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**

о выполнении практического задания № 2 по дисциплине

«Технологии аналитической обработки информации»

Выполнил:

студент группы КЭ-402 Риконина В.В.

Проверил: Юртин А. А.

Челябинск – 2025

# ЗАДАНИЕ

Требуется доработать 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). Также, требуется провести эксперименты на наборах из задания 1. В экспериментах зафиксировать значение пороговое значение поддержки (например, 10%), варьируйте пороговое значение достоверности (например, от 70% до 95% с шагом 5%).

# ВЫПОЛНЕНИЕ

Исходный код программы представлен в репозитории: https://github.com/veronirik/Analytical-information-processing-technologies

Результаты выполнения представлены в виде графиков.



Рисунок 1 – График зависимости времени выполнения от значения min\_confidence

По данному графику можно заметить, что чем больше значение min\_confidence, тем меньше времени затрачивается на работу программы. Это происходит из-за того, что при большем min\_confidence находится меньше частых наборов, вследствие чего алгоритм выполняет меньше операций, а также затрачивается меньше времени на постобработку для нахождения ассоциативных правил. Однако, разница во времени является незначительной, поскольку датасет является небольшим и зависимость от изменения min\_confidence при константном значении min\_support заметить сложно.

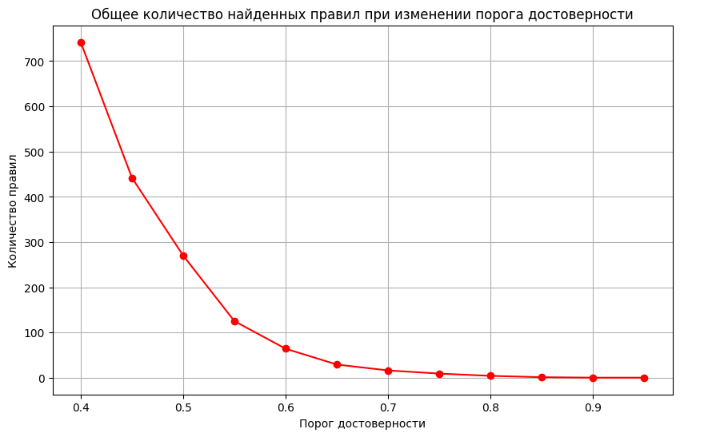


Рисунок 2 – График зависимости общего количества найденных правил от значения min\_confidence

По данному графику можно заметить, что чем больше min\_confidence, тем меньше становится различных частых наборов разной длины. Это связано с тем, что при увеличении min\_confidence алгоритм apriori начинает искать более связанные наборы продуктов, отсеивая редко встречающеся, вследствие чего

алгоритм по составлению ассоциативных правил может выявить меньшее их количество.

Это происходит из-за того, что параметр min\_confidence является вероятностью того, что если купили товар A, то купят и товар B. Следовательно, при большем значении min\_confidence, будет обнаружено меньше правил, но они будут более точными. Однако, данный параметр стоит выставлять в зависимости от объема датасета, поскольку при малом объеме данных и большом значении min\_confidence, скорее всего, будет обнаружено очень мало правил, или не будет обнаружено вовсе. Все это можно заметить на полученных выше графиках.