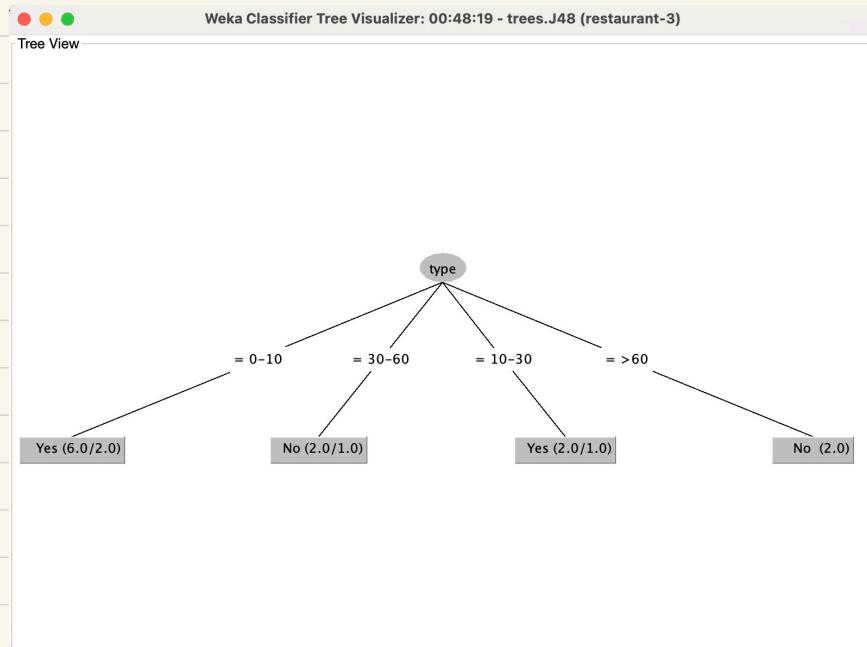


## Blatt 3

J48:

Restaurant:



Die Fehlerrate wird in der Matrix angegeben mit:

Correctly Classified Instances 8 66.6667 %

Incorrectly Classified Instances 4 33.3333 %

Kappa statistic 0.5636

Mean absolute error 0.119

Root mean squared error 0.244

Relative absolute error 50.5319 %

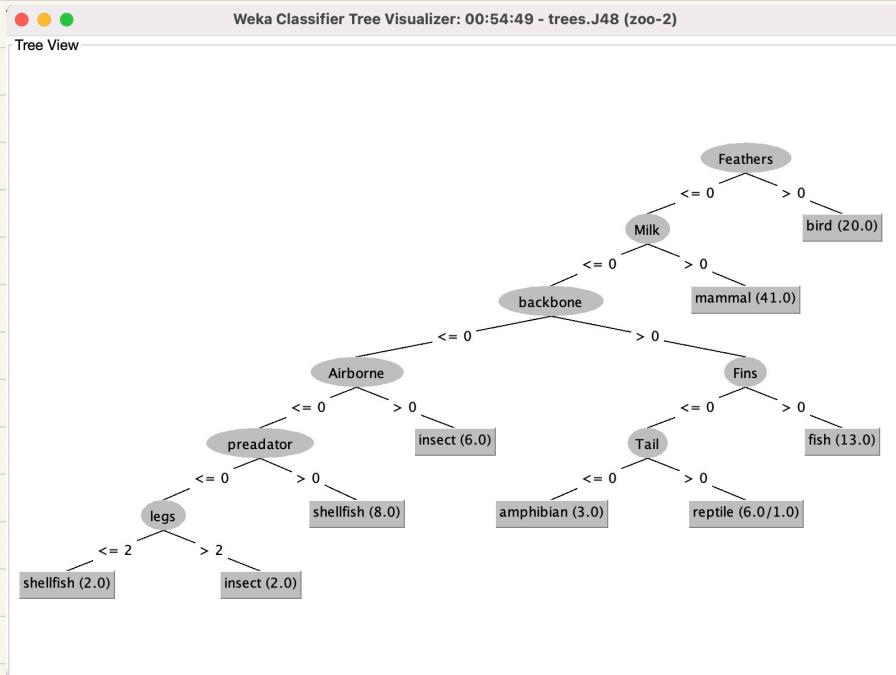
Root relative squared error 71.6089 %

Total Number of Instances 12

==== Confusion Matrix ===

a b c d e f g	<-- classified as
4 0 0 0 0 0 0	a = Yes
0 1 0 0 0 0 0	b = No
0 1 1 0 0 0 0	c = Yes
0 0 0 2 0 0 0	d = No
1 0 0 0 0 0 0	e = No
0 0 1 0 0 0 0	f = No
1 0 0 0 0 0 0	g = No

## J48 Zoo:



Die Fehlerrate wird in der Matrix angegeben mit:

Correctly Classified Instances	100	99.0099 %
Incorrectly Classified Instances	1	0.9901 %
Kappa statistic	0.987	
Mean absolute error	0.0047	
Root mean squared error	0.0486	
Relative absolute error	2.1552 %	
Root relative squared error	14.7377 %	
Total Number of Instances	101	

Datatypes in Meka:

**Numeric** sind integer und Reals.

**Nominal Values:** Haben eine nominal- spezifikation {[nominal-name1], [nominal-name2], [nominal-name3], ...}

**Strings:** Text Values. Nützlich in text-mining. Es ist möglich datasets mit string Attributen zu erstellen.

Aufgabe 3:

Id3:

Restaurant:

Correctly Classified Instances 12 100%

Incorrectly Classified Instances 0 0%

Kappa statistic 1

Mean absolute error 0

Root mean squared error 0

Relative absolute error 0%

Root relative squared error 0%

Total Number of Instances 12

==== Confusion Matrix ===

a	b	c	d	e	f	g	<-- classified as
4	0	0	0	0	0	0	a = Yes
0	1	0	0	0	0	0	b = No
0	0	2	0	0	0	0	c = Yes
0	0	0	2	0	0	0	d = No
0	0	0	0	1	0	0	e = No
0	0	0	0	0	1	0	f = No
0	0	0	0	0	0	1	g = No

Mir fällt auf, dass die Fehler Quote bei id3 deutlich deutlich geringer ist. Der Umstieg von csv auf araf hat bei mit keinen unterschied gemacht, der Wechsel der Algorithmen jedoch schon.