Two Years of Michigan MOOCs: What Analytics Tell us About Learning in these Environments

Christopher Brooks (brooksch@umich.edu)
Stephanie Teasley (steasley@umich.edu)

Learning Education & Design Lab
School of Information and Digital Education & Innovation
University of Michigan
Overview of MOOC Activity

• Michigan was an early member of the MOOC movement
• Offering courses on the Coursera platform
• Courses are video-based with discussion forums and assignments
• 13 different courses offered in 51 sessions
• Different institutions use MOOCs differently
• Michigan MOOCs show of diversity of programming at Michigan
• Keep in mind: MOOCs are “massive” like black holes
By the Numbers

Unique Learners (enrollees): 1,979,446
Participants: 922,163
  – Graded (8.6%)
    • Certificate 48,911
    • Certificate with Distinction 30,224
  – Ungraded (91.4%)
  – But, roughly 300,000 participants at least tried an assessment item
Lots of data (200+ million clickstream events)
Geographic Location

- Swap between China and Brazil from Fall 2013

Unique Learners by Country (Top 20)

- United States
- India
- China
- Brazil
- United Kingdom
- Canada
- Russia
- Ukraine
- Germany
- Spain
- Colombia
- Singapore
- France
- Australia
- Mexico
- Netherlands
- Greece
- Poland
- Italy
- Philippines
- Colombia
- Ukraine
- Greece
- Singapore
- France
- Australia
- Mexico
- Netherlands
- Greece
- Poland
- United States
- India
- China
- Brazil
- United Kingdom
- Canada
- Russia
- Ukraine
- Germany
- Spain
- Colombia
- Singapore
- France
- Australia
- Mexico
- Netherlands
- Greece
- Poland
- United States
- India
- China
- Brazil
- United Kingdom
- Canada
- Russia
- Ukraine
- Germany
- Spain
- Colombia
- Singapore
- France
- Australia
- Mexico
- Netherlands
- Greece
- Poland
- United States
- India
- China
- Brazil
- United Kingdom
- Canada
- Russia
- Ukraine
- Germany
- Spain
- Colombia
What do people do in MOOCs?

- Learning Analytics Fellows project
- What do “attendance patterns” look like in these kinds of courses?
- Top half is streaming, bottom is downloading
- Blue lines are Mondays, Red lines are assignment due dates
Relationship between Video Viewership and Grade?
Let’s back this horse up

Why are people taking MOOCs? Coursera asks intentions:

• …mastering the course material by working through the exercises and earning a certificate. 126,312 (44%)

• …learning the course material mainly by watching most of the lectures. 122,168 (42%)

• None of the above. I'm just checking out the course for now. 40,650 (14%)

But most of our participants haven’t filled this survey out...
Digging Deeper

• A grounded theory analysis suggested:
  – Personal (cognitive, emotional, social understanding, pleasure, self-confidence, curiosity, fear)
  – Employment (self-employed, under employed, horizontal and vertical movement, career change)
  – Socialization (inside the mooc, and with others in the world)
Digging Even Deeper

– Knowledge Acquisition (depth and breadth of knowledge, skill development)
– Strong relationships to formal education (current students, future students, and lifelong learners)
– Versus other modalities (fast, high quality, structured, accessible, cost)
Learn with Friends

- What is the difference on learner outcomes when they sign up with friends? \( \chi^2 \ p=0.006 \)

<table>
<thead>
<tr>
<th>Distinction</th>
<th>Normal</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friends</td>
<td>34.2%</td>
<td>5.3%</td>
</tr>
<tr>
<td>Alone</td>
<td>24.2%</td>
<td>4.5%</td>
</tr>
</tbody>
</table>

- How does discussion forum posting habits (sociability proxy) change when learners signs up with friends? (t-test, p=0.03)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friends</td>
<td>6.96</td>
<td>8.83</td>
</tr>
<tr>
<td>Alone</td>
<td>4.86</td>
<td>5.06</td>
</tr>
</tbody>
</table>
How do learners interact with friends?

- Casual discussions with family members
- Formal study group discussions with colleagues
- Ad-hoc discussions with colleagues who did not sign up for the course
- Formal professional development activities, with learners using this content directly in their teaching practice
- Collaborative lecture viewing activities,
What do Michigan Students think of MOOCs?

