Sztuczna Inteligencja

Soma Dutta

Wydział Matematyki i Informatyki, UWM w Olsztynie
soma.dutta@matman.uwm.edu.pl

Wykład - 6: Zbiory przybliżone i Zbiory rozmyte

Semestr letni 2022

Podsumowanie ostatniego wykładu

- Redukty i reguły generowane przy użyciu reduktów
- Skracanie reguł w taki sposób, aby zawierały minimalna liczbę deskryptorów
- Klasyfikacja nowego przypadku za pomocą reguł

Plany

- Klasyfikacja zbioru przykładów testowych nie jest tak prosta - czasami trzeba wykonać operację aproksymacji
- Nauczymy się przybliżonych operacji bazujących na zbiorze przybliżonym
- Istnieje inne podejście do przybliżania zbioru obiektów, które nie mają jednoznacznego opisu. Jest to znane jako teoria zbiorów rozmytych.

Przykład - 1: pierwsza metoda

	a	b	c	d	dec
o_1	0	2	1	0	0
o_2	1	2	2	1	0
o_3	2	0	2	1	1
o_4	0	2	1	1	2
u	0	0	2	1	?

- Reguły:
  $b = 2 \wedge c = 1 \wedge d = 0 \Rightarrow dec = 0$
  $b = 2 \wedge c = 2 \wedge d = 1 \Rightarrow dec = 0$
  $b = 0 \wedge c = 2 \wedge d = 1 \Rightarrow dec = 1$
  $b = 2 \wedge c = 1 \wedge d = 1 \Rightarrow dec = 2$
- Po skróceniu:
  $b = 2 \wedge d = 0 \Rightarrow dec = 0$ lub $c = 1 \wedge d = 0 \Rightarrow dec = 0$
  $b = 2 \wedge c = 2 \Rightarrow dec = 0$
  $b = 0 \Rightarrow dec = 1$
  $c = 1 \wedge d = 1 \Rightarrow dec = 2$
- $Rules(u) = \{b = 0 \Rightarrow dec = 1\}$
- Zatem u jest klasyfikowany z $dec = 1$ z maksymalnym wsparciem 1.

Przykład-2

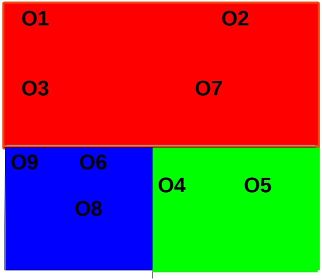
- Zakładamy, że a_1 = gorączka, a_2 = kontakt z pacjentem, a_3 = kaszel, d = Zainfekowany

	a_1	a_2	a_3	d
o_1	wysoka	bliski	średni	tak
o_2	wysoka	bliski	średni	tak
o_3	wysoka	bliski	średni	tak
o_4	więcej niż średnia	daleki	silny	nie pewne
o_5	więcej niż średnia	daleki	silny	nie
o_6	więcej niż średnia	daleki	lekki	nie
o_7	wysoka	bliski	średni	tak
o_8	więcej niż średnia	daleki	lekki	nie
o_9	więcej niż średnia	daleki	lekki	tak

- $\{\{o_1, o_2, o_3, o_7\}, \{o_4, o_5\}, \{o_6, o_8, o_9\}\}$
- $\{o_4, o_5\}$, $\{o_6, o_8, o_9\}$ mają sprzeczności, ponieważ mają elementy, które mają ten sam opis, ale różne decyzje

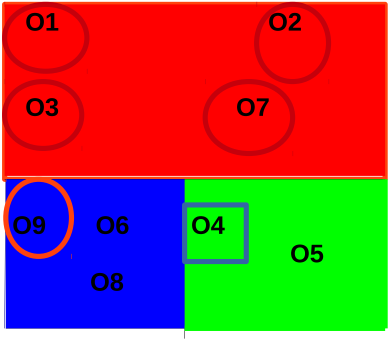
Klasy równoważności w odniesieniu do atrybutów warunkowych

- $\{\{o_1, o_2, o_3, o_7\}, \{o_4, o_5\}, \{o_6, o_8, o_9\}\}$
- Jak znaleźć reguły opisujące $\{o_4, o_5\}$ and $\{o_6, o_8, o_9\}$?



