



Technology
Forge

www.technologyforge.net

Version 0.1

Weka Tutorial 5 – Association



Mother Knows Best...



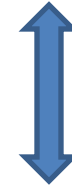
"A beer before bed means a better night's
sleep for the whole family!"





Got Beer?

Association



Market Basket Analysis



Got Diapers?

What things tend to “go together” in your market basket?





People who bought beer
also usually bought diapers.

**Example of
Market Basket
Analysis**

People who bought diapers
also usually bought beer.



Cash Register Checkout Data:

- Items in shopping cart
- Time of day
- Day of week
- Season of year
- Method of payment
- Location of store
- etc...



Associations?

Fact-Based Marketing Strategies:

- Floor plans
- Discounts
- Coupons
- etc...

What can you use associations for?



If the weather outside is:



Then to play, or not to play, that is the question!



TPONTPNom.xls dataset:

Outlook	Temp.	Humidity	Windy	Play
sunny	hot	high	false	no
sunny	hot	high	true	no
overcast	hot	high	false	yes
rainy	mild	high	false	yes
rainy	cool	normal	false	yes
rainy	cool	normal	true	no
overcast	cool	normal	true	yes
sunny	mild	high	false	no
sunny	cool	normal	false	yes
rainy	mild	normal	false	yes
sunny	mild	normal	true	yes
overcast	mild	high	true	yes
overcast	hot	normal	false	yes
rainy	mild	high	true	no



Outlook	Temp.	Humidity	Windy	Play
sunny	hot	high	false	no
sunny	hot	high	true	no
overcast	hot	high	false	yes
rainy	mild	high	false	yes
rainy	cool	normal	false	yes
rainy	cool	normal	true	no
overcast	cool	normal	true	yes
sunny	mild	high	false	no
sunny	cool	normal	false	yes
rainy	mild	normal	false	yes
sunny	mild	normal	true	yes
overcast	mild	high	true	yes
overcast	hot	normal	false	yes
rainy	mild	high	true	no

**Attribute
values
and counts**

Outlook	Temp.	Humidity	Windy	Play
sunny = 5	hot = 4	high = 7	true = 6	yes = 9
overcast = 4	mild = 6	normal = 7	false = 8	no = 5
rainy = 5	cool = 4			



Classification:

Input attributes Class

Outlook	Temp.	Humidity	Windy	Play
sunny	hot	high	false	no
sunny	hot	high	true	no
overcast	hot	high	false	yes
rainy	mild	high	false	yes
rainy	cool	normal	false	yes
rainy	cool	normal	true	no
overcast	cool	normal	true	yes
sunny	mild	high	false	no
sunny	cool	normal	false	yes
rainy	mild	normal	false	yes
sunny	mild	normal	true	yes
overcast	mild	high	true	yes
overcast	hot	normal	false	yes
rainy	mild	high	true	no

Goal: Predict whether the person will play, given weather conditions.



Outlook	Temp.	Humidity	Windy	Play
sunny	hot	high	false	no
sunny	hot	high	true	no
overcast	hot	high	false	yes
rainy	mild	high	false	
rainy	cool	normal	false	
rainy	cool	normal	true	
overcast	cool	normal	true	
sunny	mild	high	false	no
sunny	cool	normal	false	yes
rainy	mild	normal	false	yes
sunny	mild	normal	true	yes
overcast	mild	high	true	yes
overcast	hot	normal	false	yes
rainy	mild	high	true	no

For association, we look for things that “go together”

Some things that “go together”:

Outlook = sunny AND Temp. = hot go together 2 times,
 Outlook = sunny AND Temp. = cool go together only once.



