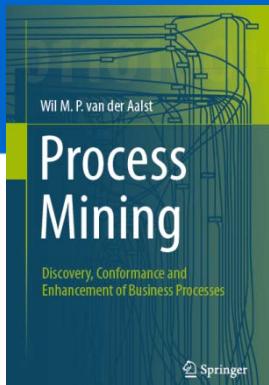


Process Mining: Data Science in Action

Association Rule Learning

prof.dr.ir. Wil van der Aalst
www.processmining.org



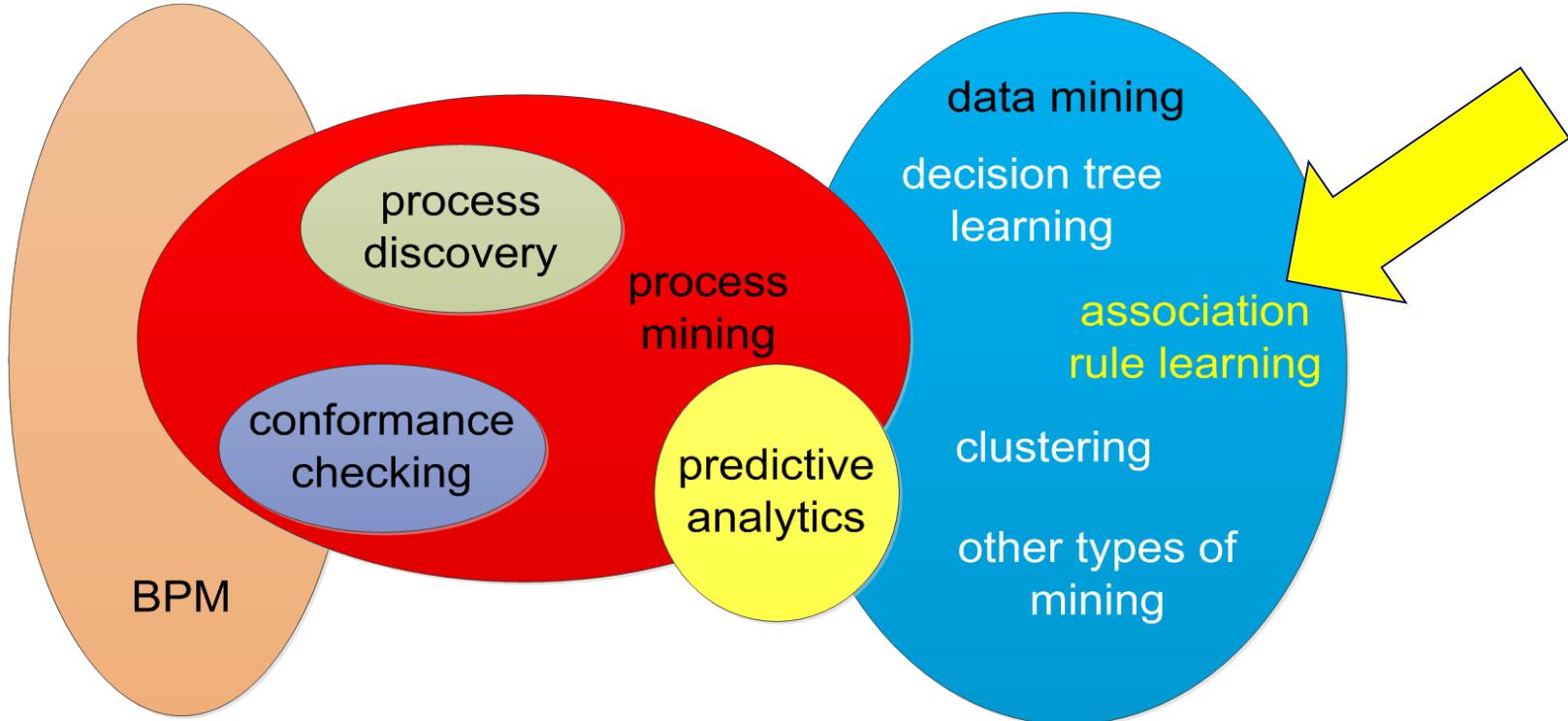
TU/e

Technische Universiteit
Eindhoven
University of Technology

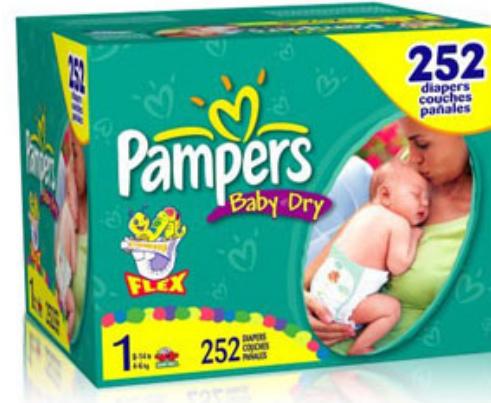
Where innovation starts

Association rule learning

(unsupervised learning: no response variable)



Urban Legend: Beer and Diapers



{beer} \Rightarrow {diapers}

Association rules: Examples

$$\{beer\} \Rightarrow \{diapers\}$$
$$\{cheese, ham, bread\} \Rightarrow \{butter\}$$
$$\{oregano\} \Rightarrow \{spaghetti, tomato sauce\}$$

All of form: $X \Rightarrow Y$



- support
- confidence
- lift
- ...

