Małe info:

-tekst podświetlony na szaro to zapis wykonania algorytmu

-wszystko po pauzie to tylko komentarze

DNF:

Dane:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | f1 | f2 | f3 | f4 | y |
| x1 | 0 | 1 | 0 | 0 | 0 |
| x2 | 1 | 1 | 0 | 0 | 1 |
| x3 | 1 | 0 | 1 | 0 | 1 |
| x4 | 0 | 1 | 0 | 1 | 0 |
| x5 | 1 | 0 | 0 | 1 | 0 |
| x6 | 0 | 1 | 1 | 0 | 1 |

**Początek algorytmu:**

h=false - ustawienie hipotezy na false

P = {x2, x3, x6} – Pozytywne przykłady (czyli tam gdzie jest 1 w kolumnie y)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | f1 | f2 | f3 | f4 | y |
| x1 | 0 | 1 | 0 | 0 | 0 |
| x2 | 1 | 1 | 0 | 0 | 1 |
| x3 | 1 | 0 | 1 | 0 | 1 |
| x4 | 0 | 1 | 0 | 1 | 0 |
| x5 | 1 | 0 | 0 | 1 | 0 |
| x6 | 0 | 1 | 1 | 0 | 1 |

**Pierwsza iteracja zewnetrznej pętli:**

r=true – zresetowanie zmiennej przchowującej tymczasowo wybrane cechy

N={x1,x4, x5} – Ustawienie przykładów negatywnych ( czyli tam gdzie 0 w kolumnie y)

**Pierwsza iteracja wewnetrznej pętli**

v1=2/1, v2=2/1, v3=2/0, v4=0/2 - Wybranie cechy (czyli dla każdego vi gdzie i to numer kolumny jeszcze nie wybranej, obliczamy ilość przykładów pozytywnych w P w której cecha ta jest równa 1. Dla przykładu v1:

|  |  |
| --- | --- |
| f1 | y |
| 0 | 0 |
| 1 | 1 |
| 1 | 1 |
| 0 | 0 |
| 1 | 0 |
| 0 | 1 |

A następnie dzielimy przez ilość przykładów negatywnych w N w których ta cecha jest równa 1

|  |  |
| --- | --- |
| f1 | y |
| 0 | 0 |
| 1 | 1 |
| 1 | 1 |
| 0 | 0 |
| 1 | 0 |
| 0 | 1 |

v1 = 2/1 = 2

max(v1,v2,v3,v4) => v3 – znalezienie najwiekszej wartości vi (w tym przypadku 2/0 jest najwiekszą wartością. Nie dzielimy przez 0 tylko zastępujemy to jakąś wartością bliską zeru np. 0.001 co daje nam wynik 2000 i jest najwieksze.

r = f3 – oznacza to wybranie cechy 3

Przykłady negatywne pokryte przez ceche f3 to {x1,x4, x5} - Pokrycie oznacza ze w kolumnie cechy w przykładach negatywnych jest 0:

|  |  |
| --- | --- |
| f3 | y |
| 0 | 0 |
| 0 | 1 |
| 1 | 1 |
| 0 | 0 |
| 0 | 0 |
| 1 | 1 |

N \ {x1,x4, x5} = zbiór pusty – usuniecie przykładów negatywnych, tych które są pokryte przez wybraną ceche, przy czym zbiór pozostał pusty więc kończymy pętle.

**Koniec wewnetrznej pętli**

h = f3 – dodanie do hipotezy wcześniej zapamietanego r

Przykłady pozytywne pokryte przez r to {x3, x6} – czyli wszędzie gdzie f3 jest równe 1 i y jest równe 1

|  |  |
| --- | --- |
| f3 | y |
| 0 | 0 |
| 0 | 1 |
| 1 | 1 |
| 0 | 0 |
| 0 | 0 |
| 1 | 1 |

P było równe: P = {x2, x3, x6} i teraz od tego będziemy odejmować to pokryte przykłady.

