

TREŚCI ZADAŃ

CZĘŚĆ 1

Na tablicy decyzyjnej zaproponowanej przez prowadzącą/prowadzącego zająć wykonać zadania od 1 do 6.

1. Określić zbiory elementarne (klasy nieodróżnialności) ze względu na cały zbiór atrybutów warunkowych.
2. Wypisać klasy decyzyjne i podać dolne i górne aproksymacje oraz obszary brzegowe tych klas ze względu na cały zbiór atrybutów warunkowych.
3. Wyznaczyć obszar pozytywny.
4. Wygenerować macierz rozróżnialności.
5. Wyznaczyć co najmniej jeden redukt i rdzeń.
6. Wygenerować reguły pewne, możliwe i przybliżone z wykorzystaniem algorytmu LEM2.

CZĘŚĆ 2

W wielu zagadnieniach data mining ważne jest określenie istotności poszczególnych atrybutów warunkowych.

Zaproponuj liczbowy współczynnik istotności atrybutu warunkowego na podstawie wpływu atrybutu na zwiększenie wielkości obszaru pozytywnego. Wyznacz ten współczynnik dla wszystkich atrybutów warunkowych z zaproponowanej przez siebie przykładowej tablicy decyzyjnej.

Kod programu

```
SIZB.R* x ind x diabetes.csv x
Source on Save
1 library(RoughSets)
2 file="C:\\Users\\HP\\Desktop\\Nowy folder\\diabetes.csv"
3 sep=';'
4 df=read.table(file=file,header=T,sep=sep)
5 dt = SF.asDecisionTable(dataset = df,decision.attr = 13,indx.nominal = c(1:12) )
6 ind = BC.IND.relation.RST(dt,c(1:12))
7 app = BC.LU.approximation.RST(dt,ind)
8 boundary = BC.boundary.reg.RST(dt, app)
9 positive = BC.positive.reg.RST(dt,app)
10 dm=BC.discernibility.mat.RST(dt, return.matrix=TRUE)
11 reducts = FS.all.reducts.computation(dm)
12 RI.LEM2Rules.RST(dt)
```

Zbiory elementarne

Name	Type	Value
ind	list [3] (S3: IndiscernibilityRelatio	List of length 3
IND.relation	list [103]	List of length 103
m 0 0 nie KIT nie nie 0...	integer [1]	59
m 0 0 nie KIT tak nie ...	integer [1]	106
m 0 0 nie KIT tak nie ...	integer [1]	94
m 0 0 nie KIT tak nie ...	integer [1]	2
m 0 0 nie KIT_IIT tak n...	integer [1]	63
m 0 0 tak KIT tak nie ...	integer [1]	80
m 0 0 tak KIT_IIT tak n...	integer [1]	49
m 0 1 tak KIT nie nie ...	integer [1]	77
m 0 1 tak KIT tak nie ...	integer [1]	90