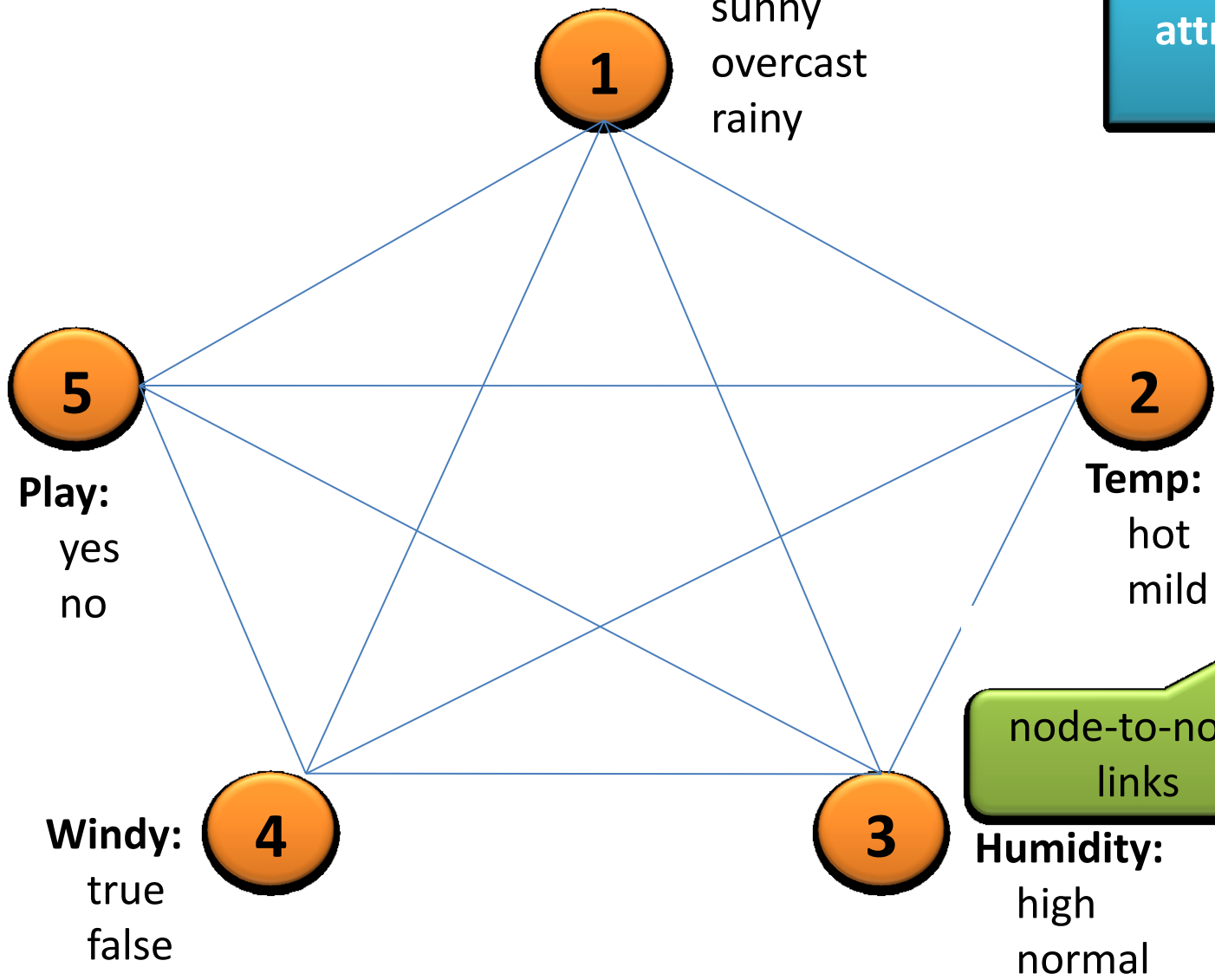
Outlook:
sunny
overcast
rainy

Two-item sets of attributes that can go together

- 1-2
- 2-3
- 3-4
- 4-5
- 5-1
- 1-3
- 2-4
- 3-5
- 4-1
- 5-2

10

node-to-node links

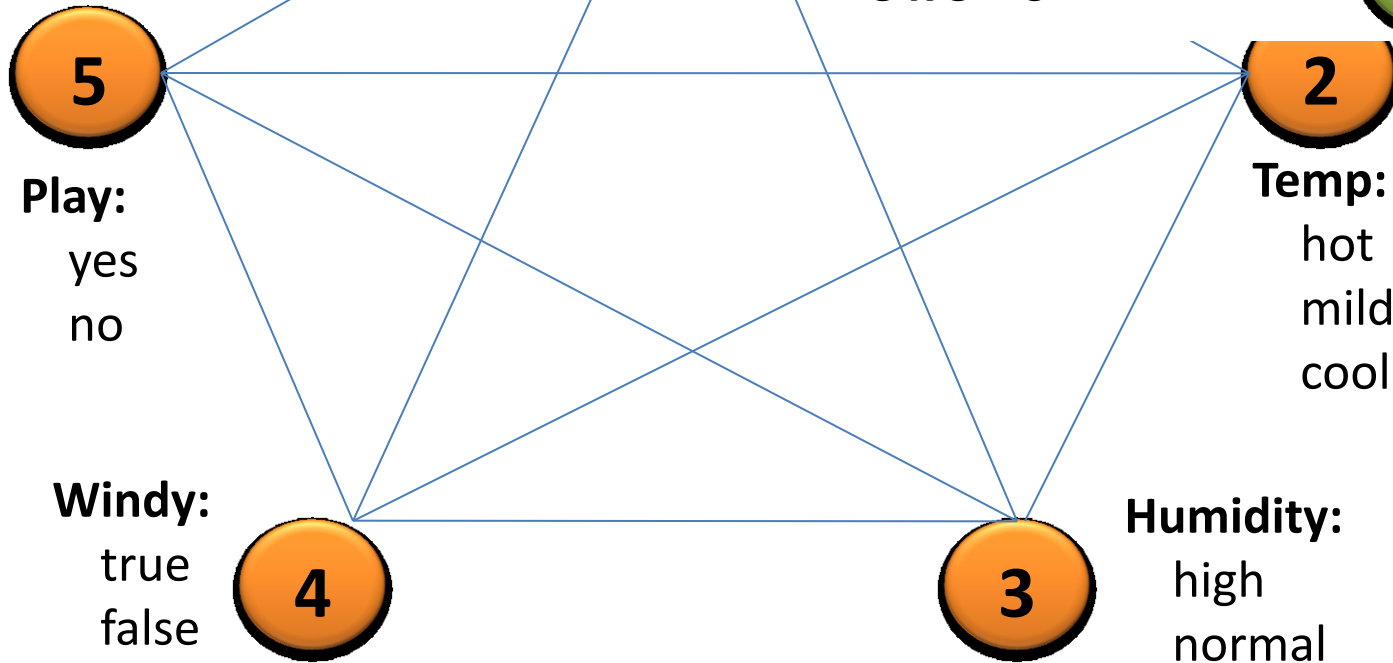


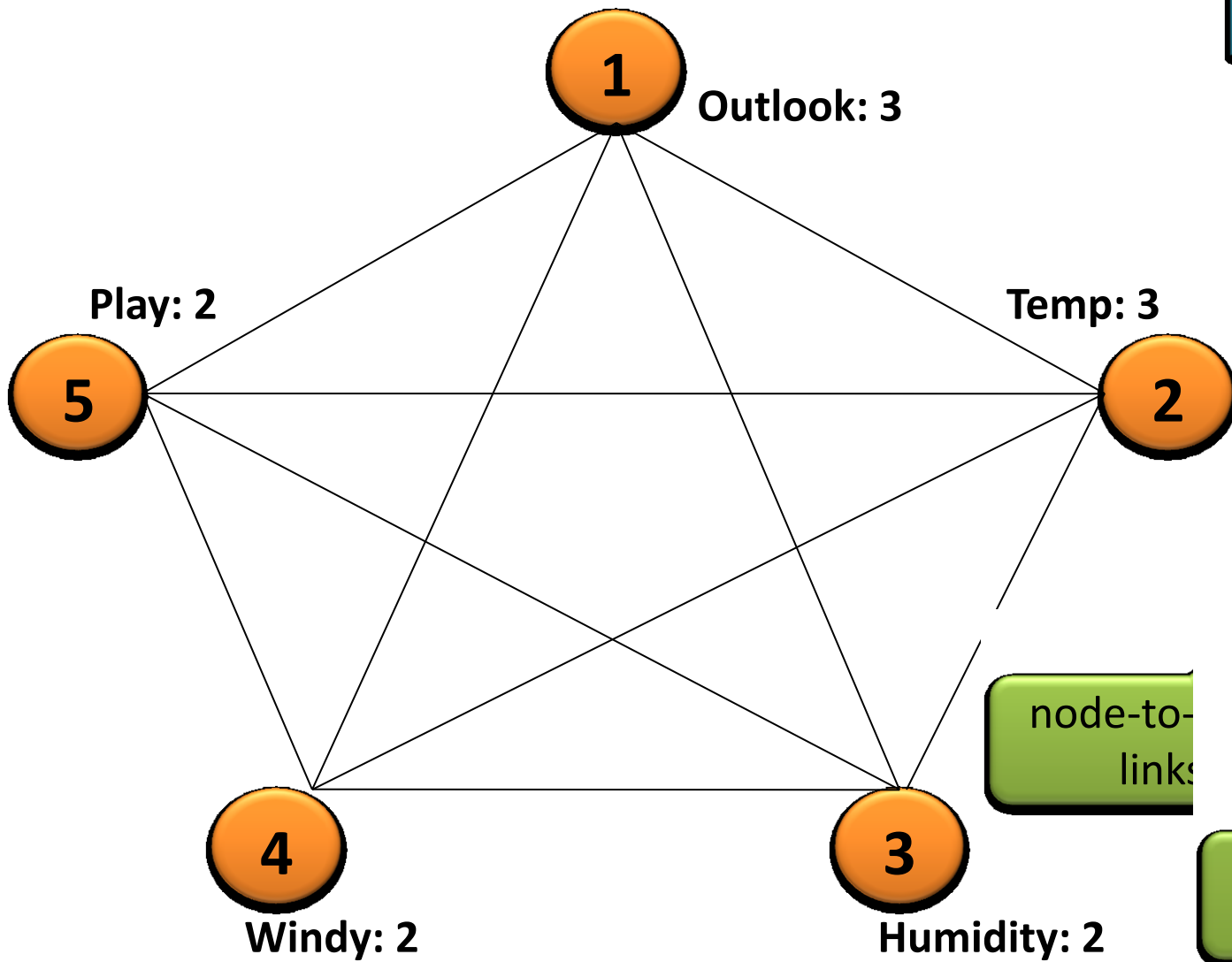
Combinations of Outlook and Temperature that can go together

