Eksploracja danych internetowych

Sprawozdanie

Ćwiczenie 1

Adam Bemski 220061

Wojciech Pełka 220090

1. Plik logów

Do wykonania ćwiczenia posłużyliśmy się plikiem logów, który stanowią dane dotyczące wszystkich zapytań http z sierpnia 1995 roku do serwera NASA Kennedy Space Center znajdującego się na Florydzie.

Format plików:
Dane w pliku są w postaci ASCII, gdzie jedna linia oznacza jedno zapytanie do serwera.
Dane dzielą się na następujące kolumny:

- **Host** – użytkownik wykonujący zapytanie
- **Timestamp** – znacznik czasowy w formacie dd/month/rrrr:hh:mm:ss
- **Request -** zapytanie do serwera
- **Kod http** – zwrócony kod odpowiedzi http z serwera
- **Ilość bajtów** – ilość bajtów otrzymanych w odpowiedzi od serwera

2. Proces przygotowania danych

1. Obcięcie pliku logów do 50000 wierszy za pomocą programu Notepad++
2. Za pomocą skryptu napisanego w języku C# z pliku zostały wybrane tylko rekordu o metodzie GET, kodzie statusu 200 oraz zostały usunięte wszystkie odwołania do plików graficznych
3. Za pomocą skryptu napisanego w języku C# zidentyfikowani zostali użytkownicy (host) oraz wyodrębnione zostały dla nich sesje o odstępie czasowym równym 30 minut.
4. Za pomocą skryptu napisanego w języku C# zostały wyznaczone najbardziej popularne strony, dla których liczba odwiedzin była większa niż 0.5%.
5. Za pomocą skryptu napisanego w języku C# została przeprowadzona transformacja koszykowa dla sesji oraz zostały nadane atrybuty takie jak czas sesji, liczba działań w czasie sesji, przeciętny czas na stronę oraz zmienne flagowe dla najpopularniejszych stron.
6. W późniejszym etapie również za pomocą skryptu zostały zmienione atrybuty sesji na typ kategoryczny dla przeciętnego czasu na stronę oraz dla czasu sesji.

3. Atrybuty sesji i użytkowników

- Użytkowników

  @ATTRIBUTE user STRING
  @ATTRIBUTE sessions INTEGER
  @ATTRIBUTE /shuttle/missions/sts-69/mission-sts-69.html {T,F}
  @ATTRIBUTE /shuttle/missions/missions.html {T,F}
  @ATTRIBUTE /shuttle/missions/sts-70/mission-sts-70.html {T,F}
  @ATTRIBUTE /ksc.html {T,F}
  @ATTRIBUTE /shuttle/countdown/ {T,F}
  @ATTRIBUTE /history/apollo/apollo-13/ {T,F}
  @ATTRIBUTE /history/apollo/ {T,F}
  @ATTRIBUTE /shuttle/missions/sts-69/ {T,F}
  @ATTRIBUTE /shuttle/missions/ {T,F}
  @ATTRIBUTE /htbin/wais.pl {T,F}
  @ATTRIBUTE /images/ {T,F}
  @ATTRIBUTE /images {T,F}
  @ATTRIBUTE /shuttle {T,F}
  @ATTRIBUTE /shuttle/ {T,F}
  @ATTRIBUTE /shuttle/missions/sts-70/ {T,F}
  @ATTRIBUTE /shuttle/technology/ {T,F}
  @ATTRIBUTE /shuttle/missions/sts-69 {T,F}
  @ATTRIBUTE /shuttle/countdown {T,F}
  @ATTRIBUTE /shuttle/missions {T,F}
  @ATTRIBUTE /software/winvn/ {T,F}
  @ATTRIBUTE /software/ {T,F}
  @ATTRIBUTE /elv/ {T,F}
  @ATTRIBUTE /shuttle/missions/sts-71/ {T,F}
  @ATTRIBUTE /software/winvn {T,F}
  @ATTRIBUTE /shuttle/technology/sts-newsref/ {T,F}
  @ATTRIBUTE /l {T,F}
  @ATTRIBUTE /history {T,F}
  @ATTRIBUTE /history/ {T,F}
  @ATTRIBUTE /history/apollo {T,F}
  @ATTRIBUTE /shuttle/mission {T,F}
  @ATTRIBUTE /shuttle/missions/sts-71 {T,F}
  @ATTRIBUTE /software {T,F}
  @ATTRIBUTE /missions/missions.html {T,F}

