

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-403
О.С. Мазжухин

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:

- сравнение быстродействия поиска правил на фиксированном наборе данных при изменяемом пороге достоверности;
- 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. Проанализируйте и изложите содержательный смысл полученного результата.

Гиперссылка на каталог репозитория с исходными текстами, наборами данных и другими материалами:

<https://github.com/LN4rkot1k/informationProcessing>

Визуализация

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

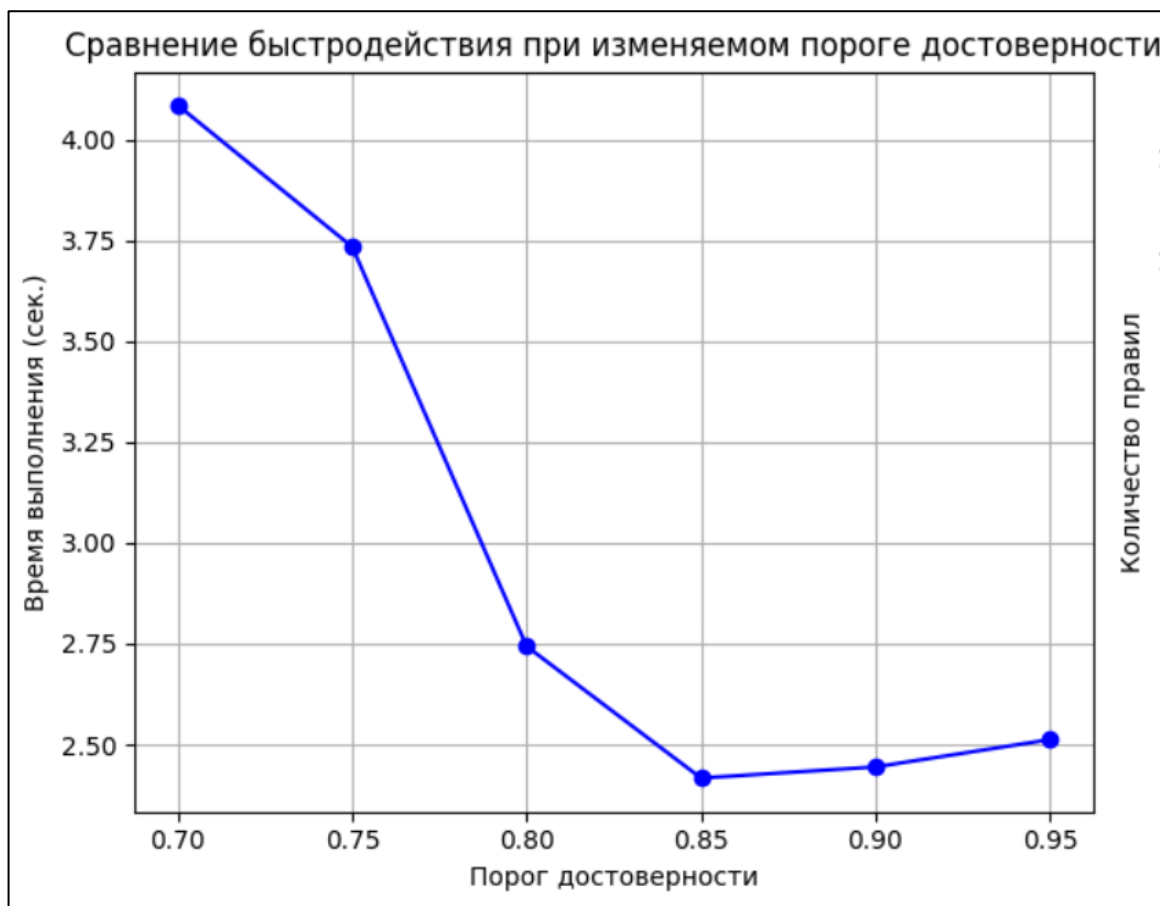


Рисунок 1 – График быстродействия при изменяемом пороге достоверности

На основе первого графика можно сделать вывод: чем выше порог достоверности, тем быстрее работает алгоритм. При низких значениях порога алгоритм обрабатывает больше возможных правил, что замедляет его выполнение. При увеличении порога достоверности количество правил сокращается, соответственно, вычисления происходят быстрее.

