Metody Inżynierii Wiedzy Wnioskowanie przybliżone, reguły, redukty - wykład 15

Adam Szmigielski
aszmigie@pjwstk.edu.pl
materiały: $ftp(public)://aszmigie/MIW$

- Systemy informacyjne,
- Zbiory przybliżone - dolna i górna aproksymacja,
- Funkcja przynależności do zbioru przybliżonego,
- Redukty,

Zbiory przybliżone

- *Zbiory przybliżone* - Zbigniew Pawlak, Andrzej Skowron (1982, 1991) - stanowią metodę analizy danych cechujących się niepewnością,
- Metoda ta stanowi podstawę do wnioskowań w sytuacjach niepewnych z niepełną wiedzą (ang. *ambiguity*),
- Formalizm tej teorii przystosowany jest do operowania informacją która jest niepełna, nieprecyzyjna lub sprzeczną.

Idea zbiorów przybliżonych

- Główna idea tej teorii opiera się na założeniu, iż każdy obiekt może być opisany w sposób ilościowy lub jakościowy poprzez pewne atrybuty.
- Obiekty które są opisywane w identyczny sposób są *nierozróżnialne* tj. pozostają w stosunku do siebie w relacji *nierozróżnialności*, która jest relacją równoważności.
- Zbiór wszystkich obiektów w pewnym uniwersum U poprzez relację nierozróżnialności $\mathcal{R}$ można podzielić na rozłączne klasy abstrakcji tej relacji.
- Suma wszystkich klas jest uniwersum U.

Relacja nierozróżnialności

- Klasę abstrakcji relacji nierozróżnialności wyznaczoną przez element x będziemy oznaczać jako:

$$[x]_{\mathcal{R}} = \{y : y\mathcal{R}x\}$$

- W teorii zbiorów przybliżonych klasy te stanowią elementarne, definiowalne pojęcia. W oparciu o nie można próbować definiować inne inne pojęcia.
- Jeśli dowolny zbiór $X \subset U$ można przedstawić jako sumę klas to wówczas nazwamy go *definiowalnym*, w przeciwnym przypadku zbiór ten jest *przybliżony*.

Reprezentacja wiedzy

- W teorii zbiorów przybliżonych obiekty opisuje się poprzez atrybuty, przy czym atrybuty mogą przyjmować pewne wartości.
- Wygodnie jest reprezentować przestrzeń obiektów opisywanych atrybutami w postaci tablicy dwuwymiarowej, gdzie w wierszach są obiekty, a w kolumnach atrybuty.

Tablica informacyjna

- Punktem wyjściowym tej teorii są dane, zgrupowane w postaci *tablicy informacyjnej*. Przykład tablicy decyzyjnej przedstawiony jest poniżej,
-

	a_1	a_2	a_3	a_4
x_1	1	0	2	3
x_2	2	0	0	1
x_3	3	2	1	1
x_4	3	0	2	2
x_5	2	1	1	4

  x_1, x_2, x_3, x_4 i x_5 - obiekty
  a_1, a_2, a_3 i a_4 - atrybuty.

- W wierszach tej tablicy umieszczone są obiekty z uniwersum $U = \{x_1, x_2, x_3, x_4, x_5\}$,
- Kolumny zawierają atrybuty ze zbioru A, opisujące obiekty $A = \{a_1, a_2, a_3, a_4\}$,
- Atrybuty mogą przyjmować pewne wartości z dziedziny V.
- *Funkcja informacji* $f : U \times A \to V$ przypisuje wartości (ze zbioru V) atrybutom (ze zbioru A), opisujących obiekty z uniwersum U.
- Funkcję przypisującą wartość atrybutu a obiektowi x oznaczać będziemy jako $a(x)$.

