

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное  
учреждение высшего образования  
**"Южно-Уральский государственный университет  
(национальный исследовательский университет)"**  
Высшая школа электроники и компьютерных наук  
Кафедра системного программирования

**ОТЧЕТ**  
**по практической работе №2**  
**«Поиск ассоциативных правил»**  
по дисциплине  
**«Технологии аналитической обработки информации»**

Выполнил: \_\_\_\_\_

студент группы КЭ-404

А.Ю. Емельянова

Проверил: \_\_\_\_\_

преподаватель

А.И. Гоглачев

Дата: \_\_\_\_\_

Оценка: \_\_\_\_\_

## **Формулировка задания**

1. Доработайте программу из задания Поиск частых наборов, чтобы она также выполняла поиск ассоциативных правил. Список результирующих правил должен выдаваться в удобочитаемом виде (антецедент→консеквент) с указанием поддержки и достоверности каждого правила. Дополнительные параметры программы: порог достоверности, способ упорядочивания результирующего списка наборов (по убыванию значения поддержки или лексикографическое).

2. Проведите эксперименты на наборах из задания 1. В экспериментах зафиксируйте значение пороговое значение поддержки (например, 10%), варьируйте пороговое значение достоверности (например, от 70% до 95% с шагом 5%).

3. Выполните визуализацию полученных результатов в виде следующих диаграмм:

1) сравнение быстродействия поиска правил на фиксированном наборе данных при изменяемом пороге достоверности;

2) общее количество найденных правил на фиксированном наборе данных при изменяемом пороге достоверности.

4. Подготовьте список правил, в которых антецедент и консеквент суммарно включают в себя не более семи объектов (разумное количество). Проанализируйте и изложите содержательный смысл полученного результата.

Гиперссылка на каталог репозитория с исходными текстами, наборами данных и другими материалами: <https://github.com/Sunless-sea/Technologies-of-analytical-information-processing>

## **Визуализация**

В данной практической работе были проведены некоторые эксперименты при изменяемом пороге достоверности и нахождении ассоциативных правил. Также были выявлены некоторые зависимости между временем обработки данных, количеством найденных правил и различными

порогами достоверности.

Сравнение быстродействия на фиксированном наборе данных при изменяемом пороге достоверности представлено на рисунке 1.

