

Eksploracja danych

Projekt 4: Asocjacyjne reguły grupowania

Celem tego zadania programistycznego jest zaimplementowanie techniki znajdowania częstych zestawów elementów przy użyciu algorytmu *Apriori* i przy użyciu tych informacji wygenerowanie reguł asocjacji, które mają nośniki i zaufanie powyżej pewnych minimalnych progów.

Plik wejściowy znajdziecie pod adresem:

<http://www.math.us.edu.pl/~pgladki/teaching/2021-2022/ed-p04.txt>

Zawiera on tekst, który wygląda następująco:

```
1 3 4
1 2 3 5
2 3 5
2 5
1 2 3 6
```

W pliku spacje oddzielają pozycje (identyfikowane przez liczby całkowite), a nowe wiersze oddzielają transakcje. Na przykład powyższy plik zawiera informacje o łącznie 5 transakcjach, a druga transakcja składa się z 4 pozycji.

Twoim zadaniem jest napisanie programu w MatLabie, który przyjmuje jako parametry minimalny nośnik, minimalne zaufanie (podane jako liczby zmiennoprzecinkowe z zakresu $[0..1]$) oraz nazwę pliku z transakcjami (których format jest taki jak w pliku `ed-p04.txt` powyżej) i tworzy wszystkie reguły asocjacji, które można wydobyć z pliku transakcji, które spełniają minimalne wymagania dotyczące nośnika i zaufania. Reguły powinny być sortowane na wyjściu najpierw według liczby zawartych w nich pozycji (w kolejności malejącej), następnie według ufności, a na końcu według ich nośnika (również w kolejności malejącej).