MP10

Added by Javari, Amin, last edited by Javari, Amin on Jul 14, 2016

Machine Problem 10 (***Optional*** - *10* points)

Purpose: Learn to extract frequent sub-graphs from a read dataset.

Instructions:

- You are required to use R in this MP. R is the leading programming language for data analytics and has many cutting-edge packages for statistics and data-mining. It is platform independent and can be used on any platform. Also, it is free and open-source and has a growing community of users. R can be downloaded from here.
- You are required to use "subgraphMining" package in this MP. The package can be downloaded from here.
- Once you downloaded the package, install it using the following command: install.packages("C:/Desktop/subgraphMining_1.0.tar.gz ", repos =NULL, type = "source ");

Requirements:

- For your answers in the Answer Document, you should include 1) the outputs 2) a brief explanation about the outputs; 3) answers to all questions.
- Put all your codes in a separate folder with the name NetId_MP11_codes. Do not use sub-folders inside this folder. All of your codes should have been successfully compiled before submission. Do not include files other than the codes you write. Put a single readme.txt file in the code folder to briefly describe the functionalities of your codes and how to run them
- Your **PDF** submission file should be at the same level as your code folder. Compress these two together into a zip file, and name it MP10.netid.zip. Submit this zip file through Compass2g.

NOTE: This is an optional MP. You should submit your MP10.netid.zip file to MP10 submission link on Compass2G. The deadline for this MP is (8/4/16).

Input:

You will use *metabolicInteractions* dataset from *subgarphMining* package. The dataset can be loaded using the following command: graph = data(metabolicInteractions).

Required Outputs:

(a) [5, L2] Extract the frequent patterns from *metabolicInteractions* dataset by employing *gspan* function (which is available in subgraphMining package). Set the minimum support as 60%, 70%, 80% and 90% in your experiments and report the outputs.

(b) [5, L3] *SubgraphMining* package has two other main functions: *SUBDUE*, *SLEUTH*. Briefly explain these functions and apply them to the input data. Report your outputs and observations. More information about the package can be found here.

No labels