Outlook:
sunny
overcast
rainy

sunny/hot
sunny/mild
sunny/cool
overcast/hot
overcast/mild
overcast/cool
rainy/hot
rainy/mild
rainy/cool

$3 \times 3 = 9$





All possible two-item sets

1-2: $3 \times 3 = 9$

2-3: $3 \times 2 = 6$

3-4: $2 \times 2 = 4$

4-5: $2 \times 2 = 4$

5-1: $2 \times 3 = 6$

1-3: $3 \times 2 = 6$

2-4: $3 \times 2 = 6$

3-5: $2 \times 2 = 4$

4-1: $2 \times 3 = 6$

5-2: $2 \times 3 = 6$

node-to-link

combinations for link



Occurrences of each pair

Outlook			Temp.			Humid.		Windy		Play	
sunny	overcast	rainy	hot	mild	cool	high	normal	true	false	yes	no
5	4	5	4	6	4	7	7	6	8	9	5

Outlook	sunny	5				2	2	1	3	2	2	3	2	3
	overcast	4				2	1	1	2	2	2	2	4	0
	rainy	5				0	3	2	2	3	2	3	3	2
Temp.	hot	4	2	2	0				3	1	1	3	2	2
	mild	6	2	1	3				4	2	3	3	4	2
	cool	4	1	1	2				0	4	2	2	3	1
Humid.	high	7	3	2	2	3	4	0			3	4	3	4
	normal	7	2	2	3	1	2	4			3	4	6	1
Windy	yes	6	2	2	2	1	3	2	3	3			3	3
	no	8	3	2	3	3	3	2	4	4			6	2
Play	yes	9	2	4	3	2	4	3	3	6	3	6		
	no	5	3	0	2	2	2	1	4	1	3	2		



		Outlook	Temp.	Humid.	Windy	Play			
		high	normal	high	normal	true	false	yes	no
		7	7	6	8	9	5		
		3	2	2	3	2	3		
		2	2	2	2	4	0		
		2	3	2	3	3	2		
		3	1	1	3	2	2		
		4	2	3	3	4	2		
	cool	4							
Humid.	high	7							
	normal	7							
W								3	3
PI								6	2

Temperature = cool and Humidity = normal occurs more often than Temperature = cool and Humidity = high

0 4

6 1

Humidity = normal and Play = yes occurs more often than Humidity = normal and Play = no



“Support” is the number of instances where a particular rule is correct.

									Humid.		Windy		Play	
			sunny	overca	rainy	hot	mild	cool	high	normal	true	false	yes	no
			5	4	5	4	6	4	7	7	6	8	9	5
Outlook	sunny	5				2	2	1	3	2	2	3	2	3
	overcast	4				2	1	1	2	2	2	2	4	0
	rainy	5				0	3	2	2	3	2	3	3	2
Temp.	hot	4							3	1	1	3	2	2
	mild	6							4	2	3	3	4	2
	cool	4							0	4	2	2	3	1
											3	4	3	4
											3	4	6	1
													3	3
													6	2

Temp. = cool AND Humid. = normal
occurs 4 times in the dataset,
so the *support* is 4.



Here are the 47 two-item sets
that have support $\Rightarrow 2$.

			Temp.			Humid.		Windy		Play				
			sunny	overcast	rainy	hot	mild	cool	high	normal	true	false	yes	no
			5	4	5	4	6	4	7	7	6	8	9	5
Outlook	sunny	5				2	2	1	3	2	2	3	2	3
	overcast	4				2	1	1	2	2	2	2	4	0
	rainy	5				0	3	2	2	3	2	3	3	2
Temp.	hot	4							3	1	1	3	2	2
	mild	6							4	2	3	3	4	2
	cool	4							0	4	2	2	3	1
Humid.	high	7									3	4	3	4
	normal	7									3	4	6	1
Windy	yes	6									3	4	3	3
	no	8									3	4	5	2
Play	yes	9												
	no	5												

Combinations with highest support:
Humidity = normal AND Play = yes
Windy = no and Play = yes



“Confidence” of the association rule
 “If Humidity = normal, then Play = yes”
 is 6/7, or 86%

			Humid.		Windy		Play				
			high	normal	true	false	yes	no			
			5	4	5	4	6	8	9	5	
Outlook	sunny	5				2	2	1	3	2	3

Combination
 Humidity = normal AND Play = yes
 occurs 6 times.

	normal	7							3	4	6	1
Windy	yes	6									3	3
	no	8									6	2
Play	yes	9										
	no	5										

Number of occurrences of
 Humidity = normal
 is 7.



Confidence of rule

If Play = yes, then Humidity = normal
is 6/9, or 67%

				Temp.		Humid.		Windy		Play		
				mild	cool	high	normal	true	false	yes	no	
sun	overcast	rain	hot	6	4	7	7	6	8	9	5	
Outlook											2	3
Temp.											4	0
Humid.											3	2
Windy											2	2
Play											4	2
high											3	1
normal											3	4
yes											6	1
no											3	3
yes											6	2
no												

Number of occurrences of
Play = yes
is 9.

Combination
Humidity = normal AND Play = yes
occurs 6 times.



THEN

		Outlook		Temp.		Humid.		Windy		Play	
		st								yes	no
IF	Outlook	sunny	5							0.40	0.60
		overcast	4							1.00	0.00
		rain								0.60	0.40
Temp.		hot								0.50	0.50
		mild								0.67	0.33
		cool								0.75	0.25
Humid.		high								0.43	0.57
		normal								0.86	0.14
Windy		yes							0.50		
		no							0.50		
Play	yes	9	0.22	0.44	0.33	0.22	0.44	0.33	0.33	0.67	
	no	5	0.60	0.00	0.40	0.40	0.40	0.20	0.80	0.20	

If Humidity = normal,
then Play = yes

If Play = yes,
then Humidity = normal

The two-item set of Humidity and Play produced two different rules.



