Predictive Analytics

\[ \frac{P(x)}{1 - P(x)} = \exp\left( \sum_{j=0}^{K} b_j x_j \right) \]

Steven Shigemoto
John Stanley
Russell Uyeno
Honolulu Community College
Overview

- What is “predictive analytics”?
- Why and how is it used?
- Issues and considerations for higher education
“Analytics is the use of data, statistical analysis, and explanatory and predictive models to gain insights and act on complex issues.” (Educause, 2012)
What is Predictive Analytics?

- Range of analytical methods
  - Focused on the future (predictive)
  - Based on (big) data (analytics)

- The nature and uses of the “predictive” part of predictive analytics vary:
  - “Models” of causality for complex problems
  - Artificial intelligence-type of targeted interaction or intervention
Data: The Essential Ingredient

The Digital Universe Growth Paradox: Falling Cost and Rising Investment

- **Cost per GB**
- **Total Investment ($T)**

Sources of Data

- Application and enrollment
  - Demographic
  - Historical
- Registration
- LMS (Laulima)
- Student ID-based activities
Modeling

- Aimed at identifying causal factors
- Based on traditional statistical modeling methods (esp. logistic regression, time series analysis)
- Theory based (sort of...)
UHM Freshman Retention Analysis

These variables account for approximately 40% of the variance in a student’s likelihood of completing a degree (Pseudo R Square = .402).

*Wald statistic (sig.)
The Wald test statistic was used to indicate strength of the variable instead of the coefficient, standardized beta. Because of the nature of the logistic regression, the coefficient is not easily interpretable to indicate strength.
Analysis: HonCC Retention

Data: Fall 2011 AtD cohort (975 students)

Question: What student characteristics are related to reenrollment in Fall 2012?
Variables Tested

- Age
- Distance from HonCC
- Sex
- Previous Credits
- LA/CTE
- Credits Attempted
- Dev English
- Dev Math
Targeted Interaction

- Aimed at anticipating needs and directing behavior at the individual level
- Algorithm-based
- Non-theoretical
Target

Kids' Home Event—spend $75, save 15% on kids’ furniture, décor, bedding & more. kids’ furniture & décor

hit the beach in style with exclusive swimwear as seen in Sports Illustrated. from Sports Illustrated

sign in new guest? my account REDcard

women men baby kids home patio furniture electronics entertainment toys health & beauty clearance more

daily deals. one day only. always free shipping.

find a store Weekly Ad GiftCards registries TargetLists

shop categories
activity gear & toys baby bath boys' clothing boys' shoes car seats diapering feeding girls' clothing girls' shoes health & safety infant carriers nursery strollers

the one-stop shop for top registry brands.
Graco Eddie Bauer Chicco Fisher-Price

only at Target. fashionable & new from Graco. new Graco

dressed up & ready for Easter. girls’ dress boys’ dress

featured deals clearance
“Take a fictional Target shopper named Jenny Ward, who is 23, lives in Atlanta and in March bought cocoa-butter lotion, a purse large enough to double as a diaper bag, zinc and magnesium supplements and a bright blue rug. There’s, say, an 87 percent chance that she’s pregnant and that her delivery date is sometime in late August. What’s more, because of the data attached to her Guest ID number, Target knows how to trigger Jenny’s habits. They know that if she receives a coupon via e-mail, it will most likely cue her to buy online. They know that if she receives an ad in the mail on Friday, she frequently uses it on a weekend trip to the store. And they know that if they reward her with a printed receipt that entitles her to a free cup of Starbucks coffee, she’ll use it when she comes back again.”