Klasy decyzyjne

- Klasy decyzyjne:
  - Tak: $X_1 = \{o_1, o_2, o_3, o_7, o_9\}$
  - Nie pewne: $X_2 = \{o_4\}$
  - Nie: $X_3 = \{o_5, o_6, o_8\}$
- Jak opisać klasy decyzyjne X_1, X_2, X_3?



Jak możemy opisać klasy decyzyjne

	a_1	a_2	a_3	d
o_1	wysoka	bliski	średni	tak
o_2	wysoka	bliski	średni	tak
o_3	wysoka	bliski	średni	tak
o_4	więcej niż średnia	daleki	silny	nie pewne
o_5	więcej niż średnia	daleki	silny	nie
o_6	więcej niż średnia	daleki	lekki	nie
o_7	wysoka	bliski	średni	tak
o_8	więcej niż średnia	daleki	lekki	nie
o_9	więcej niż średnia	daleki	lekki	tak

- Klasa decyzyjna $X_1 = \{o_1, o_2, o_3, o_7, o_9\}$
- Lower approximation (dolna aproksymacja): $\underline{X_1}_A = \cup\{[u]_A : [u]_A \subseteq X_1\} = \{o_1, o_2, o_3, o_7\}$ - Na podstawie informacji o elementach zawartej w wartościach atrybutów z A możemy stwierdzić, że elementy z pewnością należą do klasy decyzyjnej
- Upper approximation (górna aproksymacja): $\overline{X_1}_A = \cup\{[u]_A : [u]_A \cap X_1 \neq \phi\} = \{o_1, o_2, o_3, o_7, o_6, o_8, o_9\}$ - Na podstawie informacji o elementach zawartej w wartościach atrybutów z A możemy stwierdzić, że elementy być może należą do klasy decyzyjnej

Opis klasy decyzyjnej

- Zauważmy, że za pomocą A nie możemy zdefioniować jednoznacznie X_1, ponieważ istnieją pewne wystąpienia w X_1, które mają różne właściwości.
- Opis X_1 jest podzielony na trzy części: Pewne przypadki ($\underline{X_1}_A$), Możliwe przypadki ($\overline{X_1}_A$) i Graniczne przypadki ($\overline{X_1}_A \setminus \underline{X_1}_A$).
- Możemy opisać klasę decyzji X_1, czyli 'Zainfekowany = Tak', w następujący sposób.
  - Pewne przypadki : gorączka = wysoka $\wedge$ kontakt z pacjentem = bliski $\wedge$ kaszel = sliny $\Rightarrow$ zainfekowany = tak.
  - Możliwe przypadki: gorączka = więcej niż średnia $\wedge$ kontakt z pacjentem = daleki $\wedge$ kaszel = lekki $\Rightarrow$ zainfekowany = tak.

Radzenie sobie ze sprzecznymi danymi

- Generalized decision function (ugólniona funkcja decyzyjna) : $\partial_B(u) = \{dec(u') : u' \in [u]_B\}$.
- Na przykład, $\partial_A(o_4) = \{nie\ pewne, nie\}$
- Rough membership function (funkcja przybliżonego należenia): $\vec{\mu}_B(u) = \langle \mu_B^1(u), \mu_B^2(u), \ldots \mu_B^n(u) \rangle$ gdzie $\mu_B^i(u) = \frac{|[u]_B \cap X_i|}{|[u]_B|}$ dla każdej wartości decyzji i.
- Na przykład, $[o_6]_A = \{o_6, o_8, o_9\}$. Więc $\vec{\mu}_A(o_6) = \langle \frac{1}{3}, 0, \frac{2}{3} \rangle$.
- Zakładając $B = \{a_1, a_2\}$, mamy tylko dwie klasy równoważności: $\{\{o_1, o_2, o_3, o_7\}, \{o_4, o_5, o_6, o_8, o_9\}\}$.
  Więc, $[o_6]_B = \{o_4, o_5, o_6, o_8, o_9\}$ i $\vec{\mu}_B(o_6) = \langle \frac{1}{5}, \frac{1}{5}, \frac{3}{5} \rangle$.
- Niech $C = \{a_2, a_3\}$. Należy notować, że C to redukt w poprzednim znaczeniu.
- Zauważmy, że $[u]_A = [u]_C$ dla każdego elementu $u \in \{o_1, \ldots, o_9\}$.
- Tak więc $\partial_A(u) = \partial_C(u)$ oraz $\vec{\mu}_A(u) = \vec{\mu}_C(u)$ dla każdego elementu $u \in \{o_1, \ldots, o_9\}$.

