CROWDSOURCING
PERSONALIZED Online
EDUCATION

Logistics

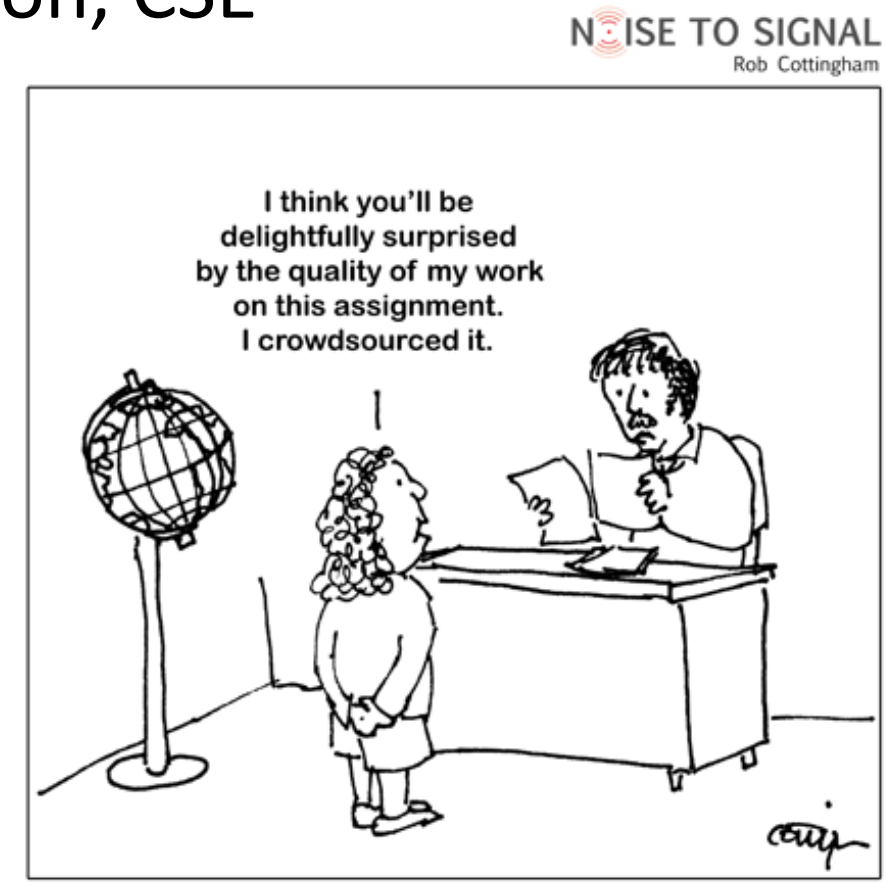
- Schedules
- Shirts
- Room keys
- Lunch
- Presentations

Thanks To…

Microsoft Research

University of Washington, CSE

Kay Beck-Benton



Why?

Education + Crowdsourcing

KNEWTON

piazza

Ask. Answer. Explore. Whenever.

Coursera

udacity

edX

MINERVAE VNIVERSITAS · sapientia critica · MMXI

Educational Goals

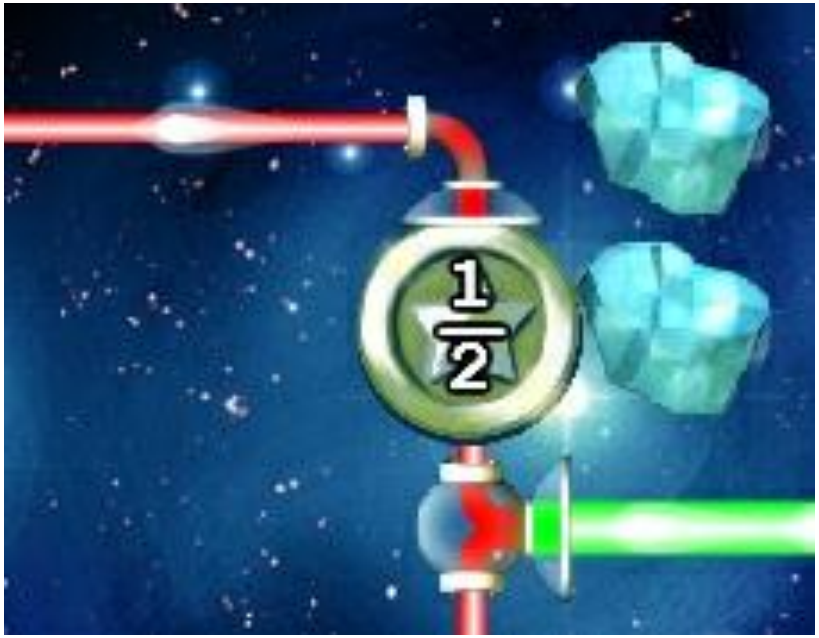
- Increasing Access: MOOCs
  - Need
  - Cost
  - Technology

