



# Engagement and Success in Online Learning: Higher Education and Beyond

Ryan Baker

@BakerEDMLab

University of Pennsylvania



I'm in trouble

# I'm in trouble

- How do you follow an act like that?

# I'll Start With a Joke

- That's always safe, right?



- “I’m nervous about my talk. How do you avoid getting butterflies in your stomach?”

- “DON’T



- “DON’T EAT



- “DON’T EAT CATERPILLARS.”







OK I feel better now

# Thank you

- For welcoming me here today
- It's a great honor to have a second opportunity to speak at one of the world's great centers for research on online learning

# In my last visit here...

- I discussed our work to model affect and disengagement using automated detectors built through educational data mining
- And how our detectors can detect constructs in middle school that predict college attendance

I can't do that again

# So this time

- I would like to tell you about some of our work to study how engagement within online learning corresponds to success in higher education and into learners' careers

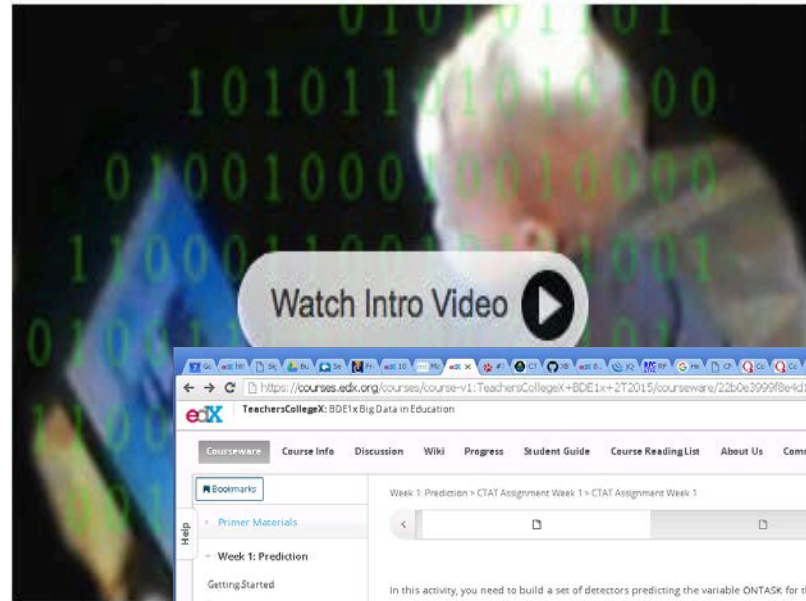
# MOOCs and online courses

TEACHERS COLLEGE  
COLUMBIA UNIVERSITY

## Big Data in Education

Education is increasingly occurring online or in educational software, resulting in an explosion of data that can be used to improve educational effectiveness and support basic research on learning. In this course, you

mining



A HISTORY OF THE UNITED STATES: 1865 TO PRESENT

### INVESTIGATION

#### Classified: The Soviet War in Afghanistan

When the USSR showed interest in acquiring Afghanistan, American forces intervened by aiding the Afghan mujahedin with supplies in an "endless" war against Soviet-backed troops. The following video, excerpted from the CNN *Cold War* documentary series, shows the difficulties of protecting countries from communist regimes.

#### Nine Years of War

Use the video and link at left to answer the questions below.

**Nine Years of War**

Why was Afghanistan of interest to the Soviet Union?

- A. Afghanistan shared a border with the Soviet Union's central Asian republics.
- B. The United States was deeply involved in Afghan politics.
- C. The Soviets wanted to stop Afghan narcotics from entering the USSR.
- D. Afghanistan was close to the Persian Gulf's oil.
- E. Afghanistan shared a border with the Soviet Union's central Asian republics AND it was close to the Persian Gulf's oil.

The civil unrest that broke out in Afghanistan in 1978 was a reaction to which of the following?

- A. the election of Nur Mohammad Taraki
- B. land reforms

32:51

Transcript

TeachersCollegeX: BDE1x Big Data in Education

Courseware Course Info Discussion Wiki Progress Student Guide Course Reading List About Us Community

Bookmarks

- Primer Materials
- Week 1: Prediction
  - Getting Started
  - Lectures
  - Activity: Rapid Miner Walkthrough
  - CTAT Assignment Week 1**  
Assignment due Aug 26, 2015 at 14:00 UTC
  - Bazaar Assignment 'Week 1
  - due Aug 26, 2015 at 14:00 UTC
  - Week 1 General Discussion
- Week 2: Diagnostic Metrics and Cross-Validation
- Week 3: Feature Engineering and Behavior Detection
- Week 4: Knowledge Inference and Knowledge Structures
- Week 5: Relationship Mining
- Week 6: Visualization
- Week 7: Clustering and

Week 1: Prediction > CTAT Assignment Week 1 > CTAT Assignment Week 1

In this activity, you need to build a set of detectors predicting the variable ONTASK for the data file CogSci-Godwinetal-2013-3.csv.

WEEK 1 CTAT ACTIVITY (EXTERNAL RESOURCE) (0.076923076 / 1.0 points)

It looks like you did not delete the student variable.

Question 1 of 13: Build a decision tree (using operator W:348 from the Weka Extension Pack) on the entire data set. What is the non-cross-validated kappa?

0.346

Question 2 of 13: That's the correct answer, but let's think about it. The kappa value you just obtained is artificially high – the model is over-fitting to which student it is. What is the non-cross-validated kappa, if you build the model (using the same operator).

0.346

# Disengagement is a problem

- For example
- Most students who register for a MOOC do not complete it (Jordan, 2013, 2014; Kizilcec et al., 2013; Khalil & Ebner, 2014; Ruby et al., 2015)



# Research Questions

- Can we determine which forms of engagement matter more, so we can provide predictive analytics to instructors about the behaviors that matter most?

# Predicting Success in Higher Education

- Online courses
- MOOCs

# Predicting success in online course

- Considerable work trying to determine which factors lead to student success in online courses

(see, for instance, Arnold & Pistilli, 2012; Wolff et al., 2013)

- Much of the published work uses demographics as predictors
  - *\*very\** important
  - but not ideal for use in predictive models driving intervention
    - harder to take rapid action to address than behavioral/engagement based predictors
- Insufficient exploration of when to use indicators – does “homework not done yet” mean the same thing at different points in the semester?

# Context

- Soomo Online Learning Platform
- Used by large online universities, both for-profit and non-profit

**A HISTORY OF THE UNITED STATES: 1865 TO PRESENT**

13 The Reagan and Bush Years, 1980–1992 / 13.5 Women and the New Right  
Questions: 0 of 4 complete (0%) | 0 of 4 correct (0%)

### Women and the New Right

The role of women was equally as important as it had been to the New Left of the 1960s. The Equal Rights Amendment (ERA), conservative institutions that reflected the gender. Reagan's nomination of Sandra Day O'Connor encouraged conservative women, but more out of hope that O'Connor would be one of the others on the Court, however, votes were cast in favor of the procedure). Feeling betrayed by the evangelicals began organizing directly.

Figure 13.4

**Get the Gist?** [Hide Study](#)

Who became the first female Supreme Court justice?

- Sonia Sotomayor
- Ruth Bader Ginsburg
- Sandra Day O'Connor
- Opha Mae Johnson
- Elena Kagan

Many conservative women were opposed to which of the following?

- gender-based divisions of labor
- abortion
- adoption
- male-dominated households
- women in leadership positions

What was the "second shift"?

- the hardship of maintaining a career in addition to family
- the wave of conservative American women moving into leadership positions
- a move by the Democratic Party to include more conservative ideals
- the growing acceptance of single women becoming mothers

**A HISTORY OF THE UNITED STATES: 1865 TO PRESENT**

## INVESTIGATION

### Classified: The Soviet War in Afghanistan

When the USSR showed interest in acquiring Afghanistan, American forces intervened by aiding the Afghan mujahedin with supplies in an "endless" war against Soviet-backed troops. The following video, excerpted from the CNN *Cold War* documentary series, shows the difficulties of protecting countries from communist regimes.

#### Nine Years of War

Use the video and link at left to answer the questions below.

**Nine Years of War**

Why was Afghanistan of interest to the Soviet Union?

- A. Afghanistan shared a border with the Soviet Union's central Asian republics.
- B. The United States was deeply involved in Afghan politics.
- C. The Soviets wanted to stop Afghan narcotics from entering the USSR.
- D. Afghanistan was close to the Persian Gulf's oil.
- E. Afghanistan shared a border with the Soviet Union's central Asian republics AND it was close to the Persian Gulf's oil.

The civil unrest that broke out in Afghanistan in 1978 was a reaction to which of the following?

- A. the election of Nur Mohammad Taraki

**soomo**  
LEARNING

# Goal

- Predict early in the course which students at-risk of not obtaining a passing grade
- Using actionable indicators that can be easily understood and used by instructors and administrators

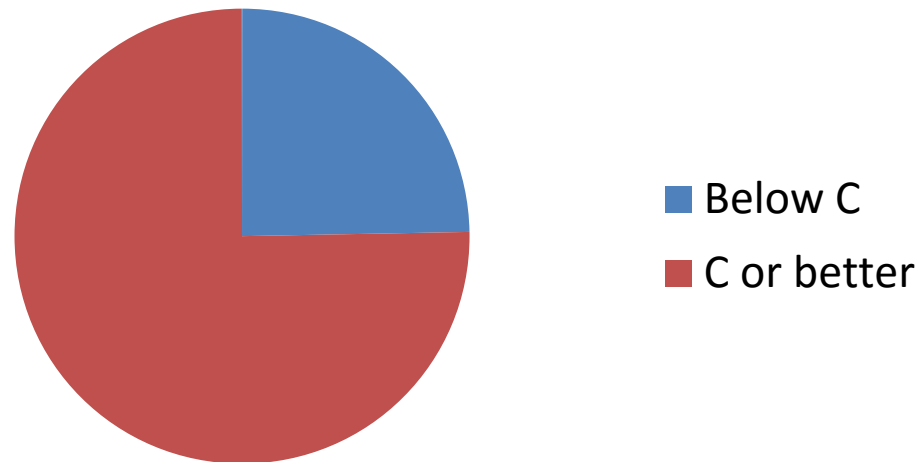
# Data set

- 4,002 students in 140 sections across 6 terms
- U.S. history
- Private non-profit university
- 2.1 M interactions with system

# Goal

- Predict who gets a C or better
- Necessary for continued financial aid

**Proportion of Students Passing**



# Findings

- Students who have not yet opened the text *before the class starts* have almost a 50% chance of getting a D or F (precision)
- and this indicator captures 70% of the students who will get a D or F (recall)



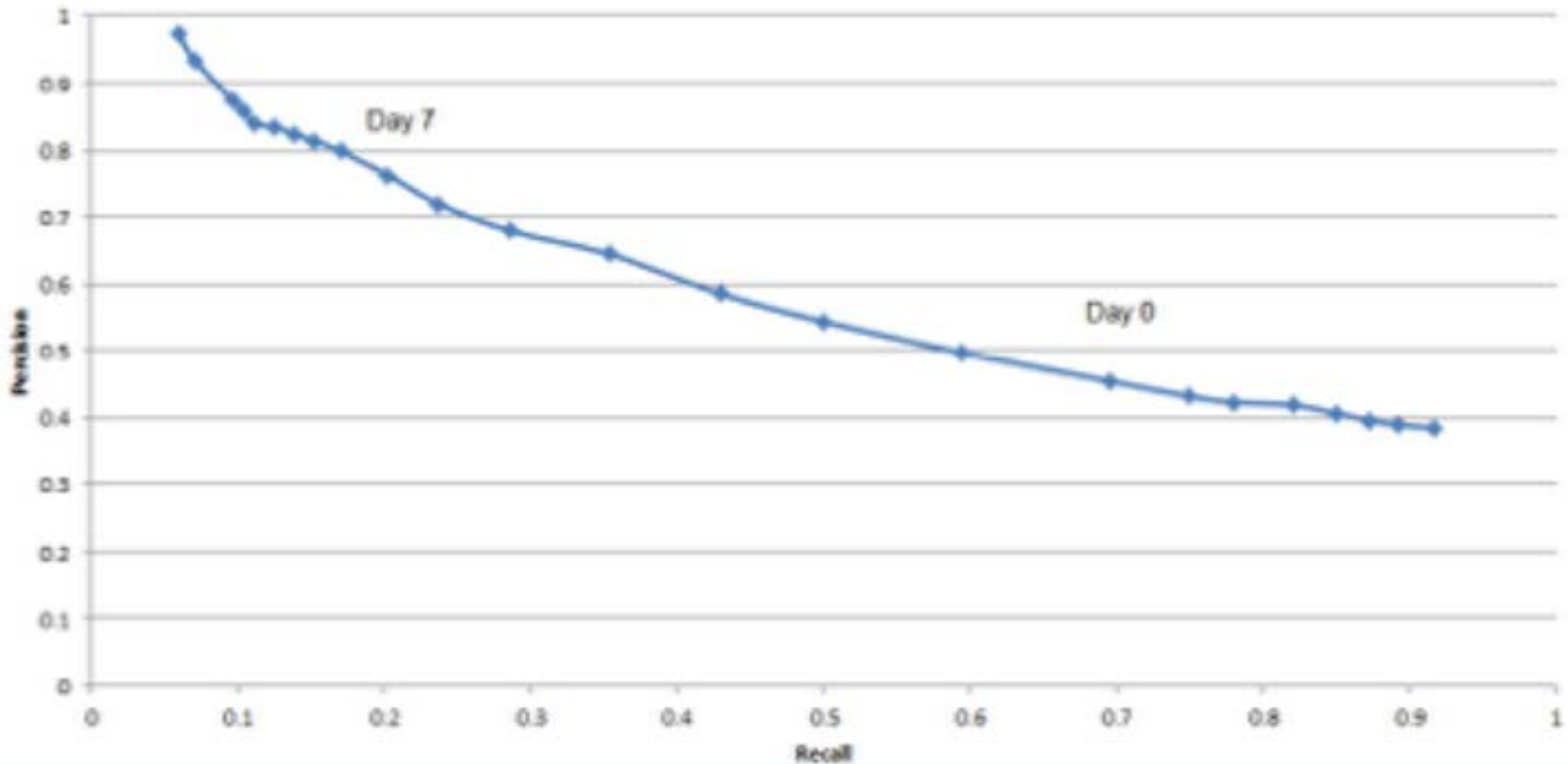
# Findings

- The same indicator – has the student opened the textbook yet
- Remains predictive one week after the class starts, but with very different metrics

# Findings

- The same indicator – has the student opened the textbook yet
- Remains predictive one week after the class starts, but with very different metrics
- Almost 80% of the students who have not opened the textbook yet by the end of the first week will get a D or F
- But only 20% of the students who will get a D or F haven't opened their textbook yet

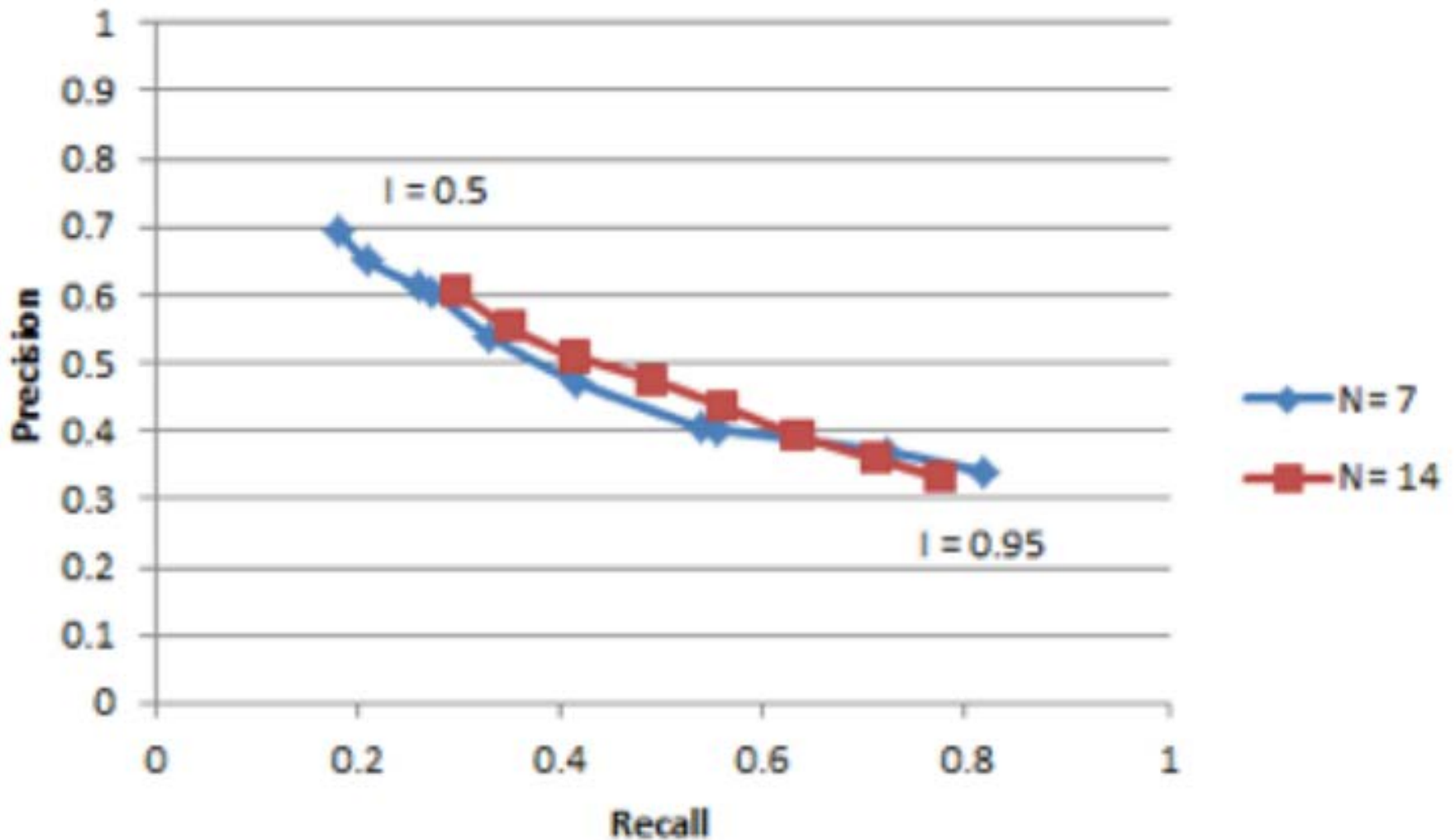
# Precision-Recall Tradeoff



# Findings

- Poor performance on early assignments is very predictive
- Half of students who get below a C on the first assignment will get a D or F for the class
- Half the students who will get a D or F for the class get below a C on the first assignment

# Precision-Recall Tradeoff



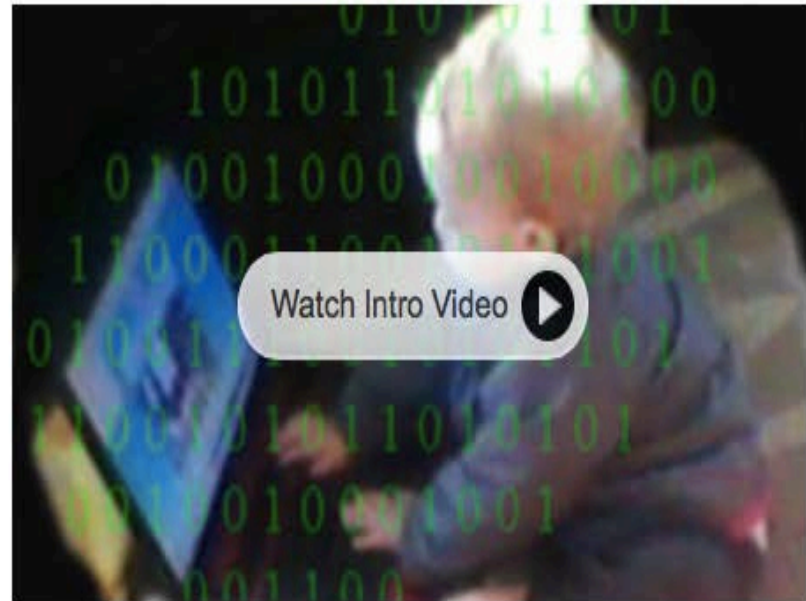
# Conclusion

- Early indicators can be very powerful
- Even if they are very simple indicators
- Provide quick indicators to instructors and student advisors of who is at-risk

# Predicting Success in MOOCs

# Big Data in Education

Education is increasingly occurring online or in educational software, resulting in an explosion of data that can be used to improve educational effectiveness and support basic research on learning. In this course, you will learn how and when to use key methods for educational data mining and learning analytics on this data.

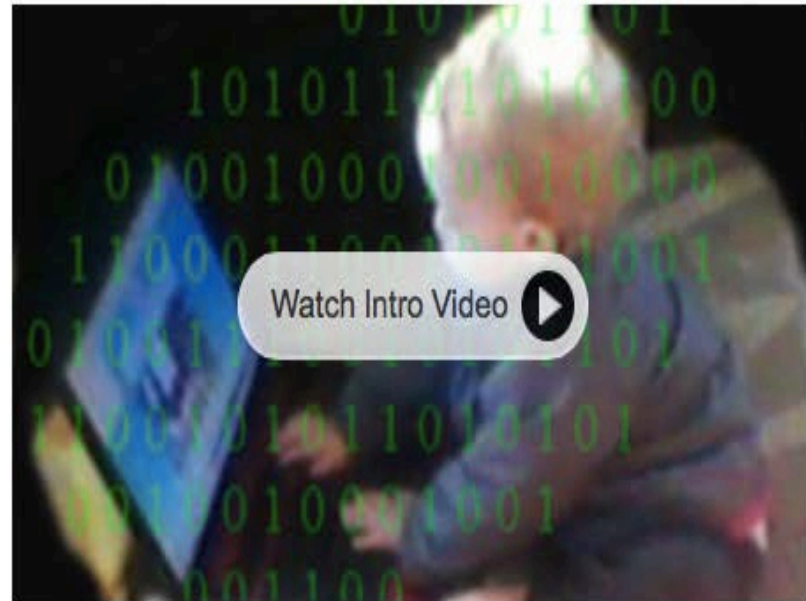


- Oct. 28, 2013 ~ Dec. 26, 2013
- <https://www.coursera.org/course/bigdata-edu>
- Content Area
  - Educational Data Mining
  - Learning Analytics
  - Theory and Application
  - Apply methods to answer research questions
  - Research design and evaluation



# Big Data in Education

Education is increasingly occurring online or in educational software, resulting in an explosion of data that can be used to improve educational effectiveness and support basic research on learning. In this course, you will learn how and when to use key methods for educational data mining and learning analytics on this data.



- July 1, 2015 ~ Sept 8, 2015
- <https://www.edx.org/course/big-data-education-teacherscollegex-bde1x>
- Content Area
  - Educational Data Mining
  - Learning Analytics
  - Theory and Application
  - Apply methods to answer research questions
  - Research design and evaluation

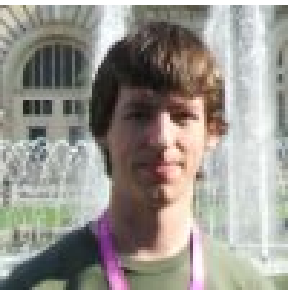
# Course staff



- Instructor



- Elle Wang, Teaching Assistant



- Luc Paquette, Head Community TA

# 2<sup>nd</sup> Iteration

- Intelligent-tutor based assignments in CTAT
- Collaborative chat in Bazaar
- Tool walkthroughs
- Enhanced lectures

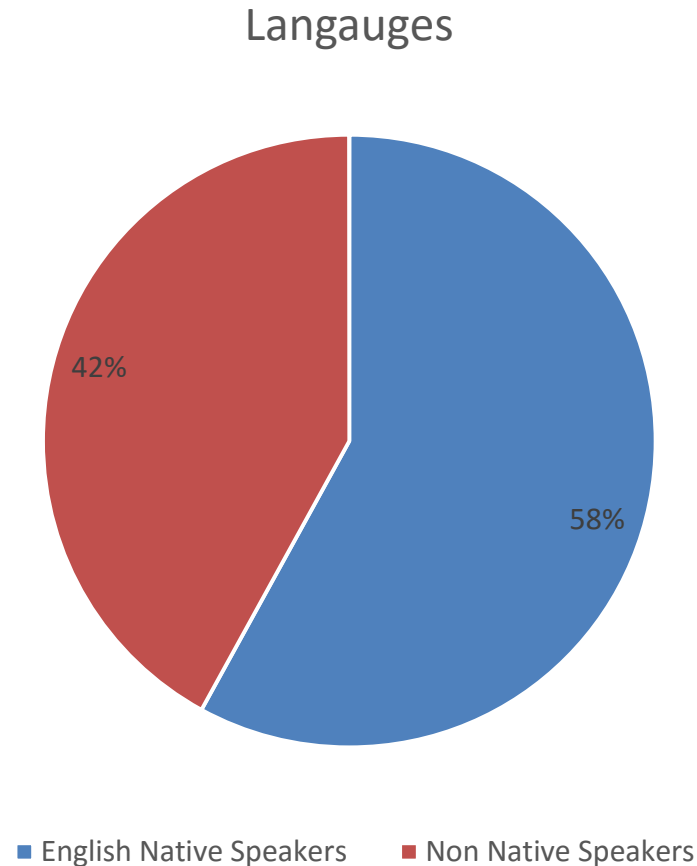


# Key components (2013 Edition)

- Videos
- Assignments
- Discussion Forums
- Self-organized study groups
  - Facebook
  - LinkedIn

# Students & Enrollment

- Over 48,000 students at official course end
- Over 106 different languages spoken



# Common Research Question

- Why do so few people complete MOOCs?

# Partial Answer

(Kizilcec et al., 2013)

- Most students who join a MOOC never have a goal of completing
- They want to learn some of the material
- Or browse in a new area
- Or many other potential motivations

# Our Group's Research Question

- What aspects of MOOC participation predict long-term participation in community of practice?



# In this context

- What characterizes the learners who choose to participate in the EDM community after taking the MOOC?

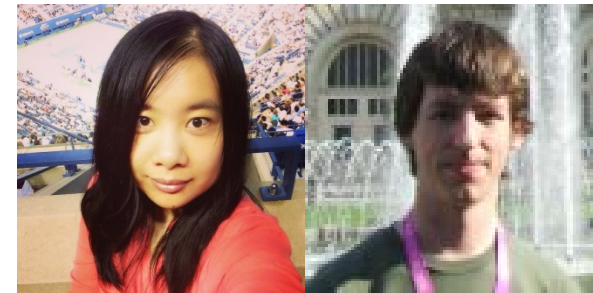
# Operationalizations

- Joining the EDM Society
- Submitting a paper to EDM conference or LAK conference

# Two rounds of analysis

- Round 1 – Summer 2014
  - Data on who joined Society during course or in first months after course
- Round 2 – Fall 2015
  - Data on who joined Society so far
  - Data on who submitted paper in 2014 or 2015

# Initial Finding (Wang et al., 2014)



# Initial Finding

## (Wang et al., 2014)

- 35 students joined EDM Society during or in first several months after class
  - Out of a total membership of 244
- 20.0% of students who joined society completed course
- 1.3% of remaining students completed course
- $\chi^2(1) = 97.438, p < 0.001$

# Indicates

- Course completion may not be the only thing that matters
- But it is clearly a strong indicator of investment in the topic area

# Second-round findings (Wang & Baker, submitted)

- 48 students joined EDM Society during or after class
- 148 students submitted papers to EDM or LAK after class



# Second-round findings (Wang et al., submitted)

- Both society joiners and paper submitters
  - Watched more lecture videos
  - Submitted more assignments
  - Read the forums more often
  - Read the course syllabus more often
- But they do not
  - Post more to the forums
  - Respond more to posts
  - Rate posts more often



# Second-round findings (Wang et al., submitted)

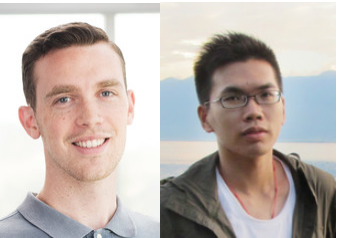
- People who submit a paper are ten times more likely to have completed (13.5%) than non-submitters (1.2%)
- People who join the society are more than ten times more likely to have completed (18.7%) than non-completers (1.3%)

# Future Work

- Study social media participation during course (e.g. Joksimovic et al., 2015) as predictor of future career participation

# Future Work

- Follow these learners forward in their career
- Ongoing collaboration with Dan Davis & Guanliang Chen



What predicts completion?

# What predicts completion?

- First week assignment performance (Zhang et al., in preparation)



# What predicts completion?

- Watching more videos (Zhang et al., in preparation)
- Downloading more videos (Zhang et al., in preparation)



# What predicts completion?

- More posts (Crossley et al., 2015)
- Shorter posts (Crossley et al., 2015)
- Linguistically more concrete posts (Crossley et al., 2015)
- Linguistically more cohesive posts (Crossley et al., 2015)
- Posting in same thread as other students who complete course (Brown et al., 2015)



# What predicts completion? (Wang & Baker, submitted)

- Students who express an intention to complete but have low grit (e.g. Duckworth et al., 2007; Duckworth & Quinn, 2009)
- Are less likely to complete the course
- Than students who have high grit and no intention of completing the course





# Future Work

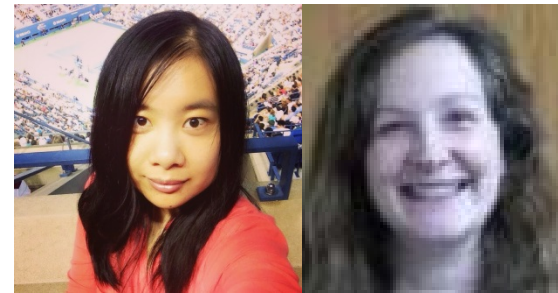
- Integrate models with different information
- Do some types of information about learning better predict learner outcomes than others?
- Which combinations of features are most powerful?

# Future Work

- Do intelligent tutor-based assignments give us additional information about learners, compared to the traditional quiz-style assignments typically used in edX and Coursera?

# Negativity Towards Instructor (Comer, Baker, & Wang, 2015)

- Verbal abuse of instructor on forums or other venues
- A significant disengaged behavior and a problem in many MOOCs



# Negativity Towards Instructor

- BDE had some of this, but actually far less than many other courses
- In other courses,
  - threats of violence towards instructors
  - sexually violent postings
  - hundreds of personal attacks towards instructors

# Negativity Towards Instructor

- Can be very upsetting to instructors, leading to disengagement from forums, other disengagement during courses, not teaching a course again, and stronger negative impacts (Parry, 2013; Freedom, 2013; McGuire, 2014; Head & Lessons, 2013; Head, 2014; Comer, 2014; Tham, 2014)

# In BDE2013

- One student repeatedly attacked instructor whenever instructor posted acknowledging error or imperfection in course
  - This student is known for doing the same thing on other courses in Coursera and Udacity
  - I read an example of this student's posting in a talk in Beijing, and someone else recognized their writing style and asked if it was *name* (and they were right!)
- Complaints about content, presentation, assignments
- Discussion about instructor's clothes and mannerisms

# Example

- “Baker is a dedicated teacher and even records video lectures while incarcerated. At least it looked like an orange prison jumpsuit in the week 7 and 8 videos...”

# Example

- “Baker is a dedicated teacher and even records video lectures while incarcerated. At least it looked like an orange prison jumpsuit in the week 7 and 8 videos...”

Alternatively, analysis is conducted using

- Pairs of words, in order, called *bigrams*
- Triplets of words, in order, called *trigrams*
  
- “Colorless green ideas sleep furiously”
- Bigrams: “Colorless green”, “green ideas”, “ideas sleep”, “sleep furiously”





# Prevalence

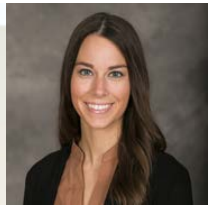
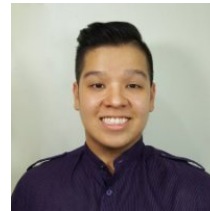
- In BDE, just 9 students out of 48,000 engaged in this type of negativity more than once
  - All male (or chose male names)
- Much higher response rate than course average on pre-course survey
- Not notably different from rest of class in terms of self-efficacy or goal orientation

# Interest in Future Work on...

- Studying how to support instructors better when negativity occurs
- “Rallying around” support anecdotally seemed to help in multiple-instructor MOOC, DALMOOC
- Banning probably unlikely to help, due to sock puppetry
  - Invisible posting might be a good alternative?

# More Ongoing Work in BDEMOOC

- Replicating 21 published findings on MOOCs with BDEMOOC data (Andres et al., in preparation)
- Through this, creating a framework for further replication of published findings on other MOOC data
- First step towards general PLeG framework that can automatically identify at-risk students and determine how to effectively intervene



# Conclusions

- BDEMOOC has been an opportunity to share research and methods for learning analytics and educational data mining with a wider audience
- It's also provided a data set that we have been able to use for a range of analyses

# Bigger Themes

- Engagement manifests in a large variety of ways in online learning
- We can detect and track engagement
- Engagement matters for long-term student outcomes!

# Learn More



[twitter.com/BakerEDMLab](https://twitter.com/BakerEDMLab)



Baker EDM Lab



[weibo.com/u/5370802148](https://weibo.com/u/5370802148)

Baker EDM Lab

See our free online MOOC “Big Data and Education”

Offered as EdX MOOC, next iteration 2017

All lab publications available online – Google “Ryan Baker”

BILL & MELINDA  
GATES *foundation*

**LearnLab**  
Pittsburgh Science of Learning Center



**Penn**  
UNIVERSITY of PENNSYLVANIA



# Extra Slides

# Emerging Work

- Use engagement data to improve instruction and learner support



# Influence design

- Determining design features associated with differences in engagement (Baker et al., 2009; Doddannara et al., 2014; Slater et al., in preparation)
- Providing real-time info to teachers when their instruction is less engaging (Carvalho et al., 2013)



# Reports on engagement to instructors and student advisors

- For Soomo learning platform
- Currently sent as weekly Excel sheets in email
- Used to determine which students to intervene for

# Guidance Counselor Reports (Ocumpaugh et al., in preparation)

all_name	Math Proficiency	Math Problem	Carelessness	Boredom	Inadequate Help-S	Biggest Factor
	Green	Yellow	Red	Green	Green	Carelessness
	Red	Yellow	Green	Yellow	Green	Math Proficiency
	Yellow	Yellow	Green	Yellow	Yellow	Math Proficiency
	Green	Yellow	Red	Green	Green	Carelessness
	Yellow	Yellow	Green	Yellow	Yellow	Math Problems Attempted
	Green	Green	Red	Yellow	Yellow	Carelessness
	Red	Yellow	Green	Yellow	Yellow	Math Proficiency
	Yellow	Yellow	Red	Green	Green	Carelessness
	Yellow	Yellow	Green	Green	Green	Math Proficiency
	Green	Green	Red	Green	Yellow	Carelessness
	Yellow	Yellow	Red	Green	Green	Carelessness
	Green	Green	Red	Green	Yellow	Carelessness
	Green	Yellow	Red	Red	Yellow	Boredom
	Green	Yellow	Red	Yellow	Yellow	Carelessness
	Green	Yellow	Red	Yellow	Yellow	Carelessness
	Green	Yellow	Red	Yellow	Green	Carelessness

# Reports to Regional Coordinators (Almeda et al., in preparation)

- Another online curriculum we work with, Reasoning Mind, deploys reports on student engagement to regional coordinators
- Enabling them to target teachers for additional support and professional development

# Eventual Goal

- Understand how engagement influences long-term success
- Use this understanding to help students succeed

We have developed models that can infer student engagement in real-time

- **Automated:** Able to make assessments about students in real-time, with no human in the loop

# We have developed models that can infer student engagement in real-time

- **Automated:** Able to make assessments about students in real-time, with no human in the loop
- **Fine-grained:** Able to make assessments about students second-by-second

# We have developed models that can infer student engagement in real-time

- **Automated:** Able to make assessments about students in real-time, with no human in the loop
- **Fine-grained:** Able to make assessments about students second-by-second
- **Validated:** Agrees with human judgment



# We have developed models that can infer student engagement in real-time

- **Automated:** Able to make assessments about students in real-time, with no human in the loop
- **Fine-grained:** Able to make assessments about students second-by-second
- **Validated:** Agrees with human judgment
- **Generalizable:** Demonstrated to apply to new students and new contexts
  - For example: validating that models function accurately across urban, rural, and suburban populations

# Detectors Built For

- ASSISTments
- Science ASSISTments/InqITS
- EcoMUVE
- SQL-Tutor
- Aplusix
- BlueJ
- Cognitive Tutors for Math, Genetics
- Reasoning Mind
- vMedic
- Newton's Playground/Physics Playground

# Inferring

- Gaming the System
- Carelessness
- Off-Task Behavior
- Boredom
- Frustration
- Confusion
- Engaged Concentration

# Detectors in Middle School Math Can Predict

- Standardized exam (Pardos et al., 2014)
- College attendance (San Pedro et al., 2013)
- College selectivity (San Pedro et al., in preparation)
- College major (San Pedro et al., 2014, 2015)



# Guidance Counselor Reports (Ocumpaugh et al., in preparation)

all_name	Math Proficiency	Math Problem	Carelessness	Boredom	Inadequate Help-S	Biggest Factor
	Green	Yellow	Red	Green	Green	Carelessness
	Red	Yellow	Green	Yellow	Green	Math Proficiency
	Yellow	Yellow	Green	Yellow	Yellow	Math Proficiency
	Green	Yellow	Red	Green	Green	Carelessness
	Yellow	Yellow	Green	Yellow	Yellow	Math Problems Attempted
	Green	Green	Red	Yellow	Yellow	Carelessness
	Red	Yellow	Green	Yellow	Yellow	Math Proficiency
	Yellow	Yellow	Red	Green	Green	Carelessness
	Yellow	Yellow	Green	Green	Green	Math Proficiency
	Green	Green	Red	Green	Yellow	Carelessness
	Green	Yellow	Red	Green	Green	Carelessness
	Green	Green	Red	Green	Yellow	Carelessness
	Green	Yellow	Red	Red	Yellow	Boredom
	Green	Yellow	Red	Yellow	Yellow	Carelessness
	Green	Yellow	Red	Yellow	Yellow	Carelessness
	Green	Yellow	Red	Yellow	Green	Carelessness



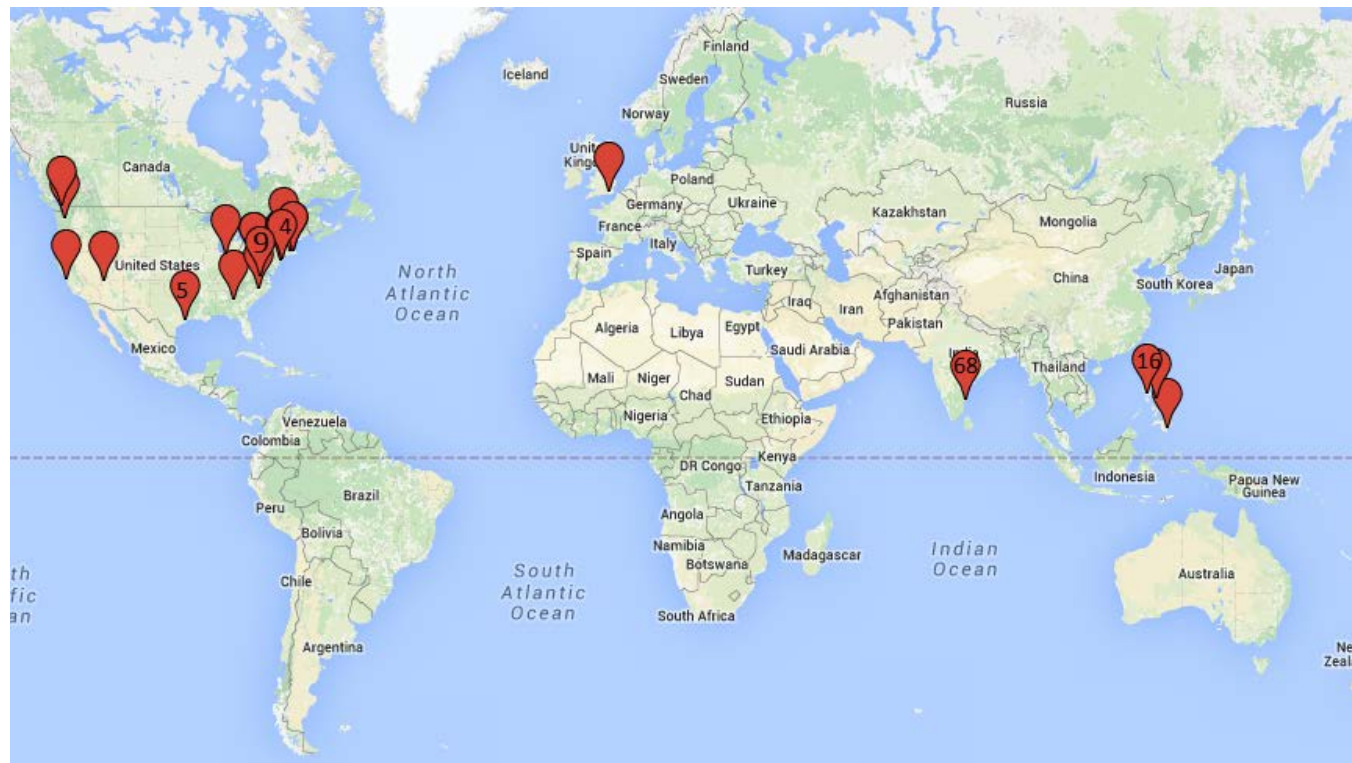
# Measuring affect and engagement in real classrooms

- Using Android app HART (Ocumpaugh et al., 2015a)
- Field observation protocol BROMP (Ocumpaugh et al., 2015b)
- Originally developed for building automated detectors (Baker et al., 2012)



# BROMP usage

- Over 150 certified coders in 4 countries
- Achieve inter-rater reliability over 0.6 with other certified coder



# BROMP usage

- Studying student engagement, curricular design, and dropout in Chennai, India public schools (Hymavathy et al., 2014, in preparation)
  - Over 100,000 field observations conducted
  - Goal of deploying 1 BROMP coder in every public school in city of 7M
- Studying student engagement in Black Rock Forest informal learning (Carvalho et al., 2003)
- Large-scale research on engagement and teacher practices in Pittsburgh-area private and charter schools (Godwin et al., 2013, under review)

