

**AC 2007-2977: USE OF A NEURAL NETWORK MODEL AND NONCOGNITIVE MEASURES TO PREDICT STUDENT MATRICULATION IN ENGINEERING**

**P.K. Imbrie, Purdue University**

# **Use of a Neural Network Model and Noncognitive Measures to Predict Student Matriculation in Engineering**

## **Abstract**

Engineering students' affective self-beliefs prior to their first year have the potential to help researchers better understand various issues related to student retention and engagement. This paper examines whether a neural network model based on student noncognitive characteristics can be used to predict student persistence in engineering, and the influence of gender in the predictive model. Eight noncognitive measures (i.e., academic self-efficacy, academic motivation, leadership, metacognition, career, type of learner (e.g., deep vs. surface), teamwork, and expectancy-value) serve as independent parameters to an artificial neural network (NN) that is used to predict student persistence within engineering school at the end of first year.

A feed-forward neural network model with back-propagation training was developed to predict third semester retention of a cohort of first-year engineering students ( $N=1,523$ ) at a large Midwestern university. The model constituted of 159 primary nodes corresponding to 8 noncognitive factors described by a 159 item instrument. The resulting model was shown to have a predicative accuracy of 82% for retained students after their first year and 30% for non-retained students. Significantly decreasing the number of inputs (i.e., only using those items that appeared to have the strongest influence) had little impact on the predicative accuracy of the retained students. However, the reduction in inputs decreased the predictive accuracy of the non-retained students by approximately 10%. Results for the same cohort also indicate that the neural network prediction rate is independent of gender.

## **Introduction**

Engineering programs typically attract the top graduates from high school in terms of grade point average (GPA) and standardized test scores, but attrition out of engineering continues to be a major issue; programs often see some of the most statistically qualified students leave engineering for other majors or drop out of college altogether. In 1975, attrition in the freshman year in engineering was about 12%, increasing to about 25% by 1990 (Beaufait, 1991). In a large study of over 25,000 students at over 300 universities, Astin (1993) found that only 47% of students who begin in engineering graduate with an engineering degree. The National Academies' report "Rising Above The Gathering Storm: Energizing and Employing America for a Brighter Economic Future" reports that undergraduate programs in science and engineering have the lowest retention rate among all academic disciplines. The National Academies describes the importance of advances in engineering and technology as crucial to the social and economic conditions for the United States to compete, prosper, and be secure in the global community in the 21st century (Augustine, 2005).

One common misconception is that students leave engineering due to lack of academic ability. Studies have shown little difference between the academic credentials of students who remain in engineering and those who leave (Besterfield-Sacre *et. al.*,1997, Seymour 1997). While there is a positive correlation between GPA and retention, GPA alone doesn't predict student attrition. Studies have shown that models incorporating cognitive variables such as student high school math and science success (Jagacinski, 1981), strong interest in science (Astin 1992), and higher confidence in basic engineering knowledge and skills (Besterfield-Sacre *et. al.*,1997) are able to establish a correlation between cognitive variables and retention, but these variables are clearly not single factors in a model to predict retention. Instead, a model using both cognitive and noncognitive – or affective – characteristics shows the greatest promise to accurately identify students who may leave engineering or who may benefit from interventions (Astin *et. al.*,1992, Felder 1993, Besterfield-Sacre *et. al.*, 1997).

In a 2002 study to investigate the predictive relationship between six variables (high school GPA, SAT math score, SAT verbal score, gender, ethnicity, citizenship status) and retention and graduation, Zhang (2002) found that high school GPA and SAT math scores were the best predictor of retention and graduation, while SAT verbal was inversely related. Gender, citizenship and ethnicity were sometimes found to be predictors, but this varied from campus to campus. Astin (*et. al.*, 1992) found that the student's record in high school was the best predictor of academic success, and performance on standardized tests also had a positive correlation. Zhang (*et. al.*, 2002) identified self-efficacy and physical fitness as positive predictors of freshman retention in a study of several cognitive and affective characteristics. These studies were valuable in identifying characteristics that were predictors for retention, but did not address multiple factors and their interaction as predictors.

