

Tato prezentace je součástí wiki-prezentace [Metoda GUHA a systém LISp-Miner](#)

Je dostupná z [této adresy](#)

Verse 11. 9. 2019

# Asociační pravidla v datech s neúplnou informací - vynechání neúplné informace

Jan Rauch

Katedra informačního a znalostního inženýrství

Vysoká škola ekonomická

# Asociační pravidla v datech s neúplnou informací

## - vynechání neúplné informace

- Matice dat s neúplnou informací
- Devítipolní tabulka
- Vynechání neúplné informace

# Matice dat s neúplnou informací

Výsledek předzpracování – matice dat  $\mathcal{M}^X$  s neúplnou informací :

$\mathcal{M}^X$	$A_1$	$A_2$	$A_3$	$\dots$	$A_P$
$o_1$	1	$X$	5	$\dots$	1
$o_2$	$X$	1	$X$	$\dots$	3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$o_{n-1}$	2	$X$	2	$\dots$	2
$o_n$	$X$	3	1	$\dots$	$X$

$X$  – kód chybějící hodnoty

# Devítipolní tabulka

$\mathcal{M}^X$	$\varphi$	$\psi$
$o_1$	1	1
$o_2$	X	1
$o_3$	0	X
$o_4$	0	0
...	...	...
$o_{n-1}$	X	0
$o_n$	1	X



$\mathcal{M}^X$	$\psi$	$\psi_X$	$\neg\psi$
$\varphi$	$f_{1,1}$	$f_{1,X}$	$f_{1,0}$
$\varphi_X$	$f_{X,1}$	$f_{X,X}$	$f_{X,0}$
$\neg\varphi$	$f_{0,1}$	$f_{0,X}$	$f_{0,0}$

# Vynechání neúplné informace

Asociační pravidlo se vyhodnocuje na základě čtyřpolní tabulky, která vznikne z devítipolní tabulky vynecháním frekvencí týkajících se neúplné informace.

$\mathcal{M}^X$	$\psi$	$\psi_X$	$\neg\psi$
$\varphi$	$f_{1,1}$	$f_{1,X}$	$f_{1,0}$
$\varphi_X$	$f_{X,1}$	$f_{X,X}$	$f_{X,0}$
$\neg\varphi$	$f_{0,1}$	$f_{0,X}$	$f_{0,0}$

$\mathcal{M}$	$\psi$	$\neg\psi$
$\varphi$	$f_{1,1}$	$f_{1,0}$
$\neg\varphi$	$f_{0,1}$	$f_{0,0}$

Součet frekvencí ve čtyřpolní tabulce při vynechání neúplné informace může být menší než součet frekvencí v devítipolní tabulce.