ASSIGNMENT 1
ADVANCED METHODS AND ANALYSIS FOR THE LEARNING AND SOCIAL SCIENCES
PROFESSOR RYAN S.J.d. BAKER
PFA AND BKT
DUE NOON, MONDAY JANUARY 30

The goal of this assignment is to build valid PFA (Performance Factors Analysis) and BKT (Bayesian Knowledge Tracing) models for the data set in Asgn1-dataset.csv

This data set involves real data used in several papers. To find out more about this data set, look at

Baker, R.S.J.d., Corbett, A.T., Roll, I., Koedinger, K.R. (2008) Developing a Generalizable Detector of When Students Game the System. *User Modeling and User-Adapted Interaction*, 18, 3, 287-314. http://users.wpi.edu/~rsbaker/USER475.pdf

You must build at least one PFA model variant and at least one BKT model variant for this data set. You may choose any variant of these models, so long as they are recognizably and validly PFA and BKT. Your model must conform to all assumptions of these approaches (warning: make sure to sort your data correctly, and remove data points that cannot be used in these approaches!)

You will have to enter the mathematical functions for PFA yourself, and can use any published variant of these mathematical formulas which Pavlik refers to as "PFA". You can use any mathematical package to do so.

You can enter the mathematical functions for BKT yourself, or you can use an existing package (either Beck et al.'s BKT-SM or Baker et al.'s Brute Force code are acceptable). You can use any variant of BKT which is claimed to be a variant of BKT by Albert Corbett, Phil Pavlik, Joseph Beck, Ryan Baker, Michael Sao Pedro, or Zachary Pardos, in a published paper.

Your hand-in must include a list of all model parameters, and the final model's predictions for the probability that each student first attempt on a problem step is correct. You must also turn in any code or Excel files (Matlab files, Maple files, etc. etc.) used in your computations. In the case of BKT-SM or Baker et al.'s Brute Force code, it is acceptable to simply state what package you used. You may also (optionally) turn in a document explaining any part of your assignment that you wish to explain. You will be graded on completeness and comprehensibility of your hand-in, and whether you correctly and validly apply the PFA and BKT models to this data.

BONUS: The student who has the best overall model (either PFA or BKT) gets the bonus. "Best" is defined as the highest value for the A' metric, computed without considering the student. Cross-validation does not need to be used for this assignment.