Niesprzeczna tabela decyzyjna

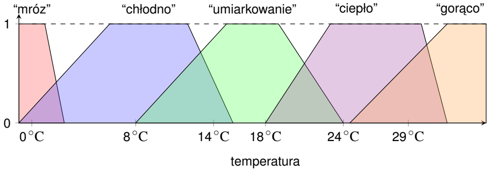
	a_1	a_2	a_3	d	∂_A	$\overrightarrow{\mu}_A$
o_1	wysoka	bliski	średni	tak	$\{tak\}$	$\langle 1, 0, 0 \rangle$
o_2	wysoka	bliski	średni	tak	$\{tak\}$	$\langle 1, 0, 0 \rangle$
o_3	wysoka	bliski	średni	tak	$\{tak\}$	$\langle 1, 0, 0 \rangle$
o_4	więcej niż średnia	daleki	silny	nie pewne	$\{nie\ pewne, nie\}$	$\langle 0, \frac{1}{2}, \frac{1}{2} \rangle$
o_5	więcej niż średnia	daleki	silny	nie	$\{nie\ pewne, nie\}$	$\langle 0, \frac{1}{2}, \frac{1}{2} \rangle$
o_6	więcej niż średnia	daleki	lekki	nie	$\{tak, nie\}$	$\langle \frac{1}{3}, 0, \frac{2}{3} \rangle$
o_7	wysoka	bliski	średni	tak	$\{tak\}$	$\langle 1, 0, 0 \rangle$
o_8	więcej niż średnia	daleki	lekki	nie	$\{tak, nie\}$	$\langle \frac{1}{3}, 0, \frac{2}{3} \rangle$
o_9	więcej niż średnia	daleki	lekki	tak	$\{tak, nie\}$	$\langle \frac{1}{3}, 0, \frac{2}{3} \rangle$

- (gorączka = więcej niż średnia) $\wedge$ (kontakt z pacjentem = daleki) $\wedge$ (kaszel = silny) $\Rightarrow$ $Zainfekowany|\partial_A = \{nie\ pewne, nie\}$
- (gorączka = więcej niż średnia) $\wedge$ (kontakt z pacjentem = daleki) $\wedge$ (kaszel = silny) $\Rightarrow$ $Zainfekowany|\mu_A\ \langle 0, \frac{1}{2}, \frac{1}{2} \rangle$

Przybliżenie przy użyciu metody zbiorów rozmytych

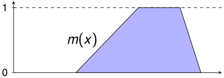
- To podejście różni się nieco od wcześniejszego. Podejście to uogólnia definicję zbioru w następujący sposób.
- Zbiór wszystkich parzystych liczb naturalnych: $E = \{2, 4, 6, 8, \ldots\} \subseteq \mathbb{N}$
- Funkcja charakterystyczna: $Ch_E : \mathbb{N} \mapsto \{0, 1\}$ jest taka, że $Ch_E(x) = 1$ jeśli $x \in E$, inaczej $Ch_E(x) = 0$.
- Jak zapisujemy zbiór wszystkich liczb rzeczywistych bliskich 2? To nie jest jednoznaczny opis i może mieć różne interpretacje.
- Możemy ten zbiór opisać za pomocą funkcji $\mu_2 : \mathbb{R} \mapsto [0, 1]$ takiej że
$$\begin{aligned} \mu_2(x) &= 1, \text{ jeśli } 0 \leq |2 - x| \leq 1 \\ &= .9, \text{ jeśli } 1 < |2 - x| \leq 2 \\ &\vdots \end{aligned}$$
- Zamiast dwóch wartości logicznych (prawda i fałsz), dopuszcza się istnienie nieskonczenie wielu wartości (odpowiadających liczbom rzeczywistym od 0 do 1)

- Tak więc dla każdego $x \in \mathbb{R}$, $\mu_2(x)$ reprezentuje stopień przynależności x do pojęcia (bliskich 2) reprezentowanego przez μ_2.
- Również można to zapisać następująco $(x, \mu_2(x))$ dla każdego $x \in \mathbb{R}$.
- Zbiór rozmyty: Zbiór rozmyty z dziedziną X jest reprezenotwany przez funkcję $f : X \mapsto [0, 1]$. Funkcję f nazywamy zbiorem rozmytym na dziedzinie X.