Support

(values between 0=bad and 1=good)

$$support(X \Rightarrow Y) = \frac{N_{X \wedge Y}}{N} = \frac{N_{X \cup Y}}{N}$$

N is the number of instances

N_X is the number of instances covering X

$N_{X \wedge Y} = N_{X \cup Y}$ is the number of instances covering both X and Y

Example

cappuccino	latte	espresso	americano	ristretto	tea	muffin	bagel
1	0	0	0	0	0	1	0
0	2	0	0	0	0	1	1
0	0	1	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	1	2	0

$\{ \text{tea}, \text{latte} \} \Rightarrow \{ \text{muffin} \}$

Frequency does not matter, only presence.

Support $\{tea, latte\} \Rightarrow \{muffin\}$

(values between 0=bad and 1=good)

cappuccino	latte	espresso	americano	ristretto	tea	muffin	bagel
1	0	0	0	0	0	1	0
0	2	0	0	0	0	1	1
0	0	1	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	1	2	0
0	0	0	1	1	0	0	0

$$support(\{tea, latte\} \Rightarrow \{muffin\}) = \frac{N_{\{tea, latte, muffin\}}}{N}$$

Confidence

(values between 0=bad and 1=good)

$$\text{confidence}(X \Rightarrow Y) = \frac{N_{X \wedge Y}}{N_X} = \frac{N_{X \cup Y}}{N_X}$$

$$\text{confidence}(\{\text{tea}, \text{latte}\} \Rightarrow \{\text{muffin}\}) = \frac{N_{\{\text{tea, latte, muffin}\}}}{N_{\{\text{tea, latte}\}}}$$

Lift

$$lift(X \Rightarrow Y) = \frac{N_{X \wedge Y}/N}{(N_X/N) (N_Y/N)} = \frac{N_{X \wedge Y}}{N_X N_Y} N$$

$$lift(X \Rightarrow Y) == \frac{N_{\{tea, latte, muffin\}} N}{N_{tea} N_{latte} N_{muffin}}$$

$lift(X \Rightarrow Y) > 1$ if X and Y are positively correlated

$lift(X \Rightarrow Y) \approx 1$ if X and Y are independent

$lift(X \Rightarrow Y) < 1$ if X and Y are negatively correlated

How are support, confidence, and lift used?

- Serve as quality criteria:
 - Can be used to **filter** rules apriori.
 - Can be used to **sort** rules.
- Important because there may be many rules.
- Typically, one is interested in rules:
 - That have a **support as high as possible** (value depends on number of items).
 - That have a **confidence close to 1**.
 - That have a **lift higher than 1** (positive correlation).

Example

100 customers buy diapers and/or beer:

- 9 customers buy just Hoegaarden
- 40 customers buy just Pampers
- 50 customer buy just Pampers and Dommelsch
- 1 customer buys Pampers, Hoegaarden and Dommelsch

Consider the
following
four rules:

- $$\{Pampers\} \Rightarrow \{Dommelsch\}$$
- $$\{Pampers\} \Rightarrow \{Hoegaarden\}$$
- $$\{Dommelsch\} \Rightarrow \{Pampers\}$$
- $$\{Pampers, Dommelsch\} \Rightarrow \{Hoegaarden\}$$

Question:

Compute support, confidence, and lift

$$\{Pampers\} \Rightarrow \{Dommelsch\}$$

100 customers buy diapers and/or beer:

- 9 customers buy just Hoegaarden
- 40 customers buy just Pampers
- 50 customer buy just Pampers and Dommelsch
- 1 customer buys Pampers, Hoegaarden and Dommelsch

Answer

$$support(\{Pampers\} \Rightarrow \{Dommelsch\}) = \frac{N_{\{Pampers, Dommelsch\}}}{N} = \frac{51}{100} = 0.51$$

$$confidence(\{Pampers\} \Rightarrow \{Dommelsch\}) = \frac{N_{\{Pampers, Dommelsch\}}}{N_{\{Pampers\}}} = \frac{51}{91} = 0.56$$

$$lift(\{Pampers\} \Rightarrow \{Dommelsch\}) = \frac{N_{\{Pampers, Dommelsch\}} N}{N_{\{Pampers\}} N_{\{Dommelsch\}}} = \frac{51 \times 100}{91 \times 51} = 1.10$$

