Smarter Education

Predictive analytics can help schools quickly identify at-risk students—so educators can intervene before it's too late

For years, marketers have used sophisticated software to track consumers' buying habits and web browsing activity, then crunch this information and—based on the data—make a series of intelligent predictions that allow them to target their sales messages much more effectively.

Now, this same technology is appearing in schools and colleges as well—and observers say it's a development that could revolutionize education.

Using predictive analytics software, teachers at

Georgia's Gwinnet County Public Schools soon will be able to see at a glance which students might need more help. Sinclaire Community College in Ohio has cut its student dropout rate in half. And the online American Public University System has watched its course completion rate steadily climb.

These breakthroughs come at a key time for U.S. education, which is under enormous pressure to innovate and provide better learning opportunities.

Smarter Education...

As school district officials struggle to meet the goal of having all students graduate from high school ready for college or a career, the challenges are significant: Operating costs are on the rise, while budgets for public institutions are shrinking. Infrastructures are aging and need costly updates. Changing demographics require that schools change, too, to meet the shifting needs of students. Performance is declining at the same time that expectations are rising. And "working harder" is simply not a sustainable option.

"When we talk with government policy makers and senior education leaders, there's a recognition that education is the differentiator for national success. Everyone recognizes that education is critical, but [schools] still get their budgets cut all the time," said Michael King, vice president of global education for IBM. "People in education really grapple with that problem."

To overcome these challenges, the education field needs new and innovative approaches. And predictive analytics is one such promising solution.

A smarter approach

Predictive analytics encompasses a variety of statistical techniques for mining information gathered across a variety of sources—for example, student information systems, library automation systems, learning management systems, and back-office enterprise systems—and analyzing those current and historical data to make predictions about the future.

The implications of predictive analytics for education are nearly endless. For instance, schools are using analytics software from companies such as IBM or SAS to track student performance over time, looking at various data points—not just test or quiz scores, but other, more subtle signals as well, such as how frequently students are logging on to an LMS, or how often they've contributed to online discussions—to identify those who are at risk of failing or dropping out. Some schools are overlaying a system that can identify these early warning signs automatically and create a customized intervention plan for students who might need it.

"Say a student comes into ninth grade with a D in eighth grade algebra. Using statistics, you can see that someone who made a D in Algebra 1 is not ready for Algebra 2. But you might also see that if they're not doing well in math, they're probably not doing well in science," said J. Alvin Willbanks, superintendent of Georgia's Gwinnett County Public Schools. "Attendance patterns can also be shown: Maybe that student was out of school 20 percent of the time. This could indicate there's a discipline problem. Using analytics can alert teachers to these things, so they can be proactive in putting interventions in place to get the student back on track."

Working with IBM, Gwinnett County is creating what is calls its digital content learning assessment support system, or eCLASS. When it is completed, eCLASS will give teachers the ability to open a page and see their class roster in a dashboard format. At a glance, teachers will get important information about their students, with the ability to drill down to see individual student results and suggestions for addressing each student's academic weaknesses.

In higher education, Sinclaire Community College used analytics to look at registered students who hadn't paid yet, perhaps because their grants or loans had not yet come through. Those students were in danger of having their registration cancelled. By doing the analysis and intervening with these students before cancellation occurred, the college was able to halve its student registration dropout rate. That allowed the school to hold onto its state funding and ultimately increase its graduation rate.

The 70,000-student American Public University System, a fully online school, uses IBM's SPSS Modeler to measure key student performance, participation, and attendance information to predict when students are in danger of dropping out, so those students can be given additional support. APUS also is part of a national initiative to study the factors influencing college dropout rates, so colleges and universities can deploy analytics technology more effectively in their student retention efforts. (See the side story "Analytics use boosts student retention.")

Ed-tech group to push for more analytics use in colleges

On the eve of its annual conference, EDUCAUSE announced plans to bring more analytics technology to campuses nationwide

Dennis Carter
Assistant Editor

Analytics programs that use complex data sets to identify struggling students, suggest the best ways to use campus budgets, and improve faculty and staff efficiency soon could be coming to more U.S. colleges and universities, thanks to a new initiative from the higher-education technology group EDUCAUSE.

EDUCAUSE, a nonprofit organization that held its annual conference Oct. 18-20 in Philadelphia, announced a three-pronged initiative to bolster the use of analytics technology in higher education.

The initiative will include a "major benchmarking study of the state of analytics in higher education" that could show how many campuses are using analytics to improve decision making and student success. It will end with a national summit for college and university leaders in fall 2012, according to the group's announcement.