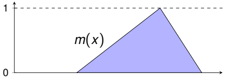


Pojęcia "ciepło" czy "gorąco" są określone w sposób nieostry: trudno jednoznacznie określić ich granice, ich zakresy mogą się częściowo pokrywać.

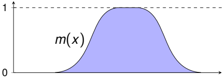
Funkcje mogą mieć kształt trapezu...

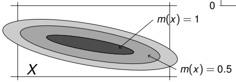


...trójkąta...



...ale też inny (np. sigmoidalny)...





...a zbiór X nie mysi być zbiorem liczb rzeczywistych

Reguły w rozmytych zbiorach

- We wcześniejszym podejściu zaczęliśmy od tabeli zawierającej przypadki i ich opisy pod względem wartości niektórych atrybutów
- gorączka = wysoka $\wedge$ kontakt z pacjentem = bliski $\wedge$ kaszel = sliny $\Rightarrow$ zainfekowany = tak.
- W aktualnie rozważanym podejściu każde zdanie ma pewną wartość, które mogą być inne niż prawda lub fałsz, można je nazwać stopniem prawdy.
- Na przykład, reprezentujmy pojęcie wysoka jak o $\mu_{wysoka} : [35, 42] \mapsto [0, 1]$ gdzie $[35, 42]$ jest dziedzinę dla temperatury gorączki. Może ta funkcja być taka, że

$$
\begin{aligned}
\mu_{wysoka}(x) &= 1 \text{ jeśli } x \geq 40 \\
&= .9 \text{ jeśli } 38 \leq x < 40 \\
&= .7 \text{ jeśli } 37 \leq x < 38 \\
&= .5 \text{ jeśli } 36 \leq x < 37 \\
&= .1 \text{ jeśli } 35 \leq x < 36
\end{aligned}
$$

- Gorączka może więc być traktowana jako zmienna, który może mieć wartości z $[35, 42]$. Jeśli gorączka ma wartość 38, stopień prawdziwości zdania 'gorączka = wysoka' jest $\mu_{wysoka}(38) = .9$

Obliczanie wartości prawdy w zdaniach prostych i złożonych

- 'gorączka = wysoka' jest przykładem prostego zdania.
- Zauważyliśmy, że każde proste zdanie jest reprezentowane przez zbiór rozmyty, czyli funkcja z odpowiedniej domeny do przedziału [0, 1]. Wartość prawdziwości zdania zależy od funkcji odpowiadającej temu rozmytemu zbiorowi.
- Np. zbiór rozmyty 'wysoka' jest reprezentowany przez funkcję μ_{wysoka} z domeny temperatury do [0, 1]. Tutaj 'wysoka' dotyczy atrybutu 'gorączka', więc jako domenę mamy przedział temperatur.
- Podobnie 'silny' jest stosowany do atrybutu 'kaszel', więc możemy przyjąć domenę reprezentującą różną intensywność kaszlu, a pojęcie 'silny' można przedstawić jako zbiór rozmyty, czyli funkcja μ_{silny}, która zwraca wartości w [0, 1].
- Przypuszczamy
  - gorączka = wysoka ma stopień prawdziwości .9,
  - kaszel = sliny ma stopień prawdziwości .7
- Jak obliczyć wartość prawdy zdania złożonego:
  (gorączka = wysoka) $\wedge$ (kaszel = sliny)

Spójniki logiczne: klasyczny scenariusz z dwiema wartościami

Spójniki logiczne: $\wedge$ (oraz), $\vee$ (lub), $\Rightarrow$ (jeśli . . . to), $\neg$ (nie)

- Niech p i q będą dwoma zdaniami.

$\wedge$	0	1
0	0	0
1	0	1

$\vee$	0	1
0	0	1
1	1	1

$\Rightarrow$	0	1
0	1	1
1	0	1

	$\neg$
0	1
1	0

- $\neg : \{0,1\} \mapsto \{0,1\}$, i dla pozostałych spójników mamy $\{0,1\} \times \{0,1\} \mapsto \{0,1\}$.
- Spójniki logiczne w kontekście zbiorów rozmytych są generalizowane z kontekstu dwuwartościowego

Logika rozmyta: Koniunkcja

- Niech α, β będą zbiorami rozmytymi na dziedzinie X. Tzn., $\mu_\alpha : X \mapsto [0,1]$ i $\mu_\beta : X \mapsto [0,1]$
- Koniunkcję α i β definiujemy jako zbiór rozmyty $\alpha \wedge \beta$ o funkcji przynależności $\mu_{\alpha\wedge\beta} : X \mapsto [0,1]$ określony wzorem $\mu_{\alpha\wedge\beta}(x) = T(\mu_\alpha(x), \mu_\beta(x)) \quad \forall x \in X$
- Funkcja $T : [0,1] \times [0,1] \mapsto [0,1]$ jest T-normą.

