NYU Data Science Community Newsletter features journalism, research papers, events, tools/software, and jobs for September 4, 2015.

Please Let us (Laura Noren, Brad Stenger) know if you have something to add to the newsletter. We are grateful for the generous financial support from the Moore-Sloan data Science Environment and NYU’s Center for Data Science.

NYU Data Science Community Newsletter Issue 017.

Data Science News

Economics Has a Math Problem
Bloomberg View, Noah Smith from September 01, 2015

A lot of people complain about the math in economics. Economists tend to quietly dismiss such complaints as the sour-grapes protests of literary types who lack the talent or training to hack their way through systems of equations. But it isn’t just the mathematically illiterate who grouse. New York University economist Paul Romer -- hardly a lightweight when it comes to equations -- recently complained about how economists use math as a tool of rhetoric instead of a tool to