AC 2008-2417: ARTIFICIAL INTELLIGENCE METHODS TO FORECAST ENGINEERING STUDENTS' RETENTION BASED ON COGNITIVE AND NON-COGNITIVE FACTORS

P.K. Imbrie, Purdue University

P.K. Imbrie is an Associate Professor in the School of Engineering Education at Purdue University. He received his B.S., M.S. and Ph.D. degrees in Aerospace Engineering from Texas A&M University. His research interests in educational research include modeling student success, modeling student team functioning, and multidisciplinary engineering education. His technical research interests include solid mechanics, experimental mechanics, nonlinear materials characterization, microstructural evaluation of materials, and experiment and instrument design.

Joe Jien-Jou Lin, Purdue University

Joe J.J. Lin is a doctoral student in the School of Engineering Education at Purdue University. He received his M.S. degree in Industrial Engineering from Purdue University and B.S.I.E. from National Tsing-Hua University in Taiwan. He worked as a production control engineer before joining Purdue. His research interests include student success models, team effectiveness, neural network, fuzzy computing, data mining and production systems.

Alexander Malyscheff, Purdue University

Alexander M. Malyscheff is a Postdoctoral Research Associate in the School of Engineering Education at Purdue University. He graduated as a Diplom-Ingenieur in Electrical Engineering from the Technical University Braunschweig (Germany) and received his M.S. in Electrical Engineering and Ph.D. in Industrial Engineering (Operations Research) from the University of Oklahoma. Before joining Purdue University he worked as a Risk Analyst in Energy Trading. His research interests include modeling of student data, team effectiveness, mathematical programming, machine learning, kernel methods, and data analysis.

Artificial Intelligence Methods to Forecast Engineering Students' Retention based on Cognitive and Non-cognitive Factors

Abstract

Engineering students' affective self-beliefs can be influential factors directly or indirectly affecting their academic success and career decision. This paper examines whether students' non-cognitive factors can be used, alone or in combination with cognitive factors, in artificial neural network (ANN) models to predict engineering student's future retention. Four ANN based retention prediction models using different combinations of non-cognitive and cognitive factors are presented. The independent variables includes survey items from nine non-cognitive constructs (leadership, deep learning, surface learning, teamwork, self-efficacy, motivation, meta-cognition, expectancy-value, and major decision) and eleven cognitive items representing student's high school academic performance. The dependent variable (i.e., the output from these models) is the student's retention status after one year.

Data from more than 4900 first-year engineering students from three freshman cohorts (2004, 2005, 2006) in a large Midwestern university were collected and utilized in training and testing these ANN prediction models. Among the four ANN models developed, the model combining 11 cognitive items and 60 selected non-cognitive items has the highest overall prediction accuracy at 71.3%, probability of detection (POD) for retained students at 78.7% and POD for not retained student at 40.5%. Removing the 11 cognitive items from this model, the overall prediction accuracy would drop slightly to 70.5%.

Results from training and testing the same model using student data from different cohorts indicate the ANN model's predictive performance is generally stable across different cohort years. Also, a model trained with earlier year (2004) freshman cohort's data has maintained its predictive power very well when tested with student data from later (2005 and 2006) cohorts.

Introduction

As Thomas Friedman described in his best selling book 'The World is Flat'[1], the world has become flatter because of the numerous new technologies and developments in the past decades. Engineers in India, China or other parts of the world today are now able and eager to compete directly with the engineers from the United States. An alarming trend over the last decade is the number of engineering graduates in U.S. continues to fail to keep pace with the increasing production of engineers from our international competitors. In the report "Rising Above The Gathering Storm: Energizing and Employing America for a Brighter Economic Future" published by the National Academies in 2005[2], it is reported that undergraduate programs in science and engineering have the lowest retention rate among all academic disciplines. The National Academies further emphasized the importance of advances in engineering and technology, and described them 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.

Since the advances in engineering and technology have such a strong impact on the future of our society, how to attract and retain students in engineering majors becomes an important topic. Every year a great number of the top graduates from high schools enter engineering programs across this country. Many of them have obtained impressive academic records during their high school education, in terms of grade point averages and standardized test scores. Still, various engineering educational studies indicate that the attrition from engineering continues to be an alarming issue[3,4]. A good number of qualified students continue to leave engineering for other majors, or leave the college completely. It was reported that the attrition in the freshman year in engineering has increased from about 12% in 1975 to 25% by 1990[3]. In a large study of over 300 universities, Astin[4] found that only 47% of freshman engineering students eventually graduate with an engineering degree. This means that more than half of these engineering-inspired young people left engineering during their college education. For educators concerned about the future of engineering education and the ultimate competitiveness of the United States, this is a problem too important to ignore.