Question:

Compute support, confidence, and lift

$$\{Pampers\} \Rightarrow \{Dommelsch\}$$
$$\{Pampers\} \Rightarrow \{Hoegaarden\}$$
$$\{Dommelsch\} \Rightarrow \{Pampers\}$$
$$\{Pampers, Dommelsch\} \Rightarrow \{Hoegaarden\}$$

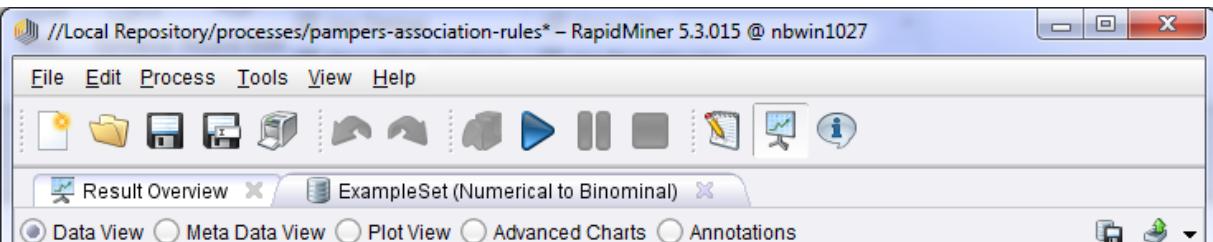
100 customers buy diapers and/or beer:

- **9 customers buy just Hoegaarden**
- **40 customers buy just Pampers**
- **50 customer buy just Pampers and Dommelsch**
- **1 customer buys Pampers, Hoegaarden and Dommelsch**

Answer	N _X	N _Y	N _{XY}	support	confidence	lift
{Pampers} \Rightarrow {Dommelsch}	91	51	51	51/100 = 0.51	51/91 = 0.56	(51*100)/(91*51) = 1.10
{Pampers} \Rightarrow {Hoegaarden}	91	10	1	1/100 = 0.01	1/91 = 0.01	(1*100)/(91*10) = 0.11
{Dommelsch} \Rightarrow {Pampers}	51	91	51	51/100 = 0.51	51/51 = 1	(51*100)/(51*91) = 1.10
{Pampers, Dommelsch} \Rightarrow {Hoegaarden}	51	10	1	1/100 = 0.01	1/51 = 0.02	(1*100)/(51*10) = 0.20

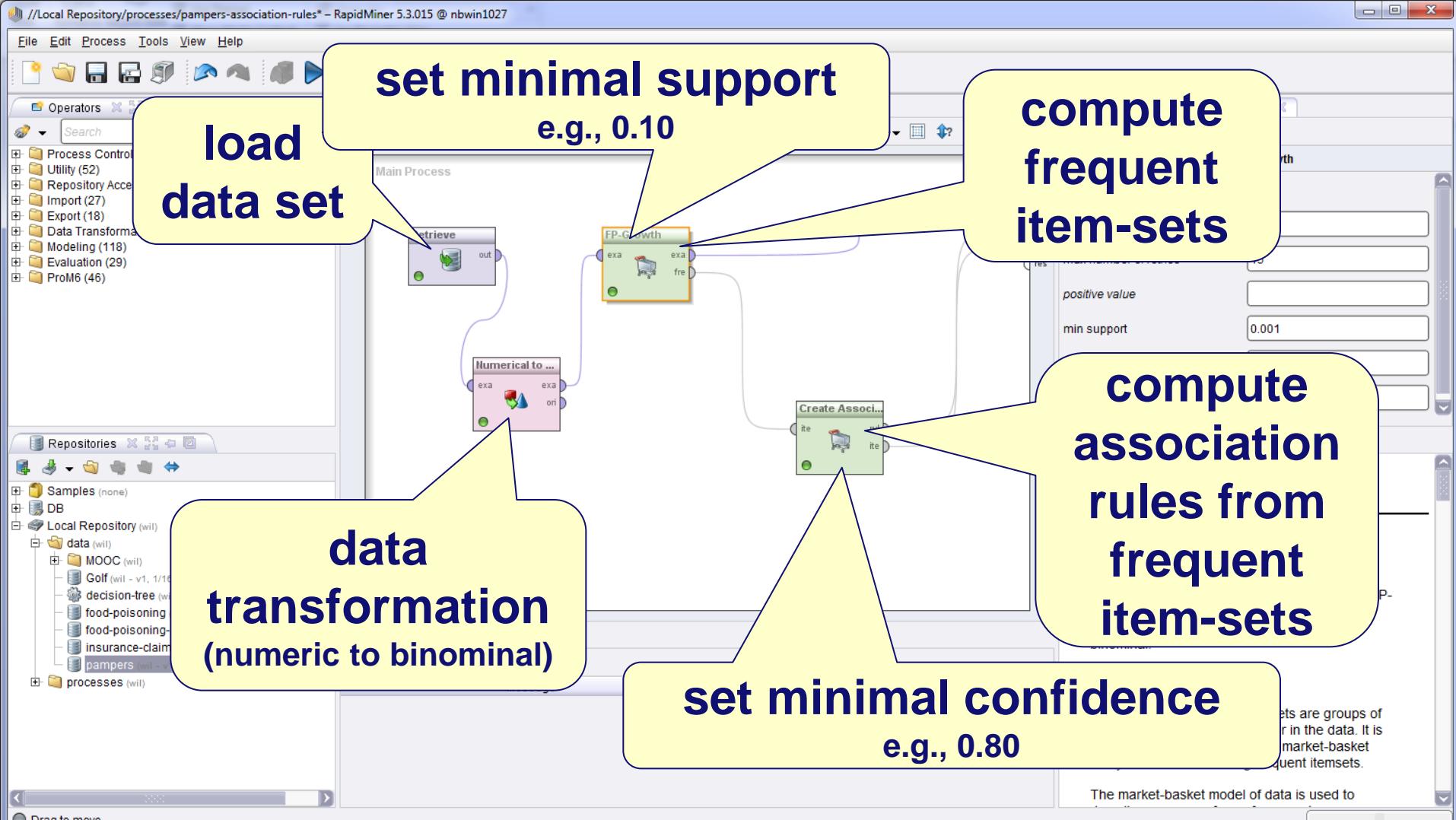
Answer	N _X	N _Y	N _{XY}	support	confidence	lift
{Pampers} \Rightarrow {Dommelsch}	91	51	51	51/100 = 0.51	51/91 = 0.56	(51*100)/(91*51) = 1.10
{Pampers} \Rightarrow {Hoegaarden}	91	10	1	1/100 = 0.01	1/91 = 0.01	(1*100)/(91*10) = 0.11
{Dommelsch} \Rightarrow {Pampers}	51	91	51	51/100 = 0.51	51/51 = 1	(51*100)/(51*91) = 1.10
{Pampers, Dommelsch} \Rightarrow {Hoegaarden}	51	10	1	1/100 = 0.01	1/51 = 0.02	(1*100)/(51*10) = 0.20

