Teaching Recommender Systems at Large Scale: Evaluation and Lessons Learned from a Hybrid MOOC

Joseph A. Konstan
J.D. Walker
D. Christopher Brooks
Keith Brown
Michael D. Ekstrand

ACM Conference on Learning at Scale
4 March 2014

UNIVERSITY OF MINNESOTA

Background

- Goals
  - Create a high-quality graduate course on recommender systems
  - Create a high-quality MOOC on recommender systems
  - Explore MOOCs broadly; department interest, university interest, dip feet in the water

Intro. to Recommender Systems

- Coursera plus “regular” graduate course
  - On-campus students had help and Q&A sessions; half were recorded for Coursera students
- 14 weeks of content (open / 6 modules / close)
- 42 lectures (average 30 minutes) plus 14 interviews with outside experts; collection of readings and references
- 7 written assignments plus 6 programming assignments
  - Software toolkit for programming recommender systems
  - Mix of programmed and peer grading
- 2 exams (multiple choice)
- 2 tracks for online students – programming/concepts
- Substantial research assessment
- Extensive outreach effort

Key Points of Exploration

- Face-to-Face + Online
  - Reaction of face-to-face students
  - Effect on online students
  - Differences
- Programming vs. Concepts
  - Can two tracks work?
- Everything about Scale
  - Effort, impact, learning ….

Measuring Student Learning

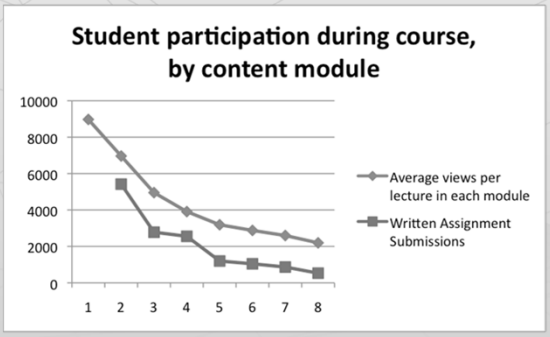
- Pre-test / Post-test Knowledge Assessment
  - Focus on concepts, algorithms, not programming

Q. What is the core idea behind dimensionality reduction recommenders?

a. To reduce the computation from polynomial to linear.
b. To strip off any product attributes so products appear simpler.
c. To reduce the computation time from $O(n^3)$ to $O(n^2)$
d. To transform a ratings matrix into a pair of smaller taste-space matrices.
e. I have no idea.

A Few Statistics

- Total Enrollment: 28,389
  - 7000 never did a single activity
  - 2195 still watching videos at the end



Why Researching MOOCs Is Hard

- Lack of motivation
- Very diverse student population
- Lack of information on students' background, aptitude, pre-course knowledge, etc.
- Non-random attrition

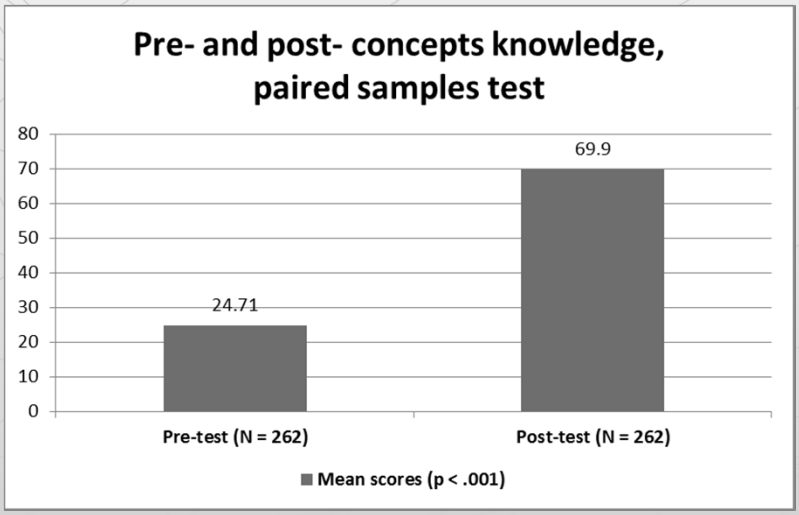
Research Questions & Design

- Do students learn in a MOOC?
  - How much?
  - Which ones?
  - What variables moderate student learning?
- How does the learning of face-to-face students in a hybrid class compare to that of fully online students?