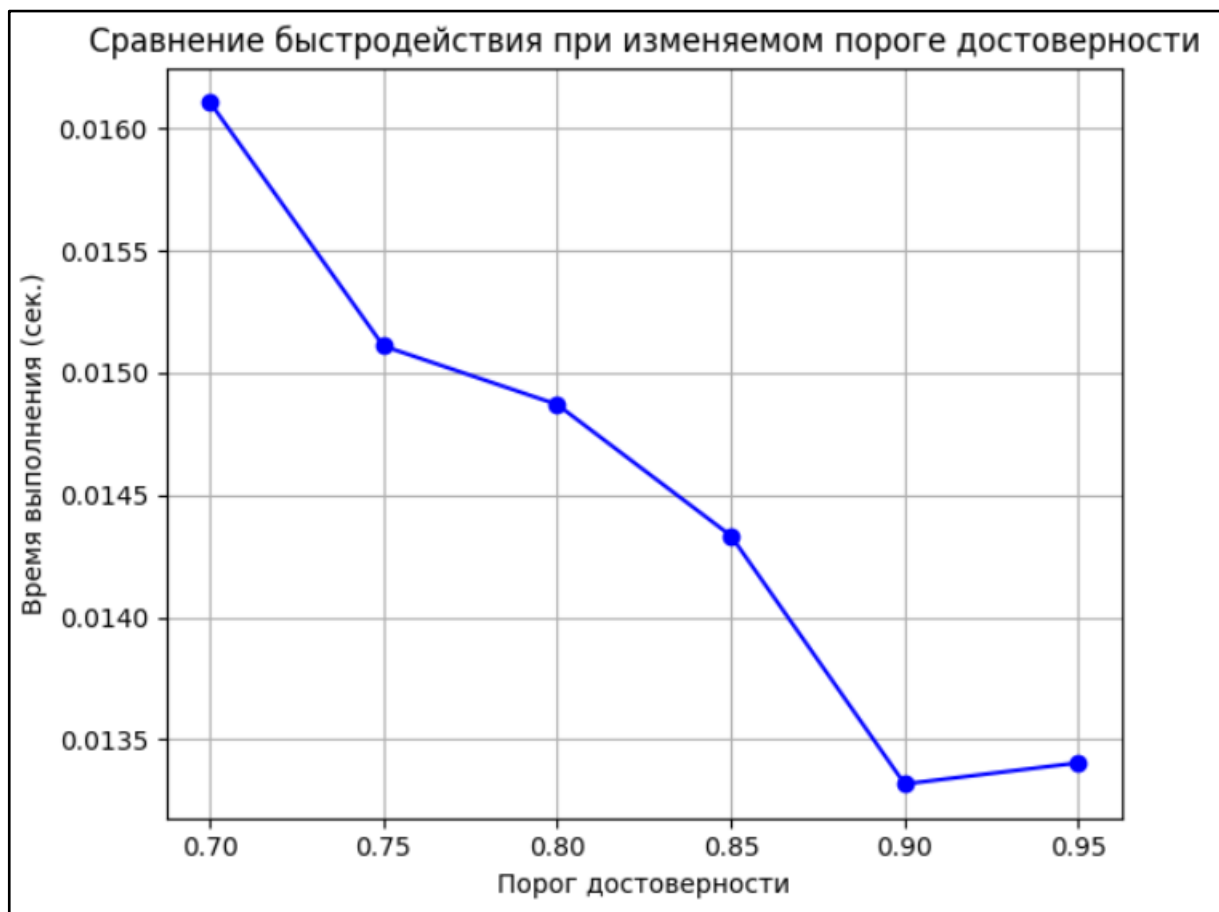


Рисунок 1 – График зависимости времени от порога достоверности

Из анализа графика можно сделать следующий вывод: при увеличении порога достоверности время выполнения уменьшается, так как при увеличении порога, количество найденных правил уменьшается, а значит необходимо обрабатывать меньшее количество правил, поэтому время работы алгоритма и сложность вычислений уменьшаются. Чем меньше правил подлежит обработке, тем ниже вычислительная сложность.

Общее количество найденных правил на фиксированном наборе данных при изменяемом пороге достоверности представлено на рисунке 2.