Example in RapidMiner

A screenshot of the RapidMiner data view window, which is part of the same process editor session. The title bar reads "//Local Repository/processes/pampers-association-rules* – RapidMiner 5.3.015 @ nbwin1027". The menu bar and toolbar are identical to the process editor. The tab bar shows "Result Overview" and "ExampleSet (Numerical to Binomial)", with "Meta Data View" currently selected. The main area displays the following text:

ExampleSet (100 examples, 0 special attributes, 3 regular attributes)

Role	Name	Type	Statistics	Range	Missings
regular	Pampers	binominal	mode = true (91), least = false (9)	false (9), true (91)	0
regular	Hoegaarden	binominal	mode = false (90), least = true (10)	false (90), true (10)	0
regular	Dommelsch	binominal	mode = true (51), least = false (49)	false (49), true (51)	0



Only one rule is found ...

//Local Repository/processes/pampers-association-rules* – RapidMiner 5.3.015 @ nbwin1027

File Edit Process Tools View Help

Result Overview Table View Graph View

Show rules matching all of these conclusions: Pampers

Min. Criterion: confidence

Min. Criterion Value:

No.	Premises	Conclusion	Support	Confidence	LaPlace	Gain	p-s	Lift	Conviction
1	Dommelsch	Pampers	0.510	1	1	-0.510	0.046	1.099	∞

support = 0.510 \geq 0.1

confidence = 1.0 \geq 0.8

$$\{Dommelsch\} \Rightarrow \{Pampers\}$$

Answer	N _x	N _y	N _{xy}	support	confidence	lift
{Pampers} \Rightarrow {Dommelsch}	91	51	51	51/100 = 0.51	51/91 = 0.56	(51*100)/(91*51) = 1.10
{Pampers} \Rightarrow {Hoegaarden}	91	10	1	1/100 = 0.01	1/91 = 0.01	(1*100)/(91*10) = 0.11
{Dommelsch} \Rightarrow {Pampers}	51	91	51	51/100 = 0.51	51/51 = 1	(51*100)/(51*91) = 1.10
{Pampers, Dommelsch} \Rightarrow {Hoegaarden}	51	10	1	1/100 = 0.01	1/51 = 0.02	(1*100)/(51*10) = 0.20

Brute force approach

To systematically generate association rules, Association rules can now be generated as follows:
typically defines two $minconf$.
1. Generate all frequent item-sets, i.e., all sets Z such that $N_z / N > minsup$ and $|Z| > 2$.

