Using Predictive Analytics to Identify Students At-Risk: A Step-by-Step Model for Easy Implementation by Small IR Offices

Jerome V. Ward, PhD
Director, Institutional Research
Cochise College
Sierra Vista, Arizona
wardj@cochise.edu

Serge Herzog, PhD
Director, Institutional Analysis
Consultant, CRDA Statlab
University of Nevada, Reno
serge@unr.edu

http://www.cis.unr.edu/IA_Web/research.aspx

Cochise College Background

- HSI college serving 2 rural border counties
- 7,500 mi² - 84 mi along Mexican border
- 2 main campuses, 4 rural centers
- ca. 5,000 fall headcount
- Educationally and financially disadvantaged pop.
- First generation, working adults, Pell – 30%
MOS Dominates Traditional FTSE

Retention Rate Relatively Stable
What Drives Enrollments?

- Began SEM planning process
- Developed SEM plan
- Created strategies – critical (high), essential (mid) and desired (low)
- One of top critical strategies – create an at-risk prediction model to assess where are students are

National Completion Agenda

- Democracy’s Colleges – Call to Action
  - Commitment to change policies, practices, cultures to enable students to progress and complete
- AACC’s Voluntary Framework of Accountability
- Gates Found. – Complete College America
- New Gainful Employment Programs reporting
  - On-time grad rate, job placement rate, median loan debt
- Pell grant award reduction – 9 yrs to 6 yrs
Student Focus

- Urgent need to channel students
  - Potential for loss of Title IV support
  - Timeframe for completion has been compressed
  - Accountability – progress and success of students
  - Accreditation issues
  - More efficient business model

Relevant Previous Research

Background for This Study

- New freshmen at rural community college in AZ
  - Fall cohorts 1999 through 2009 (over 6,000 students)
- Data sources
  - College student information system - Banner

Data Sources

<table>
<thead>
<tr>
<th>Name</th>
<th>Used</th>
<th>Not Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Age19</td>
<td>Age20plus</td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
<td>Unknown</td>
</tr>
<tr>
<td>Race/ethnic</td>
<td>Hispanic</td>
<td>BlackNatAm</td>
</tr>
<tr>
<td>Location</td>
<td>SierraFH</td>
<td>OtherCampus</td>
</tr>
<tr>
<td>HS information</td>
<td>HStoCCYrs</td>
<td>HSname</td>
</tr>
<tr>
<td>Goal/intent</td>
<td>GoalAD</td>
<td>GoalSkills</td>
</tr>
<tr>
<td>Placement testing</td>
<td>ReadingLevel</td>
<td>TestName</td>
</tr>
<tr>
<td>Financial Aid</td>
<td>UnmetNeed1K</td>
<td>Other grant</td>
</tr>
<tr>
<td>Credit load</td>
<td>Credits6</td>
<td>Credits6.5to11.5</td>
</tr>
<tr>
<td>Other Demographics</td>
<td>Displaced</td>
<td>Single parent</td>
</tr>
</tbody>
</table>
Data Analyzed for This Study

- Data elements included in the analysis
  - Student age, gender, ethnicity/race, residency
  - Campus location enrollment (we have multiple locations)
  - High school-to-college elapsed time
  - College goal (associate’s degree, skills training)
  - Pre-college reading, writing, and math level
  - Remedial placement
  - First-generation, Pell grant status
  - Unmet financial need (in $1K net of all aid)
  - First-semester credit load, online course status

At-Risk Forecasting Model

- Develop coefficients for predictors determining student fall-to-spring dropout risk
  - Logistic regression model using 1999-2008 historical cohorts as training dataset
  - Correct classification rate = 70% (baseline = 69%)
- Dropout risk scoring for new freshmen
  - linear transformation of the log odds \((p*[1-p]*\beta)\)
  - Decile grouping of scored students
  - Measures relative risk level of student
- Online reporting/dissemination of freshmen dropout risk via secure access
Data Management Tasks

• Create dummy variables for selected predictors (coded 0, 1)
• Impute missing data for select variables
  – E.g., precollege math level (via HS GPA)
• Run logit regression on training data
• Compute risk score for new freshmen, using predictor coefficients (linearized log odds), attach weights if needed
• Save output file in Excel, DB table

SPSS Syntax to Compute Risk Score

```
EXECUTE.
```
Key Predictor Variables by Statistical Significance (Wald)

- Over 12 credits/sem
- Took online courses
- Took 12 credits/sem
- Unmet financial need
- Pell grant recipient
- Multi-campus enroll
- 6.5 to 11.5 credits
- Took 6 credits/sem
- Male student, residency*

*Local, within commuting dist.

