MONTANA EARLY WARNING SYSTEM FOR DROPOUTS

PRESENTED BY ERIC MEREDITH DATA ANALYST OPI
10/21/2016
WHAT IS THE MONTANA EWS?

• A statistical model that can use readily available school, student, and other live data to identify students who are at risk of dropping out of school **before** they drop out.

• The EWS allows educators to intervene early on during the process before a student has reached the point of no return.
HOW IS THE EWS DEVELOPED?

• Compare data from dropouts to the data from high school graduates from the school years 2007-2015

• Model is found using Logistic Regression

\[ \pi(x) = \frac{e^{\alpha + \beta x_1 + \beta x_2 + \cdots + \beta x_n}}{1 + e^{\alpha + \beta x_1 + \beta x_2 + \cdots + \beta x_n}} \]

• \( \pi(x) \) is the percent chance a student will drop out of school

• Separate model is developed for each grades 6, 7, 8 and for each year of high school.
WHAT DATA IS AVAILABLE FOR THE MODEL?

• Data stored by the State.
  • Student Data
    • SIS (AIM) Data
    • Testing Data
  • School data
    • School Demographics
    • Location
  • Census Information
    • Unemployment Rates
    • Populations

• Data stored by the Schools
  • Attendance
  • Transcripts
  • Grades
  • Discipline
• Data from all Graduates and Dropouts from 2007-2015 school years at 13 school system’s in Montana.

• 13 school system’s in Montana were sampled to give a good representation of schools across the state. (roughly 11,000 students per year, or about 1/6th of the statewide students in 6-12th grades)

• Data current for each student at the end of the enrollment (whether a dropout or graduate)

• Previous term data is usually from the 3rd quarter of the year.

• This creates an assumption in the model that on average a student’s data is the same at the end of the year as it is throughout the school year.
EWS HISTORY

• Pilot Year 2012-2013 (10 School Systems involved)
  • For the 2012-2013 school year EWS Results were sent to each school once a month
  • EWS was changed and updated many times during the school year.

• 2nd Year of EWS 2013-2014
  • Model was updated during the previous summer and remained unchanged throughout the 2013-2014 school year.

• 3rd Year of EWS 2014-2015
  • New model uses less variables that OPI does not collect (9 total)

• 4th Year of EWS 2015-2016
  • Available to all schools in GEMS

• 5th Year of EWS 2016 – 2017
  • New updated model completed before start of the new school year
  • Updates to current reports and working on Intervention Report
SCHOOL SYSTEMS CURRENTLY IN EWS

- Arlee
- Belgrade
- Bozeman
- Browning
- Butte
- Columbus
- Corvallis
- Cut Bank
- Frenchtown
- Great Falls
- Havre
- Huntley Project
- Lame Deer
- Laurel
- Lewistown
- Libby
- Livingston
- Park City
- Red Lodge
- St. Ignatius
- Townsend
- Wolf Point
VARIABLES IN THE EWS MODEL

Collected by OPI

- Moved this school year (Y or N)
- Moved from out of state (Y or N)
- Repeated a grade in K-8 (Y or N)
- Age Difference (July 15 cutoff date)*
- More than 2 SS’s attended since 2007 (Y or N)
- Gender

Not Collected by OPI

- Attendance Rate
- # of Previous Term F’s
- # of Previous Term A’s
- # of Behavior Events in last 120 days
- # of Out of School Suspension Events in last 3 years
- On Track (Y or N)
- # of Credits per year
- # of Absences in last 90 days
- # of Absences in last 60 days

Over 250 Variables have been analyzed.
TWO PARTS TO A GOOD EWS MODEL

1. The Model should assign a high dropout percentage to students who end up dropping out.
   - Low dropout percentage to those that eventually graduate.
     - Can be evaluated by:
       - R squared
       - C-statistic
       - ROC Curves
       - Model AIC

2. Model should be efficient in identifying dropouts above the cut-off threshold for targeting a student as At-Risk.
   - A high percentage of At-Risk students end up being dropouts.
     - Can be evaluated by:
       - Confusion Matrix
WHEN IS A STUDENT CONSIDERED AT RISK?