In between, EDUCAUSE will work within the higher-education community to provide thought leadership and training to help schools develop the capacity to use analytics.

"Sophisticated uses of data and benchmarking have revolutionized business and industry," said Diana Oblinger, president and CEO of EDUCAUSE. "These same techniques can deepen our understanding of learner success, improve our ability to track progress, and help us make the best decisions about campus resources."

Campus technologists at the forefront of analytics use have discovered the benefits of computer-based models that show precisely how to save money and improve student grades, but Oblinger said analytics is far from mainstream use in higher education.

"Analytics holds transformative promise for education, but the field is still in the developmental stages," she said, adding that EDUCAUSE's initiative "comes at the right moment for higher education."

Analytics-based early interventions have improved course completion rates demonstrably, EDUCAUSE said. In the future, analytics will enable highly personalized learning environments characterized by adaptive testing and content displays, point-of-need help, and academic support services tailored to individual student needs.

A whitepaper published by EDUCAUSE last year described how a classroom analytics program used at a large state university helps instructors identify which students are likely to fail their class by analyzing classroom participation—including contributions to the class's online discussion board—as well as quiz grades and the number of times a student views course notes online.

The software analyzes these variables and uses a color-coded indicator to let professors know which students are most at risk. A green light shows the student is doing well in class, a yellow light indicates "possible risk" of slipping below a C, and a red light shows "elevated risk" that a student could fail the course after failing to participate in in-person and online class discussions and scoring poorly on quizzes.

The program also shows instructors which students are participating in course discussions the most, and correlates this information to class grades. Professors can send an eMail message to notify students if they appear at risk, congratulate students for


Analytics holds 'transformative promise.'

work well done, or suggest ways to get more from the class.

"Most colleges and universities collect and store vast amounts of data—in the LMS, admissions files, student and library records, and other systems," the report said. "Analytics applications can mine at least some of this data, subject it to statistical analysis, and prepare reports or data visualizations to reveal patterns, trends, and exceptions."

The new EDUCAUSE initiative is supported by a grant from the Bill & Melinda Gates Foundation.

*e*SN
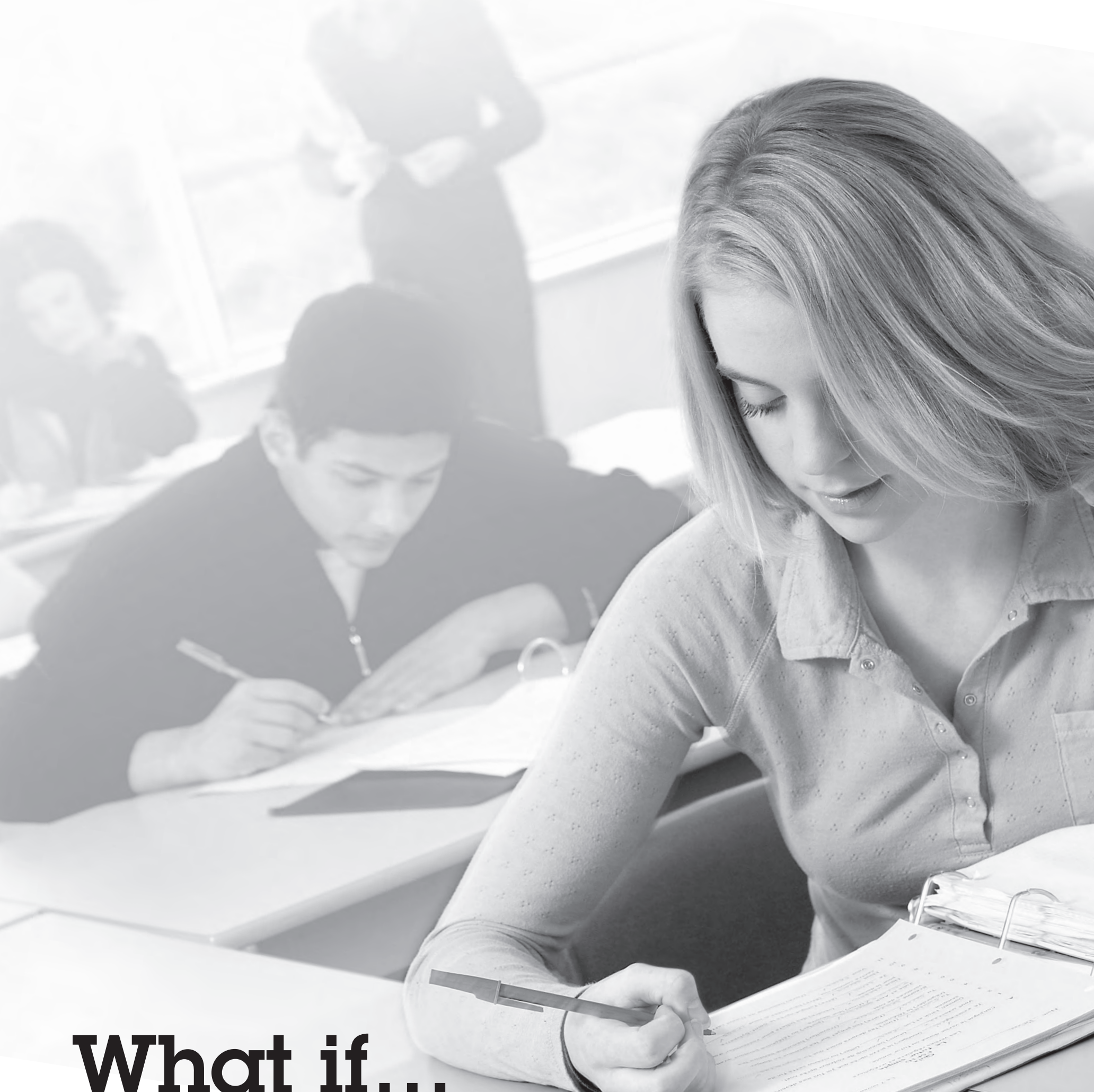
What if...