In order to address the critical topic of student attrition in engineering education, it is necessary to investigate the factors related to student retention, and purposefully develop a predictive system which can identify students with a high risk of leaving engineering early. The aforementioned predictive ability can significantly help engineering educators perform proper interventions in time to help retain these students in engineering programs. Therefore, the research question for this study is: “Can a predictive model be developed to take multiple non-cognitive factors, cognitive factors and their interaction into account and improve our prediction of students’ future retention in engineering?”

Model of student success

The undesirable fact regarding engineering students’ high attrition rate has provided the authors a powerful motivation to study the various factors that influence engineering students’ success. Figure 1 shows a Model of Students’ Success (MSS) in engineering. This MSS illustrates the potential relationships between numerous factors and outcomes associated with engineering students’ success in academics and career. The authors developed this MSS model partially based on previous studies on non-cognitive factors by Maller et al.[5] and Imbrie et al.[6] The main focus of investigation in this work is on the non-cognitive and cognitive factors, and their influences on engineering students’ retention, as highlighted in Figure 1.

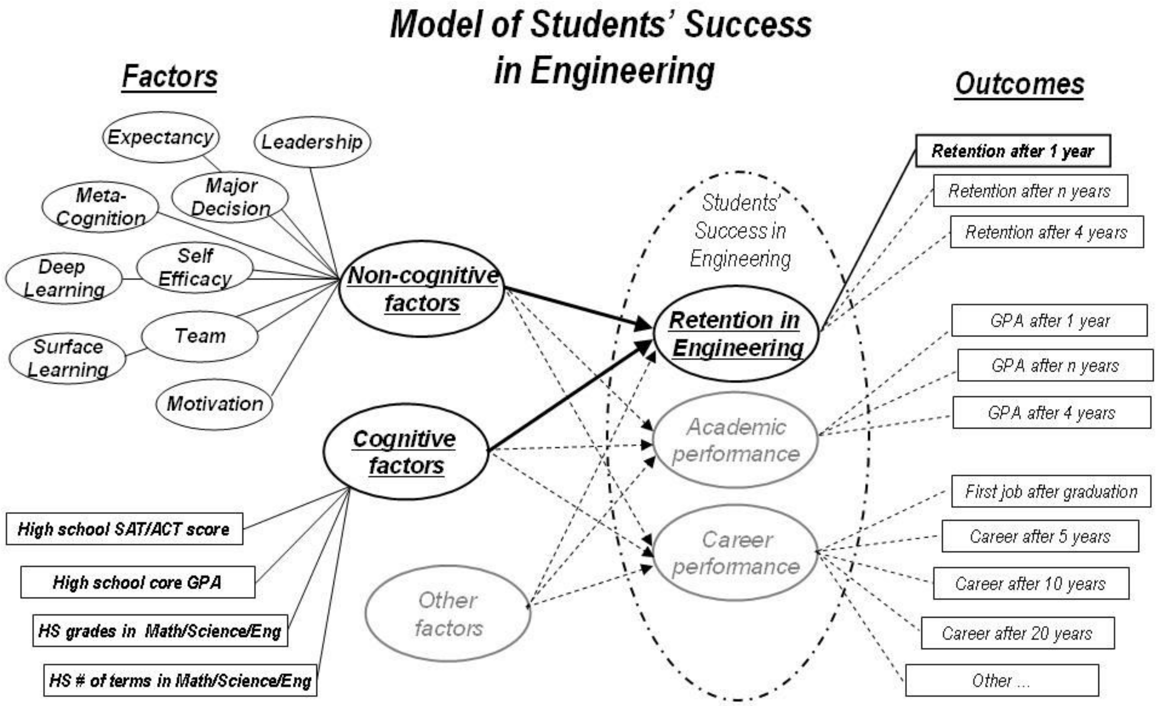

Figure 1. Model of Students' Success in Engineering

Factors affecting student retention

One common misconception about student retention is that students leave engineering largely due to lack of academic ability. Studies have found little difference between the academic credentials of students who remain in engineering and those who leave[7]. Other studies have shown that models incorporating cognitive variables such as student high school math and science success[8], and higher confidence in basic engineering knowledge and skills[7] are able to establish a correlation between cognitive variables and retention in engineering. However these variables are not strong enough to be used as single factors in a model to predict retention. Therefore, some researchers suggested a model using both cognitive and non-cognitive characteristics may provide a more promising tool to identify students who may leave engineering or who may benefit from interventions[7,9]. 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 in engineering, Zhang et al.[10] found that high school GPA and SAT math scores were the best predictor of retention and graduation, while SAT verbal was inversely related. They also identified self-efficacy and physical fitness as positive predictors of freshman retention. Astin et al.[11] found that student high school record was the best predictor of academic success, and performance on standardized tests also had a positive correlation. These studies were valuable in identifying characteristics that were predictors for retention, but did not address the interaction of multiple factors.

The Pittsburgh Freshman Engineering Attitudes Survey (PFEAS) is an instrument consisting of 50 items related to 13 student attitude and self-assessment measures[7,12]. Besterfield-Sacre et al. have used PFEAS to measure differences in student attitudes before and after the freshman year in their study on freshmen attrition from engineering programs. Another related study in recent years is the Persistence In Engineering (PIE)[13]. PIE is a survey instrument developed under the Academic Pathways Study (APS) by the Center for the Advancement of Engineering Education (CAEE). The factors studied in the PIE survey are largely related to the educational experiences student received during their college years, and some motivation and self-efficacy factors on the non-cognitive constructs. These and similar studies suggest that student attitudes and other non-cognitive characteristics may be promising factors to be incorporated into a new predictive model to predict students' persistence and retention.

