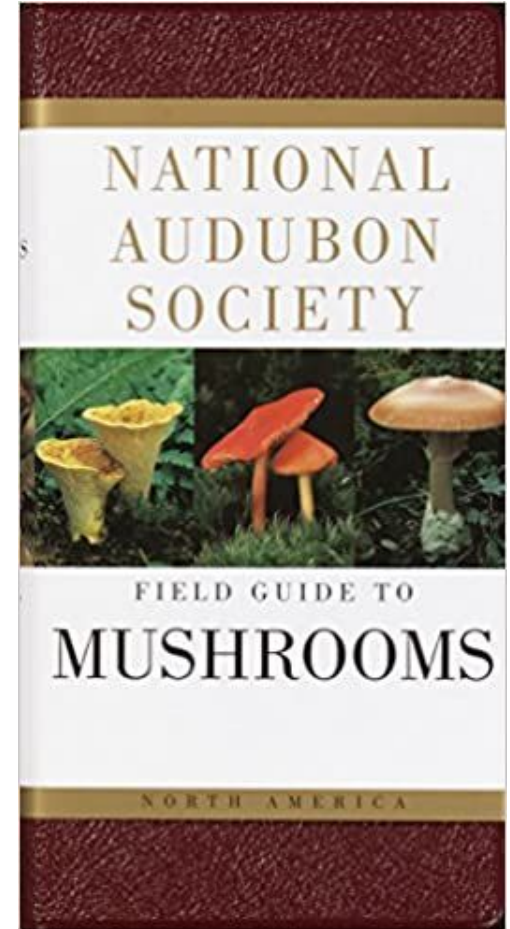
Dataset

- Dataset from: Kaggle
- Numbers of instances: 8000
- Numbers of attributes: 22

The research is from the
National Audubon Society
Field Guide to North American Mushrooms

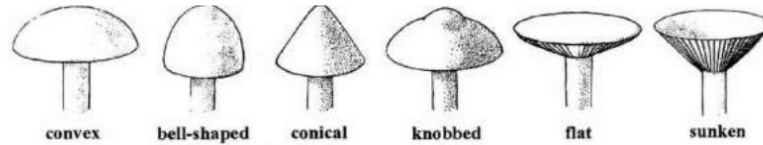
Dataset:

https://docs.google.com/spreadsheets/d/1_g487XwdE3OPiVH2Mn7sXtn0DoS21iK3m-eUQT0Q934/edit?usp=sharing

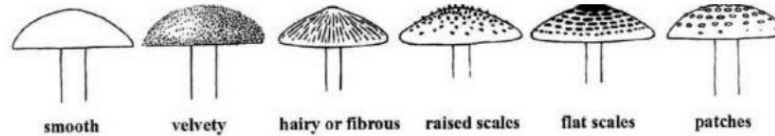


Dataset

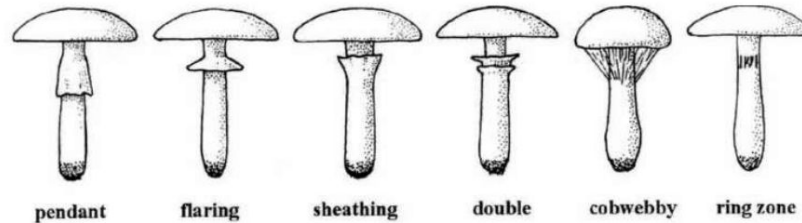
Mushroom cap shapes



Mushroom cap surfaces



Aunnular rings



Dataset

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	class	cap-shape	cap-surface	cap-color	bruises	odor	gill-attachment	gill-spacing	gill-size	gill-color	stalk-shape	stalk-root	stalk-surface-ab	stalk-surface-b	stalk-color-abo	stalk-color-belc	veil-type	veil-color
2	p	x	s	n	t	p	f	c	n	k	e	e	s	s	w	w	p	w
3	e	x	s	y	t	a	f	c	b	k	e	c	s	s	w	w	p	w
4	e	b	s	w	t	l	f	c	b	n	e	c	s	s	w	w	p	w
5	p	x	y	w	t	p	f	c	n	n	e	e	s	s	w	w	p	w
6	e	x	s	g	f	n	f	w	b	k	t	e	s	s	w	w	p	w
7	e	x	y	y	t	a	f	c	b	n	e	c	s	s	w	w	p	w
8	e	b	s	w	t	a	f	c	b	g	e	c	s	s	w	w	p	w
9	e	b	y	w	t	l	f	c	b	n	e	c	s	s	w	w	p	w
10	p	x	y	w	t	p	f	c	n	p	e	e	s	s	w	w	p	w
11	e	b	s	y	t	a	f	c	b	g	e	c	s	s	w	w	p	w
12	e	x	y	y	t	l	f	c	b	g	e	c	s	s	w	w	p	w
13	e	v	v	v	t	a	f	c	b	n	e	c	s	s	w	w	p	w
14	e					a	f	c	b	w	e	c	s	s	w	w	p	w
15	p					p	f	c	n	k	e	e	s	s	w	w	p	w
16	e	x	f	n	f	n	f	w	b	n	t	e	s	f	w	w	p	w
17	e	s	f	g	f	n	f	c	n	k	e	e	s	s	w	w	p	w

