Teach & Try: simple end-user data modelling

Advait Sarkar, Alan Blackwell, Mateja Jamnik, & Martin Spott

My research focuses on simple interfaces for advanced analytics. The goal is to offer increased analytical power to end-users without domain or computing expertise through novel interaction techniques.

Teach & Try is a spreadsheet-based interaction metaphor that can be used to build and apply sophisticated models such as neural networks, decision trees, support vector machines and linear regression.

We exploit the information in spreadsheet cell selection bounds to interpret parameters such as the training set, feature vector, and regression/prediction target. The result is a self-contained interface for using statistical models within a graphical spreadsheet environment — no programming needed.

Usability study

We had 13 participants with no prior experience in statistics or programming (including young children) each perform several tasks using the software. The participants were able to complete the tasks and showed a significant decrease in task completion time when they encountered fresh tasks, showing that the system can be easily learnt and successfully applied by such users.

We also concluded that the experience of interacting with the system leads them to acquire some understanding of the concepts underlying exploratory data analysis. For instance, when asked why the computer might make mistakes, the participants were able to identify several problems familiar to statisticians, including insufficient examples, noisy training data, incorrect model, and insufficient dimensions.

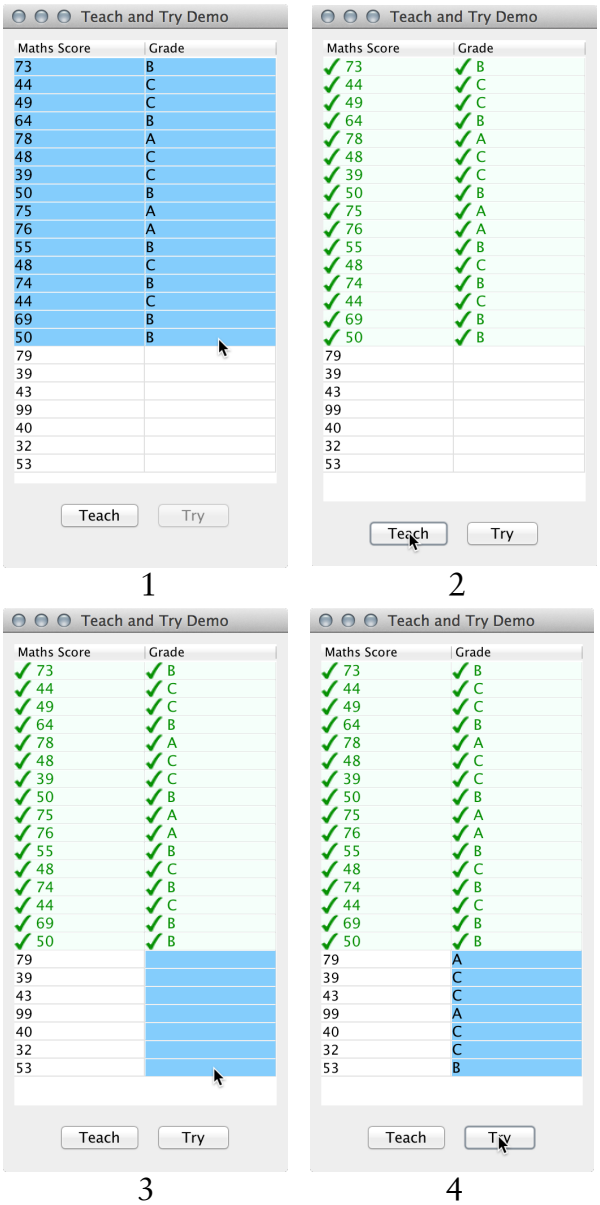

Fig.1: A 4-step sequence depicting how a maths teacher might use Teach & Try to automatically finish the task of assigning grades to students based on their scores on a recent exam. Teach & Try invisibly builds a decision tree classifier and uses it to automatically populate the empty cells.

SARKAR A., BLACKWELL A. F., JAMNIK M., SPOTT M.: Teach and try: A simple interaction technique for exploratory data modelling by end users. In *Visual Languages and Human-Centric Computing (VL/HCC), 2014 IEEE Symposium on* (July 2014), IEEE, pp. 53–56.

RAINBOW RESEARCH GROUP

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

advait.sarkar@cl.cam.ac.uk
http://www.cl.cam.ac.uk/~as2006
William Gates Building
15 JJ Thomson Avenue
Cambridge CB3 0FD