

Published Online: June 30, 2011

Schools Find Uses for Predictive Data Techniques

By Sarah D. Sparks

The use of analytic tools to predict student performance is exploding in higher education, and experts say the tools show even more promise for K-12 schools, in everything from teacher placement to dropout prevention.

Use of such statistical techniques is hindered in precollegiate schools, however, by a lack of researchers trained to help districts make sense of the data, according to education watchers.

Predictive analytics include an array of statistical methods, such as data mining and modeling, used to identify the factors that predict the likelihood of a specific result. They've long been a standard in the business world—both credit scores and car-insurance premiums are calculated with predictive analytic tools. Yet they have been slower to take hold in education.

“School districts are great at looking annually at things, doing summative assessments and looking back, but very few are looking forward,” said Bill Erlendson, the assistant superintendent for the 32,000-student San José Unified School District in California.

“Considering our economy survives on predictive analytics, it’s amazing to me that predictive analytics don’t drive public education. Maybe in education it’s considered a luxury, but it shouldn’t be; it should be a foundation for making decisions.”

Experts in predictive analytics in higher education and business say education may have a long way to go to develop the data infrastructure and staff capacity to make the tools useful on a broad scale.

“Good quantitative researchers are as hard to find in academia as Farsi linguists are for the military; we do not train enough researchers to work with these methods,” said Phil Ice, the vice president of research and development for the 90,000-student online American Public University System, in Manassas, Va. “There are plenty of numbers people, but they work in the corporate

sector, and they don't know how to apply it to the education sector. You have to understand the pedagogy, the social issues around education, and you have to understand the numbers.”

Mark D. Milliron, the deputy director for postsecondary improvement at the Seattle-based Bill & Melinda Gates Foundation, agreed. “I've heard again and again that there is a desperate need for funding data fellows and all sorts of people who can get out there in the field and do [education] research, but it's not just about doing traditional research; it's about doing more of this predictive modeling and analytics, because that's what's going to drive the scale and the implementation work to actually get people involved in the process. It's a different set of questions.”

While predictive analytics cannot say definitively whether an intervention causes a particular outcome, they show potential for K-12 education, because they can be faster and provide more practical results than traditional experimental methods such as randomized controlled trials, Mr. Ice said. “If you compare one class of 30 students to another class of 30 students that's one thing, but with this you can take an entire school district, taking all the data, not just a sample of it,” Mr. Ice said.

Predictive models might show the likelihood that a student with certain characteristics will excel in college or a teacher's credentials and instructional style will gel in a new school, but they aren't guarantees of specific results for individual teachers or children. Administrators have to walk a thin line between targeting support and changing expectations and opportunities based on predicted risks, according to Yeow M. Thum, a senior research fellow at the Northwest Evaluation Association, in Portland, Ore., who studies education growth modeling.

“Risk management and topics like that are really foreign to education research and management,” Mr. Thum said.

Predicting Graduation

One district that is taking the plunge into using predictive analytics for policymaking is San José. The district is modeling high school graduation and college-going trends based on 15 years worth of student academic, behavioral, social development and health data, as well as information on school climate from teachers, parents and students. The district is finalizing a risk-assessment protocol that identifies changing issues that contribute to a student's risk of dropping out at different grades. Each student has an individual profile, including the status of

key indicators like benchmarking and summative tests, behavior, attendance and health. Student data will be updated daily and monitored for accuracy, Mr. Erlendson said, but is not used for student report cards.

“It’s not quite a credit score, but it could be” like that eventually, as more indicators are added to the mix, he said. “Right now we’re on the cutting edge of what a student dashboard looks like. The indicators will be very clear: test scores, behavior, academics, health. The data there are going to be pretty solid. If the student is failing or is not showing up, that will be pretty obvious.”

Similarly, in 2007, Hamilton County schools, which serve 41,000 students in and around Chattanooga, Tenn., used districtwide data on student demographics, test scores, attendance and other information, comparing the graduation and dropout rates of students with different characteristics at each grade level to develop an “on-track” predictor of each student’s likelihood to graduate. Kirk Kelly, director of accountability and testing for Hamilton County schools, said individual schools have been focusing on the risk factors that carried the most weight at their schools.

For example, some elementary schools found that their students who failed early grades such as kindergarten or 1st grade ended up likelier to drop out in high school. “We found one of the factors was being over-age for their grade and separated from their peers,” Mr. Kelly said. “The student fails twice in elementary school, and then they get to high school at 16. A 16-year-old is not likely to stay in high school until they are 20.”

As a result, elementary schools began more-intensive monitoring and remediation for students at risk of failing an early grade, and the district started “adult” high school programs for students considered over-age for their grade.

It’s been four years since the district started using the tool to target interventions, Mr. Kelly said, and “the predictors have been good in how they play out. We’ve seen the tool starting to bear fruit.” The four-year graduation rate has increased steadily from 70.9 percent in 2008, the first school year of implementation, to 80.2 today, and the rate of students dropping out in each school year tumbled from 6.4 percent in 2008 to 1.8 percent today.

San José, along with the Dallas, Pittsburgh and Philadelphia school districts and the New York-based New Visions for Public Schools charter school network, are now participating in a three-

From

30 June 2011

www.edweek.org/ew/articles/2011/06/22/36analytics.h30.html?tkn=UPZFT33gpkdyX7M2yScE3ZbENPiYwX0eUrQ0&cmp=ENL-EU-NEWS1

year project analyzing school feeder systems to identify the elements at each school level that predict a student's understanding of college entrance, readiness for college content, and ability to complete a degree.

Data System Demands

Yet both Mr. Erlandson and Mr. Kelly cautioned that school districts must have a well-developed data system with several years' worth of data in multiple areas to get a nuanced prediction.

