

Master of Technology in Knowledge Engineering

Unit 7:

Developing Intelligent Systems for Performing Business Analytics

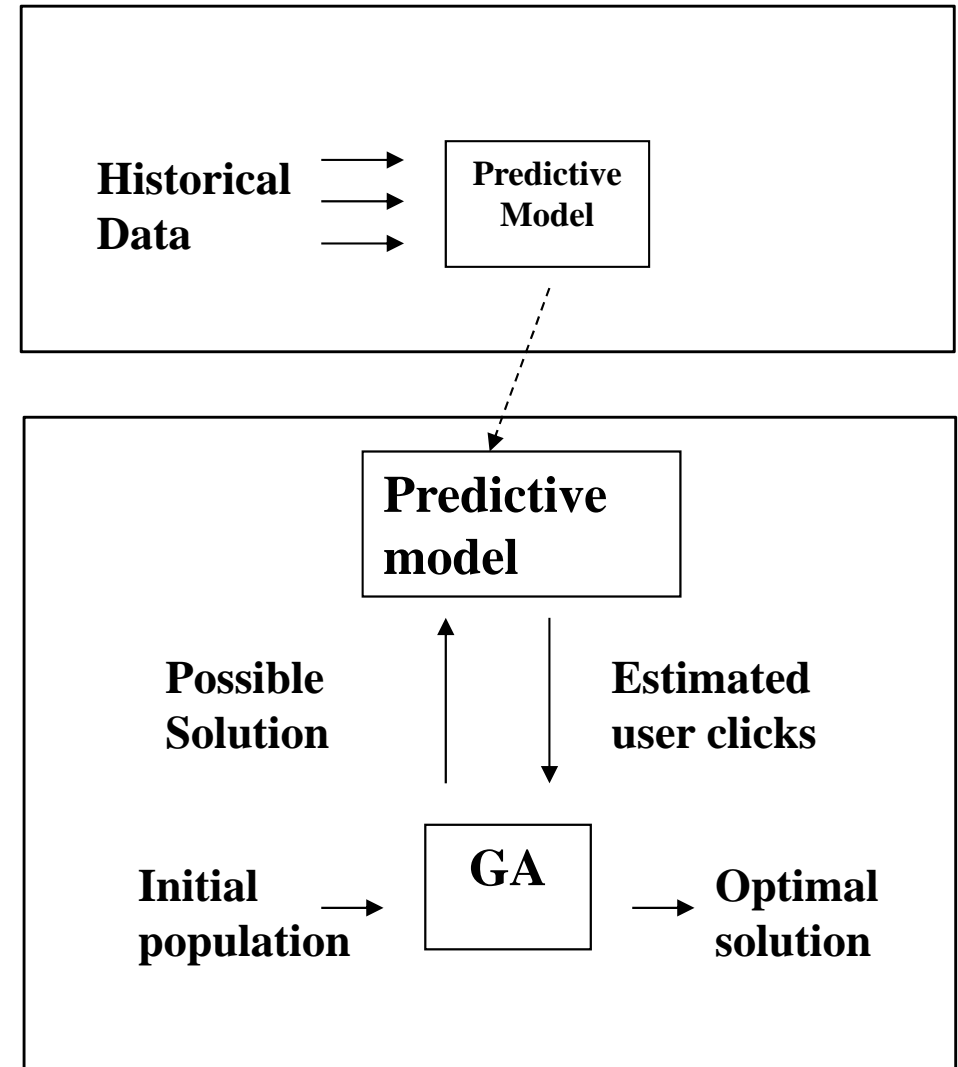
Optimisation Workshop Hints

Dr. Zhu Fangming
Institute of Systems Science,
National University of Singapore.
E-mail: isszfm@nus.edu.sg

© 2018 NUS. The contents contained in this document may not be reproduced in any form or by any means, without the written permission of ISS, NUS other than for the purpose for which it has been supplied.

Workshop 1A: One Possible Solution Architecture

- A predictive model is built using the historical records. This model is then used as a fitness function for a GA-based search for an optimal plan.
- You may adopt any other architecture you think appropriate. Your system should use at least two different KE techniques.



Suggested Steps

- **Step1: Build a predictive model to forecast the user clicks**
 - » Regression model
 - » Neural network
 - » Data pre-processing? How to evaluate the model?
- **Step2: Build a GA model to search for an optimal plan**
 - » Decide input variables
 - » Write a fitness function which makes use of the predictive model output
 - » How are the problem constraints to be handled?
 - » Experiment with learning parameters (population size, mutation rate, etc).
 - » Is there premature convergence?
 - » Run the system many times to obtain the best plan

Suggested Steps

- **Step3: Evaluate the plan using the “real” model**
 - » You must evaluate the advertising plan generated by your system to see if your system is performing well or badly. You can use the program **AdModel.exe**
 - » Is the result produced by your system different from the actual one obtained by inputting the plan into AdModel.exe? If yes, then how can your predictive or GA models be improved? Iterate from Step1 or Step2.

Software

- **Can undertake workshops using Excel Data Analysis + Excel Solver.**
- **Use Regression Analysis under Excel Data Analysis**
- **Use Excel Solver for GA**

Optimisation Workshop 1B

Part B: Utilising Background Knowledge

- While a full model is unavailable, experience suggests that
 - » the number of user clicks depends on the time of day in which the ads are displayed
 - » the day is divided into three time periods within which the number of user clicks per hour is stable
 - » the type of ads selected for each website can be modelled simply as a scale factor i.e.

User clicks achievable from ad_j for display duration D using website W =

$$(\text{dur-1} * \text{clk-1} + \text{dur-2} * \text{clk-2} + \text{dur-3} * \text{clk-3}) * \text{scalefactor}_j$$

Where $clk-i$ is the user clicks achievable per hour for the i -th time period ($i = 1$ to 3) and $dur-i$ is the time duration that ad_j ($j = 1$ to 6) is displayed on website W during the i -th time period.

Note $dur-1$, $dur-2$, $dur-3$ form a continuous time period and $dur-1 + dur-2 + dur-3 = D$

Part B: Utilising Background Knowledge

- How can your solution architecture be modified to make use of this background knowledge?
- Suggestions from the class!!!!

Suggested Steps for Workshop 1B

- Use Excel-Solver to build a model for the five websites
 - » **i.e. learn the three time-periods and their associated clicks per hour for each website and learn the scale factor for each ad.**
- Then use the built model instead of the regression model to optimise the advertising plan to meet the budget of \$300 (i.e. as in part-A).