Two-item set coverage

		Outlook			Temp.			Humid.		Windy		Play	
		sunny	overcast	rainy	hot	mild	cool	high	normal	true	false	yes	no
		5	4	5	4	6	4	7	7	6	8	9	5
Outlook	sunny	5			2	2	1	3	2	2	3	2	3
	overcast	4			2	1	1	2	2	2	2	4	0
	rainy	5			0	3	2	2	3	2	3	3	2
Temp.	hot	4	2	2	0			3	1	1	3	2	2
	mild	6	2	1	3			4	2	3	3	4	2
	cool	4	1	1	2			0	4	2	2	3	1
Humid.	high	7	3	2	2	3	4	0		3	4	3	4
	normal	7	2	2	3	1	2	4		3	4	6	1
Windy	yes	6	2	2	2	1	3	2	3	3		3	3
	no	8	3	2	3	3	3	2	4	4		6	2
Play	yes	9	2	4	3	2	4	3	3	6	3	6	
	no	5	3	0	2	2	2	1	4	1	3	2	

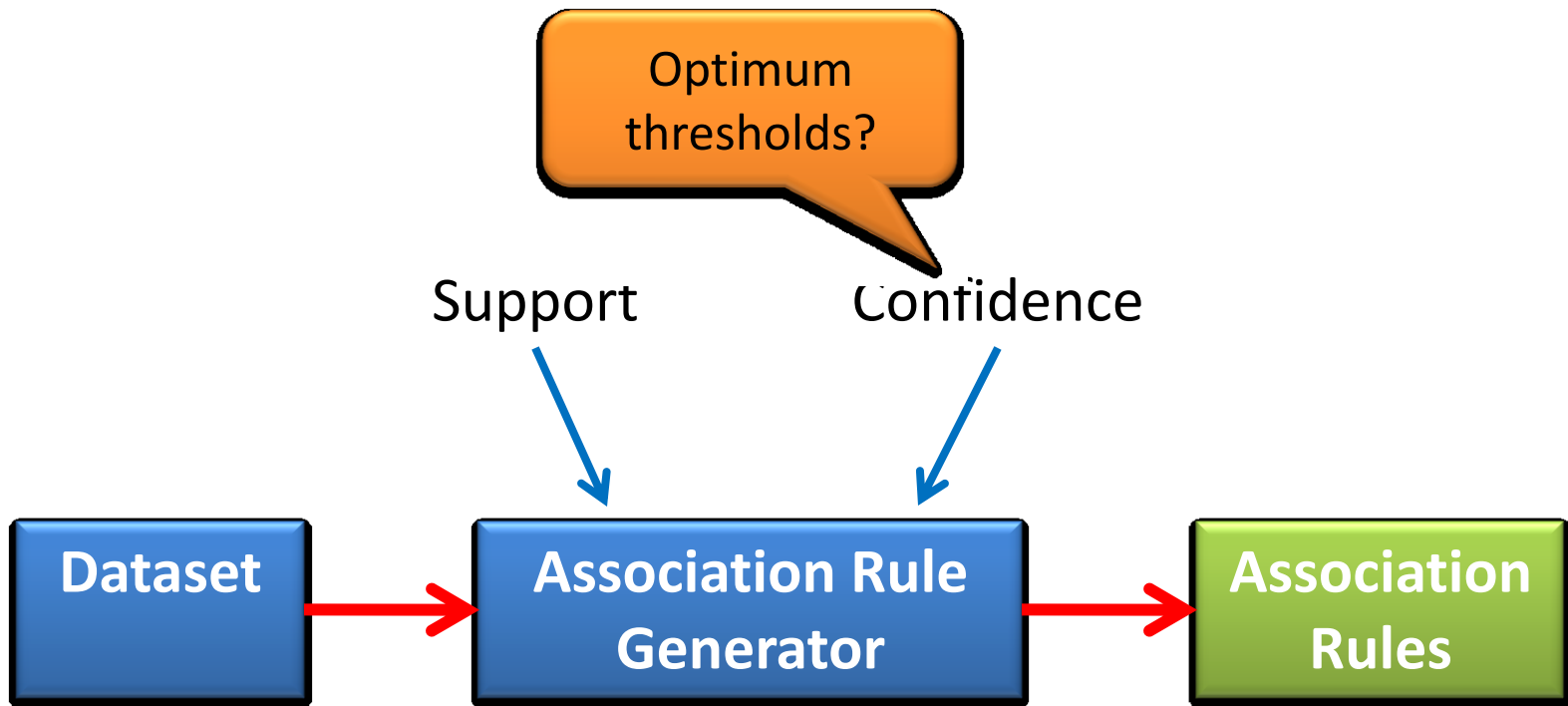
For thresholds set at:
accuracy = 100%
coverage = >2

Two-item set association rules:
If Temperature = cool,
then Humidity = normal
If Outlook = overcast,
then Play = yes

Two-item set accuracy

		THEN			Outlook			Temp.			Humid.		Windy		Play	
		sunny	overcast	rainy	hot	mild	cool	high	normal	true	false	yes	no			
		5	4	5	4	6	4	7	7	6	8	9	5			
IF Outlook	sunny	5			0.40	0.40	0.20	0.60	0.40	0.40	0.60	0.40	0.60			
	overcast	4			0.50	0.25	0.25	0.50	0.50	0.50	0.50	1.00	0.00			
	rainy	5			0.00	0.60	0.40	0.40	0.60	0.40	0.60	0.60	0.40			
Temp.	hot	4	0.50	0.50	0.00			0.75	0.25	0.25	0.75	0.50	0.50			
	mild	6	0.33	0.17	0.50			0.67	0.33	0.50	0.50	0.67	0.33			
	cool	4	0.25	0.25	0.50			0.00	1.00	0.50	0.50	0.75	0.25			
Humid.	high	7	0.43	0.29	0.29	0.43	0.57	0.00		0.43	0.57	0.43	0.57			
	normal	7	0.29	0.29	0.43	0.14	0.29	0.57		0.43	0.57	0.86	0.14			
Windy	yes	6	0.33	0.33	0.33	0.17	0.50	0.33	0.50	0.50		0.50	0.50			
	no	8	0.38	0.25	0.38	0.38	0.38	0.25	0.50	0.50		0.75	0.25			
Play	yes	9	0.22	0.44	0.33	0.22	0.44	0.33	0.33	0.67	0.33	0.67				
	no	5	0.60	0.00	0.40	0.40	0.40	0.20	0.80	0.20	0.60	0.40				





Predictive apriori association algorithm:

- Optimally combines support and confidence into **predictive accuracy**,
- Requires only that user specify number of rules generated.



Load TPONTPNom.arff dataset

Weka Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Open file... | Open URL... | Open DB... | Generate... | Undo

Filter: Choose None [Apply]

Current relation: TPONTPNom (Attributes: 5)

Attributes: All | None | Invert | Pattern

No.	Name
1	<input type="checkbox"/> Outlook
2	<input type="checkbox"/> Temp.
3	<input type="checkbox"/> Humidity
4	<input type="checkbox"/> Windy
5	<input checked="" type="checkbox"/> Play

Selected attribute: Name: Play, Missing: 0 (0%), Distinct: 2, Type: Nominal, Unique: 0 (0%)

No.	Label	Count
1	no	5
2	yes	9

Class: Play (Nom) [Visualize All]

Status: OK [Log] x 0



Weka Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Open file... | Open URL... | Open DB... | Generate... | Undo | Edit... | Save...