m 0 1 tak KIT tak nie ...	integer [1]	64	z 0 1 nie KIT tak nie 1 ...	integer [1]	60
m 0 1 tak KIT tak nie ...	integer [1]	18	z 0 1 nie KIT tak nie 1 ...	integer [1]	82
m 0 1 tak KIT tak nie ...	integer [1]	103	z 0 1 tak KIT tak nie 0 ...	integer [1]	81
m 0 2 nie KIT_IIT nie n...	integer [1]	102	z 0 2 nie KIT_IIT nie ni...	integer [1]	51
m 0 2 nie KIT_IIT nie n...	integer [1]	47	z 1 0 nie KIT nie nie 0 ...	integer [1]	19
m 1 0 nie KIT nie nie 0...	integer [1]	54	z 1 0 nie KIT tak nie 1 ...	integer [1]	25
m 1 0 nie KIT nie nie 1...	integer [1]	13	z 1 0 nie KIT_IIT nie ni...	integer [1]	62
m 1 0 nie KIT nie nie 2...	integer [1]	34	z 1 0 nie KIT_IIT tak ni...	integer [1]	1
m 1 0 nie KIT tak nie ...	integer [2]	15 65	z 1 0 nie KIT_IIT tak ni...	integer [1]	12
m 1 0 nie KIT tak nie ...	integer [1]	22	z 1 0 nie KIT_IIT tak ni...	integer [1]	76
m 1 0 nie KIT tak tak ...	integer [1]	57	z 1 0 tak KIT_IIT tak ni...	integer [1]	5
m 1 0 nie KIT_IIT tak n...	integer [1]	33	z 1 1 nie KIT nie nie 2 ...	integer [1]	58
m 1 0 nie KIT_IIT tak n...	integer [1]	75	z 1 1 nie KIT nie nie 2 ...	integer [1]	53
m 1 0 nie KIT_IIT tak n...	integer [1]	69	z 1 1 nie KIT tak nie 0 ...	integer [1]	98
m 1 0 nie KIT_IIT tak n...	integer [1]	86	z 1 1 nie KIT tak nie 0 ...	integer [1]	93
m 1 0 tak KIT_IIT nie n...	integer [1]	48	z 1 1 nie KIT tak nie 1 ...	integer [2]	17 21
m 1 0 tak KIT_IIT nie n...	integer [1]	46	z 1 1 nie KIT_IIT nie ta...	integer [1]	24
m 1 0 tak KIT_IIT tak n...	integer [2]	37 73	z 1 1 nie KIT_IIT tak ni...	integer [1]	74
m 1 1 nie KIT nie nie 0...	integer [1]	83	z 1 1 nie KIT_IIT tak ni...	integer [1]	30
m 1 1 nie KIT nie nie 2...	integer [1]	16	z 1 1 nie KIT_IIT tak ni...	integer [1]	43
m 1 1 nie KIT tak nie ...	integer [1]	91	z 1 1 nie KIT_IIT tak ta...	integer [1]	20
m 1 1 nie KIT tak nie ...	integer [1]	72	z 1 1 nie KIT_IIT tak ta...	integer [1]	70
m 1 1 nie KIT tak nie ...	integer [1]	97	z 1 1 tak KIT nie nie 0 ...	integer [1]	68
m 1 1 nie KIT tak nie ...	integer [1]	105	z 1 1 tak KIT tak nie 2 ...	integer [1]	89
m 1 1 nie KIT tak tak ...	integer [1]	79	z 1 1 tak KIT tak nie 2 ...	integer [1]	36
m 1 1 nie KIT tak tak ...	integer [1]	32	z 1 1 tak KIT_IIT tak ta...	integer [1]	28
m 1 1 nie KIT_IIT nie n...	integer [1]	9	z 1 2 nie KIT tak nie 2 ...	integer [1]	88
m 1 1 nie KIT_IIT tak t...	integer [1]	85	z 1 2 tak KIT_IIT tak ni...	integer [1]	11
m 1 1 tak KIT tak nie ...	integer [1]	52	z 2 0 nie KIT tak nie 0 ...	integer [1]	100
m 1 1 tak KIT_IIT nie n...	integer [1]	67	z 2 0 nie KIT tak nie 1 ...	integer [1]	23
m 1 2 nie KIT tak nie ...	integer [1]	41	z 2 0 nie KIT tak tak 0 ...	integer [1]	26
m 2 0 nie KIT tak nie ...	integer [1]	56	z 2 0 nie KIT_IIT nie ni...	integer [1]	101
m 2 0 nie KIT tak tak ...	integer [1]	8	z 2 0 nie KIT_IIT tak ni...	integer [1]	38
m 2 0 nie KIT_IIT nie n...	integer [1]	29	z 2 0 nie KIT_IIT tak ni...	integer [1]	4
m 2 0 nie KIT_IIT tak n...	integer [1]	44	z 2 0 tak KIT_IIT tak ni...	integer [1]	55
m 2 0 nie KIT_IIT tak n...	integer [1]	61	z 2 1 nie KIT tak nie 0 ...	integer [1]	92
m 2 0 nie KIT_IIT tak n...	integer [1]	6	z 2 1 nie KIT tak nie 1 ...	integer [1]	87
m 2 0 tak KIT tak nie ...	integer [2]	3 78	z 2 1 nie KIT_IIT tak ni...	integer [1]	7
m 2 1 nie KIT nie nie 1...	integer [1]	84	z 2 1 tak KIT tak nie 0 ...	integer [1]	99
m 2 1 nie KIT tak tak ...	integer [1]	35	z 3 0 nie KIT nie tak 0 ...	integer [1]	40
m 2 1 nie KIT_IIT tak n...	integer [1]	31	z 3 0 nie KIT_IIT nie ni...	integer [1]	39
m 2 1 nie KIT_IIT tak t...	integer [1]	104	z 3 0 tak KIT_IIT nie ni...	integer [1]	14
m 2 1 tak KIT nie nie ...	integer [1]	66	type,relation	character [1]	'equivalence'
m 3 0 nie KIT tak nie ...	integer [1]	10	type,model	character [1]	'RST'
m 3 0 nie KIT_IIT nie n...	integer [1]	27			
m 3 1 nie KIT tak nie ...	integer [1]	96			
z 0 0 nie KIT nie nie 1 ...	integer [1]	42			
z 0 0 nie KIT tak nie 0 ...	integer [1]	45			
z 0 0 nie KIT tak nie 0 ...	integer [1]	95			
z 0 0 nie KIT tak nie 0 ...	integer [1]	50			
z 0 0 nie KIT tak nie 1 ...	integer [1]	107			
z 0 1 nie KIT nie nie 0 ...	integer [1]	71			

Analiza

Zbiór elementarny zawiera elementy nierozróżnialne ze względu na przyjęty zbiór atrybutów. Na podstawie powyżej tabeli można dostrzec, że około 4% klas nieodróżnialności składa się z więcej niż jednego obiektu.

```
> View(ind)
> ind[["IND.relation"]][["m 1 0 nie KIT tak nie 0 nie 3-97 nie nie"]]
```

Np. 18-sta klasa nieodróżnialności od góry składa się z obiektów o numerach: 15, 65. Oznacza to, że obiekty mają te same wartości wszystkich atrybutów: płeć, wiek, czas, występ, infekcje, remisja i NT, itd.