- Sesji

  @ATTRIBUTE host STRING
  @ATTRIBUTE session-number INTEGER
  @ATTRIBUTE session-time INTEGER
  @ATTRIBUTE operations INTEGER
  @ATTRIBUTE average-time INTEGER
  @ATTRIBUTE /shuttle/missions/sts-69/mission-sts-69.html {T,F}
  @ATTRIBUTE /shuttle/missions/missions.html {T,F}
  @ATTRIBUTE /shuttle/missions/sts-70/mission-sts-70.html {T,F}
  @ATTRIBUTE /ksc.html {T,F}
  @ATTRIBUTE /shuttle/countdown/ {T,F}
  @ATTRIBUTE /history/apollo/apollo-13/ {T,F}
  @ATTRIBUTE /history/apollo/ {T,F}
  @ATTRIBUTE /shuttle/missions/sts-69/ {T,F}
  @ATTRIBUTE /shuttle/missions/ {T,F}
  @ATTRIBUTE /htbin/wais.pl {T,F}

@ATTRIBUTE /images/ {T,F}
@ATTRIBUTE /images {T,F}
@ATTRIBUTE /shuttle {T,F}
@ATTRIBUTE /shuttle/ {T,F}
@ATTRIBUTE /shuttle/missions/sts-70/ {T,F}
@ATTRIBUTE /shuttle/technology/ {T,F}
@ATTRIBUTE /shuttle/missions/sts-69 {T,F}
@ATTRIBUTE /shuttle/countdown {T,F}
@ATTRIBUTE /shuttle/missions {T,F}
@ATTRIBUTE /software/winvn/ {T,F}
@ATTRIBUTE /software/ {T,F}
@ATTRIBUTE /elv/ {T,F}
@ATTRIBUTE /shuttle/missions/sts-71/ {T,F}
@ATTRIBUTE /software/winvn {T,F}
@ATTRIBUTE /shuttle/technology/sts-newsref/ {T,F}
@ATTRIBUTE /l {T,F}
@ATTRIBUTE /history {T,F}
@ATTRIBUTE /history/ {T,F}
@ATTRIBUTE /history/apollo {T,F}
@ATTRIBUTE /shuttle/mission {T,F}
@ATTRIBUTE /shuttle/missions/sts-71 {T,F}
@ATTRIBUTE /software {T,F}
@ATTRIBUTE /missions/missions.html {T,F}

4. Wiersze sterujące sesji i użytkowników

- Użytkowników

120.33.med.umich.edu,1,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F
128.102.202.133,1,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F
128.104.235.113,1,F,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F
128.114.23.148,1,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F
128.119.50.139,1,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F
128.135.36.35,1,F,T,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F
128.138.169.91,1,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F
128.149.109.74,1,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F
(…)

- Sesji

120.33.med.umich.edu,1,150,4,37.5,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F
128.102.202.133,1,180,4,45,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F
128.104.235.113,1,200,3,66.67,F,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,
128.114.23.148,1,571,2,285.5,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F
128.119.50.139,1,1728,7,246.86,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F
128.135.36.35,1,4,2,2,F,T,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F
128.138.169.91,1,143,5,28.6,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,
128.149.109.74,1,92,2,46,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F
(…)

5. Wybrana Metoda Klastrowania – SimpleKMeans

Wybraną przez nas metodą klastrowania jest metoda K-średnich. Jest to metoda należąca do grupy algorytmów analizy skupień tj. analizy polegającej na szukaniu i wyodrębnianiu grup obiektów podobnych (skupień).

Zasada działania algorytmu:

1. **Ustalamy liczbę skupień** – jedną z metod ustalania jest umowny wybór ilości skupień.
2. **Ustalamy wstępne środki skupień** – tak zwane centroidy. Możemy je dobrać na kilka sposobów, jednak jedną z najczęściej stosowanych metod jest kilkukrotne uruchomienie algorytmu i wybór najlepszego modelu.
3. **Obliczamy odległości obiektów od środków skupień** – Wybór metryki wpływa na to, które z obserwacji będą uważane za podobne a które za różniące się od siebie. Najczęściej stosowaną odległością jest odległość euklidesowa. Odległość euklidesowa między dwoma punktami jest równa długości odcinka łączącego te punkty.
4. **Przypisujemy obiekty do skupień** – porównujemy odległości od wszystkich skupień i wybieramy tą która do środka ma najbliżej.
5. **Wykonujemy punkty 3,4,5 aż do zatrzymania** – czyli ustalamy ilość iteracji ile ma być wykonanych.