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

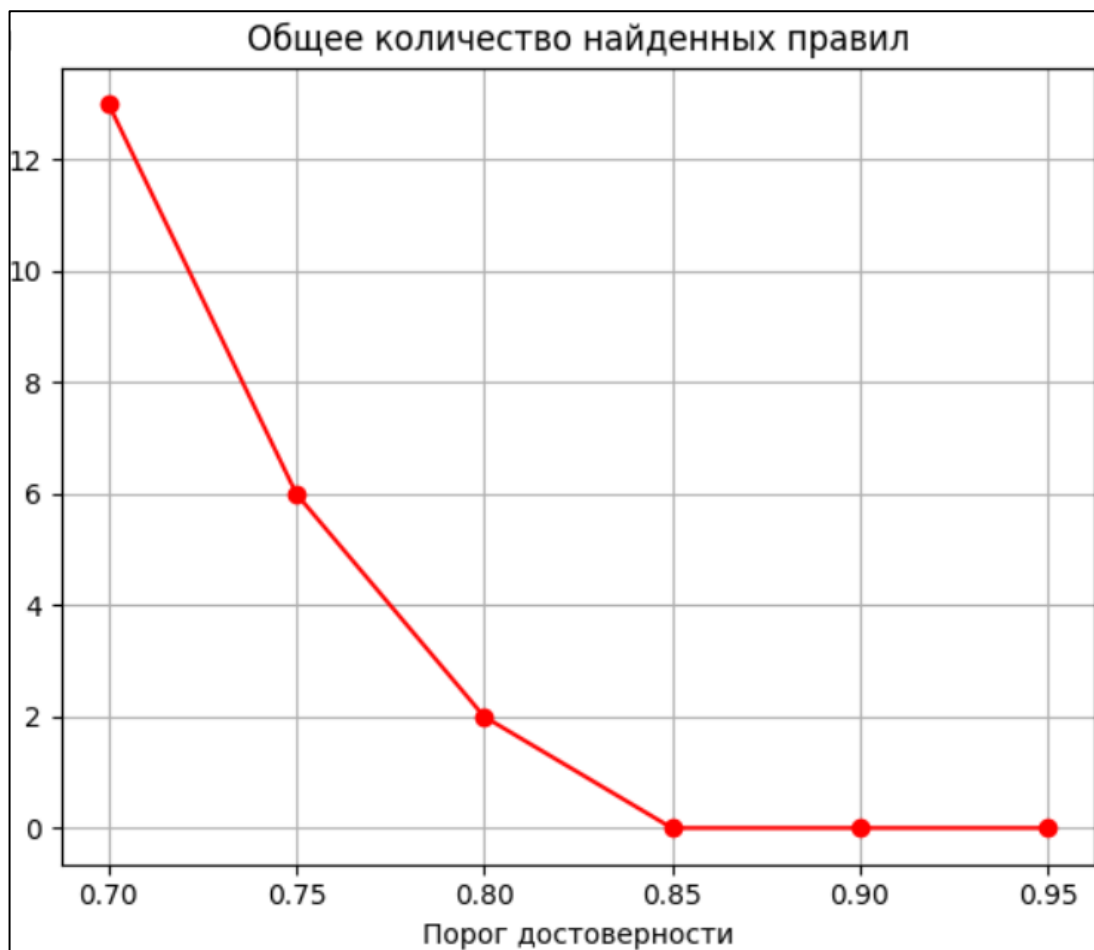


Рисунок 2 – Общее количество найденных правил

На основе второго графика можно сделать вывод о том, что при низком пороге достоверности количество найденных правил максимальное. По мере увеличения порога достоверности количество правил уменьшается, потому что более строгие требования к достоверности исключают многие правила, которые не достигают заданного уровня достоверности.

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

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

Рисунок 3 – список правил с не более 7 объектами

Глядя на датасет можно сделать вывод о том, что чаще всего встречаются правила с консеквентом минеральная вода или макароны. Это означает, что эти продукты часто берут в сочетании с другими продуктами. Глядя на достоверность, самыми надежными правилами являются правила с достоверностью 84%.

Самым распространенным правилом является эскалоп, грибной соус - макароны. Это правило имеет наибольшую поддержку.