Filter: Choose None Apply

Current relation: Relation: TPONTPNom Instances: 14

Attributes: All N

No.	Name
1	<input checked="" type="checkbox"/> Outlook
2	<input type="checkbox"/> Temp.
3	<input type="checkbox"/> Humidity
4	<input type="checkbox"/> Windy
5	<input type="checkbox"/> Play

Remove

Attribute: Outlook
 Outlook 0 (0%) Distinct: 3 Type: Nominal Unique: 0 (0%)

Label	Count
1 sunny	5
2 overcast	4
3 rainy	5

Class: Play (Nom) Visualize All

Status: OK Log x 0

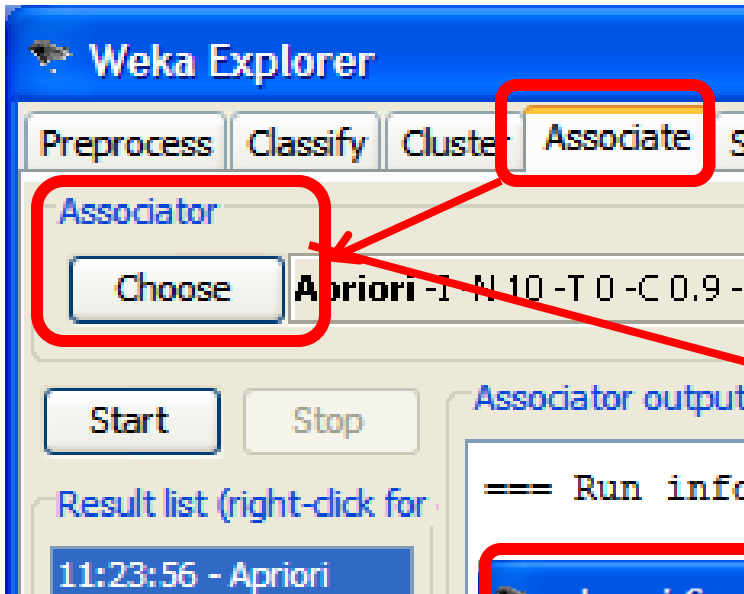
Show Outlook relative to Play

1 Outlook

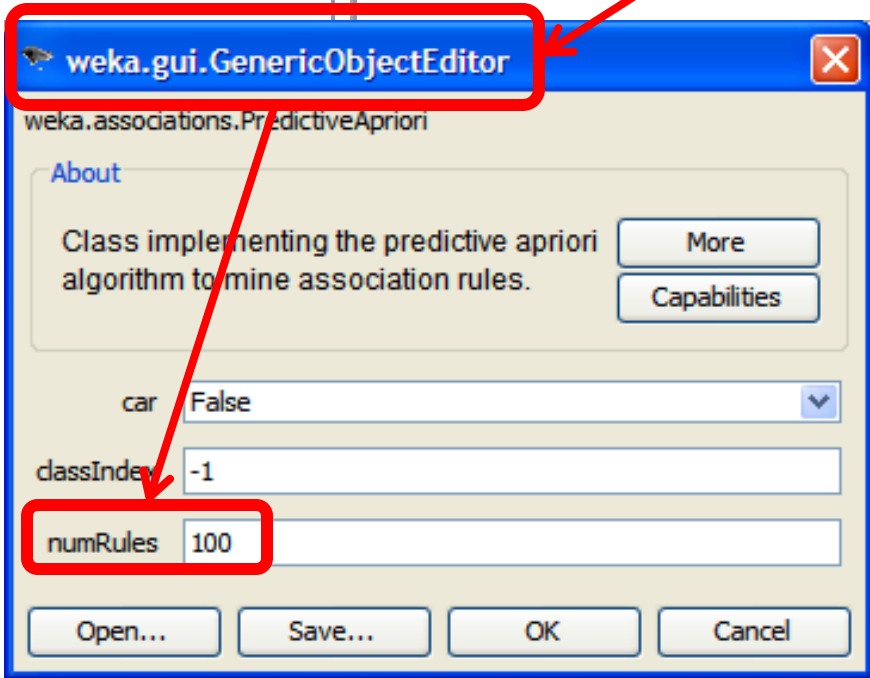
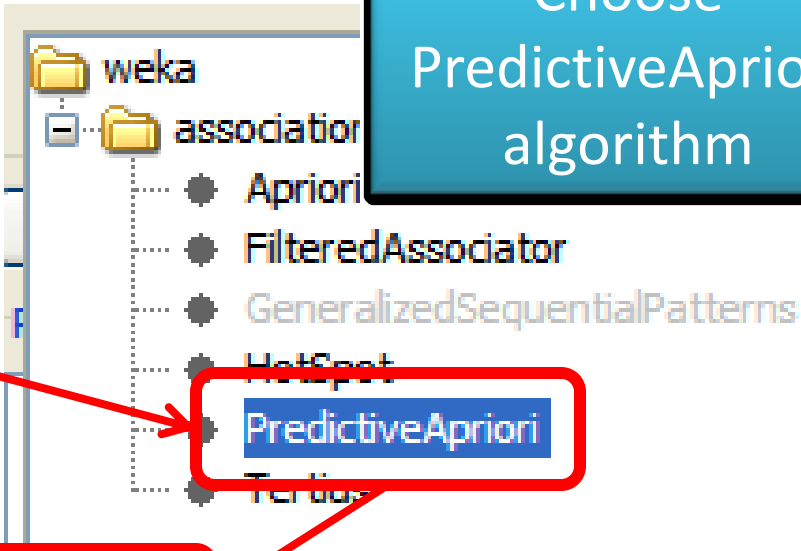
2 Temp.

Class: Play (Nom)





Choose PredictiveApriori algorithm



Best 100 rules for TPONTP dataset

1. Outlook=overcast 4 ==> Play=yes 4 acc:(0.95583)
2. Temp.=cool 4 ==> Humidity=normal 4 acc:(0.95583)
3. Humidity=normal Windy=false 4 ==> Play=yes 4 acc:(0.95583)
4. Outlook=sunny Humidity=high 3 ==> Play=no 3 acc:(0.92532)
5. Outlook=sunny Play=no 3 ==> Humidity=high 3 acc:(0.92532)
6. Outlook=rainy Windy=false 3 ==> Play=yes 3 acc:(0.92532)
7. Outlook=rainy Play=yes 3 ==> Windy=false 3 acc:(0.92532)
8. Outlook=sunny Temp.=hot 2 ==> Humidity=high Play=no 2 acc:(0.86872)
9. Outlook=sunny Humidity=normal 2 ==> Play=yes 2 acc:(0.86872)
10. Outlook=sunny Play=yes 2 ==> Humidity=normal 2 acc:(0.86872)
11. Outlook=overcast Temp.=hot 2 ==> Windy=false Play=yes 2 acc:(0.86872)
12. Outlook=overcast Windy=false 2 ==> Temp.=hot Play=yes 2 acc:(0.86872)
13. Outlook=rainy Humidity=high 2 ==> Temp.=mild 2 acc:(0.86872)
14. Outlook=rainy Windy=true 2 ==> Play=no 2 acc:(0.86872)
15. Outlook=rainy Play=no 2 ==> Windy=true 2 acc:(0.86872)
16. Temp.=hot Play=no 2 ==> Outlook=sunny Humidity=high 2 acc:(0.86872)
17. Temp.=hot Play=yes 2 ==> Outlook=overcast Windy=false 2 acc:(0.86872)
18. Temp.=mild Humidity=normal 2 ==> Play=yes 2 acc:(0.86872)
19. Temp.=mild Play=no 2 ==> Humidity=high 2 acc:(0.86872)
20. Temp.=cool Windy=false 2 ==> Humidity=normal Play=yes 2 acc:(0.86872)
21. Windy=false Play=no 2 ==> Outlook=sunny Humidity=high 2 acc:(0.86872)
22. Temp.=mild Windy=false Play=yes 2 ==> Outlook=rainy 2 acc:(0.86872)
23. Humidity=normal 7 ==> Play=yes 6 acc:(0.70415)
24. Play=no 5 ==> Humidity=high 4 acc:(0.59669)
25. Windy=false 8 ==> Play=yes 6 acc:(0.56853)