Aproksymacje

Dolna aproksymacja dla wartości atrybutów decyzyjnych "tak":

lower approximation	list [2]	List of length 2	z 1 1 tak KIT_JIT tak ...	integer [1]	28
tak	integer [56]	63 13 22 33 75 86 ...	z 1 2 tak KIT_JIT tak ...	integer [1]	11
m 0 0 nie KIT_JIT ta...	integer [1]	63	z 2 0 nie KIT tak nie...	integer [1]	100
m 1 0 nie KIT nie ni...	integer [1]	13	z 2 0 nie KIT tak nie...	integer [1]	23
m 1 0 nie KIT tak ni...	integer [1]	22	z 2 0 nie KIT_JIT nie ...	integer [1]	101
m 1 0 nie KIT_JIT ta...	integer [1]	33	z 2 0 nie KIT_JIT tak ...	integer [1]	38
m 1 0 nie KIT_JIT ta...	integer [1]	75	z 2 0 nie KIT_JIT tak ...	integer [1]	4
m 1 0 nie KIT_JIT ta...	integer [1]	86	z 2 0 tak KIT_JIT tak ...	integer [1]	55
m 1 1 nie KIT nie ni...	integer [1]	83	z 2 1 nie KIT tak nie...	integer [1]	92
m 1 1 nie KIT nie ni...	integer [1]	16	z 2 1 nie KIT_JIT tak ...	integer [1]	7
m 1 1 nie KIT tak ni...	integer [1]	91	z 2 1 tak KIT tak nie...	integer [1]	99
m 1 1 nie KIT tak ni...	integer [1]	97	z 3 0 nie KIT_JIT nie ...	integer [1]	39
m 1 1 nie KIT tak ni...	integer [1]	105			
m 1 1 nie KIT tak ta...	integer [1]	32			
m 1 1 nie KIT_JIT ni...	integer [1]	9			
m 1 1 nie KIT_JIT ta...	integer [1]	85			
m 1 1 tak KIT tak ni...	integer [1]	52			
m 1 2 nie KIT tak ni...	integer [1]	41			
m 2 0 nie KIT tak ni...	integer [1]	56			
m 2 0 nie KIT tak ta...	integer [1]	8			
m 2 0 nie KIT_JIT ni...	integer [1]	29			
m 2 0 nie KIT_JIT ta...	integer [1]	61			
m 2 0 nie KIT_JIT ta...	integer [1]	6			
m 2 1 tak KIT nie ni...	integer [1]	66			
m 3 0 nie KIT tak ni...	integer [1]	10			
z 0 0 nie KIT tak nie...	integer [1]	45			
z 0 0 nie KIT tak nie...	integer [1]	95			
z 0 0 nie KIT tak nie...	integer [1]	50			
z 0 2 nie KIT_JIT nie ...	integer [1]	51			
z 1 0 nie KIT nie nie...	integer [1]	19			
z 1 0 nie KIT_JIT tak ...	integer [1]	1			
z 1 0 nie KIT_JIT tak ...	integer [1]	12			
z 1 0 nie KIT_JIT tak ...	integer [1]	76			
z 1 0 tak KIT tak nie...	integer [1]	5			
z 1 1 nie KIT nie nie...	integer [1]	53			
z 1 1 nie KIT tak nie...	integer [1]	98			
z 1 1 nie KIT tak nie...	integer [1]	17			
z 1 1 nie KIT tak nie...	integer [1]	21			
z 1 1 nie KIT_JIT nie ...	integer [1]	24			
z 1 1 nie KIT_JIT tak ...	integer [1]	74			
z 1 1 nie KIT_JIT tak ...	integer [1]	30			
z 1 1 nie KIT_JIT tak ...	integer [1]	43			
z 1 1 nie KIT_JIT tak ...	integer [1]	20			
z 1 1 nie KIT_JIT tak ...	integer [1]	70			
z 1 1 tak KIT tak nie...	integer [1]	89			
z 1 1 tak KIT tak nie...	integer [1]	36			

Dolna aproksymacja dla wartości atrybutów decyzyjnych "nie":

