

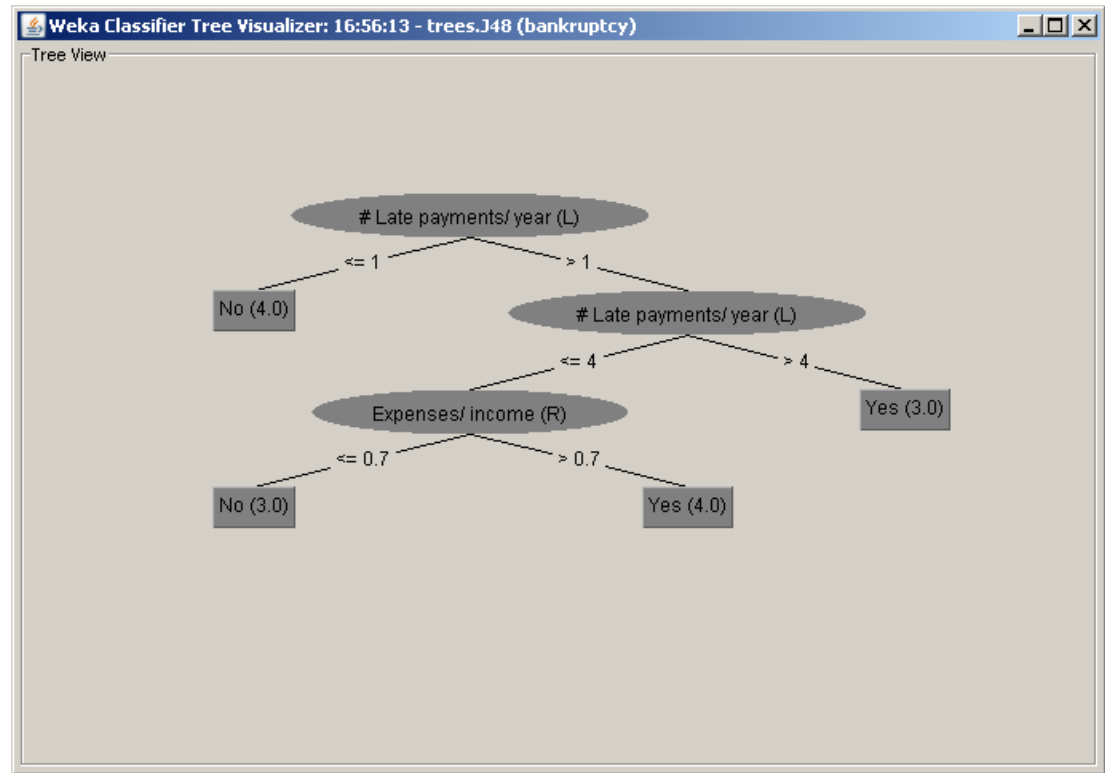
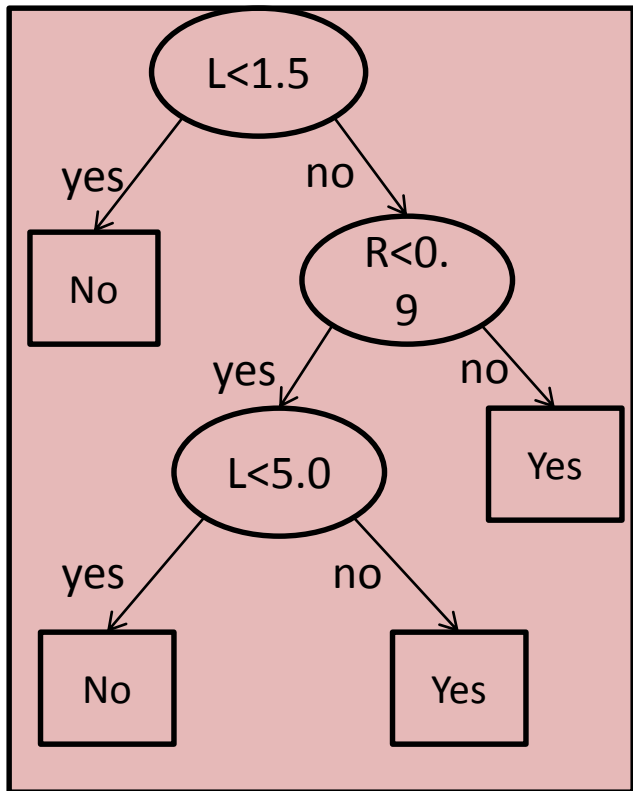
Decision trees:
numeric attributes
multi-valued attributes

Lab 1.2

Numeric values.