Agaricus lepiota data

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	class	cap-shape	cap-surface	cap-color	bruises	odor	gill-attachment	gill-spacing	gill-size	gill-color	stalk-shape	stalk-root	stalk-surface-ab	stalk-surface-b	stalk-color-abo	stalk-color-belc	veil-type	veil-color
2	p= poisonous	b= bell	f= fibrous	n= brown	t= bruises	a= almond	a= attached	c= close	b= broad	k= black	e= enlarging	b= bulbous	f= fibrous	f= fibrous	n= brown	n= brown	p= partial	n= brown
3	e= edible	c= conical	g= grooves	b= buff	f= no	l= anise	d= descending	w= crowded	n= narrow	n= brown	t= tapering	c= club	y= scaly	y= scaly	b= buff	b= buff	u= universal	o= orange
4		x= convex	y= scaly	c= cinnamon		c= creosote	f= free	d= distant		b= buff		u= cup	k= silky	k= silky	c= cinnamon	c= cinnamon		o= white
5		f= flat	s= smooth	g= gray		y= fishy	n= notched			h= chocolate		e= equal	s= smooth	s= smooth	g= gray	g= gray		y= yellow
6		k= knobbed		r= green		f= foul				g= gray		z= rhizomorphs			o= orange	o= orange		
7		s= sunken		p= pink		m= musty				r= green		r= rooted			p= pink	p= pink		
8				u= purple		n= none				o= orange		nuknown= missing			e= red	e= red		
9				e= red		p= pungent				p= pink					w= white	w= white		
10				w= white		s= spicy				u= purple					y= yellow	y= yellow		
11				y= yellow						e= red								
12										w= white								
13										y= yellow								
14																		
15																		

