

Piotr Wardecki	234128	234128@edu.p.lodz.pl
Paweł Galewicz	234053	234053@edu.p.lodz.pl
Bartosz Jurczewski	234067	234067@edu.p.lodz.pl

Zadanie 4

1. Cel zadania

Celem zadania było napisanie programu, który implementuje reguły asocjacyjne za pomocą algorytmu A-priori. Program na podstawie ostatnio przeglądanych przez użytkowników przedmiotów identyfikuje te, które regularnie występowały wspólnie w ramach pojedynczej sesji.

2. Opis implementacji

Do wykonywania zadania niezbędna była instancja *Apache Spark*. Aby ograniczyć liczbę zainstalowanych środowisk skorzystaliśmy z odpowiedniego obrazu dla Dockera [1], który zawierał także *Jupyter Notebook*, *Python* oraz *Miniconda*. Dodatkowo aby ułatwić tworzenie środowiska do kolejnych zadań i między naszymi komputerami skorzystaliśmy z narzędzia *Docker Compose* (nasz plik [2]). Do stworzenia programu bardzo pomocna była funkcja `parallelize()` która odpowiedzialna była za stworzenie rozproszonego zbioru danych na podstawie listy kolekcji. Pozwalało to na wykorzystanie paradygmatu map-reduce do rozwiązywania większej części zagadnień algorytmu.

3. Wyniki

Tabela 1. 5 reguł asocjacyjnych o największej ufności dla par

Nr	Element	Ufność
1	DAI93865 [FRO40251]	1.0
2	GRO85051 [FRO40251]	0.999176276771005
3	GRO38636 [FRO40251]	0.9906542056074766
4	ELE12951 [FRO40251]	0.9905660377358491
5	DAI88079 [FRO40251]	0.9867256637168141

Tabela 2. 5 reguł asocjacyjnych o największej ufności dla trójek

Nr	Element	Ufność
1	DAI23334 ELE92920 [DAI62779]	1.0
2	DAI31081 GRO85051 [FRO40251]	1.0
3	DAI55911 GRO85051 [FRO40251]	1.0
4	DAI62779 DAI88079 [FRO40251]	1.0
5	DAI75645 GRO85051 [FRO40251]	1.0

4. Wnioski

- Stworzone przez nas reguły asocjacyjne mają najlepszą dokładność dla trójek.
- Najczęściej występującym następnikiem wygenerowanych reguł był [FRO40251].

Bibliografia

- [1] *Jupyter Notebook Python, Spark Stack* <https://hub.docker.com/r/jupyter/pyspark-notebook>
- [2] *Plik Docker Compose do zadania 2* <https://github.com/jurczewski/PiADZD/blob/master/zad2/docker-compose.yml>