

Colleges Are Using Big Data To Predict Which Students Will Do Well--Before They Accept Them

Can predictive analytics determine which students succeed and which will fail? More universities are finding that the answer is yes.

Students at America's high schools, colleges, and universities are well into their first semesters. But while they plow through their assigned readings and write essays, administrators are turning their grades and their professors' evaluations into millions upon millions of tiny data points. Much like every other field in the world, education is embracing big data--only, this time, they're using it to determine who will thrive in college, who will fail, and who will need some extra help.

[David Wright](http://www.fastcompany.com/person/david-wright) ([//www.fastcompany.com/person/david-wright](http://www.fastcompany.com/person/david-wright)) is [Wichita State University](http://www.wichita.edu) (<http://www.wichita.edu>)'s (WSU) Associate Vice President for Academic Affairs. In his position, Wright is responsible for overseeing the vast amounts of data WSU uses to track student and faculty performance. Like a growing number of American educational institutions, Wichita State uses predictive analysis tools to optimize their offerings and steer help to students who need it.

"We know our data better than an outside agency. We know the business practices in our system better, which outside vendors don't do, and this allows us to do more with the data than them," Wright tells Co.Exist. Using data points such as a student's paper grades, the amount of hours he or she is enrolled during each semester, whether they're working part-time or full-time or not at all, the amount of assistance from family and a host of other factors, WSU can predict which students are likely to encounter problems.

“ ”

**WSU's recruitment model
had 96% accuracy identifying
“high-yield” application
prospects.**

WSU has used predictive analytics software for the past several years. According to an [IBM white paper \(http://www-01.ibm.com/software/success/cssdb.nsf/CS/STRD-8NVMCD?OpenDocument&Site=default&cty=en_us\)](http://www-01.ibm.com/software/success/cssdb.nsf/CS/STRD-8NVMCD?OpenDocument&Site=default&cty=en_us), the university's decision to implement a suite of IBM business analytics software in the school's admissions department also helped predict the success rate for incoming students. The university decided to use the company's analytics package instead of hiring external consultants to appraise incoming students. In a twist, it turns out the analytics software suite was better than the human consultants at predicting which students would succeed at Wichita State. According to IBM's data, WSU's recruitment model had 96% accuracy identifying “high-yield” application prospects compared to 82% by the external consultants.

But Wright added that integrating big data into the school's practices took delicate work in terms of preserving institutional culture. “Implementing analytics requires getting in people's business. Many units on campus that are overburdened or have low resources could see it as a burden. It means more requests for work, strangers showing up at their meetings, etc.,” he said. “But in terms of admissions, I saw they needed some information that they couldn't get their hands on through external consultants, so I told them that if I worked with them I'd help them find benefits.”

“ ”

**Implementing analytics
requires getting in people's
business.**

Katharine Frase, IBM's CTO of Global Public Sector, also notes that predictive analytics use by universities isn't limited to admissions departments. She says that big data suites help schools "ask questions they didn't know they needed to ask," and shows use in everything from scheduling classes to identifying possible at-risk students in high schools. In one example of combining big data and predictive analytics, Frase noted that big data platforms can parse students interests and past academic performance, and recommend assistance for that particular student's needs.

Another IBM executive, VP of Global Education Industry Michael King, said that, in education, big data and predictive analytics techniques are in their infancy. But he believes they will transform the industry as much as they have transformed health care. "We're only just now starting to work through what opportunities are for leveraging big data in education, and health care is a model for this. We provide prescriptive solutions to help make recommendations to clinicians around potential treatment opportunities, how to intervene with certain patients, and we see parallels in education where we can use data to personalize the educational experience," he says.

"The right set of information is everything. I think that, looking at lifelong learning and using data to help provide clearer pathways to students for a multi-institutional education plan, using tools similar to like Watson, is an important goal. We want to show how to put more tools in their hands for broad data. We can give prescriptive data to save time intervening for individual students."

Predictive analytics tools from IBM and others are starting to become more commonplace in higher education (<http://www.centerdigitaled.com/news/The-Future-of-Predictive-Analytics-Higher-Ed.html>)--sooner or later, the tools used by schools like Wichita State will become the norm.

[Image: Flickr user Mitchell Joyce]



[http://www.fastcompany.com/user/Neal-](http://www.fastcompany.com/user/Neal-UNGERLEIDER)

NEAL UNGERLEIDER

(HTTP://WWW.FASTCOMPANY.COM/USER/NEAL-UNGERLEIDER)

Fast Company journalist specializing in future technology and global innovation. [Continued \(http://www.fastcompany.com/user/Neal-ungerleider\)](http://www.fastcompany.com/user/Neal-ungerleider)

[Instagram \(http://instagram.com/Nealungeriourno\)](http://instagram.com/Nealungeriourno)

[Linkedin \(http://www.linkedin.com/in/85364308\)](http://www.linkedin.com/in/85364308)

[Twitter \(http://www.twitter.com/Nealunger\)](http://www.twitter.com/Nealunger)

[Google+ \(https://plus.google.com/u/0/112680374414703071392/posts\)](https://plus.google.com/u/0/112680374414703071392/posts)

October 21, 2013 | 8:47 AM

YOU MIGHT ALSO LIKE



FROM THE WEB

The Latin American and Hispanic Digital Opportunity: Are You Prepared?
Outbrain

<http://www.outbrain.com/blog/2012/06/the-latin-american-and-hispanic-digital-opportunity-are-you-prepared.html>



This Is What Everyday Dudes Look Like Posed As Underwear Sex Gods
Go. Go. Go.

[?]