Measures and Data

- Pre-post student surveys (demographics, motivations, etc.)
- Pre-post knowledge test (recommender systems concepts)
- Process and result metrics from Coursera

Findings: Learning gains



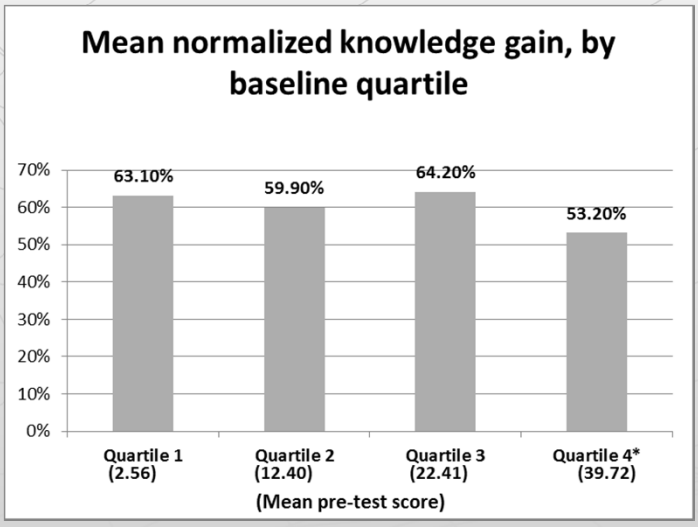
Findings: Learning Gains

- Normalized gains:

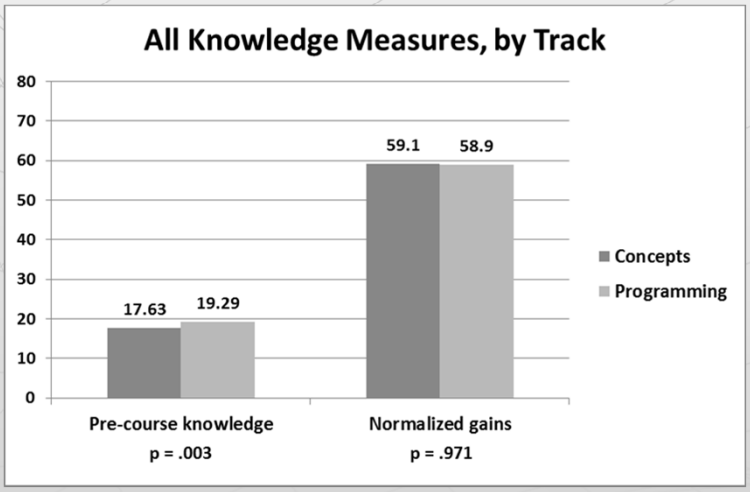
(Post-test – Pre-test / 100 – Pre-test)

- Treats learning gains as a percentage of each student's total possible gain
- Accounts for ceiling effects

Findings: Learning Gains

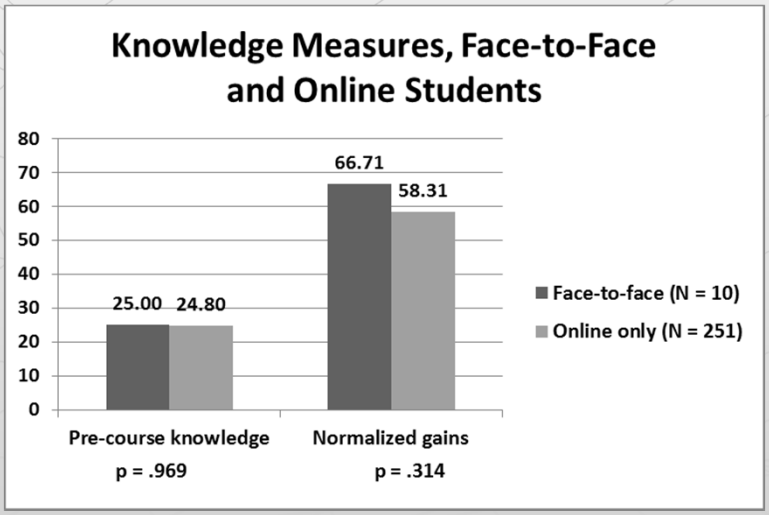


Findings: Learning Gains



(N = 4058) (N = 253)

Findings: Learning Gains



Recommender Systems @ Large Scale: Models for Predicting Completion and Success				
		Completion	*Normalized Gains*	*Final Grades*
Academic Activity	# of written assignments		+	
	# of concurrent courses	-		-
	# of programming courses	0	0	+
	# of previous MOOCs	+		0
	Programming vs concepts	0	-	+
	Programming confidence	0	0	0
Demographics	English Proficiency	0	0	+
	Location: USA	0	0	0
	Age	0	0	0
	Sex	0	0	0
Intent to complete		+	0	0
Pre-course knowledge		+	-	0
N		3326	207	181
Pseudo/Adjusted R^2		.059	.202	.516
Chi-square/*F* Test		102.18****	5.030****	13.064***