At-Risk Student Intervention

| CohortYear | PIDM | @1sttermcrs | AgeCohort | Gender | Ethnicity | PhysicalCampus | Onlinecourses | Residency | HSGradDate | HighSchool | DEGC Major | Goal | EnglishLevels | MathLevels | Remedialcourses | PellGrant | Unmetneed | FallSpring | DropoutRiskDecile | DropoutRiskPct |
|------------|------|-------------|-----------|--------|-----------|----------------|---------------|-----------|------------|------------|------------|------------|------|--------------|------------|------------------|-----------|----------|-----------|----------------|---------------|
| 2010       | 1185801 | 12           | 43 M     | White  | Benson     | No             | Commutal      | 30-Jun-85 | Utah High  | 15-May-10 | Douglas     | EMT         | Prepare for new career | 0          | 3        | 0         | 11895          | 2.05         | 10       | 100       |
| 2010       | 1185025 | 6            | 18 M     | Hispanic/L Douglas | No | Douglas     | 15-May-10 | Douglas     | 15-May-10 | Nogales Hi | AGS AGS  | 3            | 4          | 2082               | 12663      | 2.00       | 10         | 99.87          | 100          |
| 2010       | 1183814 | 3            | 22 M     | Hispanic/L Douglas | Yes | Commutal    | 1-Mar-10 GED | Douglas     | 15-May-10 | Douglas Hi | AGS AGS  | 0            | 4          | 0                  | 4151       | 2.00       | 10         | 99.74          | 100          |
| 2010       | 1184603 | 11           | 20 M     | White   | Sierra Vista | No             | Commutal      | 15-May-09 | Maine High | AGS AGS  | 0            | 4          | 0                  | 4151       | 2.00       | 10         | 99.74          | 100          |
| 2010       | 1182100 | 9            | 19 F     | Hispanic/L Nogales | No | Commutal    | 15-May-10 | Nogales Hi | AGS AGS  | 0            | 4          | 0                  | 4151       | 2.00       | 10         | 99.74          | 100          |
| 2010       | 1185661 | 5            | 19 M     | Hispanic/L Douglas | No | Douglas     | 28-May-10 | Douglas Hi | AGS AGS  | 0            | 4          | 0                  | 4151       | 2.00       | 10         | 99.74          | 100          |
| 2010       | 1184279 | 14           | 18 F     | Hispanic/L Douglas | No | Douglas     | 15-May-10 | Douglas Hi | AGS AGS  | 0            | 4          | 0                  | 4151       | 2.00       | 10         | 99.74          | 100          |
Improving the Bottom Line

• Case study at Nevada’s flagship university using this at-risk forecasting model
• Rise in freshmen retention by 4 percentage points due in part by better at-risk forecasting
  – AY 2010-11 additional net tuition revenues = $215,119
    (for 94 NV, 19 WUE, excl OS students) for one cohort in one year, without OS $!
• Incentive for student to speed up graduation
  – Opportunity cost per year in foregone earnings = $32,000 per year (published constant 2010-$)

National Media Coverage of This At-Risk Forecasting Model


Findings

% Spring Returners by Decile Rank

<table>
<thead>
<tr>
<th>Decile Rank</th>
<th>Percent Returning</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 (High Risk)</td>
<td>0.10</td>
</tr>
<tr>
<td>9</td>
<td>0.20</td>
</tr>
<tr>
<td>8</td>
<td>0.30</td>
</tr>
<tr>
<td>7</td>
<td>0.40</td>
</tr>
<tr>
<td>6</td>
<td>0.50</td>
</tr>
<tr>
<td>5</td>
<td>0.60</td>
</tr>
<tr>
<td>4</td>
<td>0.70</td>
</tr>
<tr>
<td>3</td>
<td>0.80</td>
</tr>
<tr>
<td>2</td>
<td>0.90</td>
</tr>
<tr>
<td>1 (Low Risk)</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Implementation

• Currently doing:
  – Use by Advising Office – new students required to see advisor until 12 credits accumulated
  – Using decile scores to help frame discussion with students
  – Better target those students in difficulty
  – Faculty recommend students for advising
  – Contacting students who haven't registered
Implementation

• Forming a Student Success Group to discuss strategies (i.e., FYE, end of semester follow-up, etc.)
• Implementing DegreeWorks
• Considering MAP-Works
• Holistic approach facilitate communication between advisors, faculty and students
• Goal to strengthen student progress and ultimate success

Q & A

Link to presentation:
http://www.cis.unr.edu/IA_Web/research.aspx