

# Systemy uczące się – wykład 2

dr Przemysław Juszczuk

Katedra Inżynierii Wiedzy, Uniwersytet Ekonomiczny

19 X 2018

## Podstawowe definicje

- Fakt;
- Przestanka;
- Konkluzja;
- Reguła;
- Wnioskowanie.

## Typy wnioskowania

- Wnioskowanie w przód : wnioskowanie od faktów do celu (wnioskowanie sterowane danymi).
- Wnioskowanie w tył : wnioskowanie od celu do faktów (wnioskowanie sterowane celem).
- Wnioskowanie mieszane : cechy wnioskowania w tył i w przód. Np. podział bazy wiedzy na dwie części dla wnioskowania w przód oraz w tył.

## Przykład wnioskowania

Dana jest baza wiedzy :

- R1: jeżeli „a” i „b” i „c” to „d”
- R2: jeżeli „a” i „b” to „g”
- R3: jeżeli „b” i „c” to „e”
- R4: jeżeli „a” i „c” to „f”
- R5: jeżeli „e” i „b” i „c” to „f”

Dane są fakty : „a” , „b” , „c” . Celem wnioskowania jest „f” .

## Przykład wnioskowania

Dana jest baza wiedzy :

- R1: jeżeli „a” i „b” i „c” to „d”
- R2: jeżeli „a” i „d” to „g”
- R3: jeżeli „a” i „f” to „b”
- R4: jeżeli „b” i „g” to „f”
- R5: jeżeli „a” i „e” to „f”
- R6: jeżeli „e” i „f” to „a”
- R7: jeżeli „a” i „b” to „c”

Dane są fakty : „a” i „e”. Udowodnić hipotezę „g”.

- Tablicowe przedstawienie wiedzy KRS - Knowledge Representation System.
- Tablica decyzyjna jest modyfikacją KRS.
- Definicja bazy wiedzy:

$$K = (U, R),$$

U - skończony zbiór obiektów zwany uniwersum,

R =  $\{R_1, R_2, \dots, R_n\}$  - zbiór relacji równoważnościowych nad U

- KRS to skończona tablica, w której rzędy są etykietowane przez obiekty a kolumny przez atrybuty
- na przecięciu wiersza i kolumny znajduje się wartość atrybutu danego obiektu.

Tablica: Klasyfikacja zabawek.

	<b>Kolor</b>	<b>Kształt</b>	<b>Materiał</b>	<b>Wielkość</b>	<b>Cena</b>
Miś	brązowy	owalny	plusz	duża	niska
Lalka	różowy	owalny	guma	średnia	wysoka
Samochód	czerwony	kanciasty	metal	mała	wysoka
Piłka	zielony	okrągły	plastyk	średnia	średnia

$$U = \{\{Mis\}, \{Lalka\}, \{Samochod\}, \{Pilka\}\}$$

$$A = \{\{Kolor\}, \{Kształt\}, \{Material\}, \{Wielkosc\}, \{Cena\}\}$$

Tablica: Klasyfikacja zabawek.

	Kolor	Kształt	Materiał	Wielkość	Cena
Miś	brązowy	owalny	plusz	duża	niska
Lalka	różowy	owalny	guma	średnia	wysoka
Samochód	czerwony	kanciasty	metal	mała	wysoka
Piłka	zielony	okrągły	plastyk	średnia	średnia

- Klasa abstrakcji obiektu  $x$  relacji  $R$  (oznaczamy  $[x]_R$ ) to zbiór tych obiektów z  $U$ , które są w relacji  $R$  z obiektem  $x$ .
- Przez  $U/R$  oznaczamy zbiór wszystkich klas abstrakcji  $R$ .
- **Klasy abstrakcji są rozłączne i tworzą pokrycie zbioru  $U$**
- Przez  $U/IND(B)$  oznaczmy klasy abstrakcji relacji  $IND(B)$ .



Tablica: Tablica decyzyjna.

	<b>a</b>	<b>b</b>	<b>c</b>	<b>d</b>	<b>e</b>
$X_1$	1	0	2	2	0
$X_2$	0	1	1	1	2
$X_3$	2	0	0	1	1
$X_4$	1	1	0	2	2
$X_5$	1	0	2	0	1
$X_6$	2	2	0	1	1
$X_7$	2	1	1	2	2
$X_8$	0	1	1	0	1

