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

Je dostupná z [této adresy]

Verse 6. 9. 2019

Asociační pravidla v datech s neúplnou informací - formálnější popis

Jan Rauch

Katedra informačního a znalostního inženýrství

Vysoká škola ekonomická v Praze

Asociační pravidla v datech s neúplnou informací - formálnější popis

- Hodnota asociačního pravidla v matici dat $\mathcal{M}^X$ s neúplnou informací
- Devítipolní tabulka
- Převod devítipolní tabulky na čtyřpolní
- Zabezpečené doplnění
- Třídy 4ft-kvantifikátorů
- Zabezpečená doplnění pro třídy 4ft-kvantifikátorů

Hodnota asociačního pravidla v matici dat $\mathcal{M}^X$ s neúplnou informací

- Val($\varphi \approx \psi$, $\mathcal{M}^X$) = 1
  Asociační pravidlo $\varphi \approx \psi$ je pravdivé v matici dat $\mathcal{M}^X$ s neúplnou informací, pokud je pravdivé ve všech možných doplněních matice $\mathcal{M}^X$.

- Val($\varphi \approx \psi$, $\mathcal{M}^X$) = 0
  Asociační pravidlo je nepravdivé v matici dat $\mathcal{M}^X$ s neúplnou informací, pokud je nepravdivé ve všech možných doplněních matice $\mathcal{M}^X$.

- Val($\varphi \approx \psi$, $\mathcal{M}^X$) = X
  Hodnotou asociačního pravidla je X jindy, pokud existuje doplnění $\mathcal{M}_1$ matice dat $\mathcal{M}^X$ ve kterém je pravidlo pravdivé a doplnění $\mathcal{M}_2$ matice dat $\mathcal{M}^X$, ve kterém je pravidlo nepravdivé.

Devítipolní tabulka

$\mathcal{M}^X$	φ	ψ
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

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

$$\boxed{9ft(\varphi, \psi, \mathcal{M}^X)}$$

Převod devítipolní tabulky na čtyřpolní

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

$\mathcal{M}$	ψ	$\neg\psi$
φ	$f_{1,1} + f_{1,X,a}+$ $+f_{X,1,a} + f_{X,X,a}$	$f_{1,0} + f_{1,X,b}+$ $+f_{X,0,b} + f_{X,X,b}$
$\neg\varphi$	$f_{0,1} + f_{X,1,c}+$ $+f_{0,X,c} + f_{X,X,c}$	$f_{0,0} + f_{X,0,b}+$ $+f_{0,X,d} + f_{X,X,d}$

Zabezpečené doplnění

Zabezpečené doplnění devítipolní tabulky T9

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

pro 4ft-kvantifikátor ≈ je taková čtyřpolní tabulka $\langle a_z, b_z, c_z, d_z\rangle$, pro kterou platí $\approx(a_z, b_z, c_z, d_z) = 1$ právě když $\approx(a, b, c, d) = 1$ pro každé doplnění $\langle a, b, c, d\rangle$ tabulky T9.

Třídy 4ft-kvantifikátorů

Třída 4ft-kvantifikátorů	Podmínka zachování pravdivosti
implikační kvantifikátory	$a' \geq a \wedge b' \leq b$
slabě implikační kvantifikátory	$a+b+c+d = a'+b'+'c+d' \wedge$ $\wedge\ a' \geq a\ \wedge\ b' \leq b$
dvojitě implikační kvantifikátory	$a' \geq a \wedge b' \leq b \wedge c' \leq c$
slabě dvojitě implikační kvantifikátory	$a+b+c+d = a'+b'+'c+d' \wedge$ $\wedge\ a \wedge a' \geq a \wedge b' \leq b \wedge c' \leq c$
Σ-dvojitě implikační kvantifikátory	$a' \geq a \wedge b'+c' \leq b+c$
slabě Σ-dvojitě implikační kvantifikátory	$a+b+c+d = a'+b'+'c+d' \wedge$ $\wedge\ a \wedge a' \geq a \wedge b'+c' \leq b+c$
ekvivalenční kvantifikátory	$a' \geq a \wedge b' \leq b \wedge c' \leq c \wedge d' \geq d$
slabě ekvivalenční kvantifikátory	$a+b+c+d = a'+b'+'c+d' \wedge$ $\wedge\ a' \geq a \wedge b' \leq b \wedge c' \leq c \wedge d' \geq d$
Σ-ekvivalenční kvantifikátory	$a'+d' \geq a+d \wedge b'+c' \leq b+c$
slabě Σ-ekvivalenční kvantifikátory	$a+b+c+d = a'+b'+'c+d' \wedge$ $\wedge\ a'+d' \geq a+d$
symetrické kvantifikátory	$a' = a \wedge b' = c \wedge c' = b \wedge d' = d$
s vlastností F	$(b \geq c-1 \geq 0\ \wedge$ $\wedge\ a' = a \wedge b' = b+1 \wedge c' = c-1 \wedge d' = d) \vee$ $\vee\ (c \geq b-1 \geq 0\ \wedge$ $\wedge\ a' = a \wedge b' = b-1 \wedge c' = c+1b \wedge d' = d)$