- **Typical problems:**

- **too many frequent items sets need to be checked (computational problem) and**
- **too many rules are found (interpretation problem).**

3. Output the rules found.

Smarter approaches are possible

(e.g. the well-known Apriori algorithm)

- **Observation:**

Exploited to improve efficiency and
to return only the "strongest rules".

$$Y \subseteq X \implies N_Y/N \geq N_X/N$$

- If **X is frequent (support is high enough)**, then all subsets of X need to be frequent.
- If **Y is not frequent (support is not high enough)**, then all supersets of Y are infrequent and do not need to be considered.

Italian restaurant revisited



- 50% of people eat at the restaurant.
- Menu includes: pizza margherita, pizza romana, pizza marinara, pizza capricciosa, pizza siciliana, lasagna, spaghetti carbonara, spaghetti alla diavola, vino rosso, vino bianco, birra, and espresso.



csv file loaded into the repository

Row No.	class	pizza margh...	pizza romana	pizza marin...	pizza capric...	pizza sicilia...	lasagna	spaghetti c...	spaghetti al...	vino rosso	vino bianco	birra	espresso
1	not sick	2	0	0	0	1	1	0	0	0	0	3	1
2	not sick	1	3	1	0	4	0	1	1	2	1	0	2
3	not sick	0	3	0	0	4	1	0	3	0	1	1	1
4	not sick	0	0	0	0	0	0	0	0	0	0	3	0
5	not sick	2	0	1	1	0	3	0	0	0	0	3	0
6	not sick	0	2	0	1	3	0	0	1	0	2	1	4
7	not sick	0	0	0	0	0	2	1	2	1	1	0	0
8	not sick	2	0	1	0	0	3	0	0	0	0	2	0
9	nauseous	0	1	0	0	4	1	0	1	1	1	1	2
10	not sick	0	0	0	0	0	0	0	0	0	0	1	0
11	not sick	0	1	0	0	0	3	3	1	1	0	3	0
12	not sick	0	0	0	0	0	2	2	2	1	1	1	0
13	very sick	3	0	3	0	0	1	0	0	0	0	3	1
14	not sick	1	2	0	0	3	1	0	2	2	1	0	0
15	nauseous	1	3	0	0	4	0	0	4	3	0	0	2
16	not sick	0	2	0	0	4	0	1	1	2	2	0	2
17	not sick	0	0	0	0	0	3	3	2	0	0	1	1
18	not sick	0	1	0	2	4	1	1	2	1	2	1	2
19	not sick	3	0	0	0	0	1	1	0	0	0	2	0
20	not sick	0	0	0	0	0	3	3	2	0	0	2	2

//Local Repository/processes/food-poisoning-association-rules* – RapidMiner 5.3.015 @ nbwin1027

File Edit Process Tools View Help

Result Overview FrequentItemSets (FP-Growth) AssociationRules (Create Association Rules) ExampleSet (Numerical to Binomial)

Table View Annotations

No. of Sets: 153 Size Support Item 1 Item 2 Item 3 Item 4 Item 5 Item 6

Total Max. Size: 6 1 0.610 birra
1 0.596 spaghetti all
1 0.596 espresso
1 0.540 lasagna
Max. Size: 6 1 0.520 vino rosso
1 0.516 vino bianco
Contains Item: 1 0.480 pizza sicilian
1 0.461 pizza roman;
Update View 1 0.358 pizza marghi
1 0.350 spaghetti ca
1 0.315 pizza marina
1 0.179 pizza capricc
2 0.275 birra spaghetti all
2 0.268 birra espresso
2 0.440 birra lasagna
2 0.196 birra vino rosso
2 0.193 birra vino bianco
2 0.139 birra pizza sicilian
2 0.138 birra pizza roman;
2 0.260 birra pizza marghi
2 0.261 birra spaghetti ca
2 0.222 birra pizza marina
2 0.471 spaghetti all espresso
2 0.289 spaghetti all lasagna
2 0.413 spaghetti all vino rosso
2 0.405 spaghetti all vino bianco
2 0.354 spaghetti all pizza sicilian
2 0.336 spaghetti all pizza roman;
2 0.281 spaghetti all spaghetti ca
2 0.250 espresso lasagna
2 0.413 espresso vino rosso
2 0.403 espresso vino bianco
2 0.395 espresso pizza sicilian

153 frequent item-sets with a support of at least 0.1

//Local Repository/processes/food-poisoning-association-rules* – RapidMiner 5.3.015 @ nbwin1027

File Edit Process Tools View Help

Result Overview FrequentItemSets (FP-Growth) AssociationRules (Create Association Rules) ExampleSet (Numerical to Binomial)

Table View Graph View Text View Annotations

Show rules matching all of these conclusions:

birra
spaghetti alla diavola
espresso
lasagna
vino rosso
vino bianco
pizza siciliana
pizza romana
pizza margherita
spaghetti carbonara
pizza marinara

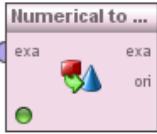
No. Premises Conclusion Support Confidence LaPlace Gain p-s Lift Convicti...