Własności T-normy

- Warunki brzegowe:
  $T(0, x) = 0$ oraz $T(1, x) = x$.
- Monotoniczność:
  Jeśli $x \leq y, \ T(x, z) \leq T(y, z)$.
- Symetria:
  $T(x, y) = T(y, x)$
- Łączność:
  $T(x, T(y, z)) = T(T(x, y), z)$

Przykładowe T-normy

- T-norma Zadeha:
  $T(x, y) = min(x, y) \quad \forall x, y \in [0, 1]$
- T-norma Mengara:
  $T(x, y) = x.y \quad \forall x, y \in [0, 1]$
- T-norma Łukasiewicza:
  $T(x, y) = max(0, x + y - 1) \quad \forall x, y \in [0, 1]$

Przykład

Przyjmujemy $X = \{x_1, x_2, x_3, x_4\}$

- Dla α równego $\{(x_1, 0.4), (x_2, 0), (x_3, 0.5), (x_4, 1)\}$ oraz β równego $\{(x_1, 0.6), (x_2, 0.5), (x_3, 0), (x_4, 1)\}$ otrzymujemy następujące wyniki dla różnych koniunkcji :
- Zadeh: $\alpha \wedge \beta$ równa się zbiorowi rozmytemu $\{(x_1, 0.4), (x_2, 0), (x_3, 0), (x_4, 1)\}$
- Mengar: $\alpha \wedge \beta$ równa się zbiorowi rozmytemu $\{(x_1, 0.24), (x_2, 0), (x_3, 0), (x_4, 1)\}$
- Łukasiewicz: $\alpha \wedge \beta$ równa się zbiorowi rozmytemu $\{(x_1, 0), (x_2, 0), (x_3, 0), (x_4, 1)\}$

Altarnetywa jako Ko-norma

- Warunki brzegowe:
  $C(0, x) = x$ oraz $C(1, x) = 1$.
- Monotoniczność:
  Jeśli $x \leq y$, $C(x, z) \leq C(y, z)$.
- Symetria:
  $C(x, y) = C(y, x)$
- Łączność:
  $C(x, C(y, z)) = C(C(x, y), z)$

Przykładowe Ko-normy

- Ko-norma Zadeha:
  $C(x, y) = max(x, y) \quad \forall x, y \in [0, 1]$
- K-norma Mengara:
  $C(x, y) = x + y - xy \quad \forall x, y \in [0, 1]$
- T-norma Łukasiewicza:
  $C(x, y) = min(x + y, 1) \quad \forall x, y \in [0, 1]$

Negacja: zbiór rozmyty

- Negację zbioru α definiujemy jako zbiór $\neg\alpha$ o funkcji przynależności $\mu_{\neg\alpha} : X \mapsto [0,1]$ określonej wzorem $\mu_{\neg\alpha}(x) = 1 - \mu_{\alpha}(x)$.
- Dla α równego $\{(x_1, 0.4), (x_2, 0), (x_3, 0.5), (x_4, 1)\}$, $\neg\alpha$ równa się: $\{(x_1, 0.6), (x_2, 1), (x_3, 0.5), (x_4, 0)\}$

Dualność: T-norm i ko-norm

- $\alpha \vee \beta = \neg(\neg\alpha \wedge \neg\beta)$
- Na przykład: $(min, max, 1 - x)$
- (T-Łukasiewicz, C-Łukasiewicz, 1-x)

Implikacja dla zbiorów rozmytych

- Jak w kontekscie klasycznym, możemy zdefinować:
  $\alpha \Rightarrow \beta = \neg\alpha \vee \beta$.
- Zadeh: $\mu_{\alpha\Rightarrow\beta}(x) = max(1 - \mu_\alpha(x), \mu_\beta(x))$
- Łukasiewicz: $\mu_{\alpha\Rightarrow\beta}(x) = min(1, 1 - \mu_\alpha(x) + \mu_\beta(x))$

Dziękuję za uwagę