**Cloud Computing and Big Data Analytics**

**HW2 report**

0760222黃昱銘

1. Brief introduction to implemented algorithm

I used ApriorAll which is mentioned in Leture4. Pseudo code as follow:

# init large 1-sequences

L[1] = init\_L1\_seq(db, support)

# generate all candidates for next level sequences

C = aprioriALL\_generate(sorted(L[1]), sorted(L[1]))

k = 2

# stop when no candidates for next level sequences

while C :

# prune the candidates that is lower than the minimum support

L[k] = pruneSupport(db, C, support)

C = aprioriALL\_generate(sorted(L[1]),L[k])

k = k + 1

1. Reference
2. ApriorAll

<https://blog.csdn.net/WeeYang/article/details/52793864>

1. Use tuple as the key of dictionary for python <https://www.developer.com/lang/other/article.php/630941/Learn-to-Program-using-Python-Using-Tuples-as-Keys.htm>
2. My suggestion to homework

I thought we should have something like autograder for private test cases, that is to say, we can upload our project to some online server and get the score back immediately. By doing so, we might detect some edge bug not in the public test cases. Although I have the same output for public test cases, I am still not sure if I could get full grade at private test cases without any retry opportunity.

1. Github link

<https://github.com/jasonyl13579/cloud_computing/tree/hw2>