

SPRAWOZDANIE Z LABORATORIUM
ANALIZA DANYCH I INTELIGENCJA OBLICZENIOWA

Asocjacje

Bartosz Borys
Grupa 5

WEATHER.NOMINAL

1. Zbiór ten obrazuje problem koligacji między pogodą, a możliwością gry na dworze.
Gdy warunki na dworze są dobre, istnieje większa szansa na to, że zabawa na dworze będzie również dobra.

2. Wyniki

outlook=rainy play=no 2 ==> windy=TRUE	2/13	1
outlook=rainy windy=TRUE 2 ==> play=no	2/13	1
temperature=hot play=no 2 ==> humidity=high	2/13	1
outlook=overcast temperature=cool 1 ==> windy=TRUE	1/13	1
temperature=cool play=no 1 ==> outlook=rainy	1/13	1
temperature=hot windy=TRUE 1 ==> humidity=high	1/13	1
temperature=hot windy=TRUE 1 ==> play=no	1/13	1
temperature=cool play=no 1 ==> windy=TRUE	1/13	1
temperature=cool windy=TRUE play=no 1 ==> outlook=rainy	1/13	1
1outlook=rainy temperature=cool play=no 1 ==> windy=TRUE	1/13	1

3. W przypadku, gdy pogoda jest deszczowa, a jednocześnie gra na dworze nie jest zalecana to pierwsza reguła daje nam informacje, że istnieje dosyć duża szansa na to, że jest wietrznie.

4. Temperature=hot humidity=high ==> windy=TRUE

Wsparcie = 1/14

Ufność = 1/3

5. Wymienić 3 różnice pomiędzy regułami asocjacyjnymi a regułami decyzyjnymi.

W regułach asocjacyjnych:

- Brak atrybutu decyzyjnego który jest wyróżniony.
- Decyzja nie jest od razu określona.
- Poprzez grupy atrybutów możliwe jest potwierdzenie występowania innych określonych atrybutów. Natomiast w decyzyjnych na podstawie wejścia wysnuwamy pewnie wnioski.

6. Wyniki

Minimalna ufność (C)	Minimalne wsparcie (M)	Liczba wszystkich możliwych reguł
0.7	0.1	19
0.7	0.2	2
0.7	0.3	1
0.8	0.1	18
0.8	0.2	1
0.8	0.3	1
0.9	0.1	17
0.9	0.2	0
0.9	0.3	0

7. Objaśnienie N,C,M

- N - Jest to ilość maksymalnych wyników. W przypadku gdzie wyników jest więcej niż N może wystąpić problem, gdyż nie zobaczymy wszystkich wyników.
- C - Określa minimalną ufność. Im mniejsza wartość tym szansa na poprawność wniosku maleje.
- M - Określa minimalne wsparcie wygenerowanej reguły.

VOTE

1. Zbiór ten stara opisać czy przynależność do systemu rządów ma wpływ pogląd polityczny i światowy jednostki.

2. Wyniki

Miara	Najlepsza reguła	Wartość miary
0	mx-missile=n Class=republican 146 ==>physician-fee-freeze=y 145	conf:(0.99)
1	physician-fee-freeze=y 177 ==>Class=republican 163	lift:(2.38)
2	physician-fee-freeze=y 177 ==>Class=republican 163	lev:(0.22)
3	Class=republican 168 ==> physician-fee-freeze=y 163	conv:(16.61)

3. Ponieważ definicja miar jest różna. Confidence sprawdza częstość występowania prawdy dla danej reguły. Natomiast Lift odnosi się do niezależności reguł, jeżeli wynik jest większy od 1 wtedy taka reguła może być dla nas atrakcyjna, ponieważ występuje wtedy zależność. Leverage sprawdza tę samą zależność skupiając się na różnicach między współczynnikami, natomiast ostatnia Conviction sprawdza czy reguły nie zostały sparowane losowo (przypadkiem). Im wyższa wartość tym mniejsza szansa, że tak było.

SUPERMARKET

1. Opisuje problem rozmieszczenia produktów i ich zawartości w różnych oddziałach sieci supermarketów.

2. Na schemacie

3. Na schemacie

4. Na schemacie

5. Najwyższa ufność: 0.921

vegetables, milk cream, frozen foods, biscuits, pet foods ==> bread and cake

6. Z wyniku można wywnioskować, że osoby kupujące produkty z lewej strony strzałki chętnie kupują produkty po prawej tak więc ustawienie pieczywa oraz ciastek blisko warzyw, lodów, mrożonek, słodkości czy karmy dla zwierząt może zwiększyć ilość sprzedanego towaru.