CAN STUDENT PROFILES DETERMINE THEIR SUCCESS UNDER DIFFERENT TEACHING METHODOLOGIES?

David P. Kopcso
Babson College
Babson Hall
Babson Park, MA 02457-0310
Kopcso@babson.edu

The initial task performed was the construction of a database containing information available when a student matriculates as well as course grades obtained while the student was at our institution. The database is a consolidation of information about students admitted between spring 2001 and fall 2005 from two disparate sources, the Admission's Office and the Registrar's Office, each of which has undergone changes in their respective record storage software during the time period in question. The Admission's data consists of a unique identifier, Student ID number, the year and term of admission, the type of application such as early decision, the status of application as of the record's creation, Admission's evaluation score of the student's high school record and recommendations on a scale of 0 to 54, the student's SAT Math and Verbal scores, and Country of Citizenship. The Registrar's data consists of multiple files which include a file for our introductory quantitative methods course, QTM 1300, containing: Student ID number, admission date, course name, course section, academic year and term, instructor id, letter grade, days and times on which the class met, as well as GPA data for students admitted fall 1999 to fall 2004. Approximately 11% of the students in the database had been exposed to a series of parameterized electronic quizzes, heretofore referred to as EQ.

Issues arose due to Babson's requirement of conforming to Family Educational Rights and Privacy Act (FERPA), http://www.ed.gov/policy/gen/guid/fpco/ferpa/index.html, about confidentiality concerns. The "bad news" is that the ability to cross reference students' performances in the classroom associated with students' names with their institutional data as described above which is associated with Student ID number is not available. The "good news" is that this investigation has precipitated the investigation of the establishment of an Institutional Review Board at Babson and the creation of a document of Informed Consent. Although conformance with FERPA has precluded the use of such variable as gender, aggregated results (which do not identify individuals) from the investigation of the baseline model built on data mining the database are reported here.

As with most data mining investigations, the majority (90 – 95%) of a data mining project consists of data cleansing. Specific issues that were addressed included but were not limited to file structure incompatibilities, software incompatibilities, filed naming inconsistencies, field data type inconsistencies, and missing and illegal data values. Time sequencing became a problem since not every student completes their college career in four years.

Other anomalies in the database that needed to be addressed before the data mining software could be utilized included:

- Some students receive advanced standing for one or more of the mathematics courses and are "waived" from that (those) course(s). Their grades are recorded as CR, but they are in a class "section" of size 1.
- Some students receive an incomplete grade for one or more of the mathematics courses. Their grades are recorded as I.
- Some data mining routines require numerical inputs while others require categorical inputs.
- Some other routines automatically transform continuous numerical inputs into discrete categories that are not readily interpretable.
- Some binary inputs (flagged as T or F) need to be redundantly entered as Indicators (0 or 1) or derived within the software.
- Some routines are influenced by the magnitude of the numerical inputs requiring continuous variable to be standardized before they can be admitted to the routine.

After the above issues were resolved, the model building investigation was undertaken utilizing the data from spring 2001 to spring 2005. Initial numerical and graphical examination of the successful completion of QTM 1300 (a grade of C or better) for those students exposed to EQ in QTM 1300 and those not exposed to EQ resulted in the following observations:

- Although students who were successful in QTM 1300 had about the same average SAT Math scores whether they were exposed to EQ or not, those who were unsuccessful and had been exposed to EQ had considerable higher average SAT Math scores.
- Similar, but not as dramatic, differences were seen with the average SAT Verbal scores.
- The average Admission's evaluation scores as well as their distribution were lower for those unsuccessful in QTM 1300 than for those who were successful although each group had similar average values whether exposed to EQ or not.
- Both the SAT Math and SAT Verbal scores for those exposed to EQ were distributed similarly to those not exposed to EQ although with a smaller variability.
- Students with lower SAT Math scores had a higher percentage of unsuccessful experiences in QTM 1300, while the same was not true of students with lower SAT Verbal scores.
- Students who took QTM 1300 three times per week (Monday, Wednesday, and Friday schedule) or were spring admits had lower SAT Math, SAT Verbal, and Admissions evaluation scores.
- International students had a similar distribution of SAT Math scores as those from the United States, but lower SAT Verbal scores.
- The Admission's evaluation scores for those exposed to EQ were

distributed similarly to those not exposed to EQ although with a smaller variability.

- There were no discernable patterns in the QTM 1300 letter grade verses SAT Math or SAT Verbal or Admissions evaluation scores whether the students were exposed to EQ or not.
- Only students who took QTM 1300 two day a week were exposed to EQ and although approximately the same number were exposed to EQ on Monday and Wednesday as were exposed to EQ on Tuesday and Thursday, a greater percentage of those who took QTM 1300 on Monday and Wednesday received EQ.
- Only students who took QTM 1300 in the first two class periods of the day or the first period after lunch were exposed to EQ with the highest percentage receiving EQ being those who had a first period class which starts at 8:00 A.M.
- The grade distributions for those exposed to EQ were similar to those for students who were not exposed to EQ.

The C5.0, CHAID, QUEST and CaRT decision tree algorithms (please see appendix for algorithm details) were able to identify rules which indicate success in QTM 1300 correctly 83%, 80%, 81%, and 82% of the time respectively. Much of the information was redundant in these four algorithms which agreed correctly 87% of the time.