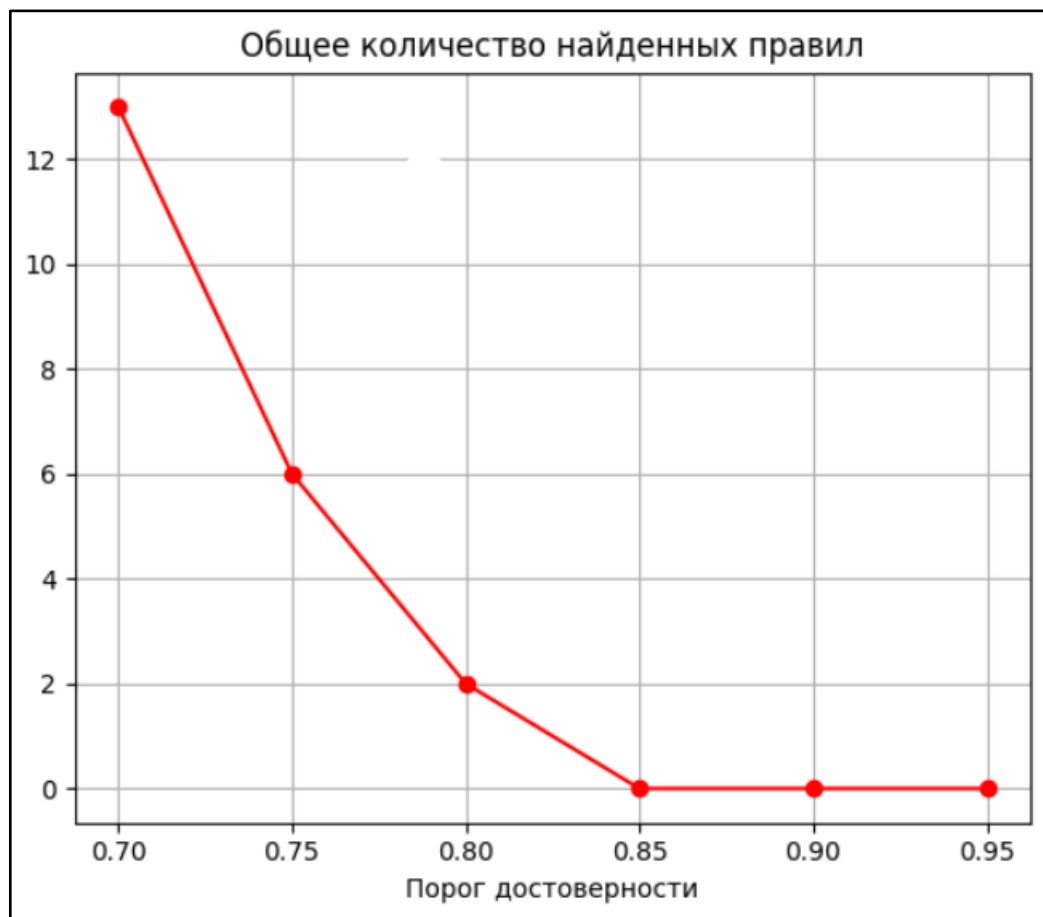


Рисунок 2 – График зависимости общего количества найденных правил от изменения порога достоверности

На графике можно заметить, что при увеличении порога достоверности количество найденных правил уменьшается, так как всё меньшее количество правил подходит под требования к достоверности, поэтому правила, не подходящие под заданный порог, исключаются.

Список правил, в которых антецедент и консеквент суммарно включают в себя не более семи объектов, представлен на рисунке 3.

	antecedents	consequents	support	confidence
0	(грибной соус, эскалоп)	(макароны)	0.004266	0.744186
2	(макароны, растительное масло, яйца)	(минеральная вода)	0.003066	0.718750
1	(суп, замороженные овощи, молоко)	(минеральная вода)	0.003066	0.766667
13	(суп, замороженные овощи, молоко)	(минеральная вода)	0.003066	0.766667
3	(шоколад, оливковое масло, замороженные овощи)	(минеральная вода)	0.002800	0.700000
4	(яйца, оливковое масло, молоко)	(минеральная вода)	0.002666	0.714286
5	(макароны, обезжиренное молоко)	(минеральная вода)	0.002533	0.730769
6	(креветки, говяжий фарш, замороженные овощи)	(макароны)	0.002533	0.791667
14	(креветки, говяжий фарш, замороженные овощи)	(макароны)	0.002533	0.791667
7	(макароны, суп, блинчики)	(минеральная вода)	0.002266	0.772727
8	(макароны, растительное масло, блинчики)	(минеральная вода)	0.002266	0.739130
15	(макароны, суп, блинчики)	(минеральная вода)	0.002266	0.772727
16	(рис, помидоры, молоко)	(макароны)	0.002133	0.800000
19	(рис, говяжий фарш, блинчики)	(минеральная вода)	0.002133	0.842105
18	(помидоры, оливковое масло, замороженные овощи)	(макароны)	0.002133	0.842105
17	(рис, говяжий фарш, блинчики)	(минеральная вода)	0.002133	0.842105
10	(рис, говяжий фарш, блинчики)	(минеральная вода)	0.002133	0.842105
12	(индейка, красное вино)	(минеральная вода)	0.002133	0.727273
11	(помидоры, оливковое масло, замороженные овощи)	(макароны)	0.002133	0.842105
9	(рис, помидоры, молоко)	(макароны)	0.002133	0.800000
20	(помидоры, оливковое масло, замороженные овощи)	(макароны)	0.002133	0.842105

Рисунок 3 – Список правил не более чем с семью объектами

В списке правил можно наблюдать самые частые зависимости между продуктами. Самыми популярными консеквентами являются макароны и минеральная вода. Это означает, что они очень часто встречаются в наборах с другими различными продуктами. Самая большая достоверность равняется 84%, это означает, что, например, вероятность того, что покупатель возьмет минеральную воду вдобавок к рису, говяжьему фаршу и блинчикам равна 84%. Самым популярным набором является грибной соус и эскалоп, к которым вдобавок берут макароны (наивысший уровень поддержки).

Сильная взаимосвязь при покупке различных продуктов дает сделать вывод о том, что можно продавать эти продукты рядом друг с другом, что увеличит вероятность того, что эти товары будут куплены, а значит увеличится прибыль.