I’m Just a Bill
The Data Science Version

Mike Sharkey, VP of Analytics
Dr. John Whitmer, Learning Analytics and Research Director

February 14, 2017
ON A FEW ADD-ONS FOR GOOD MEASURE.

© SNL Studios

This presentation leaves copyright of the content to the presenter. Unless otherwise noted in the materials, uploaded content carries the Creative Commons Attribution-NonCommercial-ShareAlike license, which grants usage to the general public with the stipulated criteria.
Analytics – Hype & Literacy

Gartner Hype Cycle

Visibility

Peak of Inflated Expectations
Plateau of Productivity
Slope of Enlightenment
Trough of Disillusionment
Technology Trigger

Time

Analytics Literacy is a Major Limiter of Ed Tech Growth

Posted on November 19, 2016 by Michael Feldstein

By Michael Feldstein

More Posts (1094)

Whatever else you think of the election, it has been the mother of all teachable moments for many of us. It has raised questions about what we thought we knew about our democracy, our neighbors, our media...and apparently learning analytics. The shock of the polls being “wrong” has raised a lot of questions about how much we can really trust data...
Machine Learning Algorithm to Predict Student Outcomes
Problem we’re trying to solve

• Identifying symptoms of risk

• What’s the probability this student will
  —pass with a C or better, or
  —attend class this week
# Machine Learning

<table>
<thead>
<tr>
<th>Regression</th>
<th>Machine Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitting</td>
<td>Learning</td>
</tr>
<tr>
<td>Coefficients</td>
<td>Importance</td>
</tr>
<tr>
<td>$r^2$</td>
<td>F1 score</td>
</tr>
<tr>
<td>Old and stuffy</td>
<td>Cool and hip</td>
</tr>
<tr>
<td><em>Makes good predictions</em></td>
<td><em>Makes good predictions</em></td>
</tr>
</tbody>
</table>

Adapted from [http://statweb.stanford.edu/~tibs/stat315a/glossary.pdf](http://statweb.stanford.edu/~tibs/stat315a/glossary.pdf)
How it Works – Plinko

© FreemaleMedia, North America Inc.

@johncwhitmer    @mjshark
Nodes, Chips, and Bins

- **Node**: A student feature/characteristic (e.g. race)
- **Chip**: A student (with known/historic data)
- **Bin**: A collection of similar students
Rinse and Repeat

http://fridayatfive.com/2016/10/diy-drinko-plinko-drinking-game-price-is-right.html

@johncwhitmer   @mjshark
Observe the Bins

- Female
- 1-14 xfer credits
- 80th %ile GPA
- Top 20% click activity
- Etc.
- ∴ 87% passed the class

- Male
- Zero xfer credits
- 60th %ile GPA
- 4th quintile click activity
- Etc.
- ∴ 58% passed the class
Results – Predictive Model(s)

Confusion Matrix for Model: Week D

Predicted

<table>
<thead>
<tr>
<th></th>
<th>Fail</th>
<th>Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail</td>
<td>1,493</td>
<td>3,012</td>
</tr>
<tr>
<td>Pass</td>
<td>970</td>
<td>18,233</td>
</tr>
</tbody>
</table>

Total Number of Student Course Records: 23,708

Performance Metrics

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>accuracy</td>
<td>0.83</td>
</tr>
<tr>
<td>precision-fail</td>
<td>0.61</td>
</tr>
<tr>
<td>recall-fail</td>
<td>0.33</td>
</tr>
<tr>
<td>f-score-fail</td>
<td>0.43</td>
</tr>
<tr>
<td>precision-pass</td>
<td>0.86</td>
</tr>
<tr>
<td>recall-pass</td>
<td>0.96</td>
</tr>
<tr>
<td>roc (area under curve)</td>
<td>0.79</td>
</tr>
</tbody>
</table>

Model Classifier Metric Definitions:

Accuracy: Overall percent predicted correctly
Precision: When we make a specific prediction (pass or fail), how often are we correct
Recall: For a specific outcome (pass or fail), how often are we correct
F-score: Weighted average of precision and recall

ROC (Area Under Curve): 1.00 is a perfect classifier, 0.50 is no better than a coin flip

Model Classifier Metric Formulas:

Accuracy: (Number of correctly predicted Pass * Actualliy Passed) + (Number of correctly predicted Fail * Actualliy Failed) / (Total Number of Predictions)

Precision Fail: (Predicted Fail * Actualliy Failed) / (Total Predicted to Fail)
Recall Fail: (Predicted Fail * Actualliy Failed) / (Total Actualliy Failed)