Occurrences of Outlook = overcast

Support for Outlook = overcast AND Play = yes

```

1. Outlook=overcast 4 ==> Play=yes 4 acc:(0.95583)
2. Temp.=cool 4 ==> Humidity=normal 4 acc:(0.86872)
3. Humidity=normal 4 ==> Play=yes 4
4. Outlook=overcast 4 ==> Play=yes 4
5. Outlook=sunny 3 ==> Play=yes 3
6. Outlook=overcast 3 ==> Play=yes 3
7. Outlook=overcast 3 ==> Play=no 3
8. Outlook=overcast 3 ==> Play=yes 3
9. Outlook=overcast 3 ==> Play=no 3
10. Outlook=overcast 3 ==> Play=yes 3
11. Outlook=overcast 3 ==> Play=no 3
12. Outlook=overcast 3 ==> Play=yes 3
13. Outlook=rainy Humidity=high 2 ==> Temp.=mild 2 acc:(0.86872)
14. Outlook=rainy Windy=true 2 ==> Play=no 2 acc:(0.86872)
15. Outlook=rainy Play=no 2 ==> Windy=true 2 acc:(0.86872)
16. Temp.=hot Play=no 2 ==> Outlook=sunny Humidity=high 2 acc:(0.86872)
17. Temp.=hot Play=yes 2 ==> Outlook=overcast Windy=false 2 acc:(0.86872)
18. Temp.=mild Humidity=normal 2 ==> Play=yes 2 acc:(0.86872)
19. Temp.=mild Play=no 2 ==> Humidity=high 2 acc:(0.86872)
20. Temp.=cool Windy=false 2 ==> Humidity=normal Play=yes 2 acc:(0.86872)

```

Confidence of If Outlook = overcast, then Play = yes is 4/4 = 100%

Predictive accuracy = 95.6%



Associator output

```
1. Outlook=overcast 4 ==> Play=yes 4    acc:(0.95583)
2. Temp.=cool 4 ==> Humidity=normal 4    acc:(0.95583)
3. Humidity=normal Windy=false 4 ==> Play=yes 4    acc:(0.95583)
4. Outlook=sunny Humidity=high 3 ==> Play=no 3    acc:(0.92532)
5. Outlook=sunny Play=no 3 ==> Humidity=high 3    acc:(0.92532)
6. Outlook=rainy Windy=false 3 ==> Play=yes 3    acc:(0.92532)
7. Outlook=rainy Play=yes 3 ==> Windy=false 3    acc:(0.92532)
8. Outlook=sunny Temp.=hot 2 ==> Humidity=high Play=no 2    acc:(0.86872)
9. Outlook=sunny Humidity=normal 2 ==> Play=yes 2    acc:(0.86872)
10. Outlook=sunny Play=yes 2 ==> Humidity=normal 2    acc:(0.86872)
11. Outlook=overcast Temp.=hot 2 ==> Windy=false Play=yes 2    acc:(0.86872)
12. Outlook=overcast Windy=false 2 ==> Temp.=hot Play=yes 2    acc:(0.86872)
13. Outlook=rainy Humidity=high 2 ==> Temp.=mild 2    acc:(0.86872)
14. Outlook=rainy Windy=true 2 ==> Play=no 2    acc:(0.86872)
15. Outlook=rainy Play=no 2 ==> Windy=true 2    acc:(0.86872)
16. Temp.=hot Play=no 2 ==> Outlook=sunny Humidity=high 2    acc:(0.86872)
17. Temp.=hot Play=yes 2 ==> Outlook=overcast Windy=false 2    acc:(0.86872)
18. Temp.=mild Humidity=normal 2 ==> Play=yes 2    acc:(0.86872)
19. Temp.=mild Play=no 2 ==> Humidity=high 2    acc:(0.86872)
20. Temp.=cool Windy=false 2 ==> Humidity=normal Play=yes 2    acc:(0.86872)
```



Input attributes

Class

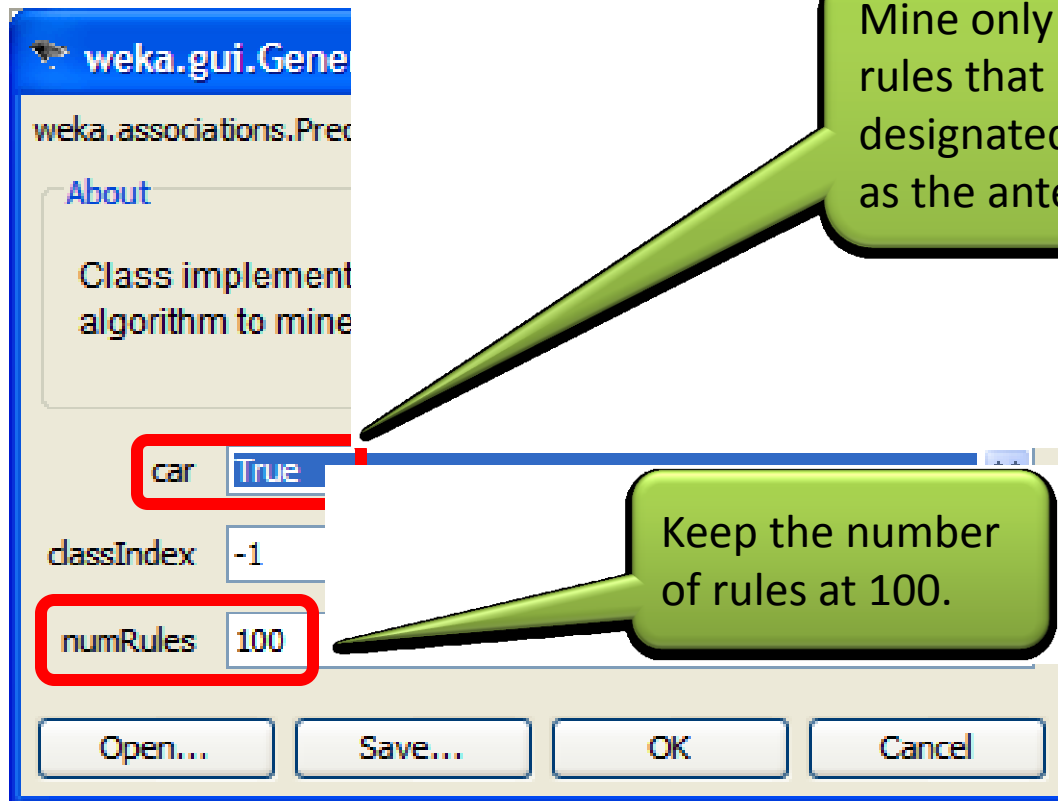
Outlook	Temp.	Humidity	Windy	Play
hot	hot	high	false	no
hot	hot	high	true	no
hot	hot	high	false	yes
mild	hot	high	false	yes
rainy	cool	normal	false	yes
rainy	cool	normal	true	no
overcast	cool	normal	true	yes
sunny	mild	high	false	no
rainy	cool	normal	false	yes
rainy	mild	normal	false	yes
rainy	mild	normal	true	yes
overcast	mild	high	true	yes
overcast	hot	normal	false	yes
rainy	mild	high	true	no

