Using Early Quizzes to Predict Student Outcomes in Online Introductory Biomedical Informatics Courses

Imgur A. Willcockson PhD1, Jorge R. Herskovic MD PhD1,2, Melanie A. Sutton PhD3, Robert E. Hoyt MD3, Craig W. Johnson PhD1, Todd R. Johnson PhD1, Elmer V. Bernstam MD1,2

1School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX;
2The University of Texas MD Anderson Cancer Center, Houston, TX;
3Medical Informatics Program, School of Allied Health and Life Sciences, University of West Florida, Pensacola, FL;
4Department of Internal Medicine, Medical School, The University of Texas Health Science Center at Houston, Houston, TX.

Background & Problem

- Need for informatics education
  - Especially introductory courses
- Relatively constrained number of qualified educational programs
- Increasing use of online education
  - Larger class sizes
  - Traditional cues not available
    - Eye contact
    - Verbal interaction
- Can use data available at time of admission [1], but may not have data for non-degree-seeking students.
- Problem: How do you identify students at risk for poor performance early?

Learning Management Systems

- Widely used
- Collect a variety of data
  - Passive data collection: Activity logs
  - Active data collection: Weekly quizzes

Hypothesis

Students at risk for non-successful completion (drop, fail or grade ≤ C) can be identified by poor performance on the first few weekly quizzes.

Methods

1. Collected weekly quiz scores for each student along with final course outcome for
   a) Foundations of Health Information Sciences I (F1) at the School of Biomedical Informatics, UT-Houston
   b) Introduction to Medical Informatics (IMI) at the University of West Florida (UWF).
2. Course outcomes (binary)
   a) Non-Successful Completion (NSC)/non-NSC
   b) Failure/not Failure
   c) Completion/non-Completion

Results

<table>
<thead>
<tr>
<th>Week(s)</th>
<th>Area under ROC Curve (AUC)</th>
<th>PPV (75% Threshold)</th>
<th>NPV (75% Threshold)</th>
<th># &quot;at risk&quot; students at SBMI of 205 total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUC</td>
<td>95% CI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.7380</td>
<td>0.6705-0.8055</td>
<td>44.3%</td>
<td>77.8%</td>
</tr>
<tr>
<td>1+2</td>
<td>0.7912</td>
<td>0.7303-0.8521</td>
<td>57.9%</td>
<td>77.8%</td>
</tr>
<tr>
<td>1+2+3</td>
<td>0.8333</td>
<td>0.7773-0.8892</td>
<td>61.9%</td>
<td>82.4%</td>
</tr>
<tr>
<td>1+2+3+4</td>
<td>0.8627</td>
<td>0.8123-0.9131</td>
<td>69.0%</td>
<td>87.1%</td>
</tr>
</tbody>
</table>

Threshold = 75% (chosen as the highest score where false positives = false negatives)

Conclusions

While representing only 6% to 8% of the total grade, the first four quizzes are highly predictive for course outcome. Using only the first two quizzes available by the UT add/drop deadline still allows prediction, but with a lower PPV. Automated prediction generalizes across institutions and compares favorably to human instructor prediction.

References


Acknowledgements

The authors thank all instructors and students who contributed data to this study. We are also grateful to Stephanie Reedy at UWF for her assistance with data collection. This work was funded in part by NCATS Grant UL1 TR000371 establishing the Center for Clinical and Translational Sciences at the University of Texas at Houston.

Corresponding author: Elmer.V.Bernstam@uth.tmc.edu