• At what dropout percentage should we be concerned about a student?
  • Depends on school
  • Depends on how many incorrect conclusions you will accept.
  • We want to be able to identify as many dropouts as we possibly can.
  • We want as many of the students as possible to be in one of the “True” boxes.
    • Small number of students in the “False” boxes.

<table>
<thead>
<tr>
<th></th>
<th>True Negative</th>
<th>False Negative</th>
<th>False Positive</th>
<th>True Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model: Graduate</td>
<td>Graduate</td>
<td>Graduate</td>
<td>Dropout</td>
<td>Dropout</td>
</tr>
<tr>
<td>Student:</td>
<td>Graduate</td>
<td>Dropout</td>
<td>Graduate</td>
<td>Dropout</td>
</tr>
</tbody>
</table>
Looking at Beginning of the Year EWS Results from 2009-2010

Only including students that had all data elements needed for the EWS. (4167 students total)

Must look at 2009-2010 to include 6th, 7th, 8th, 9th, 10th, 11th, and 12th grade students and allow time for them to graduate.

512 Dropouts from group of students that were in school 2009-2010 in the Pilot Schools

<table>
<thead>
<tr>
<th>Marked as At Risk when &gt;15%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>True Negative</strong></td>
</tr>
<tr>
<td>Model: Graduate Student: Graduate 3132 75.2%</td>
</tr>
<tr>
<td><strong>False Negative</strong></td>
</tr>
<tr>
<td>Model: Graduate Student: Dropout 131 3.1%</td>
</tr>
<tr>
<td><strong>False Positive</strong></td>
</tr>
<tr>
<td>Model: Dropout Student: Graduate 523 12.6%</td>
</tr>
<tr>
<td><strong>True Positive</strong></td>
</tr>
<tr>
<td>Model: Dropout Student: Dropout 381 9.1%</td>
</tr>
</tbody>
</table>

- Dropouts found – 74.4%
- Graduates found – 85.7%
- Accuracy – 84.3%
EWS MODEL DIAGNOSTICS

- ROC Curve and c-statistic
  - Graph of Sensitivity (True Positive Rate, % of Graduates correct) vs 1-Specificity (False Positive Rate, % of Dropouts correct)
  - Probability the model will assign a higher score to a randomly chosen dropout than to a randomly chosen graduate.
EWS MODEL DIAGNOSTICS

% Dropouts Identified

% Identified Students Dropout

True positive rate

Positive predictive value

Cutoff

Cutoff
FULL MODEL DIAGNOSTICS

- R-squared
  - Measure of the fit of the model to data
  - Works a little different with logistic regression but similar to the r squared used with linear regression
- C-statistic
  - Probability a higher dropout value is assigned to a dropout than to a graduate.

<table>
<thead>
<tr>
<th>Year</th>
<th>R squared</th>
<th>c-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>6th Grade</td>
<td>0.449</td>
<td>0.861</td>
</tr>
<tr>
<td>7th Grade</td>
<td>0.501</td>
<td>0.885</td>
</tr>
<tr>
<td>8th Grade</td>
<td>0.522</td>
<td>0.895</td>
</tr>
<tr>
<td>1st Year HS</td>
<td>0.567</td>
<td>0.910</td>
</tr>
<tr>
<td>2nd Year HS</td>
<td>0.661</td>
<td>0.943</td>
</tr>
<tr>
<td>3rd Year HS</td>
<td>0.708</td>
<td>0.968</td>
</tr>
<tr>
<td>4th Year HS</td>
<td>0.777</td>
<td>0.987</td>
</tr>
<tr>
<td>5+ Years HS</td>
<td>0.728</td>
<td>0.941</td>
</tr>
</tbody>
</table>
2015-2016 SCHOOL YEAR EWS RESULTS

• Median Dropout percentage for all students in pilot schools for 5/1/15 results was 4.5%
  • 176 Dropouts total with EWS results on 9/3/2016 (beginning of the school year)

• 121 of the dropouts had dropout percentages of greater than 15%
  • Would have been targeted as At-Risk
  • 68.8% of Dropouts would have been identified at the beginning of the school year.