F-score Fail: Weighted average of Precision Fail and Recall Fail

Precision Pass: (Predicted Pass * Actualliy Passed) / (Total Predicted to Pass)
Recall Pass: (Predicted Pass * Actualliy Passed) / (Total Actualliy Passes)

F-score Pass: Weighted average of Precision Pass and Recall Pass

@johncwhitmer  @mjshark
## Model Performance Metrics

<table>
<thead>
<tr>
<th>Metric</th>
<th>Week 0</th>
<th>0-20%</th>
<th>20-40%</th>
<th>40-60%</th>
<th>60-80%</th>
<th>80-100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>0.83</td>
<td>0.85</td>
<td>0.87</td>
<td>0.89</td>
<td>0.90</td>
<td>0.93</td>
</tr>
<tr>
<td>F-score Fail</td>
<td>0.43</td>
<td>0.49</td>
<td>0.59</td>
<td>0.68</td>
<td>0.71</td>
<td>0.79</td>
</tr>
<tr>
<td>Precision - Pass</td>
<td>0.86</td>
<td>0.87</td>
<td>0.89</td>
<td>0.91</td>
<td>0.93</td>
<td>0.96</td>
</tr>
<tr>
<td>Precision - Fail</td>
<td>0.61</td>
<td>0.67</td>
<td>0.73</td>
<td>0.76</td>
<td>0.75</td>
<td>0.79</td>
</tr>
<tr>
<td>Recall - Pass</td>
<td>0.95</td>
<td>0.96</td>
<td>0.96</td>
<td>0.95</td>
<td>0.95</td>
<td>0.96</td>
</tr>
<tr>
<td>Recall - Fail</td>
<td>0.33</td>
<td>0.39</td>
<td>0.50</td>
<td>0.62</td>
<td>0.67</td>
<td>0.79</td>
</tr>
<tr>
<td>Pr_FAIL_Ac_FAIL</td>
<td>1,493</td>
<td>5,047</td>
<td>6,719</td>
<td>9,734</td>
<td>8,165</td>
<td>12,171</td>
</tr>
<tr>
<td>Pr_FAIL_Ac_PASS</td>
<td>970</td>
<td>2,491</td>
<td>2,526</td>
<td>3,146</td>
<td>2,716</td>
<td>3,219</td>
</tr>
<tr>
<td>Pr_PASS_Ac_FAIL</td>
<td>3,012</td>
<td>7,882</td>
<td>6,716</td>
<td>6,044</td>
<td>3,976</td>
<td>3,325</td>
</tr>
<tr>
<td>Pr_PASS_Ac_PASS</td>
<td>18,233</td>
<td>53,895</td>
<td>55,175</td>
<td>64,819</td>
<td>53,546</td>
<td>71,492</td>
</tr>
<tr>
<td>ROC (Area Under Curve)</td>
<td>0.79</td>
<td>0.83</td>
<td>0.88</td>
<td>0.91</td>
<td>0.93</td>
<td>0.96</td>
</tr>
</tbody>
</table>
Results – Feature Importances

Blackboard Predict Model: Sample University

Top 15 Features

Feature Importance Rankings: Week 0

Feature Importance Rankings: Weeks 0-20%

@johncwhitmer  @mjshark

Scaling Innovation | #BbAnalytics
## Blackboard Predict Model: Sample University

### Top 15 Features

<table>
<thead>
<tr>
<th>FEATURE</th>
<th>Week 0</th>
<th>0-20%</th>
<th>20-40%</th>
<th>40-60%</th>
<th>60-80%</th>
<th>80-100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>COURSE_GRADE_PERC</td>
<td>0.00</td>
<td>0.06%</td>
<td>0.20</td>
<td>0.34</td>
<td>0.43</td>
<td>0.55</td>
</tr>
<tr>
<td>CUMULATIVE_GPA</td>
<td>0.11</td>
<td>0.11%</td>
<td>0.06</td>
<td>0.03</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>COURSE_SIZE</td>
<td>0.05</td>
<td>0.04%</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td>CAREER_CREDIT_COMPLETION_PERC</td>
<td>0.04</td>
<td>0.04%</td>
<td>0.04</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>WEEKS_UNTIL_FIRST_ACTIVITY</td>
<td>0.05</td>
<td>0.04%</td>
<td>0.03</td>
<td>0.04</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>COURSE_LEVEL</td>
<td>0.05</td>
<td>0.03%</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>WEEKS_REGISTERED_PRIOR_TO_START</td>
<td>0.05</td>
<td>0.03%</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>PREV_TERM_EARNED_CREDITS</td>
<td>0.04</td>
<td>0.02%</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>COURSE_GRADE_RANK</td>
<td>0.01</td>
<td>0.02%</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>CAREER_FAILED_CREDITS</td>
<td>0.03</td>
<td>0.02%</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>AGE</td>
<td>0.04</td>
<td>0.02%</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>SAT_MATH</td>
<td>0.03</td>
<td>0.02%</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>SAT_READING</td>
<td>0.04</td>
<td>0.02%</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>EFC</td>
<td>0.03</td>
<td>0.02%</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>CAREER_WITHDRAWN_CREDITS</td>
<td>0.03</td>
<td>0.02%</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Takeaways