Użytkownicy – dla 3 klastrów

Liczba iteracji: 3

Within cluster sum of squared errors: 1450.0278793832417

Cluster 0: 1,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

Cluster 1: 1,T,F,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

Cluster 2: 1,T,T,F,T,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

Final cluster centroids:

Attribute	Full Data (1827.0)	Cluster# 0 (1182.0)	1 (511.0)	2 (134.0)
sessions	1.2556	1.1108	1.3523	2.1642
/shuttle/missions/sts-69/mission-sts-69.html	F	F	F	T
/shuttle/missions/missions.html	F	F	F	T
/shuttle/missions/sts-70/mission-sts-70.html	F	F	F	F
/ksc.html	F	F	T	T
/shuttle/countdown/	F	F	F	T
/history/apollo/apollo-13/	F	F	F	F
/history/apollo/	F	F	F	F
/shuttle/missions/sts-69/	F	F	F	F
/shuttle/missions/	F	F	F	F
/htbin/wais.pl	F	F	F	F
/images/	F	F	F	F
/images	F	F	F	F
/shuttle	F	F	F	F
/shuttle/	F	F	F	F
/shuttle/missions/sts-70/	F	F	F	F
/shuttle/technology/	F	F	F	F
/shuttle/missions/sts-69	F	F	F	F
/shuttle/countdown	F	F	F	F
/shuttle/missions	F	F	F	F
/software/winvn/	F	F	F	F
/software/	F	F	F	F
/elv/	F	F	F	F
/shuttle/missions/sts-71/	F	F	F	F
/software/winvn	F	F	F	F
/shuttle/technology/sts-newsref/	F	F	F	F
/l	F	F	F	F
/history	F	F	F	F
/history/	F	F	F	F
/history/apollo	F	F	F	F
/shuttle/mission	F	F	F	F
/shuttle/missions/sts-71	F	F	F	F
/software	F	F	F	F
/missions/missions.html	F	F	F	F

Wnioski:

W przypadku klastra 0 o największej liczności użytkowników (1182) widać, że żaden z użytkowników nie odwiedził ani jednej popularnej strony.

W przypadku klastra 1 o średniej liczności użytkowników (511) widać, że użytkownicy odwiedzili jedną popularną stronę: */ksc.html*

W przypadku klastra 2 o najmniejszej liczności użytkowników (134) widać, że użytkownicy odwiedzili 4 popularne strony: /shuttle/missions/sts-69/mission-sts-69.html , /shuttle/missions/missions.html , /ksc.html i /shuttle/countdown/.

Użytkownicy – dla 6 klastrów

Liczba iteracji: 5

Within cluster sum of squared errors: 1114.8009201879943

Clusters

0: 1,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

1: 1,T,F,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

2: 1,T,T,F,T,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

3: 1,T,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

4: 2,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

5: 1,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

Final cluster centroids:

Attribute	Full Data (1827.0)	Cluster# 0 (228.0)	1 (413.0)	2 (116.0)	3 (114.0)	4 (90.0)	5 (866.0)
sessions	1.2556	1.2368	1	2.1207	1.1579	2.9111	1.1074
/shuttle/missions/sts-69/mission-sts-69.html	F	F	F	T	F	F	F
/shuttle/missions/missions.html	F	T	F	T	F	F	F
/shuttle/missions/sts-70/mission-sts-70.html	F	F	F	F	T	F	F
/ksc.html	F	F	T	T	F	T	F
/shuttle/countdown/	F	F	F	T	F	F	F
/history/apollo/apollo-13/	F	F	F	F	F	F	F
/history/apollo/	F	F	F	F	F	F	F
/shuttle/missions/sts-69/	F	F	F	F	F	F	F
/shuttle/missions/	F	F	F	F	F	F	F
/htbin/wais.pl	F	F	F	F	F	F	F
/images/	F	F	F	F	F	F	F
/images	F	F	F	F	F	F	F
/shuttle	F	F	F	F	F	F	F
/shuttle/	F	F	F	F	F	F	F
/shuttle/missions/sts-70/	F	F	F	F	F	F	F
/shuttle/technology/	F	F	F	F	F	F	F
/shuttle/missions/sts-69	F	F	F	F	F	F	F
/shuttle/countdown	F	F	F	F	F	F	F
/shuttle/missions	F	F	F	F	F	F	F
/software/winvn/	F	F	F	F	F	F	F
/software/	F	F	F	F	F	F	F
/elv/	F	F	F	F	F	F	F
/shuttle/missions/sts-71/	F	F	F	F	F	F	F
/software/winvn	F	F	F	F	F	F	F
/shuttle/technology/sts-newsref/	F	F	F	F	F	F	F
/l	F	F	F	F	F	F	F
/history	F	F	F	F	F	F	F
/history/	F	F	F	F	F	F	F
/history/apollo	F	F	F	F	F	F	F
/shuttle/mission	F	F	F	F	F	F	F
/shuttle/missions/sts-71	F	F	F	F	F	F	F
/software	F	F	F	F	F	F	F
/missions/missions.html	F	F	F	F	F	F	F