Instruments designed to assess freshman success include the Pittsburgh Freshman Engineering Attitudes Survey (PFEAS), consisting of 50 items relating to 13 student attitude and self-assessment measures, used to measure differences in student attitudes before and after the freshman year (Besterfield-Sacre, *et. al.* 1997, Besterfield-Sacre, 1999) The Cooperative Institutional Research Program (CIRP) Freshman Survey covers a wide variety of attributes from financial considerations to attitudes toward school to high school academic performance. A study of academic background variables and the CIRP showed that academic background variables were predictors of future grade performance, but no correlation to retention was reported (House, 2000). These and similar studies indicate that student attitudes and other noncognitive characteristics must be incorporated into any model used to predict retention.

These studies and the existing issues with attrition within engineering lead to the question: if no one characteristic has been shown to sufficiently predict student attrition, can a model be developed to take multiple factors and their interaction into account?

Data on a set of eight noncognitive variables was collected and analyzed (Maller, 2005,

Immekus, 2005). A neural network model incorporating these noncognitive variables allows the investigation into not only the predictive nature of these characteristics, but the predictive possibilities of their interaction in attrition within engineering.

### **Data collection and Instrumentation**

The sample in this study included 1,523 incoming first-year engineering students (292 females, 1,231 males) at a large Midwestern university during the 2004-2005 academic year. Ethnicity was as follows: 2.05% African American, 0.51% American Native, 10.18% Asian/Pacific Islander, 2.64% Hispanic, 82.43% Caucasian, 2.20% Other.

The students' non-cognitive measures were collected across eight scales (completed prior to the freshman year): Leadership (20 items), Deep vs. Surface Learning (20 items), Teamwork (10 items), Self-efficacy (10 items), Motivation (25 items), Meta-cognition (20 items), Expectancy-value (26 items), and Career Indecision (28 items). All Cronbach's coefficient alphas for these eight scales were  $\geq .80$ , except for the Teamwork scale ( $r=.74$ ). Multiple studies have supported the scales' construct validity based on the results of confirmatory factor analyses (Immekus *et. al.*, 2005, Maller *et. al.*, 2005). Inside these major scales, there were subscales consisting with various numbers of items; for example, under the measure academic motivation, there are four subscales: control, curiosity, career and challenge.

Students' persistence statuses were collected at the beginning of every semester following their freshman year. The investigation in this study focuses on the persistence status at beginning of third semester right after the freshman year.

### **Research Methods**

#### *Artificial Neural Networks (ANNs):*

A typical neural network model is an information processing system consisting of inputs, interconnected neurons or nodes as processing units, and output layers. New neural networks must be trained with existing data so it can learn from the examples. During the training process, the weights associated with the links between neurons are adjusted by learning or adapting through the data in repetition. Properly trained neural networks have been widely used in various prediction applications in many areas of engineering. In this study, a feed-forward neural network with back-propagation training algorithm was used to develop models for predicting freshman engineering students' persistence in engineering. The activation function utilized in these models is log-sigmoid function (Demuth, 1998). All neural network models in this study were developed using Matlab version R2006b from Math Works Inc.

The input data used in the training and testing processes are the noncognitive survey items and gender information collected from 1523 freshman engineering students during 2004-2005, as described in previous section. The dependent variable, persistence, is defined in terms of a

student's enrollment status at the start of his/her the third semester.

*Classification for the status of students' persistence*

In this study, the status of students' persistence after their first year in engineering was classified into five possible categories as described in Table 1. Students who are 'retained' in engineering fall into the first two groups: lower-division and upper-division engineering. Students who are 'not retained' are those who have transferred or left the university.

| <b>Engineering freshmen students' status after 1<sup>st</sup> year</b>                          |                                       |
|---|---------------------------------------|
| <i>Possible statuses</i>  | <i>Dichotomous statuses</i>           |
| Upper-division engineering: completed first year requirements and move to upper divisions (UE)  | <b>1. Retained in engineering</b>     |
| Lower-division engineering: still remained in the first year program (LE)                       |                                       |
| Transferred to Science or Technology schools in the same university (ST)                        | <b>2. Not retained in engineering</b> |
| Transferred to schools in the same university other than Engineering, Science or Technology (O) |                                       |
| Left the university (L)   |                                       |

Table 1. Classification for the status of students after the first year

*Prediction performance measures for dichotomous prediction:*

The performance measures considered in this study are: 1) overall prediction accuracy, 2) sensitivity, 3) specificity, 4) accuracy for "not retained" prediction, and 5) accuracy for "retained" prediction (Larpiataworn, 2003).

| Actual Persistence Status | Result of prediction |           |
|---------------------------|----------------------|-----------|
|                           | Not retained         | Retained  |
| Not Retained              | True (A)             | False (B) |
| Retained                  | False (C)            | True (D)  |

Table 2. Example classification table.

*Note: A, B, C, D represent the numbers of observations within each classification.*

The overall prediction accuracy measures the fraction of accurate predictions within the total number of all observations. Its range is 0 to 1, and perfect score is 1, which corresponds to 100% prediction accuracy. Overall prediction accuracy, is defined as:

$$\text{Overall prediction accuracy} = \frac{A+D}{A+B+C+D}. \quad (1)$$

Sensitivity measures how well the model predicts over those who are not retained. Its range is 0 to 1, with a perfect score of 1. Sensitivity equal to 1 means 100% accuracy in predicting those who were not retained in engineering program. Sensitivity, is defined as:

$$\text{Sensitivity} = \frac{A}{A+B}. \quad (2)$$

Specificity is the measure of how accurate the model predicts those who remained in engineering programs. Similarly, its range is 0 to 1 with 1 as the perfect score. Specificity, is defined as:

$$\text{Specificity} = \frac{D}{C+D}. \quad (3)$$

The following two performance measures are implemented to express the accuracy of the ‘not retained’ and ‘retained’ predictions. Accuracy for the ‘not retained’ prediction is defined as:

$$\text{Accuracy for “not retained” prediction} = \frac{A}{A+C} \quad (4)$$

Accuracy for ‘not retained’ prediction has a range of 0 to 1 with 1 as the perfect score. When accuracy for “not retained” prediction equals to 0.9, it means 90% of the students who were predicted to be ‘not retained’ do in fact leave engineering programs. Similarly, accuracy for “retained” prediction is defined as:

$$\text{Accuracy for “retained” prediction} = \frac{D}{B+D} \quad (5)$$

Accuracy for “retained” prediction has a range of 0 to 1 with 1 as the perfect score. When accuracy for “retained” prediction equals to 0.85, it means 85% of the students who received a prediction of ‘retained’ are retained in engineering programs.

## **Results and Discussion**

### *Prediction by individual non-cognitive scales vs. combination of all items*

Eight exploratory- neural network models were developed, each of which corresponded to a particular scale factor (motivation, learning types ...etc.) and included only scale items associated with the respective dimension. This was done to gain an understanding of relative weights of each item in the sub-scale. In addition, one combination model (Model 1) was developed that consisted of all the measured factors (159 items). The prediction performances of each of the exploratory models as well as Model 1 are summarized in Table 3:

| Prediction Results                   | Individual non-cognitive scale |                |              |              |              |               |               |                     | Combination  |
|--------------------------------------|--------------------------------|----------------|--------------|--------------|--------------|---------------|---------------|---------------------|--------------|
|                                      | Expectancy                     | Meta-cognition | Motivation   | Career       | Leadership   | Learning type | Self Efficacy | Team vs. Individual |              |
| <b>Overall Accuracy</b>              | <b>66.6%</b>                   | <b>70.8%</b>   | <b>69.2%</b> | <b>70.4%</b> | <b>71.8%</b> | <b>73.2%</b>  | <b>74.6%</b>  | <b>73.2%</b>        | <b>75.6%</b> |
| Sensitivity                          | 18.75%                         | 16.67%         | 25.0%        | 25.0%        | 14.58%       | 6.25%         | 7.29%         | 11.46%              | 20.83%       |
| Specificity                          | 77.97%                         | 83.66%         | 79.7%        | 81.19%       | 85.4%        | 89.11%        | 90.59%        | 87.87%              | 88.61%       |
| Accuracy for Not-retained prediction | 16.82%                         | 19.51%         | 22.64%       | 24.0%        | 19.18%       | 12.0%         | 15.56%        | 18.33%              | 30.3%        |
| Accuracy for Retained prediction     | 80.15%                         | 80.86%         | 81.73%       | 82.0%        | 80.8%        | 80.0%         | 80.44%        | 80.68%              | 82.49%       |

Table 3. Prediction results from models using inputs from individual non-cognitive scale and a combination of all items.

Table 3 shows the overall prediction accuracy rates are approximately 70% for the exploratory models, ranging between 66% and 74%. The combination model (i.e., Model 1) with all items across scales does perform better, with an overall prediction accuracy of 75%. For the performance in specificity and accuracy for retained prediction, the results are even higher, ranging from 77% to 90%. However, the other two measures, sensitivity and accuracy for not-retained prediction, were much lower than the previous three measures. This shows the current strength and weakness with regard to different performance indices in these models.

*Modifications to Model 1: using reduced set of input items across multiple scales*

After the neural network models were trained, weighting values are assigned to each link between every input variable and its directly adjacent nodes. The relative magnitude of weight values, compared with weights from other input variables, can be considered as an indication of how important this input is for this trained neural network model. Figure 1 shows the relative weight values associated with every input variable associated with Model 1 (i.e., the combination model with all of the non-cognitive items). It is clear some of the input variables do possess much higher weight values than the others in this trained NN model.

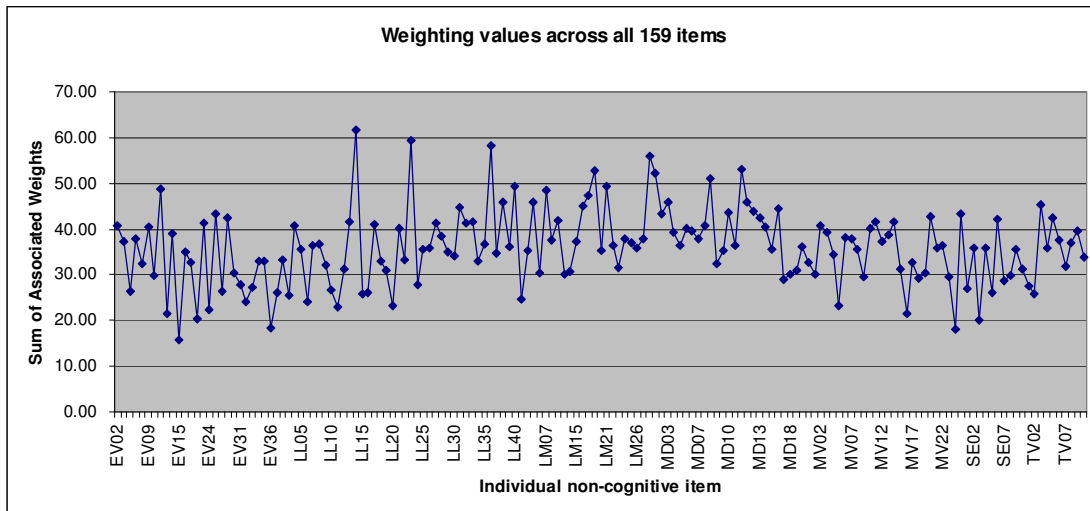


Figure 1. The summed weight values across all 159 items in the trained combination model

To investigate the possibility of reducing the size of this combination model's input data requirement (from 159 survey items), two new models with smaller sets of input items were developed based on the above weighting information. The goal is to develop models with number of input variables in a smaller range of 40 to 60, which the authors considered as a preferred size in order to reduce the survey length but also maintain the breadth of involved non-cognitive items.

In the process of determining what items to include in the new model, two levels of selection decision were considered. The first level of item selection decision was performed on individual subscales within each non-cognitive scale. That is, the new model either includes the whole subscale (consisting of more than one items), or none of the items inside that subscale. The second type of item selection decision was performed on individual items, without considering the scale or subscale to which it belongs.

The input data set for Model 2 includes subscales of items with significantly larger weight values from the previous study. The resulting subscales selected based on the weighting values are: planning (from meta-cognition scale), motivation (from career), dysfunctional belief (from career), leadership (from leadership), deep learning (from learning type) and surface learning (from learning type). Ultimately, Model 2 contains fifty nine items from six of the aforementioned subscales.

The input for a third model, Model 3, includes individual input items with higher weight values without considering their scale or subscale classification. Based on the weight information obtained from previous all-item combination model, forty nine items were selected to include in this model. After the developing, training and testing processes, comparison of prediction performances from these two new reduced-size models, as well as the previous all-item combination model were found (Table 4). The results from the new model, with selected individual items, performed better than the other new model with

selected subscales, across the five performance measures. However, it still performed slightly lower than the all-item combination model.

|                                      | Original model with all survey items (Model 1) | New model with selected sub-scales (Model 2) | New model with selected individual items (Model 3) |
|--------------------------------------|--|--|--|
| <b>Overall Accuracy</b>              | 75.6%  | 70.4%  | 72.2%  |
| Sensitivity                          | 20.83%   | 14.58%                                       | 18.75%   |
| Specificity                          | 88.61%   | 83.66%                                       | 84.9%  |
| Accuracy for Not-retained prediction | 30.3%  | 20.41%                                       | 22.78%   |
| Accuracy for Retained prediction     | 82.49%   | 81.09%                                       | 81.47%   |

Table 4. Comparison of prediction results between the original all-item model and two new models

*The influence of 'gender'*

To study the influence of the gender factor in this prediction model, a fourth model, Model 4, was developed with gender as one of the inputs in addition to the non-cognitive items. This new model is identical to Model 3, except that it had one additional input variable, namely gender. The prediction results are shown in Table 5 below.

|                                      | Model with selected individual items- without gender input (Model 3) | Model 3 with gender input |                |                  |
|--------------------------------------|--|---------------------------|----------------|------------------|
|                                      |  | All test data             | Male only data | Female only data |
| <b>Overall Accuracy</b>              | 72.2%  | 70.4%                     | 72.18%         | 63.37%           |
| Sensitivity                          | 18.75%   | 26.04%                    | 26.03%         | 26.09%           |
| Specificity                          | 84.9%  | 80.94%                    | 82.52%         | 74.36%           |
| Accuracy for Not-retained prediction | 22.78%   | 24.51%                    | 25.0%          | 23.08%           |
| Accuracy for Retained prediction     | 81.47%   | 82.16%                    | 83.28%         | 77.33%           |

Table 5. Comparison of prediction results between the models with and without gender inputs

Table 5 shows the results from new model with the gender inputs were similar to the original model (without gender inputs) when tested with data with both genders. However, when the testing data were divided into female and male subjects and processed separately with the same trained model, the female only group was found to have significantly lower overall prediction accuracy. This indicates that this neural network model can be gender sensitive and worthy of further investigation in that direction.

### **Conclusion**

This paper examines whether a neural network model based on student noncognitive characteristics can be used to predict student persistence in engineering after their second semester, and the influence of gender in the predictive model. Towards this end, NN Model 1, which was composed of all 159 items of the 8 factor noncognitive instrument, was shown to have a predicative accuracy of 82% for engineering students retained after 2 semester and 30% for non-retained students. Significantly decreasing the number of inputs to the model, using only those items which appeared to have the strongest influence (as determined by weighting factors), had little impact on the predicative accuracy of the retained students, but it did reduce the predictive accuracy of the non-retained students by approximately 10%. Future work should be concentrated on enhancing the sensitivity of the instrument to identify non-retained students. Finally, the use of a gender variable produced little-to-no measurable change in prediction rates.

### **Bibliographic information**

- Astin, A., T. Banta, P. Cross, E. El-Khawas, P. Ewell, P. Hutchings, T. Marchese, K. McClenney, M. Mentkowski, M. Miller, and E. Moran, Wright, BD, (1992). "Principles of Good Practice for Assessing Student Learning," *American Association for Higher Learning* 1992.
- Astin, Alexander W., (1993). "Engineering Outcomes," *ASEE Prism*, September 1993, pp. 27-30
- Augustine, Norman (chair), (2005). National Academies Committee on Prospering in the Global Economy of the 21<sup>st</sup> Century, *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*, National Academies Press, Washington, D.C., 2005.
- Beaufait, Fred W., (1991). "Engineering Education Needs Surgery," *Proceedings of the Frontiers in Education Conference*, September 1991, pp. 519-522.
- Besterfield-Sacre, M., C. Atman, and L. Shuman, (1997)"Characteristics of Freshman Engineering Students: Models for Determining Student Attrition in Engineering," *Journal of Engineering Education*, vol. 86, pp. 139-149, 1997.
- Besterfield-Sacre, M., M. Moreno, L. Shuman, C. Atman, (1999). "Comparing Entering

Freshman Engineers: Institutional Differences in Student Attitudes”, *proceedings of the American Society for Engineering Education National Conference*, 1999.

Demuth, H. and M Beale, (1998). *Neural Network Toolbox for Use with Matlab*, pp.5-3. Ver. 3, The Math Works Inc.,1998

Felder, R., K. Forrest, L. Baker-Ward, E. Dietz, and P. Mohr, (1993) "A Longitudinal Study of Engineering Student Performance and Retention I. Success and Failure in the Introductory Course," *Journal of Engineering Education*, pp. 15-21, 1993.

House, J., (2000). “Academic Background and Self-Beliefs as Predictors of Student Grade Performance in Science, Engineering and Mathematics," *International Journal of Instructional Media*, vol. 27, pp. 207-220, 2000.

Immekus, J., S. Maller, P.K. Imbrie, N. Wu, P. McDermott, (2005). *Work In Progress: An Analysis of Students’ Academic Success and Persistence Using Pre-College Factors” Proceedings of the Frontiers in Education Conference*, 2005.

Jagacinski, C. and LeBold, W., (1981). “A Comparison of Men and Women Undergraduate and Professional Engineers” *Journal of Engineering Education*, Vol 72, 1981, pp 213-220

Larpkiataworn, S., O. Muogboh, M. Besterfield-Sacre, L.J. Shuman, and H. Wolfe (July 2003), “Special Considerations When Using Statistical Analysis in Engineering Education Assessment and Evaluation”, *Journal of Engineering Education*, July 2003, pp. 207-215.

Maller, S., J. Immekus, P.K. Imbrie, N. Wu, P. McDermott, (2005).“Work In Progress: An Examination of Engr Student Profile Membership Over the Freshman Year”, *Proceedings of the Frontiers in Education Conference*, 2005.

Seymour E. and N. Hewitt, (1997). *Talking about Leaving: Why Undergraduates Leave the Sciences*. Boulder, CO: Westview Press, 1997.

Zhang, G., T. Anderson, M. Ohland, R. Carter and B. Thorndyke, (2002). “Identifying Factors Influencing Engineering Student Graduation and Retention: A Longitudinal and Cross-Institutional Study”, *proceedings of the American Society for Engineering Education National Conference*, 2002.