
Laboratorium

Reguły asocjacyjne

Zadanie 1

Ponad 100 lat temu zatonął Titanic. Na pokładzie znajdowały się tysiące ludzi. Niektórzy przeżyli, inni nie. Czy były jakieś zależności pomiędzy płcią, zamożnością i przeżywalnością? Twoim zadaniem jest znalezienie reguł asocjacyjnych, które odpowiedzą na te pytania. Wczytaj plik [titanic.csv](#) ze strony lub z Internetu i upewnij się, że ma tylko 4 kolumny: "Class", "Sex", "Age", "Survived".

Rozwiązanie powinno zawierać:

1. Uruchomienie algorytmu Apriori na bazie danych.
2. Wyszukanie reguł o sensownych parametrach (minimum 0.005, i ufności 0.8) i posortowanie ich wg ufności.
3. Wyszukanie najciekawszych reguł. W szczególności wskazujących, kto przeżywał a kto nie. Można tutaj próbować odfiltrować niepotrzebne reguły.
4. Zobrazowanie reguł na pomocniczych wykresach i interpretacja ich. Zaprezentuj przynajmniej jeden.

Skorzystaj z jednego z samouczków:

<https://medium.com/analytics-vidhya/association-analysis-in-python-2b955d0180c>
<https://www.codespeedy.com/apriori-algorithm-in-python/>
<https://pbpython.com/market-basket-analysis.html>
<https://towardsdatascience.com/mba-for-breakfast-4c18164ef82b>

Zadanie 2

Użyj reguł asocjacyjnych do analizy zawartości koszyka sklepowego i przedstaw swoje wyniki. Czy reguły asocjacyjne są zgodne z przewidywaniami? Czy były jakieś ciekawe zależności? Jakie minimalne suport i confidence ma sens w przypadku danej bazy danych.

Zadanie należy rozwiązać na podstawie samouczków dostępnych w Internecie, wybierając wskazaną w nich bazę danych.

Przykłady samouczków (można poszukać własnych):

- Fajna baza danych: <https://www.kaggle.com/datasets/vijayuv/onlineretail> „Online Retail Dataset”
- Samouczek: <https://towardsdatascience.com/introduction-to-simple-association-rules-mining-for-market-basket-analysis-ef8f2d613d87>
- Samouczek: <https://pbpython.com/market-basket-analysis.html>
- <https://medium.com/market-basket-analysis-mba-using-python/market-basket-analysis-mba-using-python-5017e22c62af>