Tablica decyzyjna

Czasami wyróżnia się dodatkowy atrybut d, zwany atrybutem decyzyjnym. Wartość tego atrybutu jest decyzją. *Tablica informacyjna* rozbudowuje się wówczas do *tablicy decyzyjnej* postaci

	a_1	a_2	a_3	a_4	d
x_1	1	0	2	3	2
x_2	2	0	0	1	2
x_3	3	2	1	1	1
x_4	3	0	2	2	0
x_5	2	1	1	4	3

.

Relacja *B-nierozróżnialności*

- Obiekt x_i jest nierozróżnialny z obiektem x_j ze względu na atrybuty $B \subseteq A$ wtedy i tylko wtedy gdy dla wszystkich atrybutów $a \in B$ przyjmują one te same wartości dla obu obiektów
$$\forall_{a \in B} a(x_i) = a(x_j)$$
- Relacja nierozróżnialności w odniesieniu do atrybutów ze zbioru B jest *B-nierozróżnialnością*.
- Klasę abstrakcji relacji *B-nierozróżnialności* wyznaczoną przez element x oznaczać będziemy przez $[x]_B$.

Dolna i górna aproksymacja zbioru przybliżonego

- Dla dowolnego zbioru obiektów $X \subseteq U$ i dla zbioru atrybutów $B \subseteq A$ *dolną aproksymacją zbioru* X nazywać będziemy zbiór określony jako
$$\underline{X} = \{x : [x]_B \subseteq X\}.$$
- *górną aproksymacją zbioru* X nazywać będziemy zbiór określony jako
$$\overline{X} = \{x : [x]_B \cap X \neq \emptyset\}.$$
- $$\underline{X} \subseteq X \subseteq \overline{X}.$$

Funkcja przynależności do zbioru przybliżonego

- Dla dowolnego $x \in U$ i dla niepustego podzbioru atrybutów $B \subseteq A$ funkcja przynależności do zbioru przybliżonego X (ang. *rough membership function*) określona jest jako

$$\mu_X^B(x) = \frac{|X \cap [x]_B|}{|[x]_B|}.$$

- Funkcja przynależności ma probabilistyczną interpretację jako prawdopodobieństwo warunkowe $p(X \backslash [x]_B)$.

- Wartość funkcji przynależności określa liczbowo jakie jest prawdopodobieństwo tego, że element x, który jest $B-$ nierozróżnialny, należy do zbioru X.

Reguły decyzyjne

- Tablice decyzyjną można interpretować jako zbiór reguł, tj. przypisanie obiektowi, opisanym atrybutami A pewnej decyzji d np.

$$(a_1, v_1) \wedge \ldots \wedge (a_k, v_k) \Longrightarrow (d, w),$$

  gdzie v_i jest wartością atrybutu a_i, a w jest decyzją.

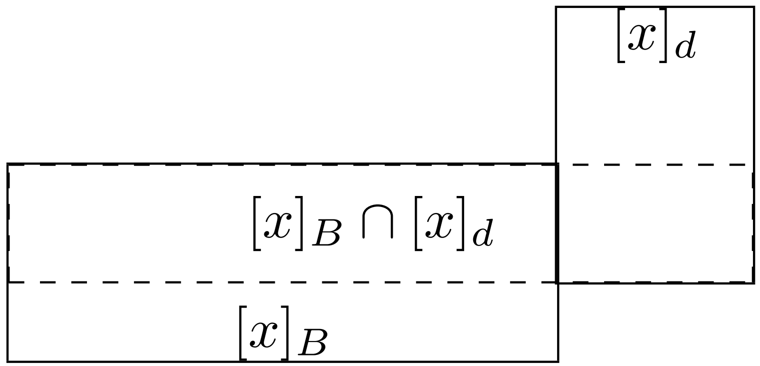
- System decyzyjny jest deterministyczny jeśli

$$[x]_A \subseteq [x]_d.$$

  gdzie $[x]_d$ oznaczymy klasę abstrakcji relacji nierozróżnialności ze względu na decyzję d wyznaczoną przez element x.