lower approximation	list [2]	List of length 2	z 0 0 nie KIT nie nie...	integer [1]	42
tak	integer [56]	63 13 22 33 75 86 ...	z 0 0 nie KIT tak nie...	integer [1]	107
nie	integer [51]	59 106 94 2 80 49 ...	z 0 1 nie KIT nie nie...	integer [1]	71
m 0 0 nie KIT nie ni...	integer [1]	59	z 0 1 nie KIT tak nie...	integer [1]	60
m 0 0 nie KIT tak ni...	integer [1]	106	z 0 1 nie KIT tak nie...	integer [1]	82
m 0 0 nie KIT tak ni...	integer [1]	94	z 0 1 nie KIT tak nie...	integer [1]	81
m 0 0 nie KIT tak ni...	integer [1]	2	z 1 0 nie KIT tak nie...	integer [1]	25
m 0 0 tak KIT tak ni...	integer [1]	80	z 1 0 nie KIT_JIT nie ...	integer [1]	62
m 0 0 tak KIT_JIT ta...	integer [1]	49	z 1 0 tak KIT_JIT tak ...	integer [1]	58
m 0 1 tak KIT nie ni...	integer [1]	77	z 1 1 nie KIT tak nie...	integer [1]	93
m 0 1 tak KIT tak ni...	integer [1]	90	z 1 1 tak KIT nie nie...	integer [1]	68
m 0 1 tak KIT tak ni...	integer [1]	64	z 1 2 nie KIT tak nie...	integer [1]	88
m 0 1 tak KIT tak ni...	integer [1]	18	z 2 0 nie KIT tak tak...	integer [1]	26
m 0 1 tak KIT tak ni...	integer [1]	103	z 2 1 nie KIT tak nie...	integer [1]	87
m 0 1 tak KIT_JIT ni...	integer [1]	102	z 3 0 nie KIT tak nie...	integer [1]	40
m 0 2 nie KIT_JIT ni...	integer [1]	47	z 3 0 tak KIT_JIT nie ...	integer [1]	14
m 1 0 nie KIT nie ni...	integer [1]	54			
m 1 0 nie KIT nie ni...	integer [1]	34			
m 1 0 nie KIT tak ni...	integer [1]	15			
m 1 0 nie KIT tak ni...	integer [1]	65			
m 1 0 nie KIT tak ta...	integer [1]	57			
m 1 0 nie KIT_JIT ta...	integer [1]	69			
m 1 0 tak KIT_JIT ni...	integer [1]	48			
m 1 0 tak KIT_JIT ni...	integer [1]	46			
m 1 0 tak KIT_JIT ta...	integer [1]	37			
m 1 0 tak KIT_JIT ta...	integer [1]	73			
m 1 1 nie KIT tak ni...	integer [1]	72			
m 1 1 nie KIT tak ta...	integer [1]	79			
m 1 1 tak KIT_JIT ni...	integer [1]	67			
m 2 0 nie KIT_JIT ta...	integer [1]	44			
m 2 0 tak KIT tak ni...	integer [1]	3			
m 2 0 tak KIT tak ni...	integer [1]	78			
m 2 1 nie KIT nie ni...	integer [1]	84			
m 2 1 nie KIT tak ta...	integer [1]	35			
m 2 1 nie KIT_JIT ta...	integer [1]	31			
m 2 1 nie KIT_JIT ta...	integer [1]	104			
m 3 0 nie KIT_JIT ni...	integer [1]	27			
m 3 1 nie KIT tak ni...	integer [1]	96			

Analiza

Dolna aproksymacja to zbiór będący sumą zbiorów elementarnych, które bez wątpienia są objęte danym nieostrym pojęciem, dlatego pewne jest, że program zdecyduje prawidłowo.

Analizując poszczególne obiekty z tabeli należące do dolnego przybliżenia nie mamy więc wątpliwości, że spełniają kryteria przynależności do danej klasy.

W dolnej aproksymacji zawierają się obiekty z jednoelementowych zbiorów elementarnych lub obiekty z wieloelementowych zbiorów elementarnych, w których wartość atrybutu decyzyjnego dla wszystkich obiektów jest taka sama.

Górna aproksymacja dla wartości atrybutów decyzyjnych "tak":