Data Collection and Instrumentation

The sample in this study included more than 4900 incoming freshman engineering students from a large Midwestern university during the 2004, 2005 and 2006 academic years. Among them, 17.02% were female, and 82.98% were male. Ethnicity was as follows: 2.17% African American, 0.48% American Native, 9.44% Asian/Pacific Islander, 2.88% Hispanic, 78.21% Caucasian and 6.81% Others.

Non-cognitive survey instruments and cognitive data

The students' non-cognitive measures were collected across nine scales in a self-reported online survey completed prior to the freshman year. This non-cognitive survey instrument was previously reported in the works by Maller et al.[5] and Immekus et al[14]. These scales are: Leadership (23 items), Deep vs. Surface Learning Types (20 items), Teamwork (10 items), Self-efficacy (10 items), Motivation (25 items), Meta-cognition (20 items), Expectancy-value (32 items), and Major decision (28 items). All Cronbach's coefficient alphas for these scales were ≥ .80, except for the Teamwork scale (r=.74)[14]. Scales may be divided into subscales with various numbers of items. Previous studies have supported the scales' construct validity based on the results of confirmatory factor analyses[5].

The following eleven cognitive items from students were also collected: overall GPA and core GPA from high school, standardized test results (SAT/ACT), average high school grades in mathematics, science, and English classes and finally the number of semesters taking mathematics, science, and English.

Persistence status

Students' persistence statuses were collected at the beginning of every semester following their freshman year. Students remaining in the lower-division and upper-division engineering programs were considered as "retained" students. The students who transferred to majors other than engineering or left the university completely were classified as "not-retained". The investigation in this study focuses on the persistence status at beginning of students' third semester. Among the 4900 students studied in this report, 82.1% of them were retained and 17.9% were not retained in engineering at the beginning of third semester.

Methods

Artificial Neural Networks (ANNs)

Artificial Neural Network (ANN) is a well developed modeling approach among the various tools within the Artificial Intelligence (AI) family. During the past decades it has been widely used in technical applications involving prediction and forecasting, especially in areas of engineering, medicine and business [15,16,17,18]. The neural network model is especially attractive for modeling complex systems because of its following favorable properties: universal function approximation capability, tolerance to noisy or missing data, accommodation of multiple non-linear variables with unknown interactions, and good generalization ability[19]. In this study, the neural network model used for predicting students' retention is a feed-forward neural network with back-propagation training algorithm (FFBP). FFBP neural network was chosen because of its strength in modeling prediction/forecast problems involving large amounts of data and relatively complex relationships between factors and outcomes [16].