Attributes

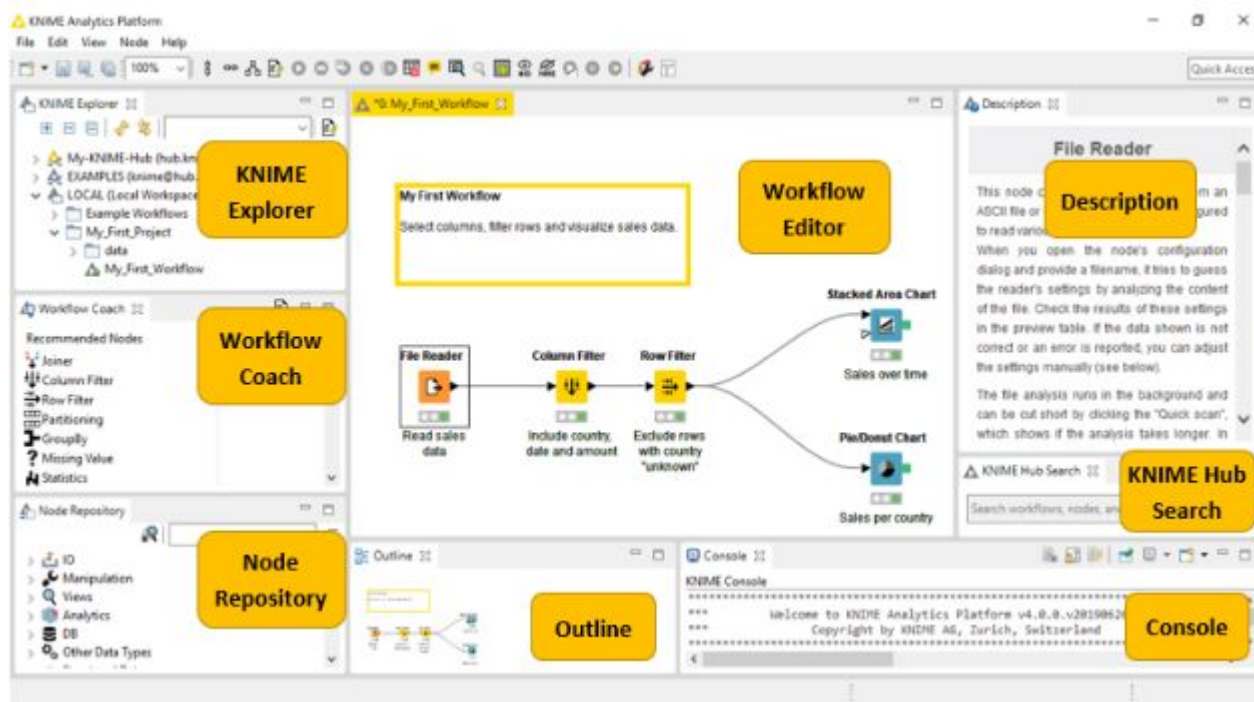
Research question

We want to identify whether mushrooms are edible

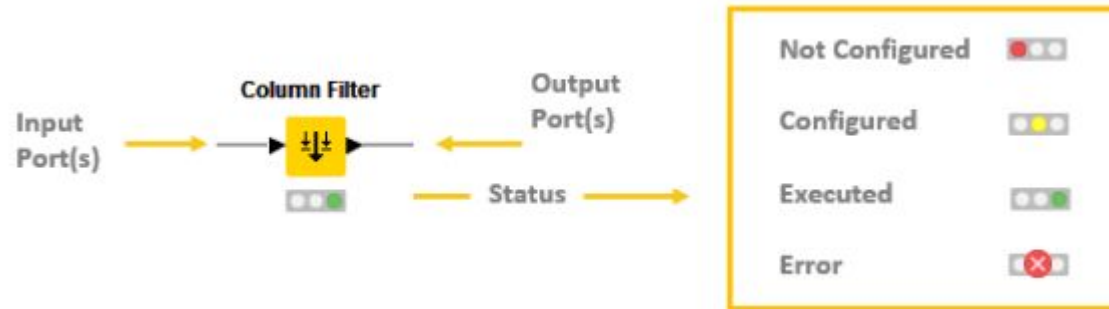
→ Find if there is any pattern with toxic/ non-toxic mushroom



ML software -KNIME



ML software -KNIME



ML software -KNIME

Recommended
Nodes

Search engine
find the node you want

The screenshot displays the KNIME Analytics Platform interface. On the left, the 'Workflow Coach' pane shows a list of 'Recommended Nodes'. The 'CSV Reader' node is highlighted in blue. Below this list, the 'Node Repository' search bar is visible. On the right, the 'CSV Reader' node configuration menu is open, showing various options like 'Configure...', 'Execute', 'Cancel', 'Reset', 'Edit Node Description...', 'New Workflow Annotation', 'Connect selected nodes', 'Disconnect selected nodes', 'Create Metanode...', 'Create Component...', 'Compare Nodes', 'Show Flow Variable Ports', 'Add File System Connection port', 'Remove File System Connection port', 'Cut', 'Copy', 'Paste', 'Undo Delete', 'Redo', 'Delete', and 'File Table'. The 'Configure...' option is highlighted with a red box. The top of the interface shows the 'KNIME Explorer' pane with a tree view of the project structure, including 'My-KNIME-Hub', 'EXAMPLES', and 'LOCAL (Local Workspace)'. The bottom of the interface shows the 'Console' pane with log messages.

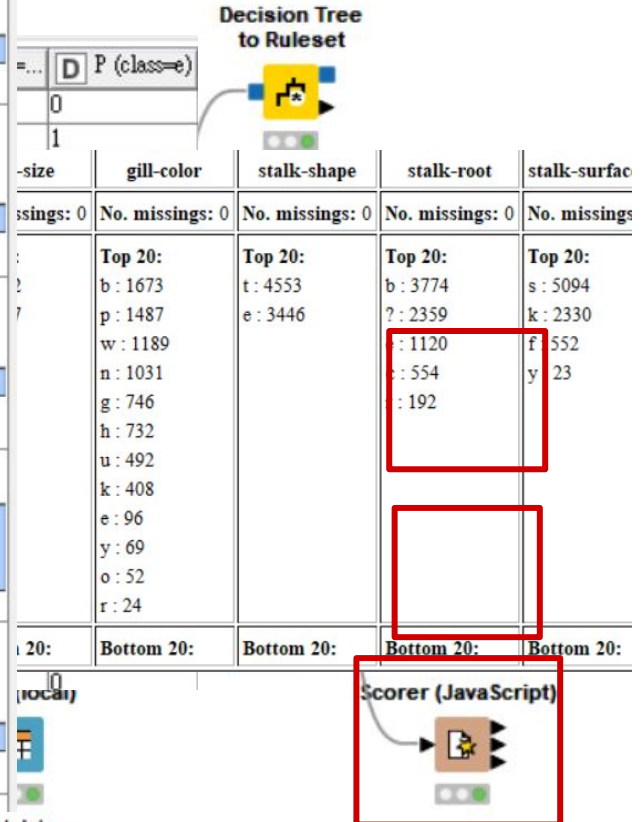
Upload your file here

Process

cap-shape	cap-surface
No. missings: 0	No. missings: 0
Top 20: x : 3634 f : 3137 k : 757 b : 435 s : 32 c : 4	Top 20: y : 3211 s : 2479 f : 2305 g : 4
Bottom 20:	Bottom 20:

Row ID	S Column	I No. miss...	Histogram
cap-shape	cap-shape	0	
cap-surface	cap-surface	0	
cap-color	cap-color	0	
bruises	bruises	0	
odor	odor	0	

Shows what rules used
to get training data



ML model - Decision Tree model

- Supervised learning.
- To solve both regression and classification problems.
- No need to normalize the data.
- Easy to understand rules.

Rule

Row ID	S Rule	D Record ...	D Number..
Row1	\$odor\$ = "p" AND TRUE => "p"	135	135
Row2	\$odor\$ = "a" AND TRUE => "e"	192	192
Row3	\$odor\$ = "l" AND TRUE => "e"	198	198
Row4	\$spore-print-color\$ = "k" AND \$odor\$ = "n" => "e"	668	668
Row5	\$spore-print-color\$ = "n" AND \$odor\$ = "n" => "e"	650	650
Row6	\$spore-print-color\$ = "h" AND \$odor\$ = "n" => "e"	25	25
Row7	\$cap-surface\$ = "s" AND \$stalk-surface-below-ring\$ = "s" AND \$spore-print-color\$ = "w" AND \$odor\$ = "n" => "e"	90	90
Row8	\$cap-surface\$ = "y" AND \$stalk-surface-below-ring\$ = "s" AND \$spore-print-color\$ = "w" AND \$odor\$ = "n" => "e"	71	69
Row9	\$cap-surface\$ = "f" AND \$stalk-surface-below-ring\$ = "s" AND \$spore-print-color\$ = "w" AND \$odor\$ = "n" => "e"	40	40
Row10	\$cap-surface\$ = "g" AND \$stalk-surface-below-ring\$ = "s" AND \$spore-print-color\$ = "w" AND \$odor\$ = "n" => "p"	3	3
Row11	\$stalk-surface-below-ring\$ = "f" AND \$spore-print-color\$ = "w" AND \$odor\$ = "n" => "e"	11	11
Row12	\$stalk-surface-below-ring\$ = "y" AND \$spore-print-color\$ = "w" AND \$odor\$ = "n" => "p"	28	23
Row13	\$stalk-surface-below-ring\$ = "k" AND \$spore-print-color\$ = "w" AND \$odor\$ = "n" => "e"	58	58
Row14	\$spore-print-color\$ = "r" AND \$odor\$ = "n" => "p"	44	44
Row15	\$spore-print-color\$ = "o" AND \$odor\$ = "n" => "e"	23	23
Row16	\$spore-print-color\$ = "b" AND \$odor\$ = "n" => "e"	18	18
Row17	\$spore-print-color\$ = "y" AND \$odor\$ = "n" => "e"	17	17
Row18	\$odor\$ = "f" AND TRUE => "p"	1,059	1,059
Row19	\$odor\$ = "c" AND TRUE => "p"	99	99
Row20	\$odor\$ = "y" AND TRUE => "p"	276	276
Row21	\$odor\$ = "s" AND TRUE => "p"	276	276
Row22	\$odor\$ = "m" AND TRUE => "p"	18	18