you could predict which students would need help making the grade?

SPSS, an IBM Company has innovative, cost effective solutions so you know which students need help...and in what areas...long before they actually fail.

Visit our free resource center: www.spss.com/academic

SPSS®
AN IBM® COMPANY

IBM®

Smarter Education...

While looking at a student's prior grades to see how well he or she will do seems rather obvious, predictive analytics gives educators the ability to look more deeply into the data, said Karen Patch, senior technical architect for SAS. "What you're able to find is hidden trends and patterns that you otherwise wouldn't be aware of," Patch said.

Predictive analysis has been used successfully in financial, insurance, retail, and other industries for years. "This is how you get targeted ads over the internet," said Alex Kaplan, national practice leader for education at IBM's Global Business Services division. "These firms are mining data and looking for patterns, such as how likely a person is to make a purchase. It's a sophisticated use of data, and it can be directly applied to education."

Sometimes the technology reveals key information that might come as a surprise, helping educators look at situations in a whole new light.

"One really big issue is how engaged students are in school," Kaplan said. You might think that the more involved in activities a student is, the more likely he or she is to fall behind academically. But Kaplan said schools are finding that the opposite is true. The more

engaged students are in school—using social collaboration tools where they can chat with each other online, accessing websites that contain curriculum materials and lesson plans, spending time in other activities such as drama or sports—the more likely they are to succeed in their studies.

"You might have students who are performing poorly academically, but who are really engaged in school, so you know they're excited about school but are maybe struggling in a certain topic," Kaplan said. "Or, you might have someone who is scoring well but is really not engaged at all. That student might actually be at risk of dropping out and would normally fall through the cracks. We're now in the position to give this information" to school leaders before it's too late, he said.

Now, Kaplan said, instead of relying solely on their intuition and observation, educators and administrators can have quantitative data to help them make better, smarter decisions. "And it's not just about the individual student; it's also about how the school thinks about instruction," he said. "Once [officials] identify patterns and trends, they might learn that everybody is struggling with the periodic table in chemistry, so it gives them a deeper insight into the instructional process."

IBM has worked with industry leaders for years, helping businesses use predictive analytics as a data-driven system for managing risk and improving their

return on investment. In fact, the company says, in one recent survey 90 percent of respondents said they had attained a positive ROI from their most successful deployment of predictive analytics—and more than half achieved a positive ROI from their least successful deployment. Now, IBM has developed what it is calling a "vision for smarter education," creating an analytics framework for schools that builds upon its expertise in analytics, business processes, and technology integration.

IBM's new framework combines predictive analytics with "intervention management" technology, which can trigger a specific intervention that is unique to each struggling student's needs and then deliver this remedial or supplemental content directly to the student. The entire process occurs through a single dashboard interface, IBM says; the solution is in beta-testing now and will be commercially available for schools and colleges early next year.

Analytics features are even appearing in popular LMS programs. Instructure, whose Canvas LMS program is available as either an open-source version that schools manage themselves or a cloud-based model hosted by the company, plans to release a version in early 2012 that includes predictive capabilities.