- Stratified semi-random sample of UM students ~800 responses (80 we knew enrolled in our MOOCs)
- Of students who had not enrolled in our MOOCs…
  - 40% had never heard of MOOCs
  - 18% had taken another institutions MOOC

Christopher Brooks & Stephanie Teasley, LED Lab
Student Opinions

- We asked the 40% how likely they would to enroll in MOOCs (large negative skew)

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Neutral</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>170</td>
<td>135</td>
<td>66</td>
</tr>
</tbody>
</table>

- We then asked them if the instructor recommended the MOOC for supplemental resources, how likely would they be to enroll

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Neutral</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>41</td>
<td>142</td>
<td>227</td>
</tr>
</tbody>
</table>

Strongly agree
I enrolled in these MOOCs to:

- Deepen my knowledge in my area of study
- Help determine what my major should be
- Broaden my understanding of other areas
- Explore an area to see if I should take a course in that area for university credit
- Supplement a face-to-face course with additional learning resources
Beyond goals, what is impact?

• Yes, it has strengthened my knowledge in my area (100%)
• Yes, it has introduced me to areas I would not otherwise get exposure to (73.1%)
• Yes, it has encouraged me to switch institutions (4.5%)
• Yes, it has helped me relate my university study to career options (22.4%)
• Yes, it has encouraged me to consider further study (e.g. to eventually apply to a graduate degree or professional program in the area) (37.3%)
Digging Deeper

- Asked whether MOOCs offered by the institution are a “benefit to students enrolled in traditional programs at that institution”
- Small majority of students said yes (375)
- More than a third (34.9%, 231 students) said not sure
- Very few said no (55)
- Just as unclear as the rest of us on the value!
Why are MOOCs a good thing for traditional students?

Cost and Social Need

- Direct and Ancillary
  - Babysitting
  - Commuting
  - Tuition

- Institutional costs
  - Opportunity to have smaller class sizes by allowing the breadth enrollees to take the course online
  - Showcase of programs to auditing students
More Reasons

Self regulation of learning
  – Development of SRL skills
  – Showcasing of SRL ability
Learning styles and approaches
  – Acknowledgement of the diversity of learning skills

“I would be too intimidated to learn a skill like coding in a face-to-face environments because I don’t have a similar skill-set or background to those studying full time...In a classroom for example, I think I’d find the gender disparity of computer science courses unnerving and a little distracting...it’s likely...that my confidence in my own ability to success would diminish in this kind of setting, simple because I’d be a female taking the course ‘for fun’ surrounded by serious students in a field that’s known for being a little competitive and very male-driven.”
• Brand recognition & preeminence
  • Some felt it strengthened the brand, some felt it diminished the value of the degree
  • How many of our learners who said they enrolled in a MOOC chose a UM MOOC?
    • Yes: 38%
    • No: 49%
    • Unsure: 13%

(18% of people who said they had not taken a UM MOOC actually did.)
(No record for over 50% of those who said they did take a UM MOOC.)
Quality

“...there isn't the connection with a professor or peers to facilitate learning and questioning...”

“the professors that teach these courses have an opportunity to engage with the material in a new way which might strengthen their face-to-face teaching. Additionally, many MOOCs have a wide user base which means that the professors need to teach for a multicultural and multinational population. This would most definitely benefit the students learning from this professor.”
How difficult is a MOOC compared to:

Average course at University of Michigan?

Average course at a US university?

Average course offered online for credit?
Value in volunteering in UM MOOCs?

- That this would provide a networking opportunity with faculty and graduate students?
- That this was a good way to understand your major better
- That interactions with MOOC students would be positive
- That over this was an opportunity worth volunteering for.
Student-led MOOC Support

- Wanted to help residential students learn through teaching
- Deepen understanding of content by helping others
- Advanced residential MBA students:
  - Created tutorial video
  - Staff live hangout “office hours”
- Made available to ~2,400 MOOC students
Online students

- Pretty small uptake from MOOC students (17% watched a video)
- Much more involvement from those in signature track (300 vs 105 watchers)
Impact on Residential Students

“...answering live questions is a great way of being drilled in concepts and practicing how to think on your feet. And hearing someone else's questions is a great way of sampling the many different approaches to any given problem, and expanding your perspective for [the] future...”

“The exercise of putting yourself in a viewer's shoes and thinking about how well something is being explained really made me consider the mechanisms of finance in a more comprehensive way.”
Other Residential MOOC Experiments

Introduction to Cataract Surgery

Understanding and Improving the US Healthcare System
Predictive Modelling

• Can we do better in understanding student success in this environment?
• Could we predict who is at-risk for students who want to obtain a certificate?