"We had to look at many things that could be predictive at different times," Mr. Erlandson said. For example, he said, "in 9th grade the most predictive factor is attendance—it really jumped out at us—but in sophomore year, the most predictive factor is grades. Public education is such a fluid thing, it's like trying to lasso an amoeba."

That complexity causes problems when trying to connect a child's experiences in elementary school with success in college 15 years later, Mr. Thum said. "We know that predictions on the near term are very useful; predictions in the far term are fraught with problems," Mr. Thum said. "In any prediction you assume that the future will be like the past, and that's a very large assumption to make."

Moreover, many predictive analytics systems in the business world, like that which provides a credit score, use the data to rank people based on the likelihood of a specific behavior, something Mr. Thum said generally is not appropriate for education.

"How people couch the accountability question will affect the results" of an analysis, he said. "Very often the folks hired into key policymaking positions probably do not have as much patience as they ought to have. Often ranking is all they are after; they're not looking at measuring student learning."

Finding Candidates

Yet sometimes, ranking can be useful, such as when a principal is trying to pick the right new teacher to fit into a school staff.

Teresa Khirallah, the senior director of Peak Preparatory, a K-12 charter school in eastern Dallas, has seen the hiring process work with and without predictive tools. When she took the administrative reins of the school four years ago, principals found new teachers by receiving

personal applications from teachers or combing through the central candidate pool for the school's parent organization, Uplift Education of Dallas.

“There was always some lag time, and you might be juggling seven to eight people. We spent a lot of time weeding people out, not really knowing from a five-minute phone conversation or email if they were the right person to continue this process,” Ms. Khirallah said. “There was just a lot of time spent on individuals who by the end of the hiring process you realized did not match the mission or would not be a good fit with your kids.”

Two years ago, with the 17-school network growing 25 percent or more a year, Uplift's Chief Executive Officer Yasmin Bhatia overhauled the network's hiring process in advance of hiring 160 teachers in a staff of 400. Uplift, working with the Emeryville, Calif.-based 3D Group, analyzed 44 different tasks that Uplift teachers perform, using a combination of surveys and interviews and classroom observations with the teachers considered to be exemplary based on qualifications, experience and recommendations.

The teachers rated each task on its importance to their daily job, the equipment or materials needed, and whether a new teacher should enter able to do the task or expect to learn on the job. Teachers also related specific examples of their biggest successes and mistakes in performing different teaching tasks.

After crunching the data, Uplift had a list of 29 ranked tasks in five teaching areas which formed the backbone of a new teacher hiring process. A potential candidate's written email and application essay questions can be rated according to those indicators, and principals follow formal interview questions, with answers also rated to make it easier to compare candidates.

“These are very specific situations we are asking candidates to describe for us; you cannot fluff our questions,” Ms. Bhatia said. “If you see that this person is only a 2.5 out of a 5 [in one of the five areas], it forces the discussion of, ‘Am I going to make the trade-off or am I going to hold off and keep looking for someone who is a better fit?’”

For principals, the change has clarified the hiring process, Ms. Khirallah said. “What I'm able to do as a principal is to spend more time talking with and training the right people, rather than spending a lot of time weeding people out.”

While Uplift is still measuring the achievement effects of the teachers hired under the new system, Ms. Khirallah said she has already seen a decline in teacher turnover. “We’ve noticed teachers are not only staying but are stronger and higher-performing teachers, and we’ve got a lot of like-minded individuals who share the same philosophy of education, so you have fewer issues at the end of the year about how people fit.”

Ms. Bhatia said the network has already started to use the initial task scores to tailor professional development based on a teacher’s areas of strength and weakness, and said the schools hope to eventually have more fine-grained data that will help match teachers to specific subjects or grade levels within schools. The May issue of the *Harvard Education Letter* noted Uplift’s predictive analytics show promise in helping districts match educators to the schools where they will be most effective in teaching.

Analyzing Personnel Policy

Hamilton County doesn’t use predictive analytics to hire or place teachers as the Uplift charters do, but it does use the tool to analyze personnel policy. For example, Mr. Kelly said he was able to track the broader-than-expected effects of teacher absenteeism on student performance throughout a building. “If you see a teacher is out and there’s no substitute, that may not just affect that teacher, but might also affect ... other teachers, because we might have to split a class and not be able to do the lessons we had planned.”

As a result, the district has moved to a 95 percent teacher attendance goal and is putting into place more support structures to account for teacher absences.

Next year, the district plans to expand the system to provide online accounts at each school so that principals and teachers can run small-scale predictive studies or simply get more frequent updated reports on their students.

Higher education has started to explore the validity of using predictive analytics for long-term goals like college success. For example, the Western Interstate Commission for Higher Education has won a \$1 million grant from the Seattle-based Bill & Melinda Gates Foundation to develop a Predictive Analytics Reporting Framework using data from more than 400,000 students in six college systems with online students: the American Public University System,

Colorado Community College System, Rio Salado College based in Tempe, Ariz., University of Hawaii System, University of Illinois Springfield, and the University of Phoenix.

The online colleges are looking at 34 different factors that may contribute to a student's ability to enter, stay and succeed in college.

"The methodology we're applying could certainly be used in the face-to-face campus and in the K-12 arena as well," said Mr. Ice, the principal investigator for the study. "I think that the K-12 environment may even be a richer source of material than the university setting; in K-12, we know so much more about what our students are doing and where they are coming from," he said.

"It takes a lot of computing power and there's some pretty high-level math," Mr. Ice said, but, "it's a matter of political will more than math or statistics." With the right resources and access to the data, he predicted, widespread use of predictive analytics could become a reality in K-12 education within four years.

Vol. 30, Issue 36