The FFBP neural network model developed for this study consists of an input layer, a hidden layer and an output layer with various numbers of neurons in each layer. The numbers of neurons in the input and output layers are determined by the number of input items and prediction outcome. In this work, depending on the models being studied, the number of input items varied from 9, 11, 60 to 71, and the number of output (prediction of retention status) is one. Determining the number of neurons in the hidden layer is more complicated. It is generally influenced by the nature of problem, such as the complexity of mapping between input and output data. For the four models with different input structures stated above, there are 9, 11, 30 and 36 hidden neurons used in hidden layers, respectively. The decision on the number of hidden neurons in each ANN model was determined by comparing performance results from extensive ANN experiments covering wide ranges of possible number of hidden neurons in the network, trained with actual student data. A general graphic illustration of applying the neural network model for the prediction of student's persistence in this work is shown in Figure 2 below.

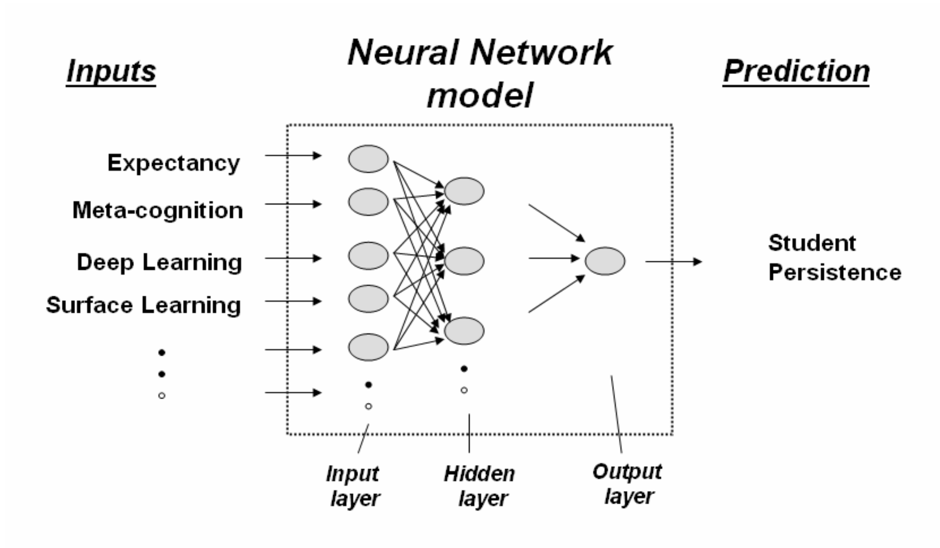

Figure 2. Using non-cognitive factors as inputs for Neural Network prediction models

Training of ANN models

New neural network models must be trained first with existing data so they can learn from the examples. A set of training data with known input and output (target) vectors is required for this process. During the back-propagation training process, weights associated with the links between neurons are adjusted in order to reduce the difference between the network's actual output and target output. This training process continues iteratively until the output results of the new network reaches a preset proximity of the desired output. After the training process is completed, a different set of data is used for testing to determine the actual performance of the trained model. In this study, Levenberg-Marquardt back-propagation training algorithm is used as training algorithm in these models[20]. The activation functions in the hidden and output neurons are both tan-sigmoidal functions. This is again determined after comparing performance results from alternative activation functions in extensive experiments. The performance function utilized is mean square error (MSE). All models were developed using Matlab version R2006b from Math Works Inc. The detailed model structures and setup parameters for these four ANN models developed are shown in table below.

Table 1. Model description

Model ID		A	B	C	D
Input factors (independent variables)		Non-cognitive factors only		Cognitive factors only	Combination of both cog. and non-cog. factors
		Average scores of each of the 9 non-cognitive constructs from the 168 item survey	Selected 60 items from the 168 item survey	11 cognitive items as described in data collection	Combination of the inputs from model B and C; totally 71 items
Output results (dependent variables)		Persistence status in engineering after one year			
Number of neurons	**Input layer (I)**	9	60	11	71
	Hidden layer (H)	9	30	11	36
	Output layer (O)	1	1	1	1
Training algorithm		Levenberg-Marquardt back propagation training algorithm			
Activation function	**Hidden layer**	tan-sigmoidal activation function			
	Output layer	tan-sigmoidal activation function			
Performance function		Mean square error (MSE)			
Size of training data sets*		900, 1050, 1050 for year 2004, 2005 and 2006			
Size of testing data sets*		600 each for all three years			

* For each year, training and testing data are two separate partitions of data without any overlapping

Prediction performance measures

The prediction performance measures considered in this study are: 1) overall prediction accuracy, 2) probability of detection (POD) for retained students, 3) probability of detection (POD) for not retained students, 4) bias for retained prediction, and 5) bias for not retained prediction. These measures will be discussed in details using the classification table below.

Table 2. Classification table for possible prediction results

		Observed (actual) status	
		Not Retained	Retained
Predicted Status	Not Retained	*Hits* *(a*)*	*False Alarm* *(b)*
	Retained	*Misses* *(c)*	*(d)*

* a, b, c, d represent the numbers of students in 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:

$$\text{Overall prediction accuracy} = \frac{a+d}{a+b+c+d}. \tag{1}$$

Probability of detection for retained student (POD Retained) measures how well the model predicts over those who are actually retained. Its range is 0 to 1, with a perfect score of 1. POD Retained equals to 1 means 100% of the retained students were predicted correctly. It is defined as:

$$\text{POD Retained} = \frac{d}{b+d}. \tag{2}$$

Probability of detection for not retained student (POD NotRetained) measures how well the model predicts over those who are actually not retained. Its range is 0 to 1, with a perfect score of 1. POD NotRetained equals to 1 means 100% of the not retained students were predicted correctly. It is defined as:

$$\text{POD NotRetained} = \frac{a}{a+c}. \tag{3}$$

The bias measures the ratio of the frequency of predicted events to the frequency of observed events. It expresses the tendency of the forecast system to over-forecast (bias > 1) or under-forecast (bias < 1) events. The range of bias is 0 to infinity, and a perfect score is 1. In this work, the Bias Retained is the ratio of number of predicted retained students over the number of actually retained students. It is defined as:

$$\text{Bias Retained} = \frac{c+d}{b+d}. \tag{4}$$

Bias NotRetained expresses the ratio of number of predicted not retained students over the number of actually not retained students. Similarly, its range can be from zero to infinity, with the perfect score as 1. It is defined as:

$$\text{Bias NotRetained} = \frac{a+b}{a+c}. \qquad (5)$$

Results and Discussion

Predicting student's persistence by different sets of cognitive and non-cognitive factors

Four artificial neural network models with different collections of cognitive and non-cognitive input items were developed as described earlier in Table 1. Model A uses the students' average scores from each of the nine non-cognitive constructs as its inputs. Model B includes 60 non-cognitive items carefully selected from our 168-item survey. They were selected based on item response theory and exploratory neural network experiments. Model C uses the eleven cognitive items previously described in data collection section as inputs. Model D incorporated the combined inputs from model B and C to create a hybrid model taking inputs from both cognitive and non-cognitive factors into the predicting process. After training and testing these four neural network models, their prediction results are presented in table below.

Table 3. Comparison of prediction results between different ANN models using data from the 2004 cohort

Model	A	B	C	D
Input Factors	Non-cognitive factors		Cognitive factors	Combination of both B & C
Description of ANN model's input data	9 Non-cognitive measures	60 Non-cognitive items	11 Cognitive items	11 cognitive and 60 Non-cog items; 71 items totally
Overall Prediction Accuracy	**67.3%**	**70.5%**	**69.7%**	**71.3%**
POD Retained	76.2%	78.3%	77.7%	78.7%
POD Not Retained	30.2%	37.9%	36.2%	40.5%
Bias Retained*	93.0%	93.2%	93.0%	93.0%
Bias Not Retained*	129.3%	128.4%	129.2%	129.3%

* The preferred score for bias here is as close to 1 (100%) as possible.

Results in Table 3 showed that the better performing models, B and D, can achieve overall prediction accuracy above 70%. The probability of detection for retained students can be as high as 78%, while probability of detection for not retained students are lower at 37% and 40%. The values in Bias Retained and Bias Not Retained indicated these models under-forecast the number of retained students slightly (7%), but over-forecast the not retained students for about 29%. The authors considered this acceptable because it is preferred to bring more high risk students into our attention than under-estimate the number, which would result in ignoring some of the students who were in need of help. Other interesting observations from the table above are discussed here. First, the results showed that the model B using only non-cognitive factors can actually predict students' retention as good as the model C using cognitive records. This is indeed a very encouraging finding. Second, for the two models A and B using only non-cognitive measures, model B with 60 non-cognitive items does perform better than model A using only averaged scores for each of the nine non-cognitive constructs. This justified the additional developing and computing efforts/costs for a larger and more complex model using individual survey items. Third, the model combining factors from both cognitive and non-cognitive categories does improve the prediction accuracy when compared with models using only cognitive or non-cognitive alone.

Predicting student's persistence for different freshman cohorts using the same neural network structure

Here we explore the reliability of prediction performance when using same ANN network structure to predict students' retention for different cohort years. Model B, with 60 non-cognitive items, was selected because of its good prediction performance without additional requirement on student's cognitive records. This ANN model was trained and tested with students from 2004, 2005 and 2006 cohort independently and results are shown in table below.

Table 4. Predicting student's persistence for different first-year cohorts using the same neural network structure within year training

Model	Model B: with 60 Non-cognitive items		
Cohort data used in training*	2004	2005	2006
Cohort data used in Testing *	2004	2005	2006
Overall Prediction Accuracy	**70.5%**	**69.5%**	**72.0%**
POD Retained	78.3%	77.5%	80.9%
POD NotRetained	37.9%	35.7%	22.8%
Bias Retained	93.2%	92.8%	94.9%
Bias NotRetained	128.4%	130.4%	128.3%

* For each year, training and testing data are two separate partitions of data without any overlapping

After training and testing independently with students from three different cohorts, the results suggest that the ANN model structure tested here is reasonable stable in their overall prediction accuracy over different cohorts, ranging from 69.5% to 72.0%. The results for the remaining four performance measures are also mostly consistent across years. The only exception is found in the POD NotRetained for year 2006, which may require further investigation by the authors. The overall results are still encouraging, because they suggested that a model structure developed and confirmed with good performance using one year's data would likely maintain its predictive performance when applied to different student populations from other cohorts/years.

Predicting future student's retention status with ANN model developed and trained by previous year's student population

One question the authors were eager to answer is: How well can an ANN model trained in previous year predict the in-coming student's retention status in future? With the data from three first-year cohorts available, we were able to investigate this possibility by training one model with year 2004 data, and test it with 2005 and 2006 cohort data in a simulated prediction experiment. Additionally, another model was also trained independently with year 2005 data and tested with 2006 students in a similar manner. The results were displayed in table below.

Table 5. Predicting future student's retention status with ANN model trained by previous year's student population

ANN model	Model B: with 60 Non-cog items				
Cohort data used in training	2004*			2005 *	
Cohort data used in testing	2004*	2005 Prediction	2006 Prediction	2005*	2006 Prediction
Overall Prediction Accuracy	**70.5%**	**70.3%**	**71.8%**	**69.5%**	**69.2%**
POD Retained	78.3%	78.1%	77.6%	77.5%	76.0%
POD NotRetained	37.9%	37.4%	40.2%	35.7%	31.5%
Bias Retained	93.2%	93.0%	88.4%	92.8%	88.4%
Bias NotRetained	128.4%	129.6%	164.1%	130.4%	164.1%

* Training and testing data are two separate sets of data without any overlapping even from the same cohort year

With the model trained with 2004 freshman students, the overall prediction accuracy for 'future students' in 2005 and 2006 cohorts have maintained very well as shown in the table. Similar prediction performance was also obtained for the model trained with 2005 data and tested with data from 2006 'future' students. These findings are again very satisfying.

Applying these prediction models to assist the academic counseling professionals and improve student retention in engineering

As discussed earlier in this paper, the engineering colleges in this country lose more than 20% of their freshman students after one year, and only less than 50% of the students will eventually complete their engineering degree. Early preventive intervention from the academic counseling professionals is a very important way to help students remaining in the engineering programs before it is too late. However, it is often difficult to identify students with high risk early enough to make the intervention meaningful and effective. With all the needed data collected and available before the students start their first semester, this artificial intelligence (AI) based prediction system will be very valuable in helping the counseling professionals focusing their efforts on the right population early, and eventually improve the overall retention in long term.

Conclusion

This paper presented artificial neural network models developed with different collections of students' cognitive and non-cognitive factors to predict student persistence in engineering after their freshman year. The prediction performance results using different input data were compared. The models developed here achieved an overall prediction accuracy around 70% or higher consistently, with the probability of detection for retained students close to 78%. However, the probability of detection for not retained students was lower and in the 40% range. Considering the fact that there are still other types of factors (such as financial issues, health condition, family reasons…etc.) that influence engineering students' persistence, it is understandable that the current models using cognitive and non-cognitive variables may only discover some of the non-persisting students, but not all of them. The authors also examined the performance robustness of using the same ANN network structure to predict student retention from different freshman cohorts, and the possibility of using a model trained with previous years' student data to predict future students' retention status. In both cases our models performed consistently satisfying across different years of data. These results are very encouraging. Future efforts will be concentrated on further improving the power of detecting the group of not retained students by 1) enhancing the ANN models further with other promising ANN techniques, and 2) incorporating fuzzy logic techniques to develop new fuzzy-neural hybrid prediction models.

Bibliography

1. Friedman, T. L. (2005). *The World Is Flat- A Brief History of The Twenty-First Century*, Farrar, Straus and Giroux, 2005.
2. Augustine, N. (chair) (2005). National Academies Committee on Prospering in the Global Economy of the 21st Century, *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*, National Academies Press, Washington, D.C., 2005.
3. Beaufait, F. W. (1991). Engineering Education Needs Surgery, *Proceedings of the Frontiers in Education Conference*, September 1991, pp. 519-522.
4. Astin, A. W. (1993). Engineering Outcomes, *ASEE Prism*, September 1993, pp. 27-30
5. Maller, S., Immekus, J., Imbrie, P. K., Wu, N. and McDermott, P. (2005).Work In Progress: An Examination of Engineering Students' Profile Membership Over the Freshman Year, *Proceedings of the Frontiers in Education Conference,* 2005.
6. Imbrie, P. K. and Lin, J.J. (2007). Use of a Neural Network Model and Noncognitive Measures to Predict Student Matriculation in Engineering, *Proceeding of the 2007 American Society for Engineering Education Annual Conference & Exposition*, 2007
7. Besterfield-Sacre, M., Atman, C. and Shuman, L. (1997). Characteristics of Freshman Engineering Students: Models for Determining Student Attrition in Engineering, *Journal of Engineering Education*, vol. 86, pp. 139-149, 1997.

8. 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

9. Felder, R., Forrest, K., Baker-Ward, L., Dietz, E. and Mohr, P. (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.

10. Zhang, G., Anderson, T., Ohland, M., Carter, R. and Thorndyke, B. (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.

11. Astin, A., Banta, T., Cross, P., El-Khawas, E., Ewell, P., Hutchings, P., Marchese, T., McClenney, K., Mentkowski, M., Miller, M., Moran, E. and Wright, B.D. (1992). *Nine Principles of Good Practice for Assessing Student Learning*, American Association for Higher Learning, 1992.

12. Besterfield-Sacre, M., Moreno, M., Shuman, L. and Atman, C. (1999). Comparing Entering Freshman Engineers: Institutional Differences in Student Attitudes, *Proceedings of the American Society for Engineering Education National Conference*, 1999.

13. Ozgur, E., D. Chachra, H. Chen, C. Rosca, L. Ludlow, S. Sheppard and K. Donaldson, A Preliminary Analysis of Correlates of Engineering Persistence: Results From A Longitudinal Study, *Proceeding of the 2007 American Society for Engineering Education Annual Conference & Exposition*, 2007

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

15. Haykin, S. (1994). *Neural Networks- A Comprehensive Foundation*, Macmillan College Publishing Company, Inc., 1994

16. Tsoukalas, L. H. and Uhrig, R. E. (1997). *Fuzzy and Neural Approaches in Engineering*, John Wiley & Sons, Inc., 1997

17. Kukar M., Konenko I., Grošelj C., Kralj K., Fettich J. (1999). Analysing and improving the diagnosis of isachaemic heart disease with machine learning. *Artificial Intelligence in Medicine* 1999;16: p. 25-50.

18. Smith K. and J. Gupta (2002), *Neural Networks in Business: Techniques and Applications*, Idea Group Publishing, 2002

19. Coit, D., B. Jackson and A. Smith (1998), Static Neural Network Process Models: Considerations and Case Studies, *International Journal of Production Research*, 1998, 36, pp. 2953-2967

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