

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

Je dostupná z [této adresy](#)

Verze 12. 9. 2019

Podmíněná asociační pravidla v datech s neúplnou informací - vynechání neúplné informace

Jan Rauch

Katedra informačního a znalostního inženýrství

Vysoká škola ekonomická v Praze

Podmíněná asociační pravidla v datech s neúplnou informací - zabezpečené doplnění

- Výchozí devítipolní tabulky
- Převod výchozích devítipolních tabulek na čtyřpolní

Výchozí devítipolní tabulky

\mathbf{M}^X/χ_1	ψ	ψ_X	$\neg\psi$
φ	$f_{1,1,1}$	$f_{1,X,1}$	$f_{1,0,1}$
φ_X	$f_{X,1,1}$	$f_{X,X,1}$	$f_{X,0,1}$
$\neg\varphi$	$f_{0,1,1}$	$f_{0,X,1}$	$f_{0,0,1}$

$9ft(\varphi, \psi, \mathbf{M}^X/\chi_1)$

\mathbf{M}^X/χ_X	ψ	ψ_X	$\neg\psi$
φ	$f_{1,1,X}$	$f_{1,X,X}$	$f_{1,0,X}$
φ_X	$f_{X,1,X}$	$f_{X,X,X}$	$f_{X,0,X}$
$\neg\varphi$	$f_{0,1,X}$	$f_{0,X,X}$	$f_{0,0,X}$

$9ft(\varphi, \psi, \mathbf{M}^X/\chi_X)$

Devítipolní tabulka $9ft(\varphi, \psi, \mathcal{M}^X/\chi_1)$ se týká trojhodnotových atributů φ a ψ v matici dat \mathcal{M}^X/χ_1 , která se skládá ze všech řádků matice \mathcal{M}^X , pro které je atribut χ pravdivý. Tedy $f_{1,1,1}$ je počet řádků matice \mathcal{M}^X pro které jsou atributy φ , ψ i χ pravdivé, $f_{1,X,1}$ je počet řádků matice \mathcal{M}^X pro které jsou atributy φ a χ pravdivé a atribut ψ nabývá hodnotu X , atd.

Devítipolní tabulka $9ft(\varphi, \psi, \mathcal{M}^X/\chi_X)$ se týká trojhodnotových atributů φ a ψ v matici dat \mathcal{M}^X/χ_X , která se skládá ze všech řádků matice \mathcal{M}^X , pro které atribut χ nabývá hodnotu X . Tedy $f_{1,1,X}$ je počet řádků matice \mathcal{M}^X , pro které jsou atributy φ a ψ pravdivé a atribut χ nabývá hodnotu X , $f_{1,X,X}$ je počet řádků matice \mathcal{M}^X , pro které je atribut φ pravdivý a atributy ψ a χ nabývají hodnotu X , atd.

Převod výchozích devítipolních tabulek na čtyřpolní

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

M^X/χ_1	ψ	ψ_X	$\neg\psi$
φ	$f_{1,1,1}$	$f_{1,X,1}$	$f_{1,0,1}$
φ_X	$f_{X,1,1}$	$f_{X,X,1}$	$f_{X,0,1}$
$\neg\varphi$	$f_{0,1,1}$	$f_{0,X,1}$	$f_{0,0,1}$

M^X/χ_X	ψ	ψ_X	$\neg\psi$
φ	$f_{1,1,X}$	$f_{1,X,X}$	$f_{1,0,X}$
φ_X	$f_{X,1,X}$	$f_{X,X,X}$	$f_{X,0,X}$
$\neg\varphi$	$f_{0,1,X}$	$f_{0,X,X}$	$f_{0,0,X}$



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

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