On their course roster, instructors will see green, yel-

Analytics use boosts student retention

Dennis Carter
Assistant Editor

Phil Ice knows numbers never lie.

Ice, the director of course design, research, and development for American Public University System (APUS), has watched retention rates at the 70,000-student online school steadily climb with the continued analysis of in-depth information that shows when a student might be on the verge of dropping out.

If a student's test scores are dropping, participation numbers are low, and disengagement is evident through various statistics, the numbers suggest that student might not last much longer at APUS.

How can professors and university officials know precisely which students are in danger of giving up on their education? One surefire strategy is to examine how many days have passed since the student last logged onto his or her course website.

If it's been a while since the online student checked the site for syllabus updates or discussion sessions, APUS's analytics system will flag the student as a potential dropout.

"We hone in on things such as a student's perception of being able to build effective community," Ice said.

APUS professors and instructors use information detailing a student's online engagement with a comprehensive survey to create a model of student retention and satisfaction, according to the university.

"We've been using analytics before it was a buzz word in higher education," Ice said.

APUS uses IBM's SPSS Modeler, which uses key student performance, participation, and attendance information in part to measure a student's social presence, along with a student's perception of online learning's effectiveness.

These two factors have proven to be reliable variables telling professors and instructors how likely a student is to drop out of classes.

APUS recently joined the Colorado Community College System, Rio Salado College, the University of Hawaii System, the University of Illinois-Springfield,
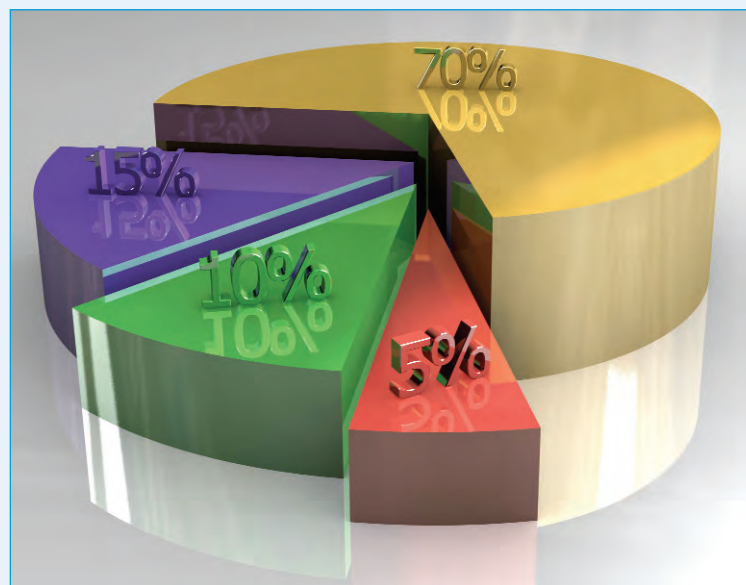
and the University of Phoenix in a nationwide initiative aiming to improve analytics and increase its use in higher education.

The Predictive Analytics Reporting (PAR) Framework project, launched last May by the educational technology group Western Interstate Commission for Higher Education's Cooperative for Educational Technologies (WCET), will examine six critical data sets as a single sample that could make analytics more effective on college campuses.

The six colleges and universities participating in the PAR project will create a pool of information from 400,000 students who will remain anonymous in the research. Analysts will use this massive collection of student records to assess factors that affect retention and learning outcomes among online students.

Ice said about nine in 10 colleges use some form of statistical analysis to determine retention and learning strategies on campus, but only 9 percent use historical data to supplement current numbers—and only about 1 percent of U.S. colleges and universities use "deep data mining" to explore why dropout rates might be rising.

"That number really needs to expand, especially around the crisis of student enrollment," Ice said. "We

need to let the data speak for themselves."

Campuses large and small have not been immune to spiking dropout rates.

University of Kansas officials worked with a data-mining company in 2010 to pinpoint strategies to keep students enrolled after a university report showed that 28.7 percent of freshmen from the fall 2007 semester have left the campus.

Kansas's 28.7 percent dropout rate among fall 2007 freshmen was significantly higher than peer institutions, said Christopher Haufler, a University of Kansas professor who chaired the school's student retention task force.

The university's average retention rate after one year is 80 percent, he said, compared with 85 percent to 90 percent at peer institutions.

Kansas's retention system uses analytics to raise real-time online flags for students who fall below a certain grade or involvement "threshold" designated by the university.