- Improved Quality
  - Materials
  - Response time
  - Personalization

Crowdsourcing Challenges

- How to crowdsource…..?
  - Image labeling, nutritional analysis of meal photo
  - Peer grading
  - Hints & explanations
  - Creation of new…
    - Problems
    - Examples
    - Animations
    - Etc…

Crowdsourcing Challenges



- Intelligent tutoring systems

  meet

- Big data analytics

  – Student modeling

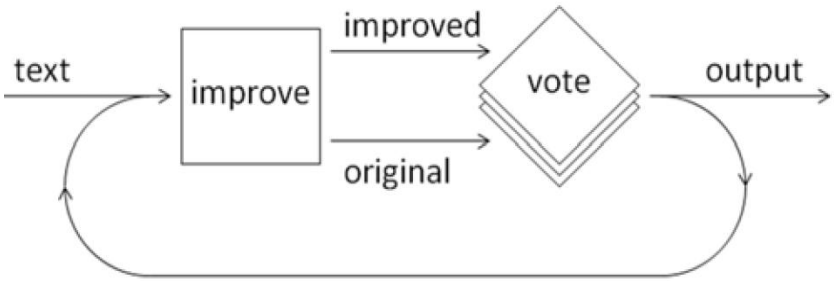
  – Real-time A/B testing for (micro) curricular changes