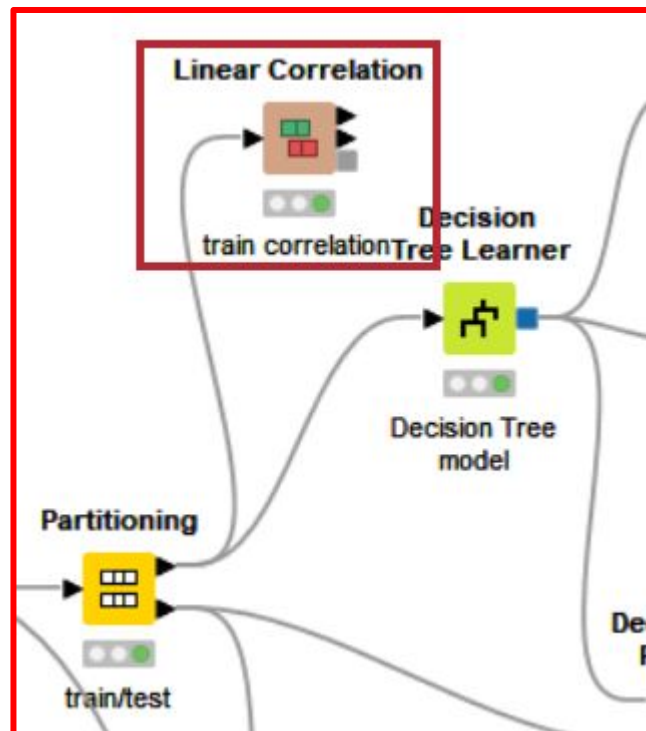
pungent

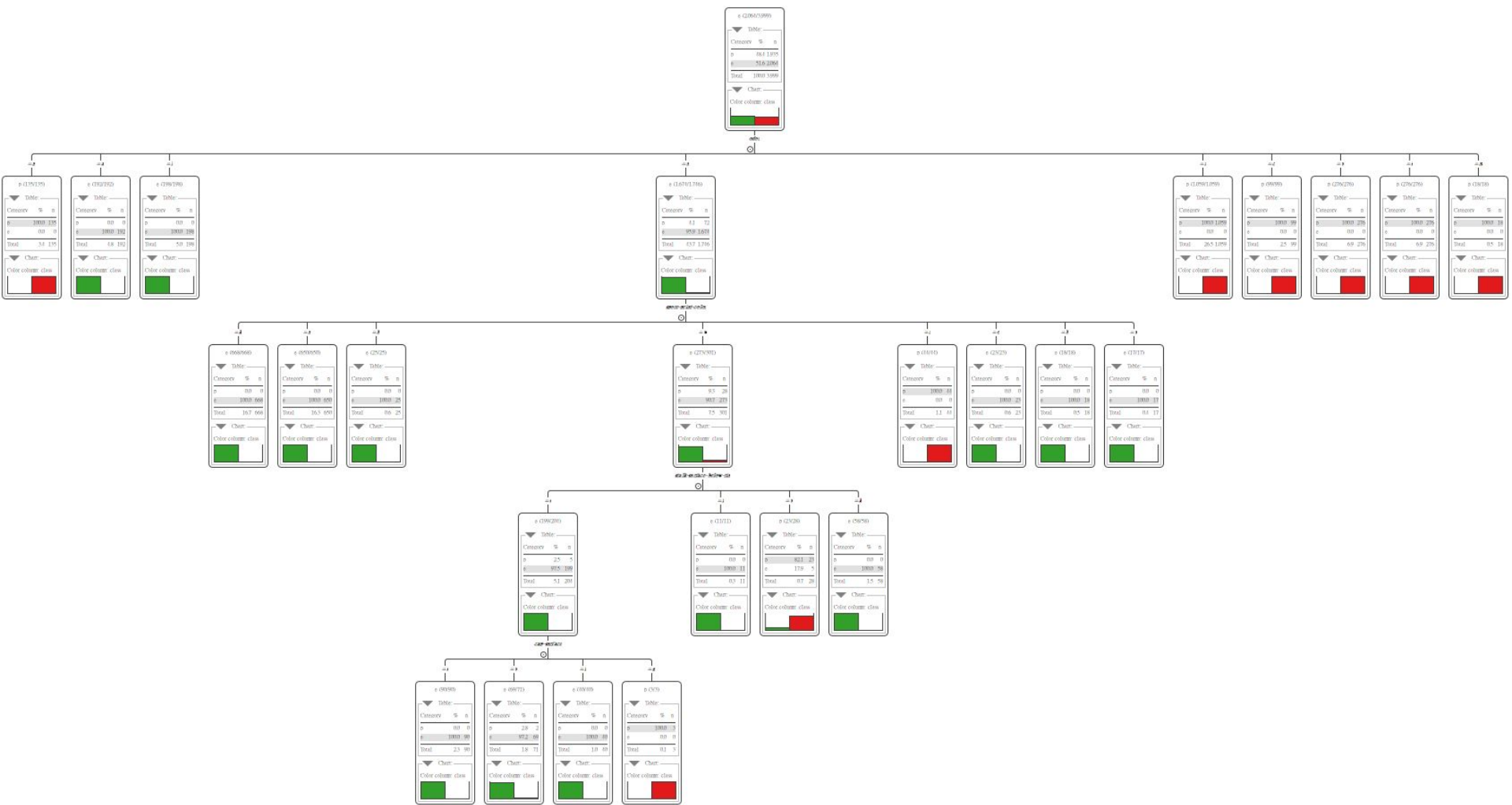
poison

black none

edible

Row107	gill-attachment	stalk-color-a...	0.97610590451557...
Row108	gill-attachment	stalk-color-b...	0.97610590451557...
Row110	gill-attachment	veil-color	0.96680173689199...
Row99	odor	class	0.96482116732567...
Row113	gill-attachment	spore-print-c...	0.86127849653227...
Row102	gill-attachment	gill-color	0.82281767825431...
Row210	stalk-color-a...	veil-color	0.81649658092772...
Row77	bruises	ring-type	0.78255306145850...
Row211	stalk-color-a...	ring-number	0.77844633154631...