Obiekty :  $X = \{1, \dots, 8\}$

Atrybuty warunkowe :  $C = \{a, b, c\}$

Atrybuty decyzyjne :  $D = \{d, e\}$

Atrybuty :  $A = C \cup D$

Wartości a :  $V_a = \{0, 1, 2\}$

Wartości b :  $V_b = \{0, 1, 2\}$

Wartości c :  $V_c = \{0, 1, 2\}$

## Skrócona metoda zero-jedynkowa

Tablica: Skrócona zero-jedynkowa

$(p$	$\wedge$	$q)$	$\rightarrow$	$(q$	$\wedge$	$p)$
	1					
1	1	1				
1	1	1		1		1
1	1	1		1	1	1
1	1	1	1	1	1	1

## Dane temporalne - definicja

- Niech  $T = t_0, t_1, \dots, t_n$  - ciąg etykiet czasu;
- $\forall_i \in T, t_i - t_{i-1} = \Delta t_i = 1$ ;

## Dane temporalne - przykład

- $t_1 : a_1 = 0.3; a_2 = 0.6; a_3 = 0.1$ ;
- $t_2 : a_1 = 0.6; a_2 = 0.2; a_3 = 0.5$ ;
- $t_n : a_1 = 0.3; a_2 = 0.3; a_3 = 0.6$ ;

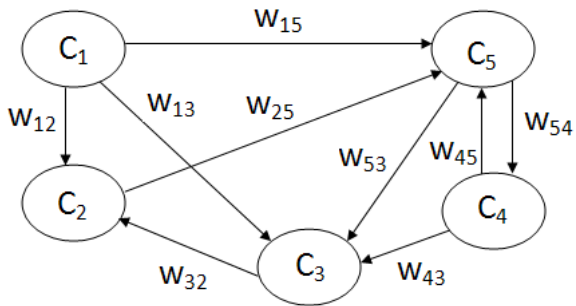
## Czym jest sieć kognitywna?

- Jedna z metod reprezentacji wiedzy stosowanych w systemach wspomagania decyzji.
- Zainspirowane biologią i psychologią.
- Korzystają z takich elementów jak : pojęcie, relacja przyczynowa.
- Mają formę grafu.
- Przy pomocy sieci kognitywnej zaprojektować można pewien proces decyzyjny, lub środowisko.

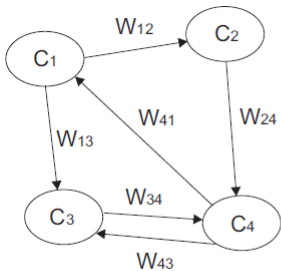
## Sieć kognitywna

$$FCM = \langle C, A, W \rangle \quad (1)$$

gdzie:  $C$  jest skończonym zbiorem pojęć,  $A$  to zbiór aktywacji pojęć ( $a_i \in [0, 1]$ ),  $W$  zbiór wartości wag  $w_{ij} \in [-1, 1]$ .



Rysunek: Sieć kognitywna



a)

	C1	C2	C3	C4
C1	0	$W_{12}$	$W_{13}$	$W_{14}$
C2	$W_{21}$	0	$W_{23}$	$W_{24}$
C3	$W_{31}$	$W_{32}$	0	$W_{34}$
C4	$W_{41}$	$W_{42}$	$W_{43}$	0

b)

Rysunek: a - reprezentacja grafowa; b - reprezentacja macierzowa

## Rozmycie sieci

- Podstawowa wersja sieci kognitywnej zakłada dwa stany : dodatni wpływ, oraz ujemny wpływ pojęć na siebie.
- FCM pozwala na określenie częściowego ujemnego, lub dodatniego wpływu.
- Rozmycie określane jest na podstawie pewnych ustalonych poziomów.

## Rozmycie

- bardzo słaby 0 – 0.2
- słaby 0.2 – 0.4
- średni 0.4 – 0.6
- silny 0.6 – 0.8
- bardzo silny 0.8 – 1.0



## Więcej o macierzach

- Macierz nie jest częścią sieci kognitywnej, tylko jej strukturą pomocniczą.
- Macierz wskazuje zależności pomiędzy pojęciami.
- Macierz jest strukturą kwadratową, gdzie liczba wierszy i kolumn równa jest liczbie pojęć.
- Każda komórka macierzy to jedno połączenie pomiędzy pojęciami.
- W przypadku braku zależności pomiędzy pojęciami, w danej komórce znajduje się 0.
- Wartość w komórce macierzy określa siłę wpływu (wagę) jednego pojęcia na inne.
- Wagi znajdują się w przedziale  $[-1, 1]$ , gdzie -1 określa wpływ ujemny, natomiast 1 dodatni.

## Uczenie sieci - problem

- Dane historyczne dla pojęć,
- Brak informacji na temat zależności pomiędzy pojęciami,
- Brak informacji na temat wag sieci,

## Uczenie sieci - problem

- Dane historyczne dla pojęć,
- Brak informacji na temat zależności pomiędzy pojęciami,
- Brak informacji na temat wag sieci,

## Zadanie

- Wykrycie zależności pomiędzy pojęciami,
- Wykrycie wartości wag pomiędzy pojęciami.