upper approximation	list [2]	List of length 2		
tak	integer [56]	63 13 22 33 75 86 ...	z 1 0 tak KJT tak nie...	integer [1] 5
m 0 0 nie KJT_JIT ta...	integer [1]	63	z 1 1 nie KJT nie nie...	integer [1] 53
m 1 0 nie KJT nie ni...	integer [1]	13	z 1 1 nie KJT tak nie...	integer [1] 98
m 1 0 nie KJT tak ni...	integer [1]	22	z 1 1 nie KJT tak nie...	integer [1] 17
m 1 0 nie KJT_JIT ta...	integer [1]	33	z 1 1 nie KJT tak nie...	integer [1] 21
m 1 0 nie KJT_JIT ta...	integer [1]	75	z 1 1 nie KJT_JIT nie ...	integer [1] 24
m 1 0 nie KJT_JIT ta...	integer [1]	86	z 1 1 nie KJT_JIT tak ...	integer [1] 74
m 1 1 nie KJT nie ni...	integer [1]	83	z 1 1 nie KJT_JIT tak ...	integer [1] 30
m 1 1 nie KJT nie ni...	integer [1]	16	z 1 1 nie KJT_JIT tak ...	integer [1] 43
m 1 1 nie KJT tak ni...	integer [1]	91	z 1 1 nie KJT_JIT tak ...	integer [1] 20
m 1 1 nie KJT tak ni...	integer [1]	97	z 1 1 nie KJT_JIT tak ...	integer [1] 70
m 1 1 nie KJT tak ni...	integer [1]	105	z 1 1 tak KJT tak nie...	integer [1] 89
m 1 1 nie KJT tak ta...	integer [1]	32	z 1 1 tak KJT tak nie...	integer [1] 36
m 1 1 nie KJT_JIT ni...	integer [1]	9	z 1 1 tak KJT_JIT tak ...	integer [1] 28
m 1 1 nie KJT_JIT ta...	integer [1]	85	z 1 2 tak KJT_JIT tak ...	integer [1] 11
m 1 1 tak KJT tak ni...	integer [1]	52	z 2 0 nie KJT tak nie...	integer [1] 100
m 1 2 nie KJT tak ni...	integer [1]	41	z 2 0 nie KJT tak nie...	integer [1] 23
m 2 0 nie KJT tak ni...	integer [1]	56	z 2 0 nie KJT_JIT nie ...	integer [1] 101
m 2 0 nie KJT tak ta...	integer [1]	8	z 2 0 nie KJT_JIT tak ...	integer [1] 38
m 2 0 nie KJT_JIT ni...	integer [1]	29	z 2 0 nie KJT_JIT tak ...	integer [1] 4
m 2 0 nie KJT_JIT ta...	integer [1]	61	z 2 0 tak KJT_JIT tak ...	integer [1] 55
m 2 0 nie KJT_JIT ta...	integer [1]	6	z 2 1 nie KJT tak nie...	integer [1] 92
m 2 1 tak KJT nie ni...	integer [1]	66	z 2 1 nie KJT_JIT tak ...	integer [1] 7
m 3 0 nie KJT tak ni...	integer [1]	10	z 2 1 tak KJT tak nie...	integer [1] 99
z 0 0 nie KJT tak nie...	integer [1]	45	z 3 0 nie KJT_JIT nie ...	integer [1] 39
z 0 0 nie KJT tak nie...	integer [1]	95		
z 0 0 nie KJT tak nie...	integer [1]	50		
z 0 2 nie KJT_JIT nie ...	integer [1]	51		
z 1 0 nie KJT nie nie...	integer [1]	19		
z 1 0 nie KJT_JIT tak ...	integer [1]	1		
z 1 0 nie KJT_JIT tak ...	integer [1]	12		
z 1 0 nie KJT_JIT tak ...	integer [1]	76		

Górna aproksymacja dla wartości atrybutów decyzyjnych "nie":

upper approximation	list [2]	List of length 2	m 3 1 nie KJT tak ni...	integer [1] 96
tak	integer [56]	63 13 22 33 75 86 ...	z 0 0 nie KJT nie nie...	integer [1] 42
nie	integer [51]	59 106 94 2 80 49 ...	z 0 0 nie KJT tak nie...	integer [1] 107
m 0 0 nie KJT nie ni...	integer [1]	59	z 0 1 nie KJT nie nie...	integer [1] 71
m 0 0 nie KJT tak ni...	integer [1]	106	z 0 1 nie KJT tak nie...	integer [1] 60
m 0 0 nie KJT tak ni...	integer [1]	94	z 0 1 nie KJT tak nie...	integer [1] 82
m 0 0 nie KJT tak ni...	integer [1]	2	z 0 1 tak KJT tak nie...	integer [1] 81
m 0 0 tak KJT tak ni...	integer [1]	80	z 1 0 nie KJT tak nie...	integer [1] 25
m 0 0 tak KJT_JIT ta...	integer [1]	49	z 1 0 nie KJT_JIT nie ...	integer [1] 62
m 0 1 tak KJT nie ni...	integer [1]	77	z 1 0 tak KJT_JIT tak ...	integer [1] 58
m 0 1 tak KJT tak ni...	integer [1]	90	z 1 1 nie KJT tak nie...	integer [1] 93
m 0 1 tak KJT tak ni...	integer [1]	64	z 1 1 tak KJT nie nie...	integer [1] 68
m 0 1 tak KJT tak ni...	integer [1]	18	z 1 2 nie KJT tak nie...	integer [1] 88
m 0 1 tak KJT tak ni...	integer [1]	103	z 2 0 nie KJT tak tak...	integer [1] 26
m 0 2 nie KJT_JIT ni...	integer [1]	102	z 2 1 nie KJT tak nie...	integer [1] 87
m 0 2 nie KJT_JIT ni...	integer [1]	47	z 3 0 nie KJT nie tak...	integer [1] 40
m 1 0 nie KJT nie ni...	integer [1]	54	z 3 0 tak KJT_JIT nie ...	integer [1] 14
m 1 0 nie KJT nie ni...	integer [1]	34		
m 1 0 nie KJT tak ni...	integer [1]	15		
m 1 0 nie KJT tak ni...	integer [1]	65		
m 1 0 nie KJT tak ta...	integer [1]	57		
m 1 0 nie KJT_JIT ta...	integer [1]	69		
m 1 0 tak KJT_JIT ni...	integer [1]	48		
m 1 0 tak KJT_JIT ni...	integer [1]	46		
m 1 0 tak KJT_JIT ta...	integer [1]	37		
m 1 0 tak KJT_JIT ta...	integer [1]	73		
m 1 1 nie KJT tak ni...	integer [1]	72	twoe.model	character [1] 'RST'
m 1 1 nie KJT tak ta...	integer [1]	79		
m 1 1 tak KJT_JIT ni...	integer [1]	67		
m 2 0 nie KJT_JIT ta...	integer [1]	44		
m 2 0 tak KJT tak ni...	integer [1]	3		
m 2 0 tak KJT tak ni...	integer [1]	78		
m 2 1 nie KJT nie ni...	integer [1]	84		
m 2 1 nie KJT tak ta...	integer [1]	35		
m 2 1 nie KJT_JIT ta...	integer [1]	31		
m 2 1 nie KJT_JIT ta...	integer [1]	104		
m 3 0 nie KJT_JIT ni...	integer [1]	27		

