Project Proposal Assignment – Behavioral Data Mining – Sp2013

Please prepare an overview of your planned project in PowerPoint or some other presentation format. Plan for a 5 minute presentation on (3/18/13 or 3/20/13). Your slides should cover:

- Background on the problem you are addressing.

- What question(s) you want to answer. This can be framed as hypotheses, or as questions to answer "as accurately as possible." Make sure the articulation of the question doesn't assume a particular solution or algorithm.

- Given the question(s) above, discuss the kind of analysis you plan to do. i.e. what models and analysis methods you will use. Try to make a clear distinction between the two.

- What dataset will you be using? How much data do you think you will need to answer your questions?

- What kinds of tools do you plan to use? Explain what you want to measure and how you will describe the results (what kinds of visualizations/representations of data you will do).

- In case you are interested in methods (i.e. in improving algorithms or inference methods, or mining software) you still need to identify one or more datasets that will be suitable for measuring the effectiveness of you methods.

Send an email asap to hzhao@cs.berkeley.edu with your team name, participants, project title and preferred presentation day.

Datasets

There are several datasets that can be used in projects that were described in class. Several of these involve a usage agreement, so please let us know asap so we can execute that. If you need help with any other dataset not discussed, please let us know.

Note: you do *not* have to use any particular system for the project. But you should make sure you have enough capacity (i.e. speed and cycles on some machine(s)) to complete your analysis.

Final Project Proposal

By Wednesday, 3/20/13, you should create a project page on the course wiki (look for "Projects" on the left menu). Please place a finalized version of you proposal (the powerpoint or a pdf version of it is fine) there.