“So Harvard research confirms: people who pass courses do more stuff than people who don’t.”

http://blogs.edweek.org/edweek/edtechresearcher/2014/03/big_data_mooc_research_breakthrough_learning_activities_lead_to_achievement.html
It’s hard to predict success in MOOCs

• MOOCs lack the diversity of data we have about residential students
  – Previous achievement (SAT/ACT, last year’s course)
  – Socioeconomic status (distance from university, first in family, wealth)
  – Gender
  – Ethnicity
  – Motivation

• Building predictive models of student achievement in learning analytics is largely done on these entry-level features

• Both frustrating and refreshing
  – Want accurate models, but want causes of issues
Solution

• Built a novel feature selection algorithm inspired by work in the text-mining community
• It looks at the pattern of engagement that a student has with course resources
• Build of historical data (last year’s course) to create day-by-day multilevel model’s (C4.5)
• Initial work is based on student certificate achievement (pass/fail) – (not the only valuable outcome variable to try and predict!)
How it works

<table>
<thead>
<tr>
<th>Resource</th>
<th>Day of Course</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Video</td>
<td></td>
</tr>
</tbody>
</table>

Daily Accesses
- Day 1: Yes
- Day 2: No
- Day 3: Yes
- Day 4: No
- Day 5: No
- Day 6: No
- Day 7: No
- Day 8: Yes
- Day 9: No

3-Day counts
- Day 1-3: Yes
- Day 4-6: No
- Day 7-9: Yes

Weekly counts
- Week 1: Yes
- Week 2: Yes

Monthly counts
- Month 1: Yes

For a 104 day long course, with three “resources” (videos, forums, quizzes) this gives us 408 features for the modelling activity.
Text mining often uses n-grams as features in a document

- A bigram (cat, good) is the number of pairs of these two words in a document, a trigam (cat, wasn’t, good), etc.
- We build engagement n-grams up to 5 gram

<table>
<thead>
<tr>
<th>Daily Accesses</th>
<th>Possible bigrams</th>
<th>Possible trigrams:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 1: Yes</td>
<td>[yes, yes]: 0</td>
<td>[yes, yes, yes]: 0</td>
</tr>
<tr>
<td>Day 2: No</td>
<td>[no, no]: 3</td>
<td>[yes, yes, no]: 0</td>
</tr>
<tr>
<td>Day 3: Yes</td>
<td>[yes, no]: 3</td>
<td>[yes, no, yes]: 1</td>
</tr>
<tr>
<td>Day 4: No</td>
<td>[no, yes]: 2</td>
<td></td>
</tr>
<tr>
<td>Day 5: No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 6: No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 7: No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 8: Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 9: No</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For a 104 day long course, with three “resources” (videos, forums, quizzes) this gives us $717$ more features for the modelling activity.
In a nutshell

We don’t have diverse set of data, but
And there is a lot of it (200 million+ clickstream events)
By pulling out patterns of resource access, we can use supervised machine learning (C4.5) techniques to build predictive models
But what if we did have entry data from students?
  – Gender & Ethnicity
  – Certification status
  – Country of origin
  – Etc.
Fliess' $\kappa$ versus Time in Days

Day of Course Offering

Activity Features Only
Demographics Features Only
Activity and Demographics Features
Results

- It is possible to create predictive models on clickstream data for MOOCs
- 3 weeks into the MOOC seems to be an interesting point for some courses
- It is computationally intensive to create these models (daily!)
- MOOC entry/demographics information doesn’t seem to add value


Next Steps

• The what of the prediction is clear, the why is not clear
  – Prediction rules are difficult for instructors to understand, reflecting latent constructs
• Does this method work in blended situations?
  – Is it predictive only for MOOCs, or for blended residential education as well?
• Can we improve model accuracy through more sophisticated modeling techniques (e.g. Bayesian methods?)
• How do we integrate this with platforms like Coursera?
Conclusions

1. Every MOOC is unique (size, domain, pedagogy, student body), and so are the research questions we can ask of it

2. Student success is a nuanced concept in MOOCs
   – What do we value at UM as measures of success?

3. Residential students largely value UM MOOC offerings
   – And there are many ways to engage out students in our MOOC program and be locally impactful

4. The breadth of data available makes for novel techniques for understanding learners

Credits:
A horses' bum: http://www.pbase.com/image/56819062