

# Today's agenda (1st class of Rec Sys)

- 1) Problem statement
- 2) Data preprocessing
- 3) Apriori Algorithm
- 4) Association Rule
- 5) Various metrics
- 6) Conclusion

Announcements: 1) Problem Solving class on Sunday 3rd Dec.

2) Module test deadline  $\rightarrow$  Dec 20

3) Dec 6 and Dec 8  $\rightarrow$  no classes

Missing lecture notes  $\rightarrow$  27 Nov, 29 Nov ?

item 1 / 2 / 3

1 → True  
0 → False

	item 1	item 2	item 3	
invoice no 1	0	1	1	(2, 3)
2	1	0	1	(1, 3)
3	0	0	1	(3)
4				
5				

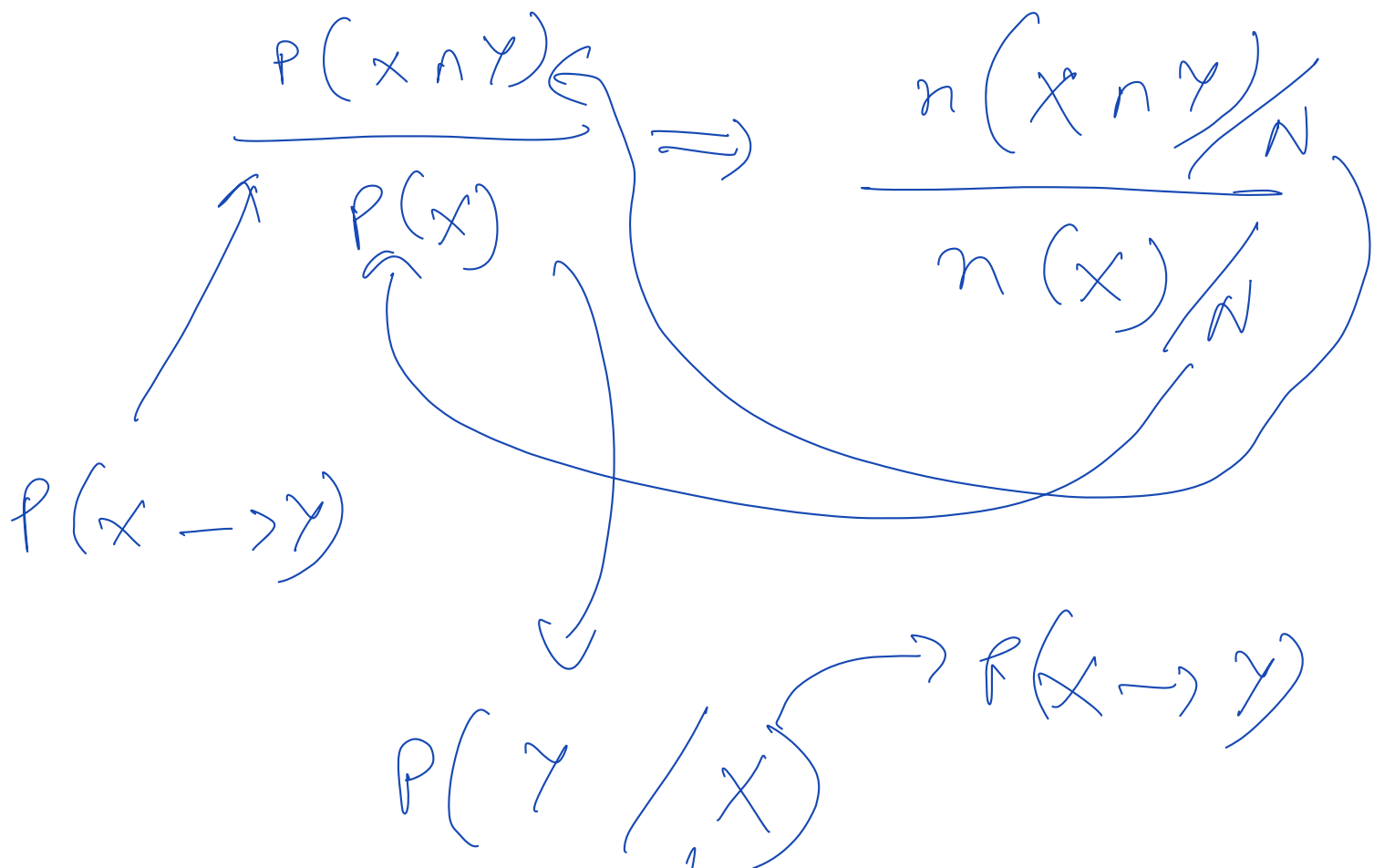
item 1	item 2	item 3
bought / not bought	1/0	1/0
1/0	↓	↓
2	2	2

$$2 \times 2 \times 2 = 2^3$$

0 , 0 , 0

$$\boxed{2^3 - 1}$$

$$2^n - 1$$



if  $X$  happens, prob. that  $Y$  will also happen

$$\boxed{P(X \cap Y) = 0.1}$$

$$P(X) = 0.11$$

→ low support

$$P(X \rightarrow Y) = \frac{P(X \cap Y)}{P(X)}$$

low support,  
high confidence

=

$$\frac{0.1}{0.11}$$

$$= \boxed{0.91}$$

high  
confidence

$$n(X \cap Y) = 1000$$

$$n(X) = 10,000$$

very high  
value

$$P(X \rightarrow Y) = \frac{n(X \cap Y)}{n(X)}$$

high - support

$$= (0.1) \rightarrow \text{low confidence}$$

high support, low confidence

$$P(X \cap Y) = P(X) \cdot P(Y)$$

independent

$$\frac{P(X \cap Y)}{P(X) \cdot P(Y)} = 1 \rightarrow \text{independence}$$

$\hookrightarrow$  lift