ADD NEW COMMENT

LOG IN

Type your comment here.

6 COMMENTS

We're talking about kids not data points. I can see the attraction of this to data scientists and IBM, but that's a crowd that never found a bunch of data they didn't like. Who needs human contact when you can dial it up in the machine.

Do you remember how much you changed from the time you applied for college and the time you succeeded in finding a career you like. There is something to be said for human experience and judgement when dealing with kids and their future. Between SAT and GPAs numbers are already overweight.

Why bother with all the college app drama when it's predetermined by the data? Just log into Google, dial the quality selected down until you hit the perfect number of candidates and send them a ticket. And who adjusts for affirmative action, diversity and how are the athletic scholarships going to be figured in? That will be a fun lawsuit. Your college will be determined by big data? Let's talk about those FB comments and the NSA data they were able to factor into the analytic. Who is going to tell you when you were rejected on the basis of mistaken identity or more likely crappy data?

There are so many things wrong with this, not the least is that most of these schools will take anyone that's ready to pay. The algorithms are for the financial aid kids. And Wichita State is the poster child for this?



[Share](#) [Link](#) [Reply](#)

DIEDRABARBER > **NEAL UNGERLEIDER** 3 MONTHS AGO

I am all for innovation and the global advances technology it is said to offer all people across multiple platforms. I am interested in the increase in access and opportunity advances in technology, I am told it holds, for disadvantaged and underrepresented communities.

I agree with some of the comments here and wonder exactly how the use of predictive analytics support or disrupt an already allusive for many, ability to access higher education. This like the

predictive analytics that are being designed to determine one's likelihood to commit a crime are problematic simply because it is born of human perspectives colored by a zeitgeist that is proven to be harmful to those with less money, less education and less privilege.

If this is the lens by which we create these innovations, we must be very careful to think about the ethical and moral dilemmas likely to arise from their use. I am not arguing against their use but to ask the we for once, look at the larger picture and think about what it is we are creating exactly.

Who exactly (race, class and gender) is gaining admittance? Who exactly (race, class and gender) is being rejected? Until we know the answers to these two questions, the use of predictive analytics in higher education admittance is highly problematic.

The future is innovative and bright but the reality is, not all of us need to wear shades.



 [Share](#) [Link](#) [Reply](#)

RICHARD TERRY > **NEAL UNGERLEIDER** 3 MONTHS AGO

Several thoughts ..

First one is to what end? To help students succeed, or to exclude students based on datapoints?

It's important not to marginalize or commoditize people. Analytics in a group is very predictive, where at the individual level it is not.

Second, and to one of Mister Wrights point, input equals output quality. If your input people are less than vigilant.. you have a problem. The several organizations I have worked with use students supplement staff

for input jobs. Input quality is not always as high a priority here as in Health Care as a comparison. Health Care data is much more disciplined, if imperfect. Data quality is weighted toward the use-case data quality.

Third, in many organizations the Big Data Analysts, are more often than not, also your Technologists. As in DBA(s) are the same persons or group. It's not a new role to DBA's of course. With Big Data the analysis can be a split brain .. again not new. With the scope change with respect to volume of data, are those limited resources prepared to expand. To keep the integrity of the Data Analysis it can be more effective to separate Analysis from Technology roles.

The word used was "infancy", and will have growing pains. To many times now I have seen conclusions reached by a group, based on a set of assumptions reached by using invalid data and or sample sizes. It can be a house of cards, or not. GIGO (old school), is still valid.

Is it a good idea? Sure. If you use metrics manage you data and process quality.



 [Share](#) [Link](#) [Reply](#)

ABAKKER > **NEAL UNGERLEIDER** 3 MONTHS AGO

What about type 2 error? a 96% success rate at choosing successful students for the college, but what about students passed over incorrectly? It seems moderately deterministic to take a student body that will already be successful and then claim to have made them so.

I'd love to know what happens to the pool of people that they fail to choose, to see how many "succeed" in the long run, in other schools as well as life in general. I'd also love to see what they define as a 4% failure.



[Share](#) [Link](#) [Reply](#)

JSPRAG > **ABAKKER** 3 MONTHS AGO

That would be an interesting piece of information.

Although it seems to me (assuming that there are more applicants than openings), that later "success" by a rejected applicant is, by itself, not indicative of a Type II. Only when a rejected applicant turns out more successful than the least successful selected applicant should we categorize his/her rejection as an error.



[Share](#) [Link](#) [Reply](#)

THE WRONG AGENCY > **NEAL UNGERLEIDER** 3 MONTHS AGO

It strikes me that Big Data does hold the promise of changing the way that academic institutions are ranked.

When you can assess the relative ability & outcomes, for each student, factoring in all environmental factors - their social & economic background, family role models, whether they have to work to keep themselves in school etc. The quality of teaching & support at the educational establishment could be determined by where they have taken the students from & where they take them to.

For, if large swathes of a student population, at a particular establishment, come from economically & socially disadvantaged backgrounds, no members of family having attended university & work their way through college - then a 2nd class degree could be a triumph.

Whereas if your core population are from more affluent, aspirational families, with the right connections, and schools behind them, then expectations and ratings should be adjusted accordingly.

However, you get the feeling that the Big Data will operate along the lines of 'Gate Keeper' - rejecting those students who aren't deemed to be 'High Yield' prospects, and in doing so, protecting the academic standing of each institution.

While some may cry 'Not True' - the fact that many universities, particularly Ivy League & Russell Group establishments, operate as massive fundraising machines - and you can see what would be the attraction of such data in helping them protect their wealth. And academic ranking.



 [Share](#) [Link](#) [Reply](#)