The Apriori algorithm which uncovers co-occurrences was used to identify potential explanatory variables to use to predict success in QTM 1300. Traditional regression models as well as logistic regression models were developed to

- Predict QTM 1300 grade based on Admission Data
- Predict fist year GPA grade based on Admission Data & QTM 1300
- Predict fist year GPA grade excluding of QTM 1300 based on Admission Data & QTM 1300
- Predict success or not in QTM 1300 based on Admission Data & EQ
- Predict Assignment To or Not in EQ

While the results of theses investigations were not as robust as was hope for, each is addressed below:

Predict QTM 1300 grade based on Admission Data

Although the Admissions data of SAT Math, SAT Verbal, and Admissions evaluation score were significant predictors of QTM 1300 grade on a twelve point scale, the percentage of variation in the QTM 1300 grade explained by the regression was 14.8% with a standard error of the estimate of 2.64, much too large to be useful.

Predict fist year GPA grade based on Admission Data & QTM 1300

SAT Math score, SAT Verbal score, Admissions evaluation score, the QTM 1300 grade on a twelve point scale, and whether or not the student was exposed to EQ as well as an interaction term between EQ exposure and SAT Math score were significant predictors of the students first-year GPA. The percentage of variation in the GPA explained by the regression was 52.4% with a standard error of the estimate of 0.3737. This model seems promising; however, the QTM 1300 grade itself is included in the GPA value being predicted!

Predict fist year GPA grade excluding of QTM 1300 based on Admission Data & QTM 1300
When the QTM 1300 grade is removed from the first-year GPA grade the explanatory variables are still significant, but the percentage of variation in the GPA explained by the regression falls to 41.3% and the standard error of the estimate increases to 0.4124. However, this model suggests an increase of 1.137 points in GPA for those exposed to EQ in QTM 1300 over those not exposed holding SAT Math score, SAT Verbal score, Admissions evaluation score, the QTM 1300 grade on a twelve point scale constant. In other words for two students with the same SAT Math score, SAT Verbal score, Admissions evaluation score, and the QTM 1300 grade on a twelve point scale the same, the student who is exposed to EQ would be predicted to have a 1.137 point higher first-year GPA then the student not exposed to EQ.

Predict success or not in QTM 1300 based on Admission Data & EQ
The logistic regression model to predict whether or not a student would be successful in QTM 1300 has SAT Math score, SAT Verbal score, Admissions evaluation score as significant explanatory variables but exposure to EQ was not found to be significant. This model correctly predicts success in QTM 1300 79% of the time.

Predict Assignment To or Not in EQ
The logistic regression model to predict whether or not a student would be successful in QTM 1300 if she/he were exposed to EQ had no significant results.

The final data mining investigation tool used was neural networks. The network learns by examining individual records, generating a prediction for each record, and making adjustments to its "weights" whenever it makes an incorrect prediction. This process is repeated many times, and the network continues to improve its predictions until one or more of the stopping criteria have been met. Caution needs to be exercised in utilizing neural networks because of their potential for over fitting and their inability to explain or have interpreted their reason for the associations they suggest.
Neural network models were developed to:

Predict success or not in QTM 1300 based on Admission Data
This model correctly predicts success in QTM 1300 81% of the time.
Predict success or not in QTM 1300 based on Admission Data & EQ
This model also correctly predicts success in QTM 1300 81% of the time. The inclusion of exposure or not to EQ does not increase the predictive power of the neural network.
Predict Assignment To or Not in EQ based on Admission Data

This model correctly predicts those who will be successful in QTM 1300 if exposed to EQ 89% of the time.

Conclusions and future work:

The most interesting and promising results are two-fold:

The regression model which indicates that exposure to EQ results in a 1.137 increase in GPA as measured on a four point scale and the neural network that has an 80% correct ability to predict if a student exposed to EQ will be successful in QTM 1300 based solely on Admissions data.

Optimism must be tempered by the fact that these results are based on the same data used to generate the models themselves and, consequently, may overstate the results. The next steps are to test them on the next year of data which were not used to build the models. Additionally, the results from student surveys will be tested to see if they will improve the predictive power of the models.

Appendix

Descriptions of the decision tree algorithms taken from SPSS Clementine data mining software documentation follow:

> "C5.0. This method splits the sample based on the field that provides the maximum information gain at each level to produce either a decision tree or a ruleset. The target field must be categorical. Multiple splits into more than two subgroups are allowed.
>
> CHAID. Chi-squared Automatic Interaction Detector uses chi-squared statistics to identify optimal splits (Kass, 1980). Exhaustive CHAID, a modification of CHAID that does a more thorough job of examining all possible splits for each predictor but takes longer to compute (Biggs et al., 1991), is also available. Target and predictor fields can be range or categorical; nodes can be split into two or more subgroups at each level.
>
> QUEST. The Quick, Unbiased, Efficient Statistical Tree method is quick to compute and avoids other methods' biases in favor of predictors with many categories (Loh and Shih, 1997). Predictor fields can be numeric ranges, but the target field must be categorical. All splits are binary.
>
> C&RT. The Classification and Regression Trees method is based on minimization of impurity measures (Breiman et al., 1984). A node is considered "pure" if 100% of cases in the node fall into a specific category of the target field. Target and predictor fields can be range or categorical; all splits are binary (only two subgroups)."