Nationwide, fewer than three-fourths of two-year career college students return to school after their first year, according to research released by the nonprofit Imagine America Foundation. Just 57 percent of public community college students return after one year, and 68 percent return after a year at a private institution, according to the research.

Other schools, such as Western Iowa Tech Community College (WITCC), use student tuition payment data to gage how likely a student is to leave school.

Educators often have a difficult time tracking the engagement level of online students, who—unlike traditional students—don't interact with their professors and fellow students every day in class. Analytics tools, like the kind used at APUS, could make that task much easier, Ice said.

"By identifying patterns of performance, we believe we can create an approach to applying predictive analytics … that will help practitioners and students alike spot barriers to success before they become problems," he said.   **eSN**

Education for a Smarter Planet

On a smarter planet, educators and government leaders must enable student success. Through a unique combination of industry experience and expertise, IBM is helping education institutions improve operational effectiveness and accelerate innovation to contribute to economic development. With deep knowledge of the personal learning paths through the educational continuum and powerful technology, IT services and business consulting, we are equipped to help provide innovative solutions to institutions, cost efficient sharing of best practices and facilitate collaborative communities. IBM has the tools, technology and the people to help you meet the challenges of today's education.

A Smarter Planet starts with a smarter education system.
Let's build a Smarter Planet.

ibm.com/education

Analytics improves effectiveness of Minnesota schools

Like most states, Minnesota contains many school districts that are quite small; one district, for example, has just 500 students. Fortunately, the state is home to TIES, a joint-powers cooperative that is owned by 46 Minnesota school districts. TIES (which is actually a school district by charter in itself, though it doesn't have any students) offers analytics packages to its members via Cognos, a company that was purchased by IBM in 2007.

TIES gives districts the opportunity to use high-powered tools they otherwise wouldn't have been able to afford, said Ben Silbergitt, the organization's director of software applications. "We help districts set benchmark target scores on local assessments, so they can use those assessments as indicators for state assessments," said Silbergitt. "This helps them identify early on whether students are on track, not just by student but by student population."

Schools then are able to pinpoint where they are effective and where they are not. "For example, they might learn how effective they are with English language learners who are below target. They might learn they are really effective with Spanish-speaking students, but not with students who have other languages spoken at home," he said. "Or maybe it's attendance: This particular population has a low attendance level and is having trouble in school. That helps [the schools] plan for student success."

TIES began using Cognos technology seven years ago, but the predictive analytics element has significantly come into play in the last five years. TIES created a data warehouse to give its member districts not just easy access to their data, but the ability to connect disparate sources of data—looking, for instance, at how attendance impacts instruction. "Enterprise software systems are not designed to get data out in order to answer those kinds of questions," Silbergitt said. "So putting that warehouse layer on top has been great in terms of answering those questions on the fly."

The districts do not have to approach TIES with their questions; TIES provides not only the software, but also the training to use it effectively. "When districts want to move forward with a certain model, we show them the process behind answering those questions, such as what types of reports they need to run," he explained.

The school districts that have taken advantage of their access to predictive analytics have found that using data in this way has made them much more effective.

"They can make decisions more quickly and are more proactive about changing their approaches. They might discover, for example, that they need to apply more of speir resources to the third grade, because they have a cohort there who hasn't grown as much as the other cohorts in the last two years. When they learn that, they know they need to shift things around. And we know that leads to better outcomes," Silbergitt said. —J.N. **eSN**

Smarter Education...

low, or red dots next to students' names, indicating how at risk they are of failing or dropping the course. Clicking on a student's name will take instructors to a page where they can see more detailed reports based on that student's grades, class participation, assignment completion, and outcomes (whether he or she has mastered the content).

At the bottom of the course roster, instructors will see a list of students who are most at risk in the class. The software also will contain dashboards for administrators to view the same information for entire departments or schools.

Getting started with predictive analytics

For school district leaders interested in using predictive analytics, the first step is to understand the kinds of data you want to analyze. Then, you must create a data warehouse that pulls this information from all

technology group EDUCAUSE noted that analytics technology is a powerful tool that can help schools "identify where and when certain investments will have the greatest benefits." But the organization cautioned that even the best data algorithms can result in misclassifications of students, "in part because such programs are based on inferences about what different sorts of data might mean relative to student success."

In other words, predictive analytics software is only as useful as the data it draws from—and the suppositions that can be made from how those data interrelate.