Classification Goal: Predict whether the person will play, given weather conditions.

Association rule:
If Humidity = normal
AND Windy = false,
then Play = yes

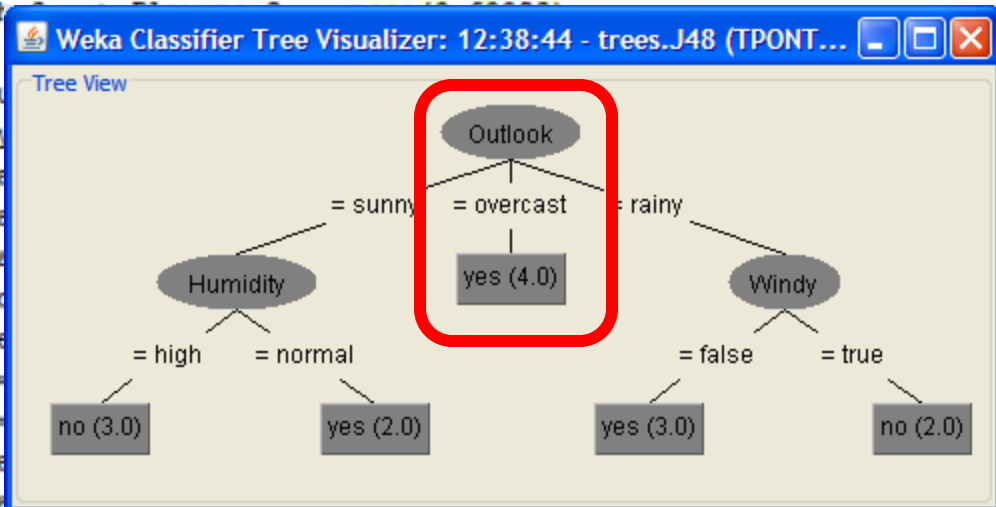
Question: Is there a linkage between classification and association?





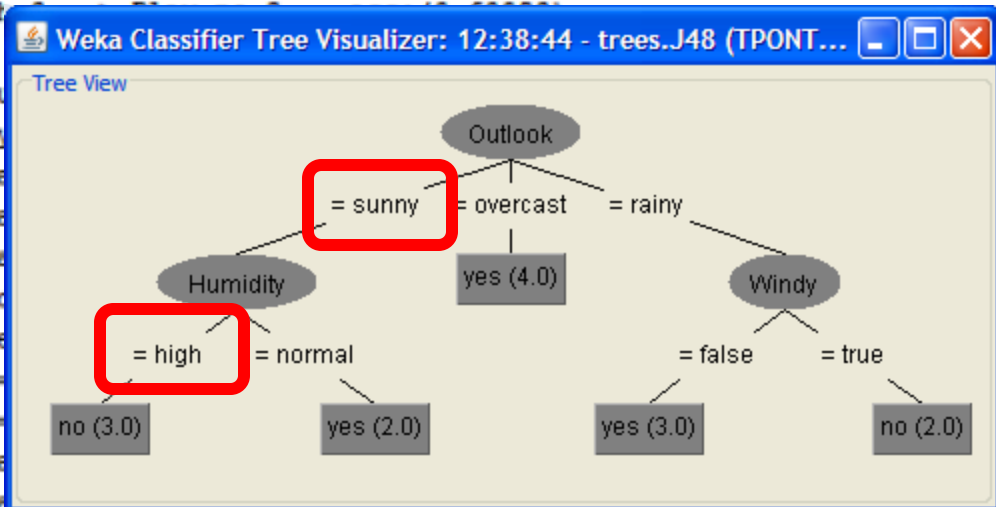
Associator output

```
1. Outlook=overcast 4 ==> Play=yes 4    acc:(0.97484)
2. Humidity=normal Windy=false 4 ==> Play=yes 4    acc:(0.97484)
3. Outlook=sunny Humidity=high 3 ==> Play=no 3    acc:(0.96246)
4. Outlook=rainy Windy=false 3 ==> Play=yes 3    acc:(0.96246)
5. Outlook=sunny Temp.=hot 2 ==> Play=no 2    acc:(0.94114)
6. Outlook=sunny Humidity=normal 2 ==> Play=yes 2    acc:(0.94114)
7. Outlook=rainy Windy=true 2 ==> Play=no 2    acc:(0.94114)
8. Temp.=mild Humidity=normal 2 ==> Play=yes 2    acc:(0.94114)
9. Temp.=cool Windy=false 2 ==> Play=yes 2    acc:(0.94114)
10. Humidity=normal 7 ==> Play=yes 6    acc:(0.78559)
11. Temp.=cool 4 ==> Play=yes 3    acc:(0.64518)
12. Windy=false 8 ==> Play=yes 6    acc:(0.62995)
13. Outlook=sunny Windy=false 3 ==> Play=no 2    acc:(0.60023)
14. Outlook=rainy Temp.=mild 3 ==> Play=yes 2    acc:(0.60023)
15. Temp.=hot Humidity=high 3 ==> Play=no 2    acc:(0.60023)
16. Temp.=mild Windy=true 3 ==> Play=no 2    acc:(0.60023)
17. Humidity=high Windy=true 3 ==> Play=no 2    acc:(0.60023)
18. Temp.=mild 6 ==> Play=yes 4    acc:(0.66667)
19. Outlook=sunny 5 ==> Play=yes 3    acc:(0.60000)
20. Outlook=rainy 5 ==> Play=no 2    acc:(0.40000)
21. Humidity=high 7 ==> Play=no 3    acc:(0.42857)
22. Temp.=hot 4 ==> Play=no 2    acc:(0.50000)
23. Temp.=hot 4 ==> Play=yes 2    acc:(0.50000)
24. Windy=true 6 ==> Play=no 3    acc:(0.50000)
25. Windy=true 6 ==> Play=yes 3    acc:(0.50000)
26. Humidity=high 7 ==> Play=yes 3    acc:(0.42857)
27. Outlook=sunny 5 ==> Play=no 2    acc:(0.40000)
28. Outlook=rainy 5 ==> Play=no 2    acc:(0.50124)
29. Temp.=mild 6 ==> Play=no 2    acc:(0.4736)
30. Windy=false 8 ==> Play=no 2    acc:(0.42451)
```



