Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 95 (2016) 375 – 382

Procedia Computer Science

Complex Adaptive Systems, Publication 6
Cihan H. Dagli, Editor in Chief
Conference Organized by Missouri University of Science and Technology
2016 - Los Angeles, CA

Application of an Artificial Neural Network to Predict Graduation Success at the United States Military Academy

Gene Lesinski[a,*], Steven Corns[b], Cihan Dagli[b]

[a]*United States Military Academy, West Point, NY 10996, USA*
[b]*Missouri University of Science and Technology, Rolla, MO 10996, USA*

Abstract

This paper presents a neural network approach to classify student graduation status based upon selected academic, demographic, and other indicators. A multi-layer feedforward network with backpropagation learning is used as the model framework. The model is trained, tested, and validated using 5100 student samples with data compiled from admissions records and institutional research databases. Nine input variables consist of categorical and numeric data elements including: high school rank, high school quality, standardized test scores, high school faculty assessments, extra-curricular activity score, parent's education status, and time since high school graduation. These inputs and the multi-layer neural network model are used to classify students as: graduates, late graduates, or non-graduates. Several neural network architectures are examined and compared by run time, minimum mean square error achieved (MSE), mean correct classification rate, precision, recall, and specificity. A multi-layer neural network with 50 hidden neurons, momentum value of 0.8, and learning rate of 0.1, with hyperbolic tangent hidden neuron activation functions was able to accurately predict graduation success and achieved the best performance with classification accuracy exceeding 95%, high recall, high precision, and high specificity. This prediction model may be used to inform admission decisions and identify opportunities for required remediation with the potential to improve graduation rates, increase student retention, reduce late graduation, and reduce first-term course failures.

Peer-review under responsibility of scientific committee of Missouri University of Science and Technology

Keywords: Neural network; backpropagation; enrollment management; student retention; classification

Peer-review under responsibility of scientific committee of Missouri University of Science and Technology
doi:10.1016/j.procs.2016.09.348

1. Introduction

All colleges and universities are concerned with graduation rates and retention of their students. Graduation and retention rates are typically used by organizations like Forbes and US News and World Report as proxy indicators of school quality which indirectly impact the institution's bottom line. Graduation and retention rates are particularly important at the United States Military Academy where a retention loss is ultimately a loss to Army officer end strength. Each year, more than 15,000 candidates, from all 50 states, apply for admission to West Point. Approximately 1,200 applicants are accepted each year and receive the equivalent of a four year full scholarship with a Government Accounting Office (GAO) estimated value of $327,000 [7]. Significant effort is applied to graduate a majority of students within four years to satisfy Army officer manning requirements. Recently there has been a spike in the number of first term course failures for entering freshmen at West Point. This has generated interest in reexamining the decision criteria and models that inform admissions decisions. Given the magnitude of commitment associated with admission and the emphasis on four year completion, it is important to closely examine and periodically revalidate the criteria used to make these important admission decisions. Accurately modeling graduation success can ultimately improve graduation rates, increase student retention, reduce late graduation, and reduce first-term course failures. An accurate prediction model can both inform admission decisions as well as identify students requiring remediation. In this research we utilize a multi-layer feedforward neural network with nine selected input variables to model and classify student graduation status to inform admission decisions and identify opportunities for required remediation.

1.1. Related Research

In studies of college graduation success, vast amounts of research are focused on identification of significant predictor variables/factors as well as different mathematical models utilizing these factors to predict successful completion of college. There are numerous studies in the literature regarding factors that may predict successful college graduation. These factors are generally divided into pre-admission and post-admissions considerations. Pre-admissions factors can be further categorized as academic and non-academic. Academic pre-admission factors often include, high school rank, high school grade point average, and standardized test scores.

1.2. Graduation Prediction Factors

Burton and Ramist found that the best combination of SAT scores to be the best predictor of graduation success.[4] Geisler and Santelices conclude that high school GPA was not only the best predictor of first year grades but also for degree completion [8]. Niu and Tienda argue that another measure of high school achievement, high school rank, is a better predictor of college performance than standardized test scores [12]. Black, et al. found a significant correlation between high school quality and student success at college and believe that high school achievement should be adjusted relative to high school quality [2]. Some examined non-academic indicators of college success include social economic status (SES), parental education, faculty references, and high school extra-curricular involvement. Several sources note the strong correlation between parental level of education and the propensity to attend college. Additionally, Nelson identified a significant relationship between parental education and student college success [11]. Willingham identified faculty references and high school activity involvement as two significant non-academic indicators of college success [13]. In this research we high school rank in conjunction with high school quality, SAT scores, parental education, high school faculty assessments, candidate activity scores, and time since high school as factors to predict college graduation success.

1.3. Mathematical Prediction Models

Within the literature there are also a wide variety of modeling approaches applied to prediction of college graduation. Bowen and Bok utilized a logistic regression model to predict graduation within six years using gender, ethnic group, SES, selectivity of the college, SAT scores and high school records as predictive factors [3]. Kanarek achieved successful results using discriminant function analyses to classify students into graduates and non-