Konflikt decyzyjny

- Jeśli każdą klasę decyzyjną można przedstawić jako sumę klas abstrakcji relacji $B-$ nierozróżnialności to wówczas system jest deterministyczny (istnieje jednoznaczne przyporządkowanie takim samym obiektom identycznych decyzji).
- Gdy istnieje taka klasa decyzyjna $[x]_d$, że $[x]_B \not\subset [x]_d \wedge [x]_B \cap [x]_d \neq \emptyset$ istnieje konflikt decyzyjny (tj. identycznym obiektom przyporządkowane są różne decyzje).



Opis reguł decyzyjnych

- **wsparcie reguły** (ang. *support*) - jest ilością przypadków, w których reguła została zastosowana, tj. $|[x]_B \cap [x]_d|$.
- **dokładność reguły** (ang. *accuracy*) acc - określa liczbowo wsparcie reguły do liczności klasy decyzyjnej.

$$acc = \frac{|[x]_B \cap [x]_d|}{|[x]_d|}$$

- **pokrycie reguły** (ang. *coverage*) - określa liczbowo wsparcie reguły do liczności klasy abstrakcji relacji B-nierozrożnialności.

$$cov = \frac{|[x]_B \cap [x]_d|}{|[x]_B|}$$

Redukty

- Z punktu widzenia "ekonomii podejmowanych decyzji" należy dążyć do eliminacji tych atrybutów, które nie mają wpływu no podejmowaną decyzję. W tym celu pomocne jest wprowadzenie pojęcia reduktu.
- Mówimy że zbiór B jest reduktem zbioru atrybutów A jeśli $B \subset A$ i oba zbiory dzielą przestrzeń U na identyczne klasy nierozróżnialności.

Redukty - przykład

- Szukamy reduktu zbioru $A = \{a_1, a_2, a_3, a_4\}$ dla poniższej tablicy informacyjnej

	a_1	a_2	a_3	a_4
x_1	1	0	2	3
x_2	4	0	0	1
x_3	3	2	1	1
x_4	1	0	2	3
x_5	2	1	1	4
x_6	2	1	1	4

- Dla zbioru atrybutów $A = \{a_1, a_2, a_3, a_4\}$ otrzymujemy następujące klasy A-*nierozróżnialności*:

$$\begin{aligned} E_1^A &= \{x_1, x_4\} \\ E_2^A &= \{x_2\} \\ E_3^A &= \{x_3\} \\ E_4^A &= \{x_5, x_6\} \end{aligned}$$

- Zbiór $B_1 = \{a_1, a_2, a_3\}$ jest reduktem, gdyż $B \subset A$ i generuje identyczne klasy nierozróżnialności co A-nierozróżnialność:

$$
\begin{aligned}
E_1^{B_1} &= \{x_1, x_4\} \\
E_2^{B_1} &= \{x_2\} \\
E_3^{B_1} &= \{x_3\} \\
E_4^{B_1} &= \{x_5, x_6\}
\end{aligned}
$$

- Najmniejszym reduktem jest zbiór $B_2 = \{a_1\}$. Klasy abstrakcji relacji B_2- nierozróżnialności mają postać

$$
\begin{aligned}
E_1^{B_2} &= \{x_1, x_4\} \\
E_2^{B_2} &= \{x_2\} \\
E_3^{B_2} &= \{x_3\} \\
E_4^{B_2} &= \{x_5, x_6\}
\end{aligned}
$$

i są identyczne jak dla A- nierozróżnialności.

Kryteria wyboru reguł

- *Długość reguły* - preferowane są reguły krótkie
- Można zredukować dużą liczbę reguł klasyfikujących dokładnie na mniejszą liczbę reguł klasyfikującą z mniejszą dokładnością.
- *wsparcie reguły* - można zrezygnować z pełnego wsparcia zyskując większą "elastyczność" reguł i zmniejszając długość reguły