Teaching Recommender Systems @ Large Scale
J.A. Konstan and J.D. Walker

UNIVERSITY OF MINNESOTA

Recommender Systems @ Large Scale: Models for Predicting Completion and Success				
		Completion	*Normalized Gains*	*Final Grades*
Academic Activity	# of written assignments		+	
	# of concurrent courses	-		-
	# of programming courses	0	0	+
	# of previous MOOCs	+		0
	Programming vs concepts	0	-	+
	Programming confidence	0	0	0
Demographics	English Proficiency	0	0	+
	Location: USA	0	0	0
	Age	0	0	0
	Sex	0	0	0
Intent to complete		+	0	0
Pre-course knowledge		+	-	0
N		3326	207	181
Pseudo/Adjusted R^2		.059	.202	.516
Chi-square/*F* Test		102.18****	5.030****	13.064***

Teaching Recommender Systems @ Large Scale
J.A. Konstan and J.D. Walker

UNIVERSITY OF MINNESOTA

Recommender Systems @ Large Scale: Models for Predicting Completion and Success				
		Completion	*Normalized Gains*	*Final Grades*
Academic Activity	# of written assignments		+	
	# of concurrent courses	-		-
	# of programming courses	0	0	+
	# of previous MOOCs	+		0
	Programming vs concepts	0	-	+
	Programming confidence	0	0	0
Demographics	English Proficiency	0	0	+
	Location: USA	0	0	0
	Age	0	0	0
	Sex	0	0	0
Intent to complete		+	0	0
Pre-course knowledge		+	-	0
N		3326	207	181
Pseudo/Adjusted R^2		.059	.202	.516
Chi-square/*F* Test		102.18****	5.030****	13.064***

Teaching Recommender Systems @ Large Scale
J.A. Konstan and J.D. Walker

UNIVERSITY OF MINNESOTA

Recommender Systems @ Large Scale: Models for Predicting Completion and Success				
		Completion	*Normalized Gains*	*Final Grades*
Academic Activity	# of written assignments		+	
	# of concurrent courses	-		-
	# of programming courses	0	0	+
	# of previous MOOCs	+		0
	Programming vs concepts	0	-	+
	Programming confidence	0	0	0
Demographics	English Proficiency	0	0	+
	Location: USA	0	0	0
	Age	0	0	0
	Sex	0	0	0
Intent to complete		+	0	0
Pre-course knowledge		+	-	0
N		3326	207	181
Pseudo/Adjusted R^2		.059	.202	.516
Chi-square/*F* Test		102.18****	5.030****	13.064***

Teaching Recommender Systems @ Large Scale
J.A. Konstan and J.D. Walker

UNIVERSITY OF MINNESOTA

Lessons Learned

- What Everyone Already Knows …
  - Good teaching at scale -> more time and effort
  - Students leave MOOCs – it is nothing personal
  - Students claim they want short (courses,videos)
  - Students really serious about “grades”
  - Students (mostly) hate peer grading
  - Multiple choice reverses the effort function
  - Demands not proportional to tuition paid!
    - But demands can be self-correcting!

Lessons Learned

- What everybody may not know
  - Long format actually does work (14 weeks, longer lectures) and many students preferred it
    - Of course, many students pick-and-choose
  - Face-to-face students overwhelmingly preferred the MOOC approach
    - Better time management
    - Better able to pace learning; rewind; language issues
  - The two-track approach worked! Possible to learn the concepts well alongside programmers

Some Unusual Experiences

- Creating a class dataset worked really well (over 5000 students contributed data)
- Personal assignment test data worked well (each student had separate test cases)
  - Some significant effort; tools needed here
- Open source infrastructure worked well to distribute tools; more than 1000 students used our toolkit to program

And the Heartwarming Tales

- The rare coin marketplace
- The Russian consultants
- Meeting students in Hong Kong

Future work …

- Peer grading and alternatives
- Reducing effort associated with later offerings
- Study learning effects on volunteer TAs

- and all sorts of ideas from this conference …

Thank you!

- To our colleagues
- To our students
- And to you …

Questions?

Teaching Recommender Systems at Large Scale: Evaluation and Lessons Learned from a Hybrid MOOC

Joseph. A. Konstan
J.D. Walker
D. Christopher Brooks
Keith Brown
Michael D. Ekstrand

ACM Conference on Learning at Scale
4 March 2014

UNIVERSITY OF MINNESOTA