Bankruptcy dataset

Theory

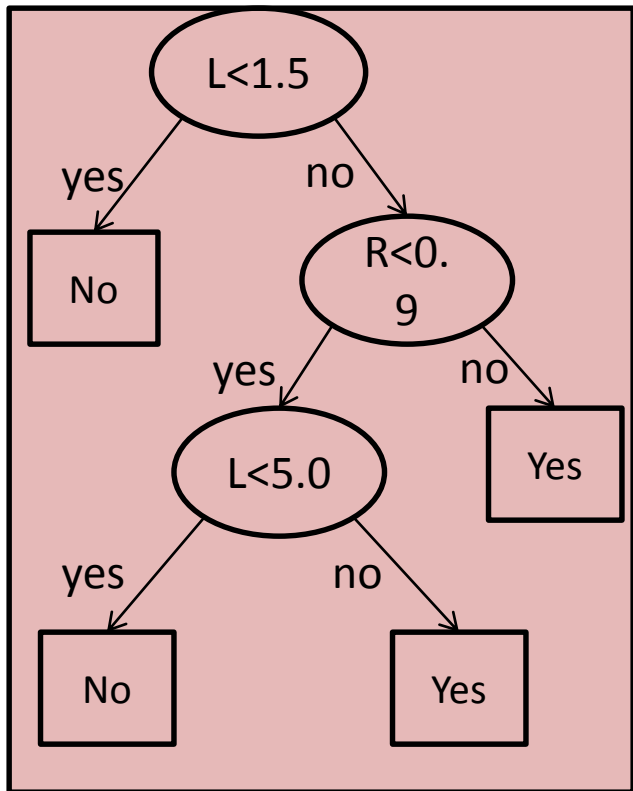


J48

Numeric values.

Bankruptcy dataset

Theory



Late payments/ year (L) < 1.5: No(4.0/0.0)

Late payments/ year (L) ≥ 1.5

Expenses/ income (R) < 0.9

Late payments/ year (L) < 5.0:

No(3.0/0.0)

Late payments/ year (L) ≥ 5.0:

Yes(3.0/0.0)

Expenses/ income (R) ≥ 0.9:

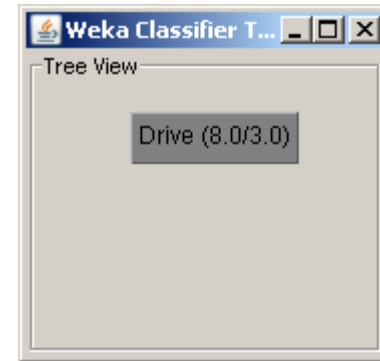
Yes(4.0/0.0)

SimpleCart

Multi-valued attributes.

My neighbor dataset

Temp	Precip	Day	Clothes	
22	None	Fri	Casual	Walk
3	None	Sun	Casual	Walk
10	Rain	Wed	Casual	Walk
30	None	Mon	Casual	Drive
20	None	Sat	Formal	Drive
25	None	Sat	Casual	Drive
-5	Snow	Mon	Casual	Drive
27	None	Tue	Casual	Drive
24	Rain	Mon	Casual	?

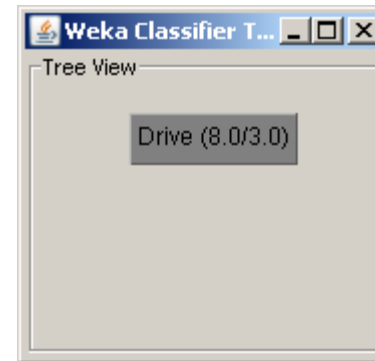


J48

Multi-valued attributes.

My neighbor dataset

Precip	Day	Clothes	
None	Fri	Casual	Walk
None	Sun	Casual	Walk
Rain	Wed	Casual	Walk
None	Mon	Casual	Drive
None	Sat	Formal	Drive
None	Sat	Casual	Drive
Snow	Mon	Casual	Drive
None	Tue	Casual	Drive
Rain	Mon	Casual	?



J48

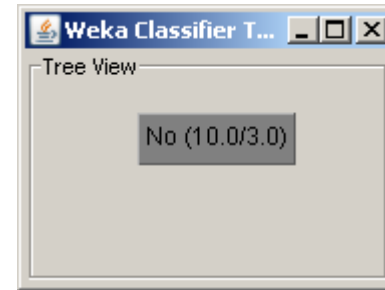
Day = Fri: Walk
Day = Sun: Walk
Day = Wed: Walk
Day = Mon: Drive
Day = Sat: Drive
Day = Tue: Drive

ID3 with Temp removed

What happened here?