Analiza

Do górnego przybliżenia należą obiekty, których nie można wykluczyć, że są objęte danym nieostrym pojęciem, dlatego nie jest pewne, że program zdecyduje prawidłowo. Jest to spowodowane tym, że

należą do klas abstrakcji mających niepuste przecięcie z danym nieostrym pojęciem. Są zatem nierozróżnialne z pewnymi obiektami należącymi do tego pojęcia.

Warto zwrócić uwagę na fakt, że dolna aproksymacja zawiera się w górnej. Na tej podstawie można stwierdzić, że górne przybliżenie jest zbiorem większym i patrzącym na obiekty w "szerszej perspektywie".

Obszar brzegowy

Name	Type	Value
boundary	list [3] (S3: BoundaryRegion, list)	List of length 3
boundary.reg	integer [0]	
degree.dependency	double [1]	0
type.model	character [1]	'RST'

Analiza

Obszarem brzegowym nazywamy te obiekty, co do których nie mamy pewności, czy reprezentują dane pojęcie. W danym przypadku różnica między górnym a dolnym przybliżeniem wyszła równa 0, oznacza to, że w danym zbiorze nie występuje obszar niespójności.

Obszar pozytywny

Name	Type	Value
positive	list [3] (S3: PositiveRegion, list)	List of length 3
positive.reg	integer [107]	1 2 3 4 5 6 ...
degree.dependency	double [1]	1
type.model	character [1]	'RST'

```
> View(positive)
> positive[["positive.reg"]]
[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39
[40] 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78
[79] 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107
> positive[["degree.dependency"]]
[1] 1
> positive[["type.model"]]
[1] "RST"
```

Analiza

Obszar pozytywny jest sumą dolnych przybliżeń przez co do tej grupy należą wszystkie obiekty dla których mamy pewność, że należą do pewnej wartości klasy decyzyjnej.

Dla powyższych obiektów obszar pozytywny dał nam wynik pokrywający się z całością zbioru startowego, co oznacza niesprzeczność i ostrość zbioru.

Wygenerowana macierz

```
> dm[["disc.mat"]]
      V1      V2      V3 V4 V5 V6 V7 V8 V9 V10 V11 V12 V13 V14 V15 V16 V17
1      NA      NA      NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
2 Plec, Wiek, Rodzaj, HbA1c, NT      NA      NA      NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
3 Plec, Wiek, Wystep, Rodzaj, NT      NA      NA      NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
4      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
5      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
6      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
7      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
8      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
9      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
V18 V19 V20 V21 V22 V23 V24 V25 V26 V27 V28 V29 V30 V31 V32 V33 V34 V35 V36 V37 V38 V39 V40 V41 V42 V43 V44 V45 V46 V47 V48 V49 V50 V51 V52 V53 V54 V55 V56 V57
1 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
2 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
3 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
4 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
5 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
6 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
7 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
8 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
9 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
V58 V59 V60 V61 V62 V63 V64 V65 V66 V67 V68 V69 V70 V71 V72 V73 V74 V75 V76 V77 V78 V79 V80 V81 V82 V83 V84 V85 V86 V87 V88 V89 V90 V91 V92 V93 V94 V95 V96 V97
1 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
2 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
3 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
4 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
5 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
6 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
7 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
8 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
9 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
V98 V99 V100 V101 V102 V103 V104 V105 V106 V107
1 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
2 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
3 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
4 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
5 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
6 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
7 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
8 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
9 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
[reached "max" / getOption("max.print") -- omitted 98 rows ]
```

Przy pomocy wygenerowanej macierzy jesteśmy w stanie wyznaczyć rdzeń i redukty.

