CPSC 473
FINAL PROJECT

2018-10-29 Riley Wium, 230090227

For my term project I intend to implement the DHP algorithm in java, with the ability to change the hashing function and see how it effects run-time (basically the example from class). The program will take in a database much like the first apriori program, and will output a similar result but the program will also include a tool for analyzing the output files. Specifically the output will be a txt or csv file containing the frequent patterns with their supports, the hash function used, and the time taken for each k.

A graphical interface will appear when you start the program from here you should be able to import the database file you want, then a field for selecting or inputting hash functions, then will be an option for naming your output file. I have already picked out two papers for how I should implement utilizing the different hash functions, but if that fails for some reason there will at least be a field where you can input the amount of buckets you want the hash function to have. The second part of this program takes one or more output files from the first part, the user can then select what type of chart they would like to generate and what the parameters should be (there will at least be time taken vs. number of buckets).

If I have extra time to work on this project I will decide if I should make, a way to more accurately modify hash functions, a small paper showing some of my findings and comparisons to some of my research, or add in memory usage into my data. One of the papers I am currently looking at discusses a modified version of DHP so if I find it interesting enough there is a small chance I could allow for the choice between classic DHP and this new version. I plan on doing this project solo.