CSE601 FALL 2017

PROJECT 1 - PART 2

ASSOCIATION ANALYSIS

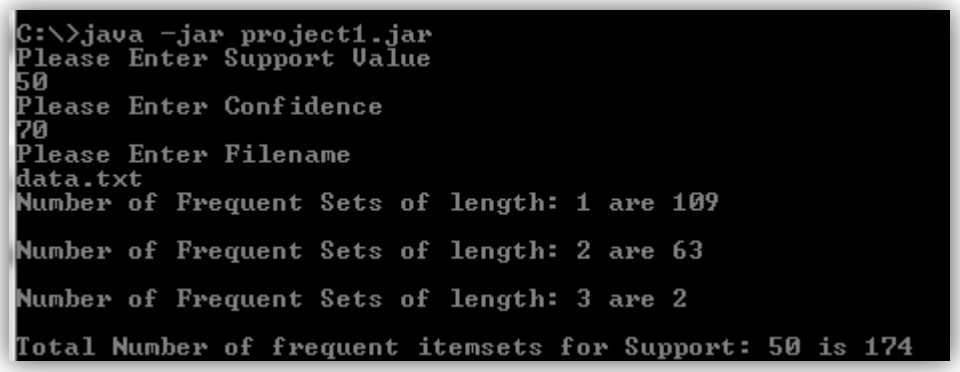
README

The folder 'Code' contains a JAR file of the compiled code 'project1.jar' , a data file 'data.txt', and another folder 'Source Code' which contains the source code files.

TO RUN: Open command prompt inside the 'Code' folder and run the following command:

java -jar project1.jar

- The prompt will ask you to enter support and confidence values. Please enter an integer value.
- For example if you want to run with minSupport as 50%, enter 50. If minSupport is 70%, then enter 70. Same goes for minConfidence.
- The prompt will then ask you for the file containing the data. Type in 'data.txt', and press enter.
- You can now see the result number of frequent sets generated for each length on the screen, and the result will also be output to a text file 'FreqSets_suppX_confY.txt' in the same folder as the JAR file.



You will now be prompted to select further options;

1. Enter D to display all the rules generated, and also output them to a file 'rules.txt'
2. Enter S to display result of the given sample queries, and output them to file 'SampleQueryOutput.txt'
3. Enter Q to enter a new query and then press enter after entering the query to display it's result and also output the result to a file 'QueryOutputTimestamp.txt'. The format of templates for entering queries is given below.
4. Enter X to exit the program.

Template Query Format for new queries:

Template 1:

- To enter a Template 1 query, The query should be of the form RULE|BODY|HEAD HAS ANY|1|NONE OF (ITEM1,ITEM2,ITEM3)
- For example, you want to query for Rules having 1 of G59_Up, the query should be "**RULE HAS 1 OF (G59_Up)"**

```
RULE HAS 1 OF (G59_Up)
Rules for Query: RULE HAS 1 OF (G59_Up)
Rule 1 :[G59_Up] ==> [G88_Down]
Rule 2 :[G59_Up] ==> [G82_Down]
Rule 3 :[G87_Up] ==> [G59_Up]
Rule 4 :[G72_Up] ==> [G59_Up, G82_Down]
Rule 5 :[G28_Down] ==> [G59_Up]
Rule 6 :[G10_Down] ==> [G59_Up]
Rule 7 :[G82_Down] ==> [G59_Up, G72_Up]
Rule 8 :[G13_Down] ==> [G59_Up]
Rule 9 :[G59_Up] ==> [G13_Down]
Rule 10 :[G72_Up, G82_Down] ==> [G59_Up]
Rule 11 :[G96_Down] ==> [G59_Up, G72_Up]
Rule 12 :[G72_Up] ==> [G59_Up]
Rule 13 :[G82_Down] ==> [G59_Up]
Rule 14 :[G59_Up, G82_Down] ==> [G72_Up]
Rule 15 :[G59_Up, G72_Up] ==> [G82_Down]
Rule 16 :[G88_Down] ==> [G59_Up]
Rule 17 :[G59_Up, G72_Up] ==> [G96_Down]
Rule 18 :[G6_Up] ==> [G59_Up]
Rule 19 :[G59_Up, G96_Down] ==> [G72_Up]
Rule 20 :[G59_Up] ==> [G72_Up]
Rule 21 :[G32_Down] ==> [G59_Up]
Rule 22 :[G59_Up] ==> [G96_Down]
Rule 23 :[G38_Down] ==> [G59_Up]
Rule 24 :[G72_Up, G96_Down] ==> [G59_Up]
Rule 25 :[G96_Down] ==> [G59_Up]
Rule 26 :[G1_Up] ==> [G59_Up]
Total Count of Rules := 26
```

If query is Body has ANY of (G59_Up, G10_Down), the query to enter will be "**BODY HAS ANY OF (G59_Up,G10_Down)**"

Please keep in mind that all the words in the query will be in CAPITAL, except for the attributes G59_Down.
The attributes should be entered as GColNumber Up|Down. For multiple attributes, you need to separate them by a comma.

Template 2:

- To enter a Template 2 query, the query should be of the form **SizeOf(BODY|HEAD|RULE) >= NUMBER**.
- For example, you want to query for rules having size of BODY greater than or equal to 2, then the query would be "**SizeOf(BODY) >= 2**"

- The keyword SizeOf is important to be entered as is.

Template 3:

- You can combine 2 queries of templates 1 and 2, in template 3 using the keywords AND and OR.
- Suppose you want to combine 2 queries of template 1 and 2, by the OR keyword, then the query would be
  BODY HAS ANY OF (G10_Down) OR SizeOf(HEAD) >= 2

  If you want to combine with AND, the query would be
  BODY HAS ANY OF (G10_Down) AND SizeOf(HEAD) >= 2

- Likewise, you can combine queries of similar templates as well, for example combining 2 queries of template 1:
  BODY HAS ANY OF (G10_Down) AND HEAD HAS 1 OF (G59_Up)
  and 2 queries of template 2 combined
  SizeOf(BODY) >= 1 OR SizeOf(HEAD) >= 2