Associator output

1. Outlook=overcast 4 ==> Play=yes 4 acc:(0.97484)
2. Humidity=normal Windy=false 4 ==> Play=yes 4 acc:(0.97484)
3. Outlook=sunny Humidity=high 3 ==> Play=no 3 acc:(0.96246)
4. Outlook=rainy Windy=false 3 ==> Play=yes 3 acc:(0.96246)
5. Outlook=sunny Temp.=hot 2 ==> Play=no 2 acc:(0.94114)
6. Outlook=sunny Humidity=normal 2 ==> Play=yes 2 acc:(0.94114)
7. Outlook=rainy Windy=true 2 ==> Play=no 2 acc:(0.94114)
8. Temp.=mild Humidity=normal 2 ==> Play=yes 2 acc:(0.94114)
9. Temp.=cool Windy=false 2 ==> Play=yes 2 acc:(0.94114)
10. Humidity=normal 7 ==> Play=yes 6 acc:(0.78559)
11. Temp.=cool 4 ==> Play=yes 3 acc:(0.64518)
12. Windy=false 8 ==> Play=yes 6 acc:(0.62995)
13. Outlook=sunny Windy=false 3 ==> Play=no 2 acc:(0.60023)
14. Outlook=rainy Temp.=mild 3 ==> Play=yes 2 acc:(0.60023)
15. Temp.=hot Humidity=high 3 ==> Play=no 2 acc:(0.60023)
16. Temp.=mild Windy=true 3 ==> Play=no 2 acc:(0.60023)
17. Humidity=high Windy=true 3 ==> Play=no 2 acc:(0.60023)
18. Temp.=mild 6 ==> Play=yes 3 acc:(0.50124)
19. Outlook=sunny 5 ==> Play=yes 3 acc:(0.50124)
20. Outlook=rainy 5 ==> Play=no 2 acc:(0.50124)
21. Humidity=high 7 ==> Play=no 3 acc:(0.42451)
22. Temp.=hot 4 ==> Play=no 2 acc:(0.42451)
23. Temp.=hot 4 ==> Play=yes 2 acc:(0.42451)
24. Windy=true 6 ==> Play=no 2 acc:(0.42451)
25. Windy=true 6 ==> Play=yes 2 acc:(0.42451)
26. Humidity=high 7 ==> Play=no 3 acc:(0.42451)
27. Outlook=sunny 5 ==> Play=no 2 acc:(0.42451)
28. Outlook=rainy 5 ==> Play=no 2 acc:(0.50124)
29. Temp.=mild 6 ==> Play=no 2 acc:(0.4736)
30. Windy=false 8 ==> Play=no 2 acc:(0.42451)

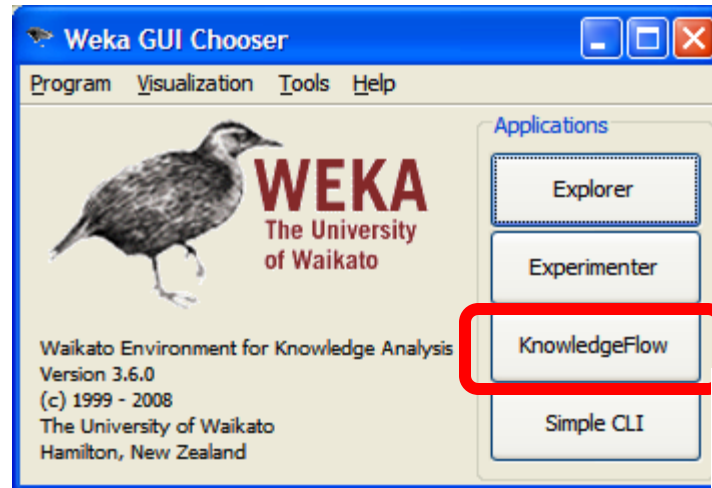


What's the difference between classification and association?

- In one sense, association is like classification, except that instead of considering just one attribute as the class attribute, it considers every attribute and every combination of attributes as a class, and then performs a form of classification on all possible classes.
- Unlike classification, association does not necessarily consider all attributes when attempting to create rules, but operates on sub-set combinations of attributes.
- Association only works on nominal attributes, not numerical attributes, although numerical attributes can be discretized.



The Weka Experimenter Environment supports efficient automation of algorithm testing.



Topic of the next Weka tutorial



Weka Documentation:

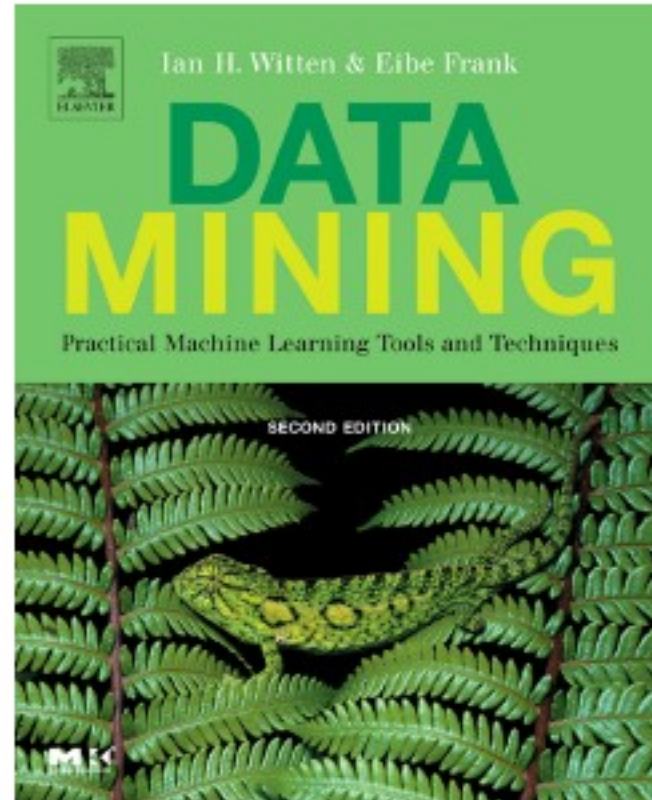
Data Mining: Practical Machine Learning Tools and Techniques (Second Edition)

Ian H. Witten, Eibe Frank

Morgan
Kaufmann
June 2005
525 pages
Paper
ISBN 0-12-
088407-0



Eibe Frank and Ian
Witten





Contact the Author:

**Mark Polczynski, PhD
The Technology Forge**

mhp.techforge@gmail.com

THANKS!

