The Online Revolution: Education at Scale

Daphne Koller & Andrew Ng
Stanford University & Coursera
Availability
Affordability

Price Changes Since 1985

- College Tuition & Fees: 559% Increase
- Medical Care
- Gas
- All consumer items

Source: Bureau of Labor Statistics
In jobs requiring a college degree: 55.6%
In jobs not requiring a college degree: 22%
Not working: 22.4%

Tom Friedman summarized better than anyone what I’m about to tell you: Big breakthroughs are what happens when what is suddenly possible meets what is desperately necessary. We’ve talked about the desperately necessary. What about the suddenly possible?
What is suddenly possible is to use technology to offer education at scale. This started with 3 courses that Stanford opened to the world in the fall. For example, the machine learning class was taught by my colleague Andrew Ng. Andrew’s ML class is one of the larger at Stanford, with 400 students. Andrew’s public machine learning class had an enrollment of over 100K students. So for Andrew to reach the same audience by teaching his Stanford class, he would have to teach it for ... 250 years.
Coming from a middle class family from a small town in India. Never had the luck and guidance to reach Stanford for education. Guess what? God has sent the opportunity right across my door step! Heartfelt thanks to the great team and teachers who made this happen!

(Aakash Goswami)
I'm a single mom 39 years old with two college boys... I've been trying to get back to college for my masters and want to learn more about computers. Looking forward to this class!

(Jenny Ramirez)
High academic standards
It turns out that people like to get great content from the best instructors ... for free. Since mid-February, when our website opened, we have accumulated more than 600K students from 190 countries. We have close to 1.5 million enrollments in 41 courses, across a range of disciplines. In the 15 courses that have already launched, we have 14 million video views and 6 million quiz submissions.
Real Course

Course Begins

# users on site

Timeline

Real Course
VIDEO-BASED INSTRUCTION
ADD GRAPH?
Writing on tablet appears on monitor

Webcam to record instructor

Lecture Recording
Let’s talk about some of the key elements that went into these courses. First, when we design content for the online environment, we can remove the constraints of classroom scheduling, and abandon the monolithic 1hr lecture. The content can be divided into short, coherent modules of 8-12 minutes each, which can be traversed in different ways by different students. Some students might benefit from some additional preparation. Others might want to go into depth in certain topics. So we can finally move away from the one-size-fits-all model of education into a much more personalized experience.
ASSESSMENTS
Testing Improves Learning

“Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping.”
• Multiple-choice (radio button, checkbox)
• Numerical answers
• Short answer
• Math
• Structured outputs: Programming and modeling assignments
• Peer assessment

Multiple forms of assessment
Expression equivalence (e.g., $a^2 - b^2 = (a+b)(a-b)$)

Theorem proving
• Programming assignments
  – Support for any language: R, Python, Matlab, Java, ...
  – Advanced assessments: timing, memory use, test set performance, ...
  – Spell-check, Kinect™ action recognition
• Modeling assignments (e.g., electrical circuits)
• Excel spreadsheets (data analysis, financial models, ...)

Grading structured outputs
Question 8

I-maps. I-maps can also be defined directly on graphs as follows. Let $I(G)$ be the set of independencies encoded by a graph $G$. Then $G_1$ is an I-map for $G_2$ if $I(G_1) \subseteq I(G_2)$.

Which of the following statements about I-maps are true? You may select 1 or more options (or none of them, if you think none apply).

- A graph $K$ is an I-map for a graph $G$ if and only if $K$ encodes exactly the same independencies as $G$.
- I-maps are Apple's answer to Google Maps.
- An I-map is a function $f$ that maps a graph $G$ to itself, i.e., $f(G) = G$.
- A graph $K$ is an I-map for a graph $G$ if and only if $K$ and $G$ are identical, i.e., they have exactly the same nodes and edges.
- A graph $K$ is an I-map for a graph $G$ if and only if all of the independencies encoded by $K$ are also encoded by $G$. 

Personalization and Mastery
Open-Ended Work
Peer Grading Workflow

Cf. Calibrated Peer Review™
(Chapman, 2001)
Peer Grading

**Evaluation criteria & Grading rubric**

Grade within 40 points

- **Scoring questions**
  - 0-20 points
  - 20-30 points

  - Can you make alternative prototypes of two ideas? Include both prototypes?
  - If both prototypes are too similar, make a rough cut of both ideas. Then create computer models for the final prototype?

- **Evaluation**

  - Photos of your prototypes

  - Did the student make alternative prototypes of two ideas? Include both prototypes?
  - If both prototypes are too similar, make a rough cut of both ideas. Then create computer models for the final prototype?

  - Photos of your prototypes

  - Did the student make alternative prototypes of two ideas? Include both prototypes?
  - If both prototypes are too similar, make a rough cut of both ideas. Then create computer models for the final prototype?

**Evaluation**

- Did the student make alternative prototypes of two ideas? Include both prototypes?
- If both prototypes are too similar, make a rough cut of both ideas. Then create computer models for the final prototype?

**Aggregate score:** 72%

- **Comments:**
  - Your prototypes were at the right level of familiarity.
  - They clearly got a lot of effort, but the assignment asked for more emphasis on the final prototype.
  - I notice a bit of confusion about what parts of your first prototype is yours, as the photo shows a good initial prototype overall with details for views that don't impact the face of the A4.
In one of the questions, it is suggested that we "Use the unix command line utilities". How do I solve this on a non-unix OS? I have never worked with this, so I am at a complete loss what to do. Thanks in advance.

If you want the true Unix experience, running a virtual machine is better than using cygwin/gnuwin, imo.

