CSCI-B565: Data Mining

Homework Assignment # 6

Assigned: 04/15/2016 *Due: 04/29/2016, 11:59pm, through Oncourse*

One project report, 50 points in total. Good luck!
Prof. Predrag Radivojac, Indiana University, Bloomington

Project report. (50 points) Submit your project report. The format of the project proposal is listed below.

Instructions for writing a report

1. Project title.
2. Names of the team members and their Indiana University emails.
   (a) If the project is under the guidance of a particular Indiana University faculty member, list the professor (or professors) working with you.
3. Objectives and significance (1-3 paragraphs)
   (a) Describe what the goal of the project is, why is it important, and your motivation for doing it.
4. Background (1-2 pages)
   (a) Introduce all important concepts and background information.
   (b) Describe previous work on this problem.
   (c) Describe what makes your work particularly interesting.
5. Methods (2-4 pages)
   (a) Describe your data and how you obtained it.
   (b) Describe your methodology: give flowcharts, diagrams, pseudocode or formulas where appropriate.
   (c) Describe evaluation strategy.
6. Results (5-6 pages)
   (a) Present here all your results and findings
   (b) Describe experiments done for each figure and table.
   (c) Provide headers for all your tables and make sure it is easy to understand what your results are.
7. Conclusions (1 page)
   (a) Summarize what the major conclusions are. If your ideas worked, try to provide rationalizations as to why. If not, discuss what may have gone wrong. In either case discuss what could be done in the future to improve or further improve this project.
8. Individual tasks (1-3 paragraphs)

   (a) This part of the writeup is individual and team members need to write their own summary.
   (b) Describe what each member of the team did. Be specific about the roles of everyone.
   (c) If the project was a part or extension of some other project you have carried out, explain what was new in this project.

9. References
   (a) List books, scientific papers, web sites etc. that you referenced in the proposal body.

Technical details:

1. Use 1 inch margins from each side.
2. Use 11pt or 12pt font size.
3. Use standard font types such as Times New Roman, Arial, or Latex default fonts.
4. Use 1.5 line spacing that will allow for our comments during grading.

Homework Directions and Policies

Submit a single package containing all answers, results and code. You submission package should be compressed and have extension .zip. In your package there should be a single pdf file named main.pdf that will contain answers to all questions, all figures, and all relevant results. Your solutions and answers must be typed[1] and make sure that you type your name and IU username (email) at the beginning of the file. The rest of the package should contain all code that you used. The code should be properly organized in folders and subfolders, one for each question or problem. All code, if applicable, should be turned in when you submit your assignment as it may be necessary to demo your programs to the associate instructors. Use Matlab, Python, R, Java, or C/C++.

Unless there are legitimate circumstances, late assignments will be accepted up to 5 days after the due date and graded using the following rules:

on time: your score $\times$ 1

1 day late: your score $\times$ 0.9

2 days late: your score $\times$ 0.7

3 days late: your score $\times$ 0.5

4 days late: your score $\times$ 0.3

5 days late: your score $\times$ 0.1

For example, this means that if you submit 3 days late and get 80 points for your answers, your total number of points will be $80 \times 0.5 = 40$ points.

All assignments are individual, except when collaboration is explicitly allowed. All the sources used for problem solution must be acknowledged; e.g., web sites, books, research papers, personal communication with people, etc. Academic honesty is taken seriously! For detailed information see Indiana University Code of Student Rights, Responsibilities, and Conduct.

[1] We recommend Latex; in particular, TexShop-MacTeX combination for a Mac and TeXnicCenter-MiKTex combination on Windows. An easy way to start with Latex is to use the freely available Lyx. You can also use Microsoft Word or other programs that can display formulas professionally.