

IDENTIFY ALIENS

REFERENCE CODE:

<https://github.com/IRS-MR/S-MR-Workshop/tree/master/S-MR-Workshop3/knowledge-discovery-identify-aliens>

3.3 KNOWLEDGE DISCOVERY IDENTIFY ALIENS

Aliens



Not aliens



Training Data

SN	Triangle	Antenna	Teeth	Eyes	Alien
1	1	3	1	2	TRUE
2	1	3	0	2	TRUE
3	1	3	1	2	TRUE
4	1	3	0	3	TRUE
5	1	2	1	2	FALSE
6	0	3	0	3	FALSE
7	1	6	0	2	FALSE
8	0	3	0	2	FALSE

Which one is alien?

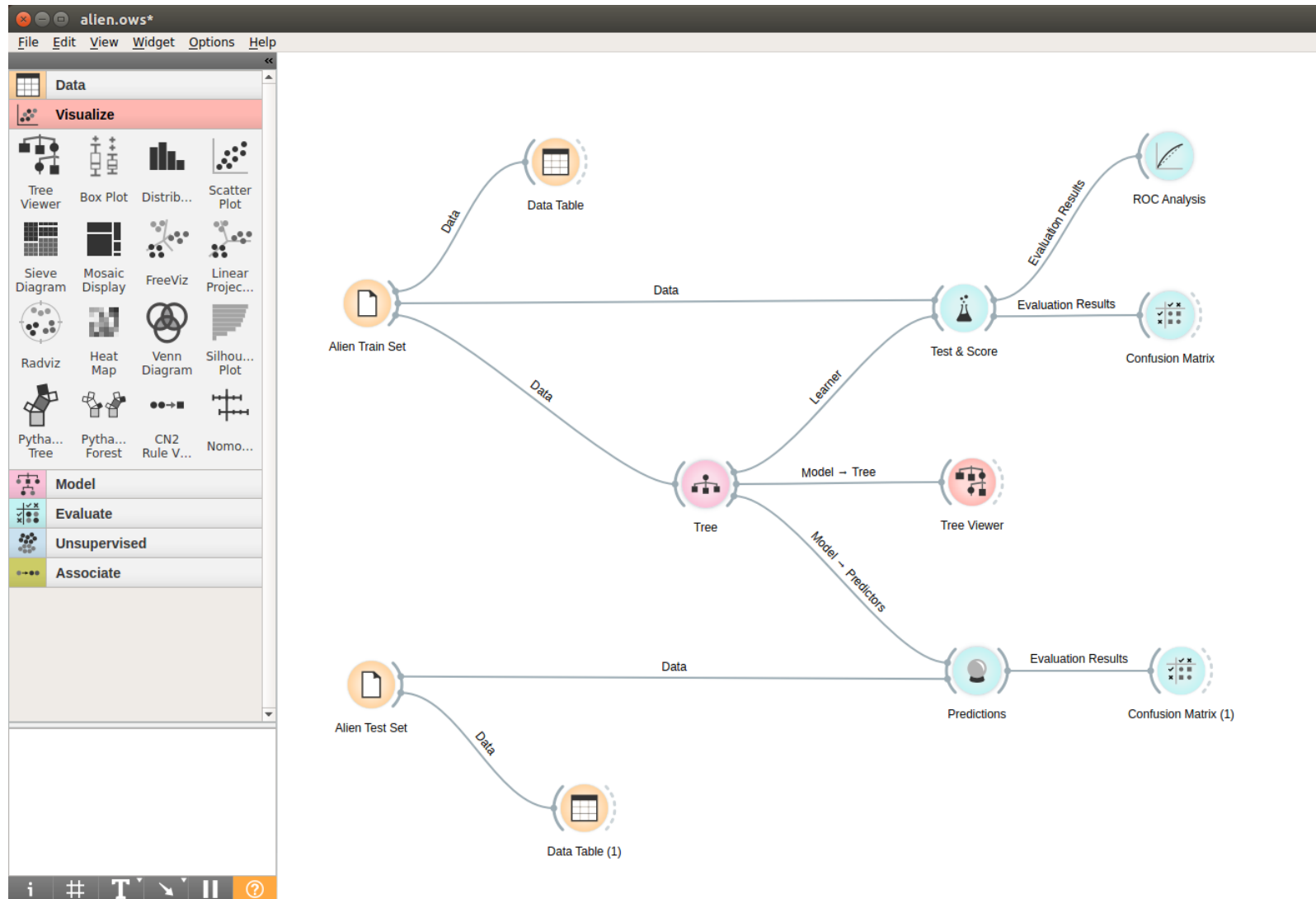


A **B** **C** **D** **E**

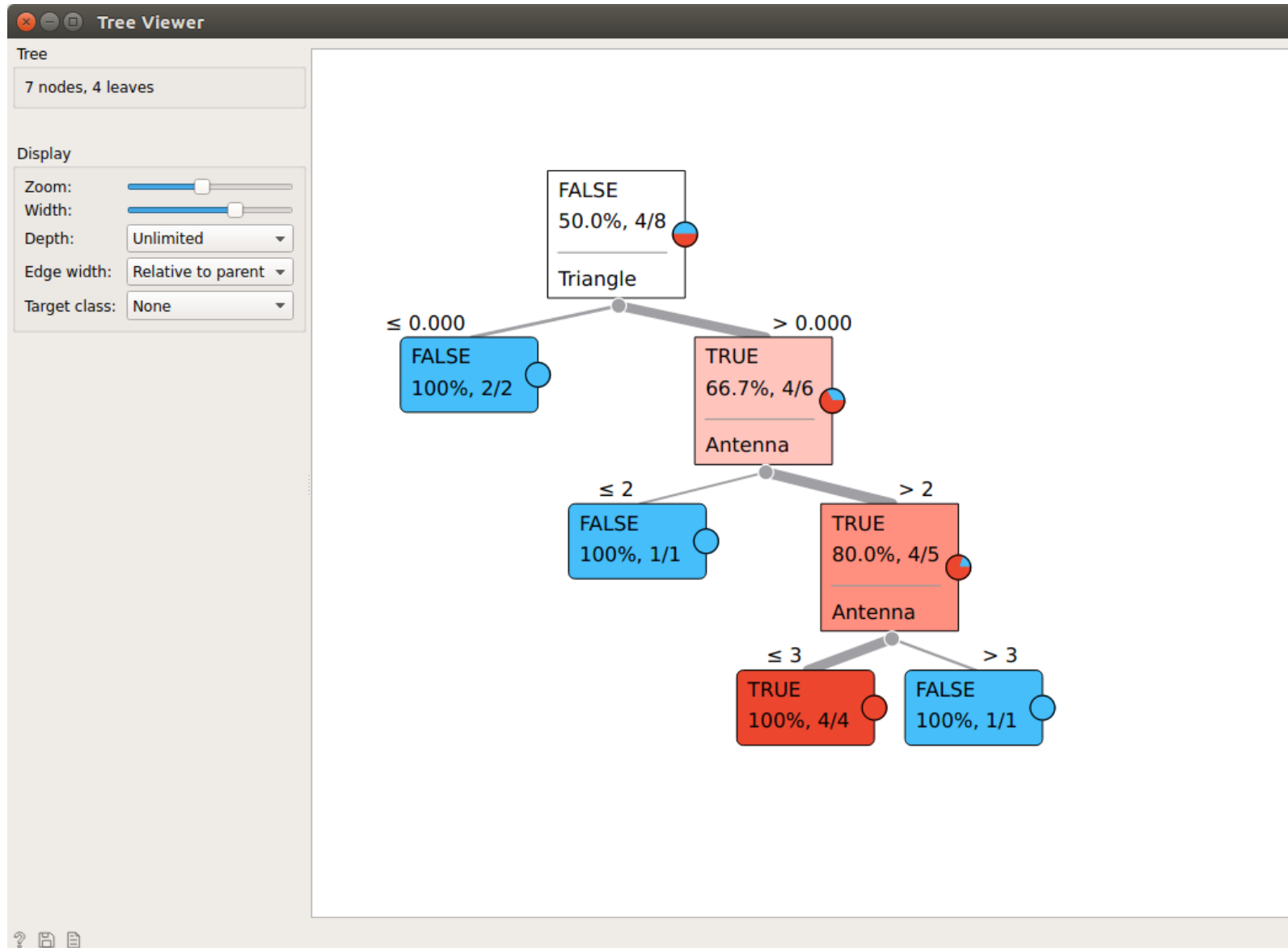
Test Data

SN	Triangle	Antenna	Teeth	Eyes	Alien
A	1	2	0	2	FALSE
B	3	2	1	2	FALSE
C	1	4	0	2	FALSE
D	1	3	0	2	TRUE
E	0	3	0	2	FALSE

3.3 KNOWLEDGE DISCOVERY IDENTIFY ALIENS



3.3 KNOWLEDGE DISCOVERY IDENTIFY ALIENS



3.3 KNOWLEDGE DISCOVERY IDENTIFY ALIENS

Aliens



Not aliens



Which one is alien?



