Applications in Predictive Analytics

Prediction as a Path to Better Educational Outcomes

Northwestern University, Masters of Science in Predictive Analytics

Applications in Predictive Analytics

Prediction as a Path to Better Educational Outcomes

Shifts in the global economy have resulted in an increasing demand for so-called "knowledge workers," highly educated people able to do complex jobs. With salaries and job prospects reflecting this evolution, college enrollment has increased in recent decades. However, while college enrollment has climbed, timely degree attainment has remained a challenge for higher education institutions (Bound, Lovenheim, & Turner, 2007). According to the National Center for Education Statistics, less than 60% of first-time, full-time students who began seeking a bachelor's degree at a four-year institution in fall 2006 completed that degree within six years (National Center for Educational Statistics, 2014).

Predictive analytics offers an opportunity for educational institutions to more effectively tackle this problem. At the institutional level, this can mean modeling student retention, progression, and graduation to identify risk factors and target interventions. In the classroom, educators can use predictive analytics to grade student work more quickly and with less bias while students might use adaptive learning platforms that provide differentiated instruction tailored to each student's particular strengths, weaknesses, and needs.

While many tools and technologies are coming on the market to enhance the educational experience, good management practice demands that institutions do their due diligence to clearly understand the opportunities for such efforts to have the greatest impact. Fortunately, higher education institutions have large repositories of student data available to be analyzed, including not just basic enrollment information, but academic histories, and rich student demographics. These data lend themselves well to analytics that can support stakeholders throughout the system, all ultimately in support better academic outcomes for students.

Methods

Efforts to analyze student data utilize a wide array of analytic methods aimed at ultimately understanding the constellation of factors that affect the target outcome, be that retention, performance in a given course, or graduation. A 2012 initiative, the Predictive Analytics Reporting Framework (PAR), discussed many of these methods (Ice, et al., 2012).

Data Review

While not an analytic method, data review is a crucial part of the analytical process, as it helps to identify issues of data quality and completeness. Incomplete or bad data can make certain analyses difficult or produce misleading results. This is an especially important step when combining data from various sources, where seemingly similar variables may be defined differently.

Descriptive statistics

The first step for any analysis includes a basic overview of the composition of the data set, such as characteristics of the student body being analyzed. Summaries of demographics provide the context for the deeper analyses some insight on possible trends.

Regression Analysis

Linear and logistic regression provide insight in to the relationships between the various independent variables, such as demographic characteristics, economic background, or academic history and a dependent variable, such as grade matriculation for a given group of students.

CHAID Analysis

Chi-Square Automatic Interaction Detection is a form of a decision tree technique useful for finding the most salient differences between groups of students, enabling development and exploration of sub-groups within the broader population.

Modeling

To enable findings to be put to use, the findings of the analysis are condensed in to one or more predictive models. These models use a selection of variables to allow for the user to input a set of information and obtain a predicted result. In the case of student retention, such a model might produce a retention likelihood score used to prioritize access to resource intensive interventions or might produce an assignment into a specific intervention category.

Management

The value of any analytic effort is ultimately determined by the degree to which it informs practice. In the field of education, there potential consumers of the outputs of predictive analytics at every level.

The PAR framework provides an example of how analytics can guide industry level collaboration. The initiative was funded by the Bill & Melinda Gates foundation and led by a unit of the Western Interstate Commission for Higher Education (WICHE), described as “a regional compact of fifteen Western states that began operations in 1953 to facilitate resource sharing among the higher education systems in the West.” Armed with a common understanding of the problems related to retention, progression, and completion, institutions can develop specific interventions that are more likely to succeed.

On a more applied-level, analytics can nudge students down a course of study that sets them up for success. Tristan Denley (2014) was part of a team that developed course a recommendation system, Degree Compass, designed “to nudge students toward course selections in which the data suggest they would have the most productive success, but using an interface that would minimize choice overload.”

In a 2012 article in the Journal of Asynchronous Learning Networks, Anthony Picciano discusses a set of examples of predictive analytics being used in practice to support student success in the classroom. At Rio Salado Community College, the Progress and Course Engagement (PACE) system was built on the finding that three factors predicted student success, defined as course completion with a “C” or better, with 70% accuracy. The system produces reports for instructors to guide intervention with students.

By better utilizing information that already exists in student management systems, universities can empower management, educators, and students in service of better educational outcomes.

References

Bound, J., Lovenheim, M. F., & Turner, S. (2007). *Understanding the Decrease in College Completion Rates and the.* Ann Arbor: Population Studies Center, University of Michigan Institute for Social Research.

Denley, T. (Winter 2014). How predictive analytics and choice architecture can improve student success. *Research & Practice in Assessment*, p61.

Ice, P., Diaz, S., Swan, K., Burgess, M., Sharkey, M., & Sherrill, J. (2012). The par framework proof of concept: initial findings from a multi-institutional analysis of federated postsecondary data. *Journal of Asynchronous Learning Networks*, p63.

National Center for Educational Statistics. (2014, May). *Institutional Retention and Graduation Rates for Undergraduate Students*. Retrieved from http://nces.ed.gov/programs/coe/indicator_cva.asp

Picciano, A. G. (Aug 2012). The evolution of big data and learning analytics in American higher education. *Journal of Asynchronous Learning Networks*, p9.