1	espresso, lasagna	birra, spaghetti carbonara	0.126	0.502	0.900	-0.375	0.060	1.919	1.482
2	espresso, spaghetti carbonara	vino rosso	0.106	0.505	0.914	-0.315	-0.003	0.970	0.968
3	espresso, lasagna	birra, spaghetti alla diavola	0.127	0.508	0.901	-0.374	0.058	1.849	1.474
4	espresso	vino rosso, vino bianco, pizza siciliana, pizza romana	0.303	0.509	0.817	-0.888	0.106	1.534	1.362
5	lasagna	spaghetti carbonara	0.275	0.509	0.828	-0.805	0.086	1.455	1.325
6	espresso	vino rosso, vino bianco, pizza romana	0.305	0.512	0.818	-0.805	0.105	1.526	1.362
7	espresso, spaghetti carbonara	birra, spaghetti alla diavola, lasagna	0.108	0.513	0.915	-0.313	0.064	2.470	1.628
8	spaghetti carbonara	birra, spaghetti alla diavola, lasagna	0.180	0.514	0.874	-0.520	0.107	2.473	1.630
9	spaghetti alla diavola	espresso, vino rosso, vino bianco	0.307	0.516	0.819	-0.884	0.103	1.506	1.358
10	espresso	spaghetti alla diavola, vino rosso, vino bianco	0.307	0.516	0.819	-0.884	0.106	1.525	1.367
11	spaghetti alla diavola	espresso, vino bianco, pizza siciliana	0.307	0.516	0.819	-0.884	0.099	1.479	1.345
12	espresso	spaghetti alla diavola, vino bianco, pizza siciliana	0.307	0.516	0.819	-0.884	0.111	1.566	1.385
13	spaghetti alla diavola	espresso, pizza siciliana, pizza romana	0.308	0.517	0.820	-0.883	0.093	1.432	1.323
14	espresso	spaghetti alla diavola, pizza siciliana, pizza romana	0.308	0.517	0.820	-0.883	0.112	1.570	1.389
15	spaghetti alla diavola	vino bianco, pizza siciliana, pizza romana	0.309	0.518	0.820	-0.883	0.092	1.423	1.320
16	pizza marinara	lasagna, pizza margherita	0.164	0.519	0.885	-0.466	0.091	2.266	1.604
17	birra vino rosso	vino bianco	0.102	0.520	0.922	-0.289	0.001	1.009	1.010
18	birra spaghetti alla diavola, lasagna	espresso, spaghetti carbonara	0.108	0.521	0.918	-0.307	0.064	2.470	1.647
19	spaghetti alla diavola	vino rosso, vino bianco, pizza romana	0.108	0.521	0.822	-0.880	0.089	1.402	1.314
20	espresso	spaghetti alla diavola, vino rosso, vino bianco, pizza romana	0.108	0.521	0.822	-0.880	0.111	1.556	1.391
21	spaghetti alla diavola	vino bianco, pizza romana	0.311	0.523	0.822	-0.880	0.088	1.396	1.311
22	spaghetti alla diavola	vino rosso, vino bianco, pizza siciliana	0.312	0.524	0.822	-0.879	0.101	1.480	1.356
23	spaghetti alla diavola	espresso, vino rosso, vino bianco, pizza romana	0.312	0.525	0.823	-0.879	0.095	1.440	1.337
24	espresso, lasagna	spaghetti alla diavola, vino rosso, vino bianco, pizza romana	0.311	0.525	0.905	-0.369	0.061	1.869	1.513
25	spaghetti alla diavola	espresso, vino rosso, pizza siciliana	0.313	0.525	0.823	-0.879	0.102	1.483	1.360
26	espresso	spaghetti alla diavola, vino rosso, pizza siciliana	0.313	0.525	0.823	-0.879	0.114	1.574	1.403
27	vino rosso	spaghetti alla diavola, vino rosso, vino bianco, pizza romana	0.221	0.526	0.838	-0.767	0.124	1.827	1.503
28	birra vino bianco	spaghetti alla diavola, vino rosso, vino bianco, pizza romana	0.101	0.527	0.923	-0.285	0.001	1.013	1.014
29	pizza margherita	birra, pizza marinara	0.189	0.527	0.875	-0.527	0.109	2.373	1.646
30	vino rosso	spaghetti alla diavola, espresso, vino bianco, pizza romana	0.275	0.528	0.838	-0.766	0.124	1.823	1.504
31	spaghetti alla diavola, espresso, lasagna	vino rosso	0.101	0.528	0.924	-0.281	0.001	1.015	1.017
32	spaghetti alla diavola	vino rosso, pizza romana	0.315	0.529	0.824	-0.876	0.091	1.405	1.324
33	vino bianco	spaghetti alla diavola, espresso, vino rosso, pizza siciliana, pizza romana	0.274	0.531	0.840	-0.758	0.123	1.812	1.507