TaxCheating dataset

ID	Refund	Marital status	Taxable income	Cheat
1	Yes	Single	125 K	No
2	No	Married	100 K	No
3	No	Single	70 K	No
4	Yes	Married	120 K	No
5	No	Divorced	95 K	Yes
6	No	Married	60 K	No
7	Yes	Divorced	220 K	No
8	No	Single	85 K	Yes
9	No	Married	75 K	No
10	No	Single	90 K	Yes



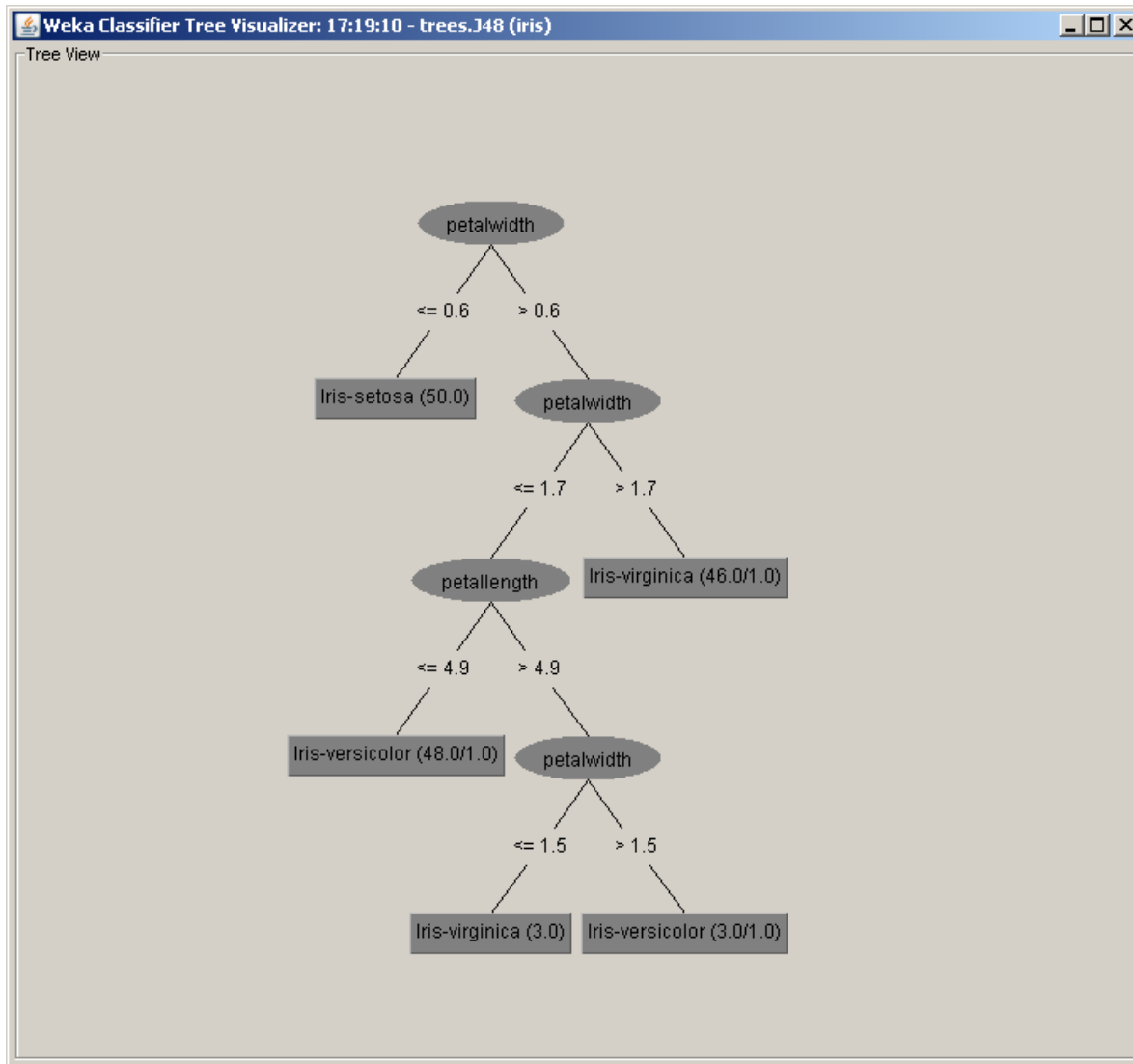
CART Decision Tree
: No(7.0/3.0)