Wnioski:

W przypadku klastra 5 o największej liczności użytkowników (866) widać, że żaden z użytkowników nie odwiedził ani jednej popularnej strony.

W przypadku klastra 1 o średniej liczności użytkowników (413) widać, że użytkownicy odwiedzili jedną popularną stronę: */ksc.html*

W przypadku klastra 4 o najmniejszej liczności użytkowników (90) widać, że użytkownicy odwiedzili jedną popularną stronę: */ksc.html*

Sesje – dla 3 klastrów

Liczba iteracji : 10

Within cluster sum of squared errors: 2240.243112148035

Clusters:

0: 4,194,2,97,F,F,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

1: 2,139,4,34.75,F,F,F,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

2: 7,189,4,47.25,F,F,F,T,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

```
Final cluster centroids:
                                                  Cluster#
Attribute                                Full Data         0         1         2
                                          (2294.0)  (489.0)  (1450.0)  (355.0)
=============================================================================================
session-number                             2.9024    3.4254    2.7317    2.8789
session-time                             463.5641  1220.182  182.7869  568.1859
operations                                 5.0663    7.7321    3.7945    6.5887
average-time                             109.1155  276.5702   52.0037  111.7259
```

Wnioski:

W przypadku klastra 1 o największej liczności użytkowników (1450) widać, że dla każdego z użytkowników średni czas trwania sesji to około 183 sekundy. Średnia liczba operacji na sesję to 3.7, natomiast średni czas spędzony na stronie około 52 sekundy.

W przypadku klastra 0 o średniej liczności użytkowników (489) widać, że dla każdego z użytkowników średni czas trwania sesji to około 1220 sekund czyli około 20 minut. Średnia liczba operacji na sesję to około 7.7, natomiast średni czas spedzony na stronie około 277 sekund, czyli około 4.6 minuty.

W przypadku klastra 2 o najmniejszej liczności użytkowników (355) widać, że dla każdego z użytkowników średni czas trwania sesji to około 568 sekund, czyli około 9 minut. Średnia liczba operacji na sesję to około 6.6, natomiast średni czas spędzony na stronie około 112 sekund, czyli około 2 minut.

Sesje – dla 6 klastrów

Liczba iteracji: 7

Within cluster sum of squared errors: 1499.1434104211758

Clusters:

0: 4,194,2,97,F,F,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

1: 2,139,4,34.75,F,F,F,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

2: 7,189,4,47.25,F,F,F,T,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

3: 8,74,2,37,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

4: 1,170,2,85,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

5: 1,254,4,63.5,F,F,T,F,T,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F,F

Final cluster centroids:

Attribute	Full Data	Cluster# 0	1	2	3	4	5
	(2294.0)	(559.0)	(178.0)	(160.0)	(106.0)	(1191.0)	(100.0)
session-number	2.9024	3.4132	2.5618	2.1938	17.8585	1.5743	1.75
session-time	463.5641	561.5403	427.691	457.9875	468.934	392.0999	834.1
operations	5.0663	4.6762	5.0281	6.3	4.0189	4.7767	9.9
average-time	109.1155	154.1826	96.0328	71.9551	133.5691	92.9598	106.4277

Wnioski:

W przypadku klastra 4 o największej liczności użytkowników (1191) widać, że dla każdego z użytkowników średni czas trwania sesji to około 393 sekundy, czyli około 6.5 minuty. Średnia liczba operacji na sesję to około 4.8, natomiast średni czas spędzony na stronie około 93 sekundy, czyli niewiele ponad 1.5 minuty.