Min. Criterion: confidence

Min. Criterion Value: 0.1

set minimal support to
0.3 (rather than 0.1)



Parameters Context

FP-Growth

find min number of itemsets

positive value

0.3

min support

-1

max items

must contain

Help Comment

FP-Growth (RapidMiner Core)

Synopsis

//Local Repository/processes/food-poisoning-association-rules* – RapidMiner 5.3.015 @ nbwin1027

File Edit Process Tools View Help

Result Overview **FrequentItemSets (FP-Growth)** **AssociationRules (Create Association Rules)** **ExampleSet (Numerical to Binomial)**

Table View **Graph View** **Text View** **Annotations**

Show rules matching **all of these conclusions:**

espresso
vino rosso
vino bianco
pizza siciliana
pizza romana

No. Premises Conclusion Support Confidence LaPlace Gain p-s Lift Convict...

1	espresso, pizza siciliana, pizza romana	vino bianco	0.326	0.904	0.974	-0.396	0.140	1.752	5.023	
2	vino rosso, vino bianco, pizza romana	espresso, pizza siciliana	0.303	0.904	0.976	-0.368	0.171	2.291	6.310	
3	vino rosso, vino bianco	pizza siciliana	0.354	0.303	0.904	0.973	-0.429	0.166	1.886	5.443
4	vino rosso, pizza siciliana, pizza romana	espresso	0.330	0.906	0.975	-0.398	0.113	1.521	4.307	
5	espresso, vino rosso, pizza romana	vino bianco, pizza siciliana	0.303	0.906	0.976	-0.366	0.172	2.301	6.464	
6	spaghetti alla diavola, vino rosso, vino bianco	espresso	0.307	0.908	0.977	-0.369	0.106	1.525	4.412	
7	vino rosso, vino bianco, pizza romana	espresso	0.305	0.909	0.977	-0.366	0.105	1.526	4.435	
8	espresso, vino rosso, pizza romana	vino bianco	0.305	0.911	0.978	-0.365	0.132	1.766	5.440	
9	vino rosso, pizza siciliana, pizza romana	vino bianco	0.332	0.912	0.976	-0.396	0.144	1.767	5.477	
10	vino bianco, pizza siciliana, pizza romana	vino rosso	0.332	0.912	0.977	-0.396	0.143	1.753	5.455	
11	vino rosso, vino bianco, pizza siciliana, pizza romana	espresso	0.303	0.914	0.979	-0.361	0.106	1.534	4.694	
12	espresso, pizza siciliana, pizza romana	vino rosso	0.330	0.914	0.977	-0.392	0.142	1.757	5.585	
13	espresso, vino rosso, pizza siciliana	vino bianco	0.324	0.915	0.978	-0.384	0.141	1.773	5.673	
14	vino rosso, vino bianco, pizza siciliana	espresso	0.324	0.915	0.978	-0.384	0.113	1.536	4.738	
15	espresso, pizza siciliana	pizza romana	0.361	0.915	0.976	-0.428	0.179	1.984	6.330	
16	spaghetti alla diavola, pizza romana	espresso, pizza siciliana	0.308	0.917	0.979	-0.364	0.175	2.324	7.313	
17	spaghetti alla diavola, pizza romana	vino bianco, pizza siciliana	0.309	0.919	0.980	-0.363	0.176	2.334	7.484	
18	vino rosso, pizza romana	vino bianco, pizza siciliana	0.364	0.919	0.977	-0.428	0.182	1.994	6.673	
19	espresso, vino bianco, pizza siciliana, vino bianco	vino rosso, pizza siciliana	0.303	0.919	0.980	-0.357	0.133	1.782	6.007	
20	espresso, vino bianco, pizza romana	vino rosso, pizza siciliana	0.303	0.922	0.981	-0.355	0.173	2.326	7.704	
21	spaghetti alla diavola, vino rosso, vino bianco	pizza siciliana	0.312	0.922	0.980	-0.365	0.150	1.922	6.667	
22	vino bianco, pizza siciliana, vino bianco	vino rosso, pizza siciliana	0.324	0.924	0.979	-0.424	0.182	2.005	7.123	
23	espresso, vino bianco, pizza romana	vino bianco, pizza siciliana	0.326	0.926	0.982	-0.353	0.134	1.780	6.524	
24	spaghetti alla diavola, pizza romana	espresso	0.311	0.927	0.982	-0.360	0.111	1.556	5.520	
25	spaghetti alla diavola, pizza romana	vino bianco, pizza siciliana	0.311	0.927	0.982	-0.360	0.138	1.798	6.664	
26	espresso, vino bianco, pizza siciliana, vino bianco	vino rosso	0.324	0.928	0.981	-0.374	0.142	1.783	6.638	
27	espresso, vino bianco, pizza siciliana, pizza romana	vino rosso	0.303	0.930	0.983	-0.349	0.134	1.787	6.862	
28	spaghetti alla diavola, pizza romana	vino rosso, pizza siciliana	0.312	0.930	0.982	-0.359	0.179	2.348	8.665	
29	spaghetti alla diavola, espresso, pizza siciliana	vino bianco	0.307	0.930	0.983	-0.353	0.137	1.804	6.951	
30	spaghetti alla diavola, pizza siciliana	vino bianco	0.329	0.931	0.982	-0.379	0.147	1.804	6.968	
31	spaghetti alla diavola, pizza siciliana	pizza romana	0.329	0.931	0.982	-0.379	0.166	2.018	7.756	
32	spaghetti alla diavola, vino bianco, pizza siciliana	espresso	0.307	0.933	0.983	-0.352	0.111	1.566	6.000	
33	espresso, vino rosso, pizza siciliana	pizza romana	0.330	0.933	0.982	-0.378	0.167	2.023	8.013	