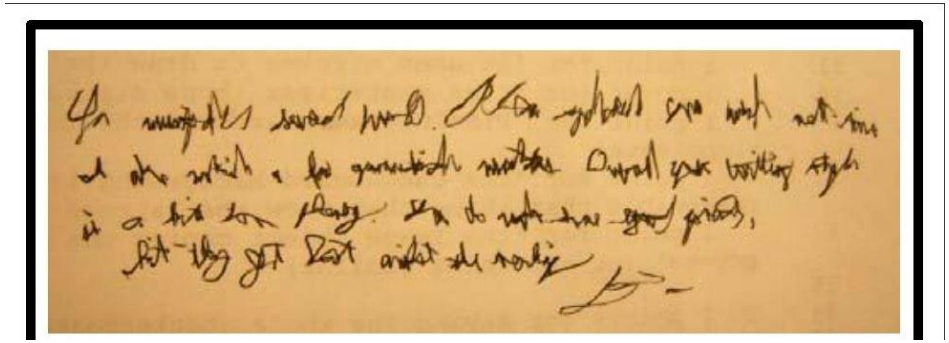
  – Personalization & engagement

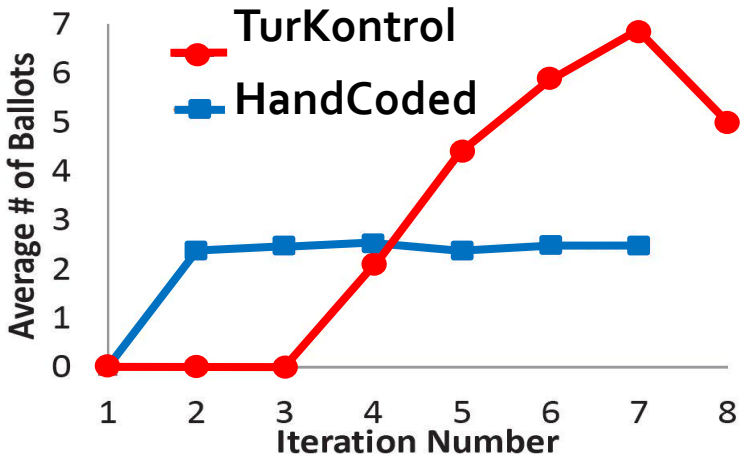
  15% Udacity; 8% MIT; 12-28% Coursera

Crowdsourcing Challenges

Workflow design & control







Course = very complex workflow

Crowdsourcing Challenges

- Novel form of crowd
- Motivation?

Motivation / Incentives

- Money

amazonmechanical turk
beta
Artificial Artificial Intelligence

Worker

mobileworks beta

oDesk®
Changing How the World Works.

Motivation / Incentives

- Mone

- Altruism, Esteem & Community

Motivation / Incentives

- Money

- Altruism, Esteem & Community

- Entertainment

Motivation / Incentives

- Money
- Altruism, Esteem & Cor
- Entertainment
- Self-Interest

Motivation / Incentives

- Money
- Altruism, Esteem & Community
- Entertainment
- Self-Interest
- Education

Crowds of …?

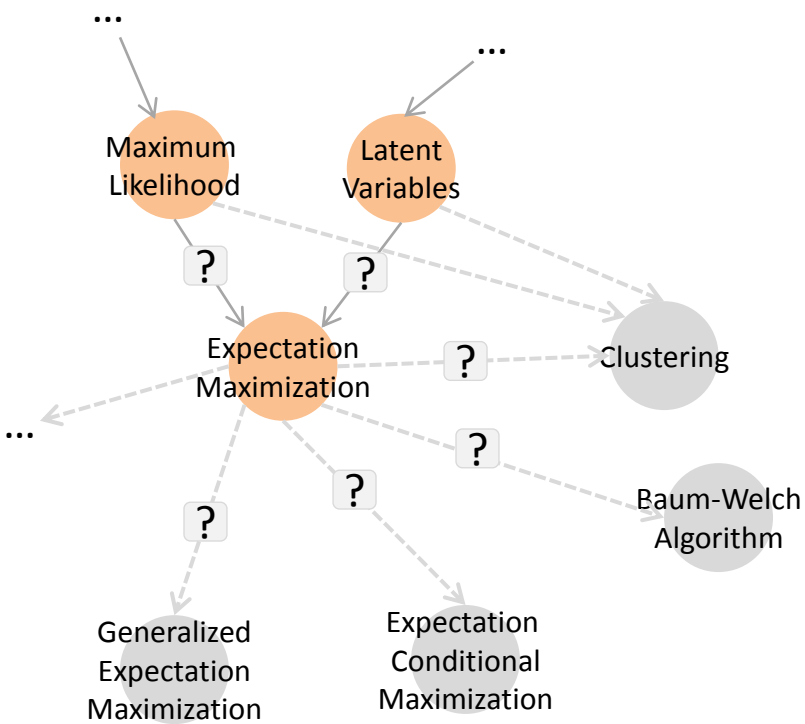
- Students
  - Peer grading
- Graduates
  - Forum monitoring
- Teachers

Machine Reading & Education

- Classify pages (syllabus, problem set…)
- Identify topics, prerequisites, hierarchy
- Extract content…