Další informace o třídách 4ft-kvantifikátorů jsou zde.

Zabezpečená doplnění pro třídy 4ft-kvantifikátorů (1)

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

Pokud $\approx$ je slabě implikační nebo implikační, pak každá čtyřpolní tabulka

$\mathcal{M}$	ψ	$\neg\psi$
φ	$f_{1,1}$	$f_{1,0} + f_{1,X} + f_{X,X} + f_{X,0}$
$\neg\varphi$	$f_{0,1} + f_{X,1} + f_{0,X,c}$	$f_{0,0} + f_{0,X,d}$

taková, že $f_{0,X,c} + f_{0,X,d} = f_{0,X}$ je zabezpečené doplnění tabulky $9ft(\varphi, \psi, \mathcal{M}^X)$ pro $\approx$. Neformálně lze říci, že frekvenci a ze čtyřpolní tabulky necháme stejnou a frekvenci b uděláme co největší.

Zabezpečená doplnění pro třídy 4ft-kvantifikátorů (2)

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

Pokud $\approx$ je slabě Σ – dvojitě implikační, Σ – dvojitě implikační, slabě Σ – ekvivalenční nebo Σ – ekvivalenční, pak každá čtyřpolní tabulka

$\mathcal{M}$	ψ	$\neg\psi$
φ	$f_{1,1}$	$f_{1,0} + f_{1,X} + f_{X,X,b} + f_{X,0}$
$\neg\varphi$	$f_{0,1} + f_{X,1} + f_{0,X} + f_{X,X,c}$	$f_{0,0}$

taková, že $f_{X,X,b} + f_{X,X,c} = f_{X,X}$ je zabezpečené doplnění tabulky $9ft(\varphi, \psi, \mathcal{M}^X)$ pro $\approx$. Neformálně lze říci, že frekvence a a d ze čtyřpolní tabulky necháme stejné a součet frekvencí $b + c$ uděláme co největší.

Zabezpečená doplnění pro třídy 4ft-kvantifikátorů (3)

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

Pokud $\approx$ je symetrický 4ft-kvantifikátor s vlastností F, pak čtyřpolní tabulka

$\mathcal{M}$	ψ	$\neg\psi$
φ	$f_{1,1}$	$f_{1,0} + f_{1,X} + f_{X,X,b} + f_{X,0}$
$\neg\varphi$	$f_{0,1} + f_{X,1} + f_{0,X} + f_{X,X,c}$	$f_{0,0}$

taková, že $f_{X,X,b} + f_{X,X,c} = f_{X,X}$ a zároveň absolutní hodnota rozdílu frekvencí $|(f_{1,0} + f_{1,X} + f_{X,X,b} + f_{X,0}) - (f_{0,1} + f_{X,1} + f_{0,X} + f_{X,X,c})|$ je minimální možná, je zabezpečené doplnění tabulky $9ft(\varphi, \psi, \mathcal{M}^X)$ pro $\approx$. Neformálně lze říci, že frekvence a a d ze čtyřpolní tabulky necháme stejné a absolutní hodnotu rozdílu frekvencí b a c uděláme co nejmenší.