Row133	gill-size	gill-color	0.77727909117451...
Row219	stalk-color-b...	ring-number	0.77691964869516...
Row84	odor	gill-size	0.76787548311239...
Row249	spore-print-c...	class	0.749952883665368
Row94	odor	ring-number	0.74231805895382...
Row238	ring-number	ring-type	0.71330829810540...
Row95	odor	ring-type	0.69820033201148...
Row212	stalk-color-a...	ring-type	0.67600010504300...





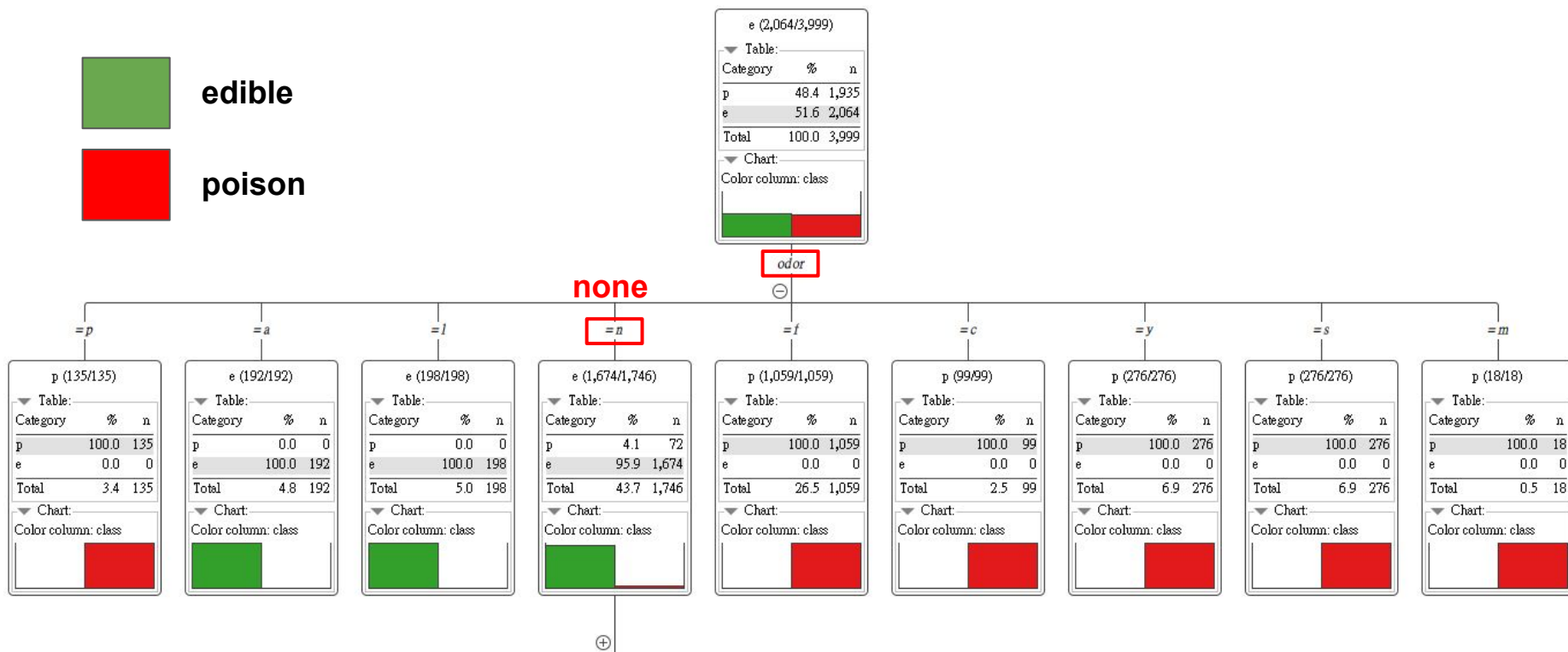
Decision Tree



edible



poison



Results

Scorer View

Confusion Matrix

Rows Number : 4000	e (Predicted)	p (Predicted)	
e (Actual)	2066	10	99.52%
p (Actual)	2	1922	99.90%
	99.90%	99.48%	