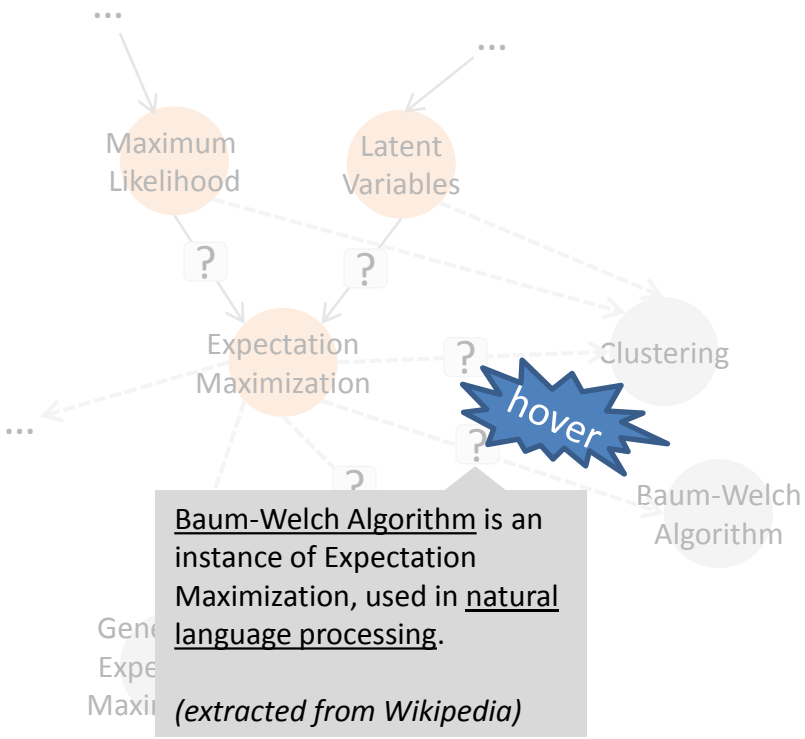
- Create a educational resource to drive traffic

Noticed an error?



Topic: Expectation Maximization

Noticed an error?



In statistics, an **expectation–maximization (EM) algorithm** is an iterative method for finding maximum likelihood or maximum a posteriori (MAP) estimates of parameters in statistical models, where the model depends on unobserved latent variables. The EM iteration alternates between performing an expectation (E) step, which computes the expectation …

 (source: Wikipedia)

EM is an iterative optimization method to estimate some unknown parameters T, given measurement data U. However, we are not given some "hidden" nuisance variables J, which need to be integrated out. In particular, we want to maximize the posterior probability of the parameters T given the data U, marginalizing over J …

 (source: F. Dellaert)

The **EM algorithm** is used to approximate a probability function (p.f. or p.d.f.). **EM** is typically used to compute maximum likelihood estimates given incomplete samples. Let X be the set of hidden variables, Y the set of observable variables, …

Topic: Expectation Maximization

Noticed an error?

P. Niyogi
0:59 h
Source: videolectures.net
1423 views ★★★★★

E. Candes
1:15 h
Source: videolectures.net
783 views ★★★★★

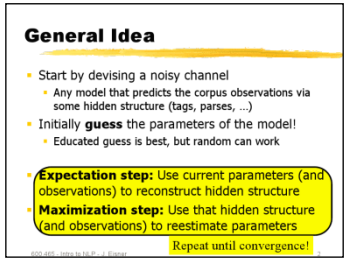
A. Anandkumar
1:02 h
Source: videolectures.net
345 views ★★★★★

B. Schölkopf
1:18 h
Source: videolectures.net
214 views ★★★★★

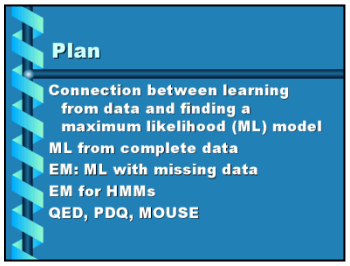
S. Kpotufe
0:44 h
Source: videolectures.net
14 views ★★★★★

Topic: Expectation Maximization

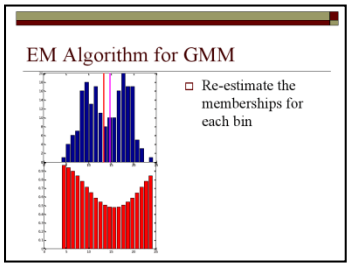
Noticed an error?