Oracle VirtualBox is great, and you can find many preconfigured virtual machine image on Virtualboxes.org. Choose Ubuntu, if you are new to the Linux world. That one also has Python 2.7 included, while many other distributions like Debian still have Python 2.8.

Open a terminal window by entering "terminal" in the dashboard. You can also install "Gnome Terminal", then you can always open and close a terminal by pressing F12.

Assuming you are using windows, you can

- install cygwin, or
- install gnuwin32, or
- run live Linux in virtualbox (e.g. ubuntu live cd)
Ordering for assigning factors to cliques in `ComputeInitialPotentials`. The order of assignment of factors to cliques should happen in the order cliques are given to you at the end of the `CreateCliqueTree` function. Each factor should be assigned to the first clique that contains the variables in the factor, where ordering of the cliques is given by `C.nodes` (C is the argument for `ComputeInitialPotential` function).

For example: in function `ComputeInitialPotentials`, the argument C has a field nodes. Now let's say the contents of C.nodes are:

```
C.nodes[1] = [1 2]
C.nodes[2] = [2 3]
```

And your factors are [1, 2, 3]. So [1] and [2] should be assigned to the 1st clique. Even though [2] can be assigned to the second clique, for the purpose of this assignment we are going to assign [2] to the first clique that contains it.

Order of Variables in Cliques. You should use `CliqueTree.nodes` to get the ordering for your variables and those nodes are in numerical order.

Empty cliques. It is possible that you may end up with cliques with no factors assigned to them. If that is the case, set the initial potential to 1 for all variable assignments for that clique.

CliqueTree.Calibrate (for max-sum). If you are having problems with this part, but your code is otherwise correct for sum-product message passing, make sure that your FactorMaxMarginalization works properly with logspace-potentials.

Clique Potential. If you have a clique over variables {1 2 3} with only one factor assigned to it, say [1], then you should assume that there's an initial potential over {1 2 3} with all 1s and multiply it.

This is the implementation we have and it doesn't really affect the answer because you will end up multiplying stuff about all your variables in the clique. If a variable does not appear in any factor, then it shouldn't be in any clique. However, if a variable is in a clique and just the factors assigned to the clique don't contain the variable then there must be some other clique that contains that variable, and to which the factor is assigned. By running intersection property that variable has to be in the subset so you will get messages for it.

Comments

- Thanks a lot! It clarified several things for me. However, when you mention empty cliques, apart from the initial potential set to 1: what variable \( j \) means, field \( \cdot \cdot \cdot \) should we put 0? An empty vector \( j \)? And the cardinality? Because this has an effect on the amount of values

- (all non-zero, as you stated.)
Fellow students on these forums really gave the sense that I wasn’t just sitting in my office working on it by myself. The spirit of cooperation and information sharing has been far more than any “non-virtual” course I ever took.  

(Sanjaya Kumar)
Student Study Groups

San Francisco HCI-Class Study Group

Vietnam, Czechoslovakia, Nigeria, Miami, Austin, Texas, Minnesota, Greece, Nepal, Kenya

Guatemala, Arabic speaking students, Athens, Nepal, Kenya

A multilingual universal study group
First lectures posted

STATISTICS & ANALYTICS

Coursera
Wrong answers submitted for machine learning class programming assignment

ON HOLD
For each forum question $q$, estimate:

*Increase in probability of transitioning to correct response given that a user viewed question $q$*

(focusing only one “cluster” of incorrect responses at a time)
Question 830: 98 views and 1949 not viewed.
- Fraction transitioned to correct if viewed: 0.643
- Fraction transitioned to correct if not viewed: 0.34
- Fraction improvement by viewing: 1.893

What helped the most?

computeCost
12
I have obtained ans = 32.079 for the computeCost function in ex1 but when I submit 2)
Computing Cost (for one variable) [ computeCost.m ] I am told === Sorry, your answer was incorrect.
How can this be?
Tags: exercise 1

14
Something to consider is the matrix X. Remember that a column of '1's are added so this will
change the way you are indexing values in X. I had the same answer and was also failing until I
considered this.

[Jon Huang]

The 2 Sigma Problem
College is a place where a professor’s lecture notes go straight to the students’ lecture notes, without passing through the brains of either.

—Mark Twain

The mind is not a vessel that needs filling, but wood that needs igniting.

—Plutarch

from Ian Kidd's translation of Essays
"Improved Learning in a Large-Enrollment Physics Class."
• Just-in-time teaching
• Small group problem solving
• Higher-level discussion of material
• Presentation of real-world applications
• Optional interactive sessions better attended than lectures in standard televised courses
• Enhanced interaction, with immediate feedback
• Individual tailoring of flow and pace through content
• Less threatening environment for students
• Detailed analytics to improve courses
• Time for meaningful engagement between students and faculty, students and peers
• Interaction and creative problem solving are the real value of top universities

Benefits to on-campus teaching
What would we get if we could offer a free high-quality education to everyone? First, it establishes education as a basic human right, so that anyone with the motivation and the ability would have the opportunity to get the skills that they need to make a better life for themselves, their families and their communities. Second, it enables lifelong learning. It’s a shame that for most of us, learning stops when we finish our formal education. With these amazing courses, we would always have the opportunity to explore new directions, whether to expand our minds or to make a change in our lives. Finally, it opens the door to a wave of innovation. Because talent can be found everywhere. Maybe the next Albert Einstein or the next Steve Jobs is living in some remote village in Africa. With access to education, he or she can come up with the next big idea, and help make the world a better place for all of us.