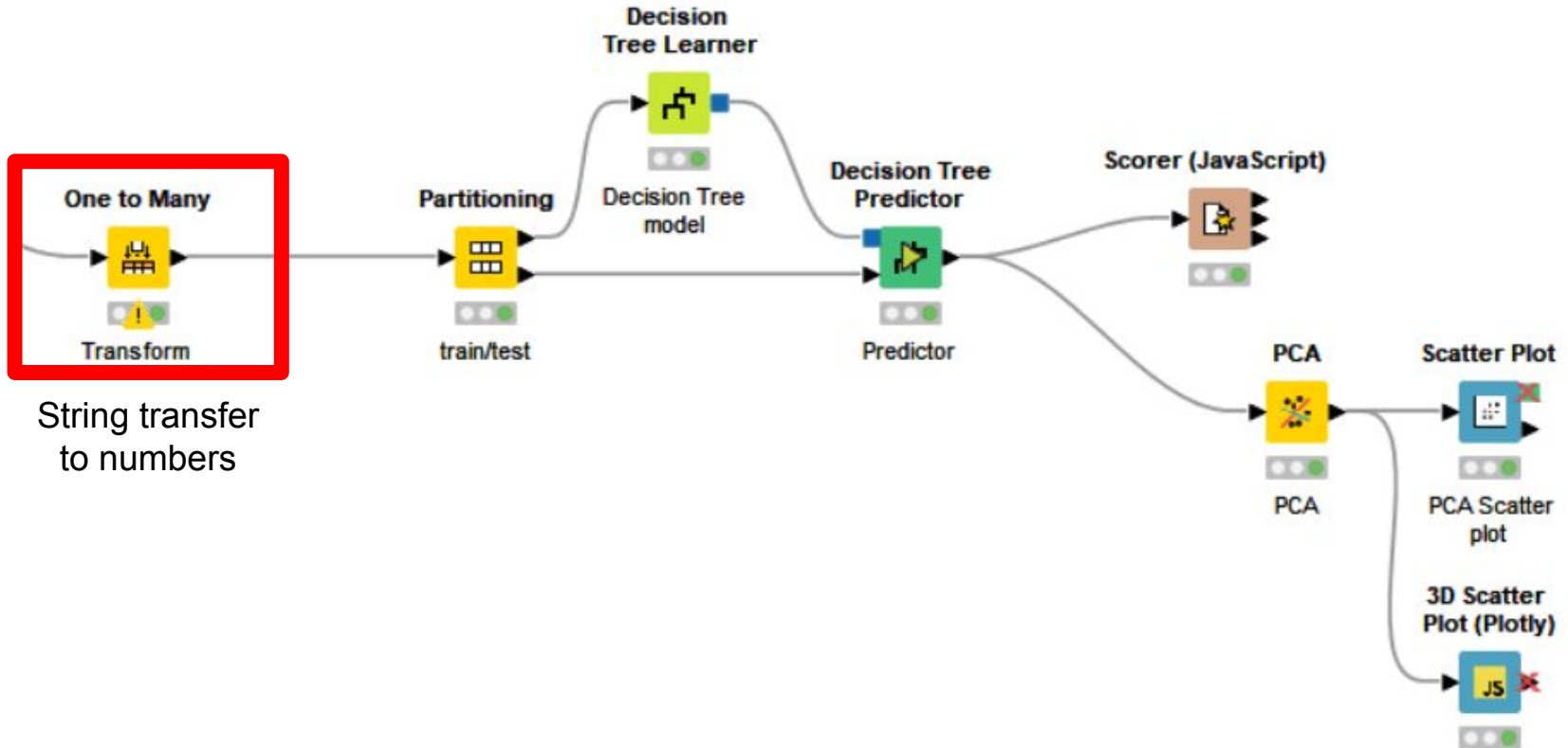
Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
e	2066	2	1922	10	99.52%	99.90%	99.52%	99.90%	99.71%
p	1922	10	2066	2	99.90%	99.48%	99.90%	99.52%	99.69%