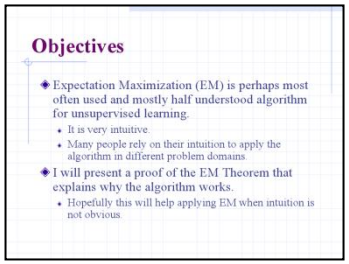
Jason Eisner
Source: ...
435,342 views ★★★★★



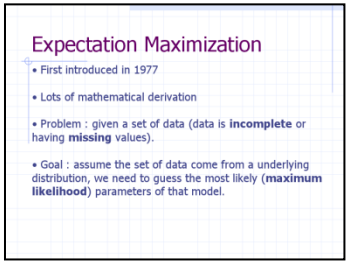
Michael Littman
Source: ...
43,456 views ★★★★★



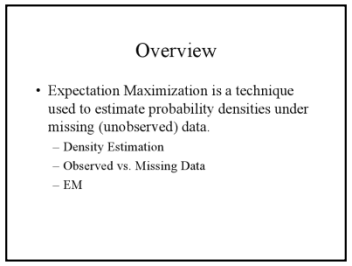
Rong Jin
Source: ...
12,003 views ★★★★★



Dekang Lin
Source: ...
235 views ★★★★★



Dekang Lin
Source: ...
123 views ★★★★★



Markus Weber
Source: ...
55 views ★★★★★

Line Fitting: Write a function that fits a line to data. Note that Weiss describes a method for doing this using weighted least squares, which essentially only looks at error in the y direction. This fits the examples below, in which noise is added to ...

Difficulty:
Usefulness:
Comments (3)

(Source: ...)

The two local newspapers, the C-Gazette and the U-Gazette, publish n articles everyday. The article length in the newspapers is distributed based on the Poisson distribution with parameter λ. To clarify, for a non-negative integer x ...

(Source: ...)

Difficulty:
Usefulness:
Comments (23)

Consider data, D =2,3,1, sampled from a two-dimensional (separable) distribution, p(x1 , x2) = px1 (x1)px2 (x2), with ...

(Source: ...)

Difficulty:
Usefulness:
Comments (1)

Wednesday, July 18

8:00-9:00 am	Breakfast (Owens A/B)
9:00-9:30 am	Opening Remarks (**Dan Weld**)
9:30-10:00 am	**Jennifer Widom**: *From 100 Students to 100,000*
10:00-10:30 am	Coffee break
10:30-11:30 am	Perspectives from Educational and Cognitive Psychology • **Michelene Chi**: *Two Approaches to Enhancing Online Learning: Dialogue Videos and Engagement Activities* • **Ken Koedinger**: *Crowdsourcing Cognitive Models for Assessment, Tutoring, and In-Game Support*
11:30 am-1:00 pm	Lunch (Owens B)
1:00-2:45 pm	Insights from Research on Intelligent Tutoring Systems • **Kurt Van Lehn**: *Toward socially intelligent tutoring systems: Of the crowd, for the crowd* • **Sumit Gulwani**: *Intelligent Tutoring Systems: A Logical Perspective* • **Emma Brunskill**: *Student Variability and Automated Instructional Policies*
2:45-6:00 pm	Outside activities: hiking, biking, canoeing
6:15-7:30 pm	Dinner and discussions (Owens B)
7:30-8:30 pm	**Luis von Ahn**: *DuoLingo: Learn a Language for Free while Helping to Translate the Web*

Jennifer Widom

- Fletcher Jones Professor and Chair of CS at Stanford.
- Bachelor's from the Indiana University School of Music in 1982
- Ph.D. in CS from Cornell University in 1987.
- Research Staff at IBM Almaden for 6 years
- She is an ACM Fellow,
  - member of the National Acad Engr,
  - American Academy of Arts & Sciences;
  - Received ACM SIGMOD Edgar F. Codd Innovations Award in 2007
  - Guggenheim Fellow in 2000;
- Her research interests span many aspects of nontraditional data management, including crowdsourcing
- Taught one of the first MOOCs
  - 90k accounts, 25k work, 6500 successes