Rdzeń i redukty

```
> reducts[["decision.reduct"]]
$reduct1
A feature subset consisting of 9 attributes:
Plec, Wiek, Czas, Wystep, Rodzaj, Remisja, HbA1c, NT, TG

$reduct2
A feature subset consisting of 9 attributes:
Plec, Wiek, Czas, Wystep, Rodzaj, Remisja, HbA1c, Mc, TG

$reduct3
A feature subset consisting of 9 attributes:
Plec, Wiek, Czas, Wystep, Rodzaj, Remisja, HbA1c, Ch, TG

$reduct4
A feature subset consisting of 10 attributes:
Plec, Wiek, Czas, Wystep, Rodzaj, Infekcje, HbA1c, NT, Mc, TG

$reduct5
A feature subset consisting of 10 attributes:
Plec, Wiek, Czas, Rodzaj, Infekcje, HbA1c, NT, Mc, Ch, TG

$reduct6
A feature subset consisting of 10 attributes:
Plec, Wiek, Czas, Wystep, Rodzaj, Infekcje, HbA1c, Mc, Ch, TG
```

Analiza

W powyższym przypadku otrzymano sześć reduktów. Analizując otrzymane wyniki okazuje się, że nie wszystkie atrybuty są potrzebne w procesie podejmowania decyzji. W trzech pierwszych przypadkach potrzeba tylko dziewięć atrybutów do poprawnego przydzielania obiektów do poszczególnych klas, zaś w pozostałych przypadkach potrzeba ich aż 10.

Podsumowując, można więc powiedzieć, że redukty ograniczają liczbę atrybutów do minimum, niezbędnego do poprawnego przydzielania obiektów do poszczególnych klas.

core	character [6]	'Czas' 'HbA1c' 'Plec' 'Rodzaj' 'TG' 'Wiek'
discernibility.type	character [1]	'RST'
type.task	character [1]	'computation of all reducts'
type.model	character [1]	'RST'

Analiza

Rdzeń jest częścią wspólną wszystkich reduktów systemu informacyjnego. W danym przypadku zbiorem atrybutów niezbędnych dla zachowania rozróżnialności obiektów w systemie są: Czas, HbA1c, Plec, Rodzaj, TG, Wiek.

Algorytm LEM

```
> rss
A set consisting of 35 rules:
1. IF Czas is 0 and Plec is m and NT is tak THEN Mikroalbuminuria is tak;
   (supportSize=5; laplace=0.8571)
2. IF NT is nie and Infekcje is tak and Plec is z and Wiek is 1 and Czas is 1 THEN Mikroalbuminuria is tak;
   (supportSize=9; laplace=0.9091)
3. IF Czas is 0 and HbA1c is 1 and Wiek is 2 THEN Mikroalbuminuria is tak;
   (supportSize=7; laplace=0.8889)
4. IF Wystep is nie and NT is nie and Rodzaj is KIT_IIT and Plec is z THEN Mikroalbuminuria is tak;
   (supportSize=12; laplace=0.9286)
5. IF Remisja is nie and TG is nie and Infekcje is tak and HbA1c is 0 and Plec is z and Rodzaj is KIT and NT is nie THEN Mikroalbuminuria is tak;
   (supportSize=6; laplace=0.875)
6. IF Wiek is 1 and Plec is m and Rodzaj is KIT and Infekcje is nie and Czas is 1 THEN Mikroalbuminuria is tak;
   (supportSize=2; laplace=0.75)
7. IF Wiek is 1 and Plec is m and NT is tak THEN Mikroalbuminuria is tak;
   (supportSize=5; laplace=0.8571)
8. IF Wystep is nie and Plec is m and HbA1c is 1 and Mc is 3-97 and Czas is 0 THEN Mikroalbuminuria is tak;
   (supportSize=8; laplace=0.9)
9. IF HbA1c is 0 and NT is tak and Plec is z and Rodzaj is KIT_IIT THEN Mikroalbuminuria is tak;
   (supportSize=2; laplace=0.75)
10. IF Ch is tak and Wystep is nie and Rodzaj is KIT_IIT and TG is nie THEN Mikroalbuminuria is tak;
    (supportSize=8; laplace=0.9)
... and 25 other rules.
```

Algorytm LEM2 generuje reguły decyzyjne, które pozwalają podejmować decyzje klasyfikacji dla całkiem nowych obiektów spoza tych w tabeli decyzyjnej.

Ogólna analiza

- Dane do zadania składały się z 12 atrybutów warunkowych i jednego decyzyjnego.
- Przybliżenia dolne zawierały obiekty, w których jesteśmy pewni decyzji.
- Do górnej aproksymacji trafiły obiekty, co do których nie było pewności.
- Obie aproksymacje posiadały takie same obiekty, dlatego różnica tych zbiorów tj. obszar brzegowy wynosi 0.
- Obszar pozytywny dał wynik pokrywający się z całością zbioru startowego co oznacza niesprzeczność i ostrość zbioru.
- Wyznaczono 6 reduktów i 6-cio elementowy rdzeń, który jest iloczynem reduktów.
- Rdzeń jest niezbędny do prawidłowego sklasyfikowania każdego obiektu z danych.
- Algorytm Lem2 wygenerował 35 reguł. Można zauważyć że po redukcji zbioru atrybutów warunkowych liczba reguł nie zmieniła się, gdyż do klasyfikacji obiektów używa się tylko atrybutów zawartych w reduktach.