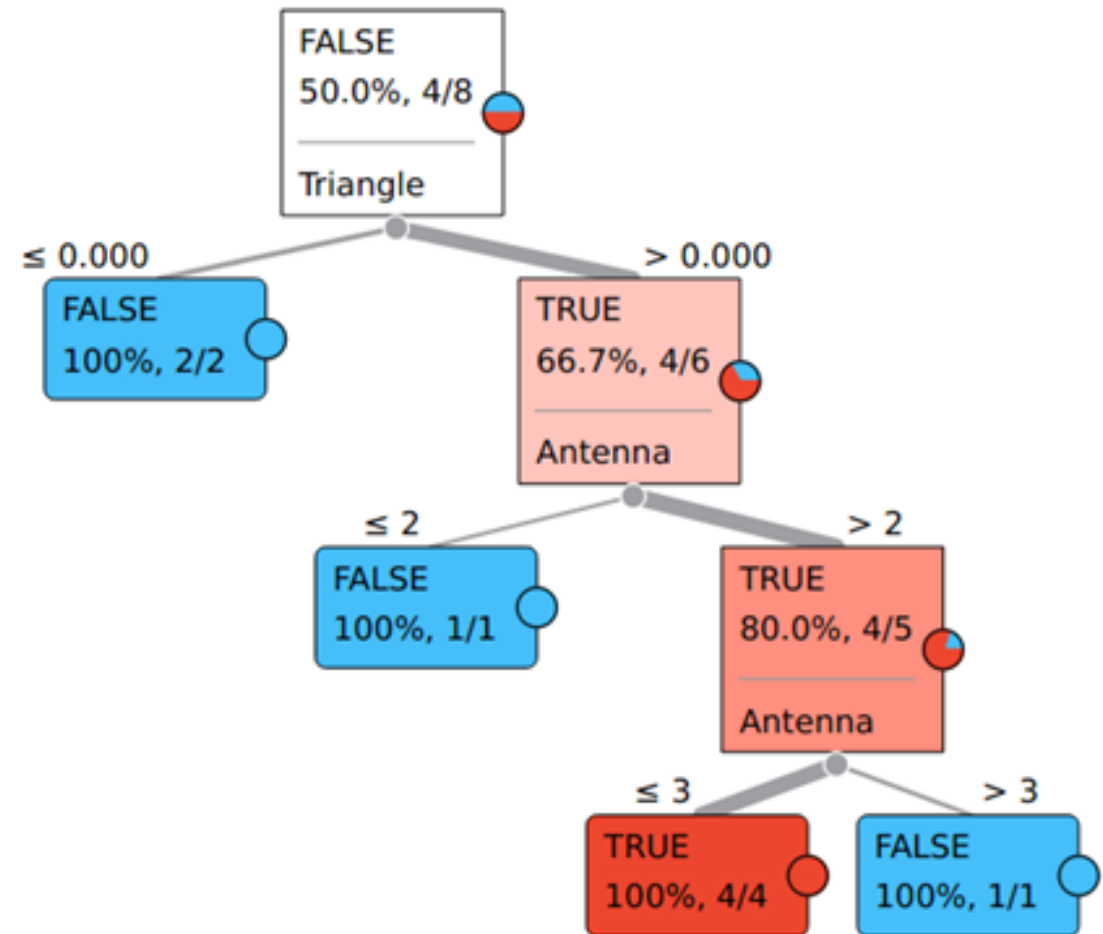
A

B

C

D

E



3.3 KNOWLEDGE DISCOVERY IDENTIFY ALIENS

Info

Data: 5 instances.
Predictors: 1
Task: Classification

Restore Original Order

Show

☒ Predicted class
☒ Predicted probabilities for:

FALSE
TRUE

☒ Draw distribution bars

Data View

☒ Show full dataset

Output

☒ Original data
☒ Predictions
☒ Probabilities

Tree

1 1.00 : 0.00 → FALSE

2 1.00 : 0.00 → FALSE

3 1.00 : 0.00 → FALSE

4 0.00 : 1.00 → TRUE

5 1.00 : 0.00 → FALSE

Alien	Triangle	Antenna	Teeth	Eyes
FALSE	1	2	0.000	2.000
FALSE	3	2	1.000	2.000
FALSE	1	4	0.000	2.000
TRUE	1	3	0.000	2.000
FALSE	0	3	0.000	2.000

Confusion Matrix (1)

Tree

Output

☒ Predictions ☒ Probabilities
☒

Send Automatically

Show: Number of instances

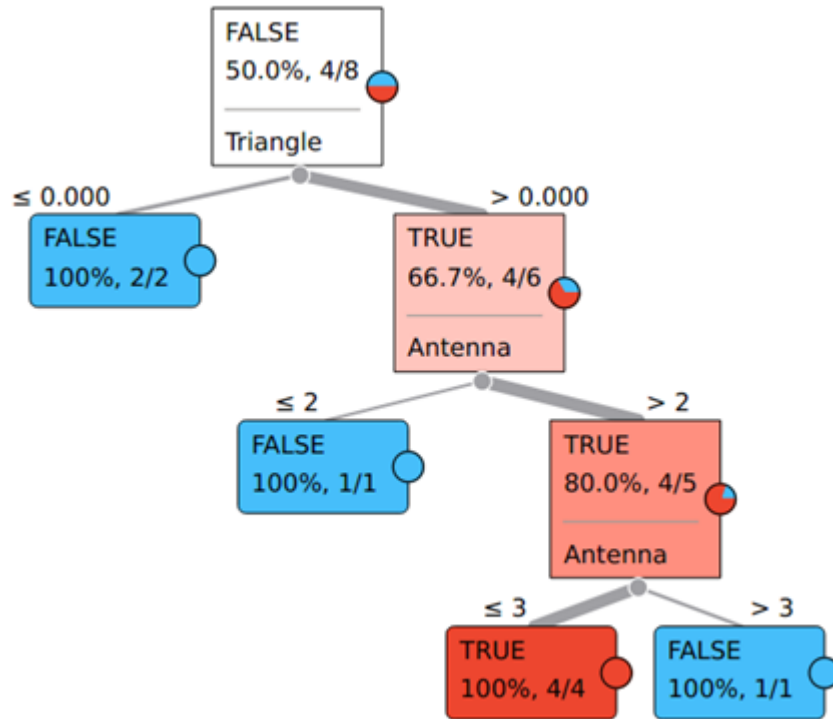
		Predicted		
		FALSE	TRUE	Σ
Actual	FALSE	4	0	4
	TRUE	0	1	1
Σ		4	1	5

Select Correct

Select Misclassified

Clear Selection

3.3 KNOWLEDGE DISCOVERY IDENTIFY ALIENS



Test Data

SN	Triangle	Antenna	Teeth	Eyes	Alien
A	1	2	0	2	FALSE
B	3	2	1	2	FALSE
C	1	4	0	2	FALSE
D	1	3	0	2	TRUE
E	0	3	0	2	FALSE
F	2	3	0	2	FALSE

		Predicted		
		FALSE	TRUE	Σ
Actual	FALSE	4	1	5
	TRUE	0	1	1
Σ		4	2	6

	Tree	Alien	Triangle	Antenna	Teeth	Eyes
1	1.00 : 0.00 → FALSE	FALSE	1	2	0.000	2.000
2	1.00 : 0.00 → FALSE	FALSE	3	2	1.000	2.000
3	1.00 : 0.00 → FALSE	FALSE	1	4	0.000	2.000
4	0.00 : 1.00 → TRUE	TRUE	1	3	0.000	2.000
5	1.00 : 0.00 → FALSE	FALSE	0	3	0.000	2.000
6	0.00 : 1.00 → TRUE	FALSE	2	3	0.000	2.000

Which one is alien?



A

B

C

D

E



F

The 'black swan': unseen before in historical data (no scenario/representation in training data)

END OF LECTURE NOTES