What The Numbers Don’t Tell you. (A Learning Analytics Case)

Brid Lane
DBS Research Day 2016
High Level Projects
Require High Level Resources


- 21 classification algorithms (including regression, clustering, pattern mining, neural nets, etc.) to predict student marks
- Instructors can detect high-risk students in time
- Instructor can decide what Moodle tools to promote:
  - activity on quizzes \(\rightarrow\) positive impact on grades
- What worked varied a lot
- Predictions of 65% accuracy with the Moodle log data
- Should integrate offline data also (e.g. students openness to learning, class participation, etc.) but these are difficult to measure.
Predicting “At Risk” Students from log data

Date:
Friday, 12 September, 2014
What have I got?

Am I wasting my time putting content on Moodle?

Do the students even look at it?

Is the content helping them at all in terms of assessment results?

What might their activity patterns on Moodle tell me about their assessment results?

The Context
Research Questions

• What are students’ access patterns to content given to them on Moodle?
  – What content are students accessing on Moodle – lectures notes & other content
  – When are they accessing it
  – How often are they accessing content

• How do these access patterns match with assessment achievement?
Are numbers enough?

LA metrics on their own are weak.
E.g. number of LMS logins is a trivial measure of learning progression
   Access /= learning
Counting hits and activities not as good as counting what the students do with them.

Not convinced that the dashboard stats indicate any student engagement or enhancement of the learning experience.
Just based on easily collectable data and no pedagogy
“The number of times a student accesses their LMS is a better predictor of student success than their high school GPA.”

John Whitmer, Ed.D.
Director, Platform Analytics and Research

#EDU15
...with due respect to interpreting what we find
# Data Interpretation

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Learning Analytics needs to go beyond the numbers,</td>
<td>• A weakness of LA is the possible <em>misinterpretation</em> of output</td>
</tr>
<tr>
<td>• Aspects of <strong>human judgement</strong> required</td>
<td>• Human judgement can be faulty</td>
</tr>
<tr>
<td>• Need to interpret the data <strong>in context</strong> before making decisions based on it.</td>
<td>• What about conflicting outputs?</td>
</tr>
</tbody>
</table>
Methodology

MBA Cloud Computing Students

September ➔ January ➔ April
## Student Profile

<table>
<thead>
<tr>
<th></th>
<th>September Intake</th>
<th>January Intake</th>
<th>April Intake</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Class size</strong></td>
<td>10</td>
<td>18</td>
<td>6</td>
</tr>
<tr>
<td><strong>Geographical makeup</strong></td>
<td>1 German, 1 Nigerian, 8 Indian</td>
<td>1 Jordan, 1 Malawi, 16 Indian</td>
<td>6 Indian</td>
</tr>
<tr>
<td><strong>Average age</strong></td>
<td>28</td>
<td>27</td>
<td>25</td>
</tr>
<tr>
<td><strong>Male / Female Ratio</strong></td>
<td>2 / 10</td>
<td>3 / 18</td>
<td>0 / 6</td>
</tr>
<tr>
<td><strong>Performance on the Module</strong></td>
<td>6 Passed on First Attempt</td>
<td>13 Passed on First Attempt</td>
<td>5 Passed on First Attempt</td>
</tr>
</tbody>
</table>
The Myth of the Full-Time Student

The East → West Transition

The Tech → Business Transition
What’s on the 3 Moodle Pages

Each student’s log file was downloaded

Items of interest isolated

Statistics run for the viewing figures for each of the above items.
Findings

What are students’ access patterns to content given on Moodle?
Average % of (1) all content items & (2) Lectures, viewed per cohort

- Average Sem A
- Average Sem B
- Average Sem C

% of all available items viewed
% of Lecture items viewed
How do these access patterns match with assessment achievement?
Moodle % Lectures →

The **percentage** of the 8 lecture files that were viewed

Moodle Total Lecture Views →

The **total number of hits** on the 8 lecture pdf files. Students can look at an individual file multiple times.

<table>
<thead>
<tr>
<th>Correlation</th>
<th>Sem A</th>
<th>Sem B</th>
<th>Sem C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Grade vs Moodle % Lectures =</td>
<td>-0.51</td>
<td>-0.64</td>
<td>-0.51</td>
</tr>
<tr>
<td>Overall Grade vs Moodle Total Lecture Views =</td>
<td>0.30</td>
<td>0.70</td>
<td>0.01</td>
</tr>
</tbody>
</table>
A deep breath taken for Interpretation

The more you study

The lower your grades
Before panic sets in...

Students who complete one but not the other can cause distortive overall grades.

Class numbers so small here that even one or two such students can cause much distortion.
Just the Coursework Assignments...

Cohort A:
  – Two groups who both performed well at merit level
  – The third group didn’t reach a pass grade

Cohort B:
  – One group with a plagiarism problem (a zero grade)
  – The other groups in this cohort ranged from high 30s to mid-60s

Cohort C:
  – Just 2 groups, both achieving a passing grade only

<table>
<thead>
<tr>
<th></th>
<th>Sem A</th>
<th>Sem B</th>
<th>Sem C</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA Grade vs Moodle % Lectures =</td>
<td>0.61</td>
<td>-0.53</td>
<td>-0.55</td>
</tr>
<tr>
<td>CA Grade vs Moodle Total Lecture Views =</td>
<td>0.56</td>
<td>0.54</td>
<td>0.45</td>
</tr>
</tbody>
</table>
Just the Exams...

Cohort A:
- No real correlation

Cohort B:
- A notable suggestion that accessing lecture notes is detrimental to the student grades.
- Yet, the highest correlation in terms of Total Lecture Views.

Cohort C:

<table>
<thead>
<tr>
<th></th>
<th>Sem A</th>
<th>Sem B</th>
<th>Sem C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exam Grade vs Moodle % Lectures =</td>
<td>0.10</td>
<td>-0.71</td>
<td>-0.47</td>
</tr>
<tr>
<td>Exam Grade vs Moodle Total Lecture Views =</td>
<td>-0.18</td>
<td>0.58</td>
<td>-0.08</td>
</tr>
</tbody>
</table>
Keep Digging –
What is Accessed & When

12 teaching weeks → Study week → Exam week

Mid-term reading week:
- Week 8 for cohort A (+ 2 weeks for Christmas)
- Week 5 for cohort B
- Cohort C didn’t have a reading week
Number of times lecture notes were viewed per week

The week before the exams

Weeks of the term

Number of hits on lecture notes slides
Conclusion

How do these access patterns match with assessment achievement?

– Too little too late for exams
– By-passed for assignments

Inconsistent
Conclusion

• At institutional level, managers typically see summary stats for a range of modules... and can jump to conclusions too fast
• They don’t see the rich explanatory detail. “My head is on fire, my feet are in ice but, on average, I’m all right”.
• Moodle gives an opportunity to break down the numbers within a module

• **But** the interpretation is critical.
• Metrics need to interpreted by the instructor – someone familiar with the class and how it is structured and run.
• Only they can they make informed judgements based on the metrics.
References