That's why it's important to have as complete a picture as possible of a student's academic history. And that's the motivation behind states' efforts to create longitudinal data systems that can follow a student's progress from pre-kindergarten through college graduation and on to the workforce—a development that IBM's King finds encouraging.

"If you can track a student across multiple schools, you have a common view of the student that can help drive understanding," he said.

Predictive analytics software can be a powerful tool, but it's only as useful as the data it draws from.

available sources, including student information systems, learning management systems, enterprise software, and other areas. And the data need to be both consistent and trustworthy. "Having this information all in one place gives schools an enormous amount of leverage," IBM's Kaplan said.

The next step is to think about the key performance indicators you want to be able to compare. Some commercial analytics software might have certain options built into the system, while others might give users the flexibility to design a nearly unlimited number of queries themselves. Whatever option you choose, you should recognize that your needs are likely to evolve as you dig deeper into using the system.

Education leaders also should consider what actions, if any, they want certain indicators to trigger, such as sending a message about tutoring options to a student who is considered at risk of failing. Again, some software programs might contain certain built-in intervention choices, while others might not.

In an April 2010 whitepaper called "7 Things You Should Know About Analytics," the higher-education

Other countries are starting to look at how data systems can help their educational systems fuel economic growth, by identifying students coming through the pipeline as individuals with particular skill sets that can supply the country's needs in various areas. By taking a "P80" view of students—that is, a single longitudinal view of each student from pre-kindergarten through age 80—policy makers and education leaders are able to look at the "supply chain" of people moving into the workforce and discern whether there are enough teachers or nurses, for example.

In order to compete globally, this is the direction the United States needs to take as well, King argues: "It's important that our educational leaders and policy makers recognize that that's where we're heading. It's a completely different competitive dynamic. We have to start thinking not just about our schools, but … how we transform our state systems by bringing together a P80 view of each student." **eSN**

Jennifer Nastu is a freelance writer who writes frequently about education and technology.

Texas system provides early warnings to help stem dropouts

Randy Sumrall, chief information officer for the Education Service Center (ESC) Region 10 in eastern North Texas, knows the importance of case management for individual students—that is, looking at where at-risk students are in their education, discovering where intervention needs to take place, deciding what that intervention will be, and giving them the additional instruction and follow-up they need.

So he was happy when, in 2005, the Technology and Data Services division of ESC Region 10 partnered with IBM to provide thought leadership to its client districts on the use of data warehousing and business intelligence to become more data-driven, and he was hired to lead the initiative.

ESC Region 10 is one of 20 education service agencies for the state of Texas. It provides administrative, instructional, and information services to 80 districts and 32 charter schools. ESC Region 10 and IBM built Empower, an educational data warehouse that could draw from multiple silos of data, such as student performance, student demographics, finance, and human resources, and provide comparisons across multiple school districts. The system runs on technology from IBM Cognos.

ESC Region 10 piloted the system with nine districts and began production with 15 districts. The biggest element the districts looked at was student performance. For example, said Sumrall, one district gave its principals a "snapshot" of their students multiple times through the year to pinpoint the students' specific needs.

But at that point, Region 10 discovered the piloted districts were all it could handle. "Without charging unaffordable fees, we could not scale the system," said Sumrall. "The data were much more diverse than we expected, with no data standard, even within a single student or business database vendor."

At that time, the Texas Education Agency and the Michael and Susan Dell Foundation announced their intention to create a state data warehouse based on a data standard to be developed. The state data warehouse would provide student performance dashboards throughout the state.

"We knew immediately that this would become the solution we were pursuing," Sumrall said, and they transferred their intellectual property and assistance to the new Texas Student Data System. Now, Sumrall and the ESC Region 10 are helping to support two districts in a limited production of the Texas Student Data System. One solution stemming from the state system is the Reveal Dropout Early Warning System. It uses data from the state data warehouse to provide a comprehensive view of students at risk of dropping out, automatically and dynamically.

"At Region 10, we believe the data warehouse is only a part of the needed school information solution. There are four major areas where school information needs to be managed, and they are interconnected," said Sumrall. "They are data warehousing and reports on student and teacher performance, predictive analysis of potential instructional and intervention solutions, management of content that is correlated to state standards with qualified built-in assessments, and case management to unify the efforts of multiple specialists that wrap around the needs of the individual students."

As the Texas Student Data System is being developed, ESC Region 10 is helping its districts understand what data they have, what data they want, and what they want to use this information for. "If you want to be a data-driven district, you need to know what you want to collect. Often there's a total disconnect between getting the data and knowing how to use it," explained Sumrall. —J.N. *e*SN