J48

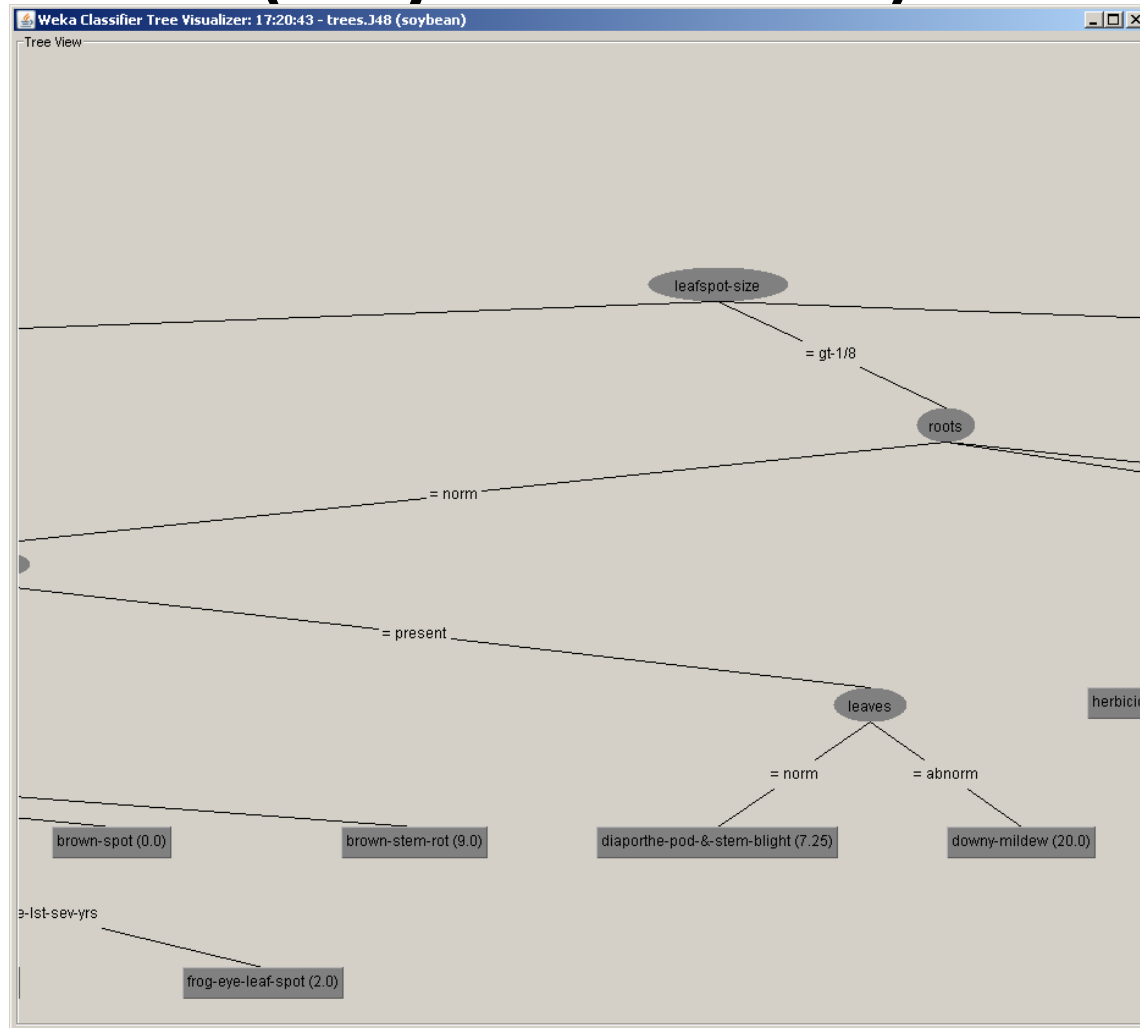
Id3

Taxable income = 125 K: No
 Taxable income = 100 K: No
 Taxable income = 70 K: No
 Taxable income = 120 K: No
 Taxable income = 95 K: Yes
 Taxable income = 60 K: No
 Taxable income = 220 K: No
 Taxable income = 85 K: Yes
 Taxable income = 75 K: No
 Taxable income = 90 K: Yes

Iris species classification (Iris.arff)



Soybean diseases classification (soybean.arff)



Soybean dataset

- Diagnosing soybean diseases
- 35 attributes with a small set of possible values

Diseases:

diaporthe stem canker, rhizoctonia root rot etc.

- 97.5% correct as compared to 72% correct by a human expert