Wnioski

- Teoria zbiorów przybliżonych pozwala na identyfikowanie redundancji danych w tablicach informacyjnych oraz jej eliminację – tworzenie reduktów.
- Klasą decyzyjną nazywamy zbiór obiektów o zbliżonej złożoności obliczeniowej potrzebnej do podjęcia decyzji.
- Aproksymacja zbioru określa nam dokładny sposób klasyfikowania obiektów naszego zbioru danych. Dolne przybliżenie zawiera obiekty pewne co do swojej decyzji, zaś górne przybliżenie niepewne – może, ale nie musi mieć zawsze jednoznacznej decyzji.
- W celu precyzyjnego i konkretnego opisanie relacji pomiędzy obiektami występującymi w bazie wiedzy, stosuje się redukcję liczby atrybutów opisujących owe relacje.
- Algorytm Lem2 pozwala nam na wygenerowanie reguł decyzyjnych z aproksymacji tablicy decyzyjnej. Jeżeli są to dolne aproksymacje generuje on reguły pewne, jeśli górne - reguły niepewne (Zamiast aproksymacji można podać zbiór brzegowy, wtedy zostaną wygenerowane reguły przybliżone).
- Teoria zbiorów przybliżonych świetnie nadaje się do reprezentacji pojęć niespójnych.
- Pozwala redukować wiedzę do niezbędnego minimum.
- Pozwala generować reguły minimalne – które skracają drogę wnioskowania.
- Dziedziny, w których teoria ta została zastosowana to nie tylko medycyna czy biznes (bankowość, badania rynku) ale również rozpoznawanie mowy, sieci neuronowe czy ogólnie mówiąc sztuczna inteligencja

Część 2

Eksploracja danych (data mining) to jeden z etapów procesu uzyskiwania wiedzy z baz danych, przeznaczony do badania dużych zasobów danych w poszukiwaniu regularnych wzorców oraz systematycznych współzależności pomiędzy zmiennymi.

	temperatura	ból głowy	osłabienie	nudności	kaszel	czy covid
1	1	0	1	0	1	1
2	0	0	0	1	2	0
3	1	1	0	0	-1	0
4	0	1	1	0	2	1
5	0	1	0	0	-2	0
6	1	1	1	0	1	1
7	1	1	0	1	2	1
8	1	1	0	1	2	0
9	0	0	1	0	0	0
10	0	0	1	0	0	1
11	1	1	0	0	-2	1
12	1	1	0	0	-2	1
13	0	0	1	1	2	0
14	1	1	0	0	1	1

Rys. Tablica decyzyjna.

Obszar pozytywny danych:

> View(positive)

```
> positive[["positive.reg"]]
```

```
[1] 1 2 3 4 5 6 13 14
```

```
> positive[["degree.dependency"]]
```

```
[1] 0.5714286
```

```
> positive[["type.model"]]
```

```
[1] "RST"
```

Obliczanie wpływu atrybutów warunkowych na wyniki

```
$decision.reduct
$decision.reduct$reduct1
A feature subset consisting of 3 attributes:
temperatura, ból_głowy, kaszel

$decision.reduct$reduct2
A feature subset consisting of 3 attributes:
temperatura, nudności, kaszel

Score
[1] "kaszel"      "temperatura"

$discernibility.type
[1] "RST"

$type.task
[1] "computation of all reducts"

$type.model
[1] "RST"

attr(,"class")
[1] "ReductSet" "list"
```

Na podstawie powyższych danych można zauważyć, że rdzeń zawiera takie atrybuty jak „kaszel” oraz „temperatura”. Dane atrybuty mają więc znaczny wpływ na klasyfikację obiektów.

Propozycja wyliczania współczynnika istotności dla atrybutów warunkowych:

- Jeżeli zmiana wartości atrybutu na przeciwny, zmieni wartość atrybutu decyzyjnego , dodawane są do atrybutu „punkty istotności”.
- Następnie po wykonaniu owej operacji na wszystkich obiektach, dzielę punkty istotności przez sumę obiektów.

Po wykonaniu tych operacji otrzymano:

- Współczynnik istotności dla atrybutu „temperatura”: 0.43
- Współczynnik istotności dla atrybutu „ból_głowy” : 0.21
- Współczynnik istotności dla atrybutu „osłabienie”: 0
- Współczynnik istotności dla atrybutu „nudności”: 0.64
- Współczynnik istotności dla atrybutu „kaszel”: 0.71





Okazuje się, że kaszel jest najważniejszym atrybutem warunkowym. Zmiana wartości na przeciwną przyczyniłaby się do zmiany wartości atrybutu decyzyjnego. Wartość osłabienie jest atrybutem neutralnym. Nie wpływa ono na wykrycie Covid’a. Jego usunięcie nie

zmieniłoby wyniku w klasyfikacji. Można więc powiedzieć, że nie występuje też ono w żadnym redukcje.

Wnioski

Dzięki data eksploracji danych możliwe jest przewidywanie i klasyfikacja obiektów w przyszłości, dlatego też predykcyjny data mining jest bardzo popularny, natomiast aby go wykonywać prawidłowo potrzebne jest określenie istotności poszczególnych atrybutów warunkowych.

Pliki:

-  grypa.csv- zawiera dane do zadania 2 czy_covid
-  diabetes.csv
-  Zad2_SI.R
-  SI2B.R