• Most had much higher percentages in the EWS.
  • Median Dropout Percentage of 158 dropouts was 56.2%
  • 51 of the 176 dropouts had over 90%
GEMS EWS RESULTS

- http://gems.opi.mt.gov/StudentCharacteristics/Pages/EarlyWarningSystemOverview.aspx

- EWS Results only available in GEMS Secure
  - Must get a login and access rights to the page.

- 3 Reports in GEMS
  - School Report
  - Student Summary Report
  - Student Detail Report
SCHOOL LEVEL REPORT

- Available for every school/district you have access to
  - School or district wide results to see numbers of students being identified.

- Can compare results by Grade

- Can compare to Statewide average results

- Will display results for the last 2 EWS runs
### STUDENT SUMMARY REPORT

- Lists EWS results for every student in your district/school in an excel file (other formats available)

* Names, School, and Data provided in the report is fictitious

<table>
<thead>
<tr>
<th>SC</th>
<th>School Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCD</td>
<td>Early Warning System School</td>
</tr>
<tr>
<td></td>
<td>Anderson Joel DJHDFIEF 4 12 99.8% Attendance Grades</td>
</tr>
<tr>
<td>ABCD</td>
<td>Early Warning System School Smith Maria JDUHJH 4 12 0.1%</td>
</tr>
<tr>
<td>ABCD</td>
<td>Early Warning System School Lackey Edin GFSFWFD 3 11 9.6%</td>
</tr>
<tr>
<td>ABCD</td>
<td>Early Warning System School Underman Hal KJJHYGVX 3 11 6.1%</td>
</tr>
<tr>
<td>ABCD</td>
<td>Early Warning System School Grossman Keith JSUWEHDBH 2 10 3.9%</td>
</tr>
<tr>
<td>ABCD</td>
<td>EarlyWarning System School Player Joe IJUJHUUUS 2 10 0.4%</td>
</tr>
<tr>
<td>ABCD</td>
<td>EarlyWarning System School Stein Thomas ODJEHDYST 1 09 70.2%</td>
</tr>
<tr>
<td>ABCD</td>
<td>EarlyWarning System School Caligher Mary DSYDHEGD 1 09 1.8%</td>
</tr>
<tr>
<td>ABCD</td>
<td>EarlyWarning System School Thompson Jess UDJEHGD 0 08 81.6%</td>
</tr>
<tr>
<td>ABCD</td>
<td>EarlyWarning System School Banby Shane MSJDEYDG 0 08 8.3%</td>
</tr>
<tr>
<td>ABCD</td>
<td>EarlyWarning System School Smith Jane NSHDEYRG 0 07 76.5%</td>
</tr>
<tr>
<td>ABCD</td>
<td>EarlyWarning System School Anderson Mike MKNJBHGC 0 07 13.7%</td>
</tr>
<tr>
<td>ABCD</td>
<td>EarlyWarning System School Abbott Megan HUGYFTDRE 0 06 50.2%</td>
</tr>
<tr>
<td>ABCD</td>
<td>EarlyWarning System School Cornrow Mike KDHSTDGXC 0 06 18.3%</td>
</tr>
</tbody>
</table>

- Attendance Grades |
- Off Track |
- Mobility |
- Previous Dropout |
- Previous Prob. |
- Behavior Odds |
- Attendance Odds |
- Grades Odds |
- Mobility Odds

* * *
STUDENT LEVEL REPORT

- Available for every student enrolled in your school
- Displays all data used by the EWS Model
- Graphically displays the following:
  - Dropout Probability
  - Grades Risk Factor
  - Attendance Risk Factor
  - Behavior Risk Factor
  - Mobility Risk Factor
- Will display results for up to the last 12 EWS results
- Attendance Risk Factor Example
  - Based on grades alone, the odds of this student dropping out is 11.18 times the odds of an average student, with all other factors held constant
  - Above 1.25 all risk factors are flagged
- * All names and data in report are fictitious *