Min. Criterion: confidence

Min. Criterion Value:

//Local Repository/processes/food-poisoning-association-rules* – RapidMiner 5.3.015 @ nbwin1027

File Edit Process Tools View Help

Show rule
all of the
espresso
vino rosso
vino bianco
pizza siciliana
pizza romana

R
Table
Flowchart
Process
Report
Graph
Information

nvicti...
23
10
43
07
64
12
35
40
77
55
94
85
73
38
30
13
84
73
07
04
67
23
24
20
64
38
52
65
07
04
67
23
24
20
64
38
52
65

$$R_1 = \{espresso, pizza\ siciliana, pizza\ romana\} \Rightarrow \{vino\ bianco\}$$
$$support(R_1) = 0.326$$

$$R_2 = \{vino\ rosso, vino\ bianco, pizza\ romana\} \Rightarrow \{espresso, pizza\ siciliana\}$$
$$support(R_2) = 0.303$$

$$R_3 = \{vino\ rosso, vino\ bianco\} \Rightarrow \{pizza\ siciliana\}$$
$$support(R_3) = 0.354$$
$$confidence(R_3) = 0.904$$
$$lift(R_3) = 1.886$$

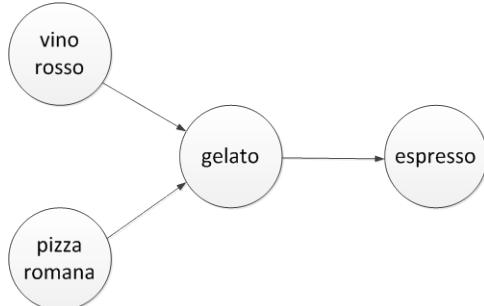
Min. Criterion:	confidence	Min. Criterion Value:	0.000						
29	spaghetti alla diavola, espresso, pizza siciliana	vino bianco	0.307	0.930	0.983	-0.353	0.137	1.804	6.951
30	spaghetti alla diavola, pizza siciliana	vino bianco	0.329	0.931	0.982	-0.379	0.147	1.804	6.968
31	spaghetti alla diavola, pizza siciliana	pizza romana	0.329	0.931	0.982	-0.379	0.166	2.018	7.756
32	spaghetti alla diavola, vino bianco, pizza siciliana	espresso	0.307	0.933	0.983	-0.352	0.111	1.566	6.000
33	espresso, vino rosso, pizza siciliana	pizza romana	0.330	0.933	0.982	-0.378	0.167	2.023	8.013

Other types of pattern mining

- **Sequence mining**

$\langle \{vino\ rosso\}, \{pizza\ romana, pizza\ siciliana\}, \{gelato\} \rangle \Rightarrow \langle \{espresso\} \rangle$

- **Episode mining**



Data mining
techniques do not
consider end-to-end
process models!

Part I: Preliminaries

Chapter 1

Introduction

Chapter 2

Process Modeling and Analysis

Chapter 3

Data Mining



Part III: Beyond Process Discovery

Chapter 7

Conformance Checking

Chapter 8

Mining Additional Perspectives

Chapter 9

Operational Support

Part II: From Event Logs to Process Models

Chapter 4

Getting the Data

Chapter 5

Process Discovery: An Introduction

Chapter 6

Advanced Process Discovery Techniques

Chapter 10

Tool Support

Chapter 11

Analyzing “Lasagna Processes”

Chapter 12

Analyzing “Spaghetti Processes”

Part IV: Putting Process Mining to Work

Chapter 13

Cartography and Navigation

Chapter 14

Epilogue

Wil M. P. van der Aalst

Process Mining

Discovery, Conformance and Enhancement of Business Processes

Springer