P = P \ {x3, x6} = { x2} – P nie puste więc następna iteracja pętli

**Druga iteracja zewnetrznej pętli:**

Należy zwrócić uwagę, że P w tej iteracji jest równe P = { x2} więc przy wyborze cech będziemy brać tylko ten jeden przykład pozytywny.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | f1 | f2 | f3 | f4 | y |
| x1 | 0 | 1 | 0 | 0 | 0 |
| x2 | 1 | 1 | 0 | 0 | 1 |
| x3 | 1 | 0 | 1 | 0 | 1 |
| x4 | 0 | 1 | 0 | 1 | 0 |
| x5 | 1 | 0 | 0 | 1 | 0 |
| x6 | 0 | 1 | 1 | 0 | 1 |

r=true – zresetowanie zmiennej przchowującej tymczasowo wybrane cechy

N={x1,x4, x5} –Ustawienie przykładów negatywnych ( czyli tam gdzie 0 w kolumnie y)

**Pierwsza iteracja wewnetrznej pętli**

v1=1/1, v2=1/2, v3=0/0, v4=0/2 - Pamiętamy ze obliczamy dla P = { x2} i N={x1,x4, x5}

Dla przykładu v1:

Pozytywne:

|  |  |
| --- | --- |
| f1 | y |
| 0 | 0 |
| 1 | 1 |
| 1 | 1 |
| 0 | 0 |
| 1 | 0 |
| 0 | 1 |

Negatywne:

|  |  |
| --- | --- |
| f1 | y |
| 0 | 0 |
| 1 | 1 |
| 1 | 1 |
| 0 | 0 |
| 1 | 0 |
| 0 | 1 |

czyli v1=1/1

max(v1,v2,v3,v4) => v1

r = f1

Przykłady negatywne pokryte przez ceche f3 to {x1,x4} - Pokrycie oznacza ze w kolumnie cechy w przykładach negatywnych jest równe 0:

|  |  |
| --- | --- |
| f1 | y |
| 0 | 0 |
| 1 | 1 |
| 1 | 1 |
| 0 | 0 |
| 1 | 0 |
| 0 | 1 |

N = N \ {x1,x4} = {x5} – Usuniecie z tablicy N przykładów negatywnych pokrytych przez wybraną ceche i jako że ta cecha nie pokrywała wszystkich przykładów przchodzimy do 2 iteracji wewnętrznej pętli.

**Druga iteracja wewnetrznej pętli**

v2=1/0, v3=0/0, v4=0/1 - Pamiętamy ze obliczamy dla P = {x2} i N={x5} oraz ze cecha f1 już została wybrana.

max(v2,v3,v4) => v2

r = f1 ^ f2 - Dorzucamy drugą wybarną cechę do zmiennej r

Przykłady negatywne pokryte przez cechę f2 to {x5} - Pokrycie oznacza ze w kolumnie cechy w przykładach negatywnych jest równe 0:

|  |  |
| --- | --- |
| f2 | y |
| 1 | 0 |
| 1 | 1 |
| 0 | 1 |
| 1 | 0 |
| 0 | 0 |
| 1 | 1 |

N = N \ {x5} = pusty – Usuwamy pokryte negatywne przykład i dostajemy pusty zbiór negatywnych, oznacza to koniec pętli wewnetrznej

**Koniec wewnetrznej pętli**

h = f3 v f1 ^ f2 – dodanie do hipotezy wcześniej zapamietanego r (wcześniej h było równe f3 teraz dodajemy znowu r powstałe w wewnętrznej pętli i miedzy tym znak alternatywy „v”

Przykłady pozytywne pokryte przez r to {x2} – czyli wszędzie gdzie f1 i f2 jest równa 1 i y jest równe 1:

|  |  |  |
| --- | --- | --- |
| f1 | f2 | y |
| 0 | 1 | 0 |
| 1 | 1 | 1 |
| 1 | 0 | 1 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 0 | 1 | 1 |

P było równe: P = {x2 } i teraz od tego będziemy odejmować to pokryte przykłady.

P = P \ {x2, x6}= zbiór pusty – P puste więc koniec algorytmu

**Wynik:**

Hipoteza: f3 v f1 ^ f2