## Uczenie sieci - problem II

Potrzebny jest algorytm, który w sposób automatyczny potrafi:

- Określić zbiór pojęć danej sieci,
- Znaleźć zależności pomiędzy nimi,
- Obliczyć wpływ poszczególnych pojęć na siebie.

## Uczenie sieci

Znane są dwie główne metody uczenia rozmytych sieci kognitywnych:

- Uczenie z wykorzystaniem wiedzy eksperta z danej dziedziny.
- Automatyczne generowanie sieci z danych historycznych.

## Metoda klasyczna

Pierwsza opisywana metoda opiera się na wykorzystaniu pomocy ekspertów dziedzinowych. Zadaniem ekspertów jest:

- Określenie kluczowych pojęć.
- Wskazanie relacji pomiędzy pojęciami.
- Ustalenie siły wpływu poszczególnych pojęć na siebie.

## Wnioskowanie w FCM

$$C_i(t+1) = \gamma(\sum_{j=1}^n w_{ij} \cdot C_j(t))$$

$$C_1(t+1) = 0.9 \cdot 0.9 + 0.8 \cdot 0.8$$

$C_1 = 1.45$ , a wartość  $C$  musi należeć do przedziału  $[0, 1]$ .

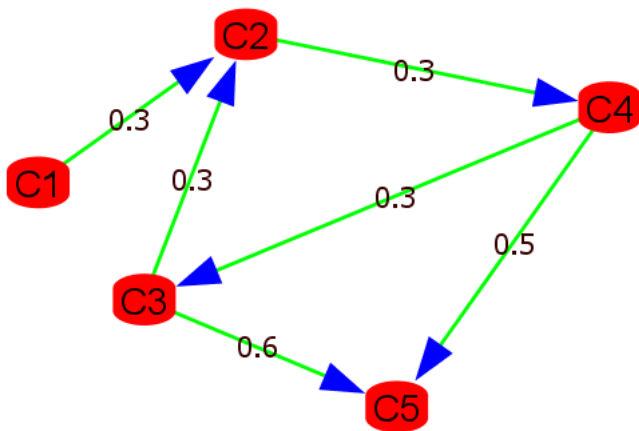
$\gamma$

$\gamma$  pełni rolę funkcji normalizującej wartość pojęcia do przedziału  $[0, 1]$ .

## Coś o normalizacji

Wartość każdego pojęcia zmieniana jest tak, aby pasowała do przedziału  $[0, 1]$ . Dokonać można tego za pomocą specjalnej funkcji zwanej funkcją sigmoidalną:

$$f(x) = \frac{1}{1+e^{-\alpha x}}$$



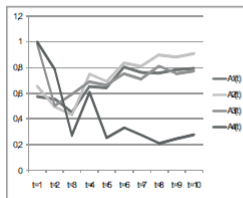
Rysunek: Rozmyta sieć kognitywna

$$f = \frac{1}{(t_e - 1) \cdot n} \cdot \sum_{t=t_s}^{t_e} \sum_{i=1}^n |a_i(t) - a'_i(t)|^p, \quad (2)$$

where:

- $t_l$  dolna granica okna czasowego oraz indeks początkowy serii danych;
- $t_u$  górna granica okna czasowego oraz indeks końcowy serii danych;
- $n = \text{card}(C)$  liczba pojęć;
- $p$  parametr sterujący procesem uczenia  $p = 1$ ,
- $a_n(t)$  obserwowana wartość  $i$ -tego pojęcia w chwili czasu  $t$
- $a'_n(t)$  obserwowana wartość wygenerowana przez FCM.





	$A_1$	$A_2$	$A_3$	$A_4$
t=1	0.57609	0.65286	0.98	1
t=2	0.5563	0.49524	0.5	0.78788
t=3	0.45	0.43	0.59	0.27273
t=4	0.65434	0.75056	0.69	0.60607
t=5	0.6413	0.69	0.66	0.25212
t=6	0.80435	0.83928	0.75	0.33333
t=7	0.76304	0.81	0.71	0.21273
t=8	0.756	0.9	0.81	0.21212
t=9	0.78319	0.88381	0.75	0.24697
t=10	0.7963	0.9112	0.77	0.28234

Rysunek: Atraktor chaotyczny

## Zadanie klasyfikacji – dane wejściowe

Zestaw danych w dowolnym formacie, w którym uwzględniony został zestaw atrybutów warunkowych oraz atrybut decyzyjny.