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
99.70%	0.30%	0.994	3988	12

PCA

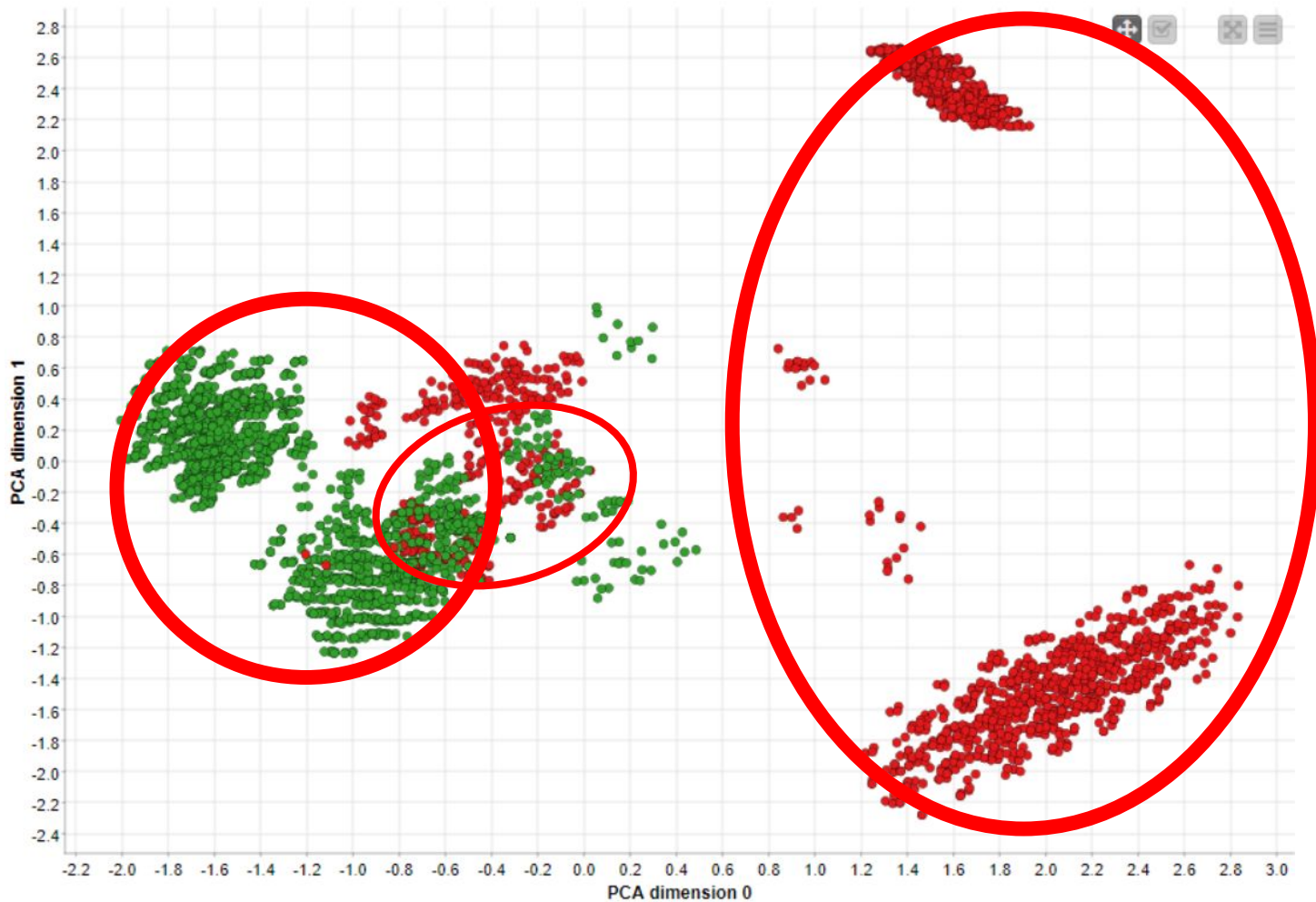


String transfer
to numbers

Row ID	S cap-shape	S cap-surf...	S cap-color
Row0	x	s	n
Row1	x	s	y
Row2	b	s	w
Row3	x	y	w
Row4	x	s	g
Row5	x	y	y
Row6	b	s	w
Row7	b	y	w
Row8	x	y	w
Row9	b	s	y
Row10	x	y	y
Row11	x	y	y
Row12	b	s	y
Row13	x	y	w
Row14	x	f	n
Row15	s	f	g
Row16	f	f	w

Row ID	S class	I x_cap-shape	I b_cap-s...
Row0	p	1	0
Row1	e	1	0
Row2	e	0	1
Row3	p	1	0
Row4	e	1	0
Row5	e	1	0
Row6	e	0	1
Row7	e	0	1
Row8	p	1	0
Row9	e	0	1
Row10	e	1	0
Row11	e	1	0
Row12	e	0	1
Row13	p	1	0
Row14	e	1	0
Row15	e	0	0
Row16	e	0	0

PCA



Conclusion

City Rumor:

Brightly colored mushrooms are poisonous.

Result:

NO!!!!!! Odor is the most related to whether it is poisonous or not.

Division of labor

Division of content:

The project was discussed, research, planned, and executed with the participation of all our team members. 沒有人划水