• What’s the question you are trying to predict (e.g. pass class with a C or better)
• Need enough historical data
• Define the features/characteristics

• What will you do with the results?
Cluster Analysis to Uncover Meaningful Patterns between Data Points

Problem: Identifying Common LMS Course Designs

1. **Desired outcome**: meaningful indicators of student performance based on student LMS activity / behaviors

2. **Hard because**: variation in how LMS is used; consequently the “effectiveness” of student time as predictor of course grade

3. **Potential solution**: course design - activities & content faculty include in their LMS course shells. Academic literature has heuristic models for course design, little large-scale empirical analysis

4. **Approach**: analyze distributions of student time spent using different tools (e.g. if students spent lots of time in test, course design was test-heavy).

5. **DS approach**: Use “K means cluster analysis” to find patterns

---

**Effect Size: Student Time as Predictor of Grade**

(n=7,648 courses, p<.05, 22% of sample)
Metaphor: K Means Cluster Analysis as Optimizing Restaurant Planning

• Cluster analysis: method to find patterns between explicitly unrelated data points using 1-n factors

• Example
  – What are the best locations to plan new restaurants closest to office buildings?
  – What are the age ranges for people working near these restaurants? Incomes?

@johncwhitmer   @mjshark
Restaurant Selection

1. Choose # of “clusters”
Restaurant Selection

1. Choose # of “clusters”
Restaurant Selection

1. Choose # of “clusters”

2. Assume starting “center” locations (centroids)

Can conduct with multiple variables; just minimize distance between center & observations in clusters
Restaurant Selection

1. Choose # of “clusters”

2. Assume starting “center” locations (centroids)

3. Assign data points to closest centers

Can conduct with multiple variables; just minimize distance between center & observations in clusters
Restaurant Selection

1. Choose # of final restaurants (clusters). Plot your data points (need means, doesn’t have to be grid)

2. Assume starting “center” locations

3. Assign observations to nearest “center” location

4. **Create new “center” location based on minimum distances of assignments to mean**

Can conduct with multiple variables; just minimize distance between center & observations in clusters
Restaurant Selection

1. Choose # of “clusters”
2. Assume starting “center” locations (centroids)
3. Assign data points to closest centers
4. Move “centers” to average of data points
5. Rinse and repeat

Can conduct with multiple variables; just minimize distance between center & observations
Results: LMS Course Design Archetypes
Assessment Time Clustering Results

- Spring 2016 BB Learn Data, North America
- Convenience sample: 70k courses, 3.4M learners, 927 institutions
- Filtered for instructional courses (using grades, with activity, within reasonable bounds): final data 20k courses, 665k learners
Discussion Forum Results

Distribution of Proportion of Time in Discussions by Cluster
## Distribution of Courses by Type

<table>
<thead>
<tr>
<th>Course Archetype</th>
<th>Count (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Supplemental</strong>&lt;br&gt;Content-heavy. Low interaction.</td>
<td>9,909 (53%)</td>
</tr>
<tr>
<td><strong>Complementary</strong>&lt;br&gt;One-way communication through content, announcements, and gradebook.</td>
<td>4,588 (24%)</td>
</tr>
<tr>
<td><strong>Social</strong>&lt;br&gt;High peer-to-peer interaction through discussion board.</td>
<td>2,130 (11%)</td>
</tr>
<tr>
<td><strong>Evaluative</strong>&lt;br&gt;Heavy use of assessments.</td>
<td>1,832 (10%)</td>
</tr>
<tr>
<td><strong>Holistic</strong>&lt;br&gt;High LMS activity with balanced use of assessments, content, and discussion.</td>
<td>351 (2%)</td>
</tr>
</tbody>
</table>
Implications

• K means clustering is as simple as planning a restaurant in the best location possible

• Start with meaningful data that is well distributed

• Clusters are basis for further analysis of relationships

• Application: course archetypes are instructive and form basis for recommendations for effective practices in driving LMS adoption
Thank you!

Mike Sharkey
Vice President, Analytics
mike.sharkey@blackboard.com
@mjshark

John Whitmer
Director of Education & Research
john.whitmer@blackboard.com
@johnwhitmer