W przypadku klastra 0 o średniej liczności użytkowników (559) widać, że dla każdego z użytkowników średni czas trwania sesji to około 562 sekundy, czyli około 9 minut. Średnia liczba operacji na sesję to około 4.7, natomiast średni czas spędzony na stronie około 154 sekundy,d czyli około 2.5 minuty.

W przypadku klastra 5 o najmniejszej liczności użytkowników (100) widać, że dla każdego z użytkowników średni czas trwania sesji to około 834 sekundy czyli około 14 minut. Średnia liczba operacji na sesję to około 10, natomiast średni czas spedzony na stronie około 107 sekund czyli około 2 minut.

Można więc zauważyć, że im większa liczność danych dla klastra tym średni czas trwania sesji maleje.

6. Metoda znajdowania reguł asocjacyjnych

Wybraną przez nas metodą znajdowania reguł asocjacyjnych jest moteda Apriori. Metoda asocjacyjna (odkrywania asocjacji) jest jedną z najpopularniejszych metod eksploracji danych, polegającą na analizowaniu zbioru atrybutów z bazy danych pod kątem występowania w nim powtarzających się zależności. Wykorzystywane jest to w analizie koszykowej, marketingu krzyżowym, projektowaniu katalogów itd.

Dla celów zadania plik został poddany dyskretyzacji.

- Czas sesji został podzielony na : **Short** jeśli czas znajduje się w przedziale 1 – 368 sekund. **Average** jeśli czas znajduje się w przedziale 369 – 938 sekund. **Long** jeśli czas jest większy niż 938 sekund.
- Średni czas na stronę został podzielony na : **Small** jeśli czas znajduje się w przedziale 0.5 – 42 sekund. **Medium** jeśli czas znajduje się w przedziale 43 – 89 sekund. **Large** jeśli czas jest większy niż 938 89.
- Liczba operacji na stronę została podzielona na : **Small** jeśli liczba operacji na stronie znajduje się w przedziale 1 – 16 operacji. **Medium** jeśli liczba operacji na stronie znajduje się w przedziale 17 – 32 operacji. **Large** jeśli liczba operacji na stronie jest większa niż 32.

Wyniki reguły przeprowadzonej w programie Weka zostały przedstawione poniżej.

```
Apriori
=======

Minimum support: 0.15 (344 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 17

Generated sets of large itemsets:

Size of set of large itemsets L(1): 7

Size of set of large itemsets L(2): 9

Size of set of large itemsets L(3): 3

Best rules found:

 1. session-time=short average-time=medium 358 ==> operations=small 358    <conf:(1)> lift:(1.04) lev:(0.01) [13] conv:(13.58)
 2. session-time=short 1359 ==> operations=small 1356    <conf:(1)> lift:(1.04) lev:(0.02) [48] conv:(12.89)
 3. session-time=short average-time=small 854 ==> operations=small 851    <conf:(1)> lift:(1.04) lev:(0.01) [29] conv:(8.1)
 4. average-time=large 858 ==> operations=small 851    <conf:(0.99)> lift:(1.03) lev:(0.01) [25] conv:(4.07)
 5. session-time=long average-time=large 370 ==> operations=small 363    <conf:(0.98)> lift:(1.02) lev:(0) [7] conv:(1.75)
 6. operations=small average-time=small 871 ==> session-time=short 851    <conf:(0.98)> lift:(1.65) lev:(0.15) [335] conv:(16.91)
 7. operations=small session-time=long 372 ==> average-time=large 363    <conf:(0.98)> lift:(2.61) lev:(0.1) [223] conv:(23.29)
 8. average-time=small 895 ==> operations=small 871    <conf:(0.97)> lift:(1.01) lev:(0) [9] conv:(1.36)
 9. session-time=average 500 ==> operations=small 479    <conf:(0.96)> lift:(1) lev:(-0) [-2] conv:(0.86)
10. average-time=small 895 ==> session-time=short 854    <conf:(0.95)> lift:(1.61) lev:(0.14) [323] conv:(8.69)
```

Wnioski:

Na powyższym obrazku widać, że reguła apriori została przeprowadzona poprawnie. Zostały wyznaczone wspólne zachowania dla różnych sytuacji. Np. według wygenerowanych wyników możemy stwierdzić, że kiedy czas sesji jest krótki to liczba operacji na stronie jest mała (reguła 2). Kiedy liczba operacji jest mała to średni czas spędzony na stronie jest mały (reguła 6) itd. Dzięki regule apriori, możemy badać poszczególne przypadki i zachowania użytkowników na stronie.