## Zadanie klasyfikacji – dane wyjściowe

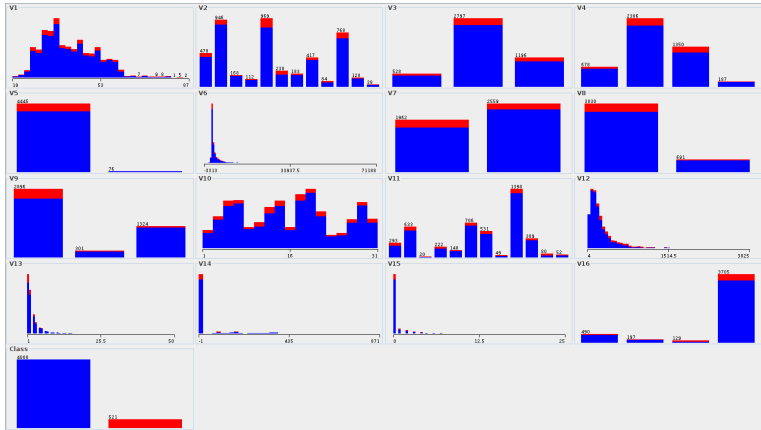
Model pozwalający wskazać nowym danym wyjściowym określoną wartość atrybutu decyzyjnego na podstawie obserwacji dostępnych z danych wejściowych.

## Predykcja klasy decyzyjnej

Zadaniem klasyfikatora jest predykcja pewnej nieznannej wartości atrybutu decyzyjnego danego obiektu.

## Proces klasyfikacji

- zbiór trenujący i testujący;
- klasyfikator budowany jest na podstawie zbioru trenującego;
- ocena jakości klasyfikacji bazuje na zbiorze testującym;
- wynikiem klasyfikacji może być zbiór reguł, budowa klasyfikatora;
- współczynnik dokładności jako miara oceny klasyfikacji.



Popularne miary jakości klasyfikacji:

- dokładność,
- czułość,
- precyzja.

# Macierz błędu

Informacje zawarte w macierzy dotyczą klasy decyzyjnej obiektu. Na podstawie macierzy możliwe jest wskazanie wartości miar oceny klasyfikacji takich jak dokładność, precyzja, czy czułość.

	Przewidziana pozytywna	Przewidziana negatywna
Pozytywne przykłady	True positive ( <i>TP</i> )	False negative ( <i>FN</i> )
Negatywne przykłady	False positive ( <i>FP</i> )	True negative ( <i>TN</i> )

Klasyczna miara oceny jakości klasyfikacji. Opisuje liczbę obiektów, które zostały sklasyfikowane poprawnie, do wszystkich obiektów.

$$ev_{acc}(T, S) = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

	Przewidziana pozytywna	Przewidziana negatywna
Pozytywne przykłady	True positive ( <i>TP</i> )	False negative ( <i>FN</i> )
Negatywne przykłady	False positive ( <i>FP</i> )	True negative ( <i>TN</i> )

Miara stosowana w przypadku klasyfikacji binarnej. Wskazuje stosunek liczby obiektów zaklasyfikowanych do klasy  $P$  i do wszystkich obiektów, które powinny zostać zaklasyfikowane do  $P$ .

$$ev_{rec}(T, S) = \frac{TP}{(TP + FN)}$$

	Przewidziana pozytywna	Przewidziana negatywna
Pozytywne przykłady	True positive ( $TP$ )	False negative ( $FN$ )
Negatywne przykłady	False positive ( $FP$ )	True negative ( $TN$ )



Pozwala wyznaczyć liczbę niepoprawnych prób klasyfikacji obiektów z klasy  $N$  do klasy  $P$ . Jest to stosunek obiektów poprawnie sklasyfikowanych do  $P$  do wszystkich obiektów przypisanych do tej klasy.

$$ev_{prec}(T, S) = \frac{TP}{(TP + FP)}$$

	Przewidziana pozytywna	Przewidziana negatywna
Pozytywne przykłady	True positive ( $TP$ )	False negative ( $FN$ )
Negatywne przykłady	False positive ( $FP$ )	True negative ( $TN$ )

## Metody klasyfikacji

- drzewa decyzyjne;
- klasyfikator Bayesa;
- sieci neuronowe;
- algorytm k-nn;
- metaheurystyki.

## Klasyfikacja a predykcja

- w przypadku klasyfikacji chodzi nam o wskazanie klasy decyzyjnej dla określonego obiektu, przy czym wartości atrybutu decyzyjnego są dyskretne;
- predykcja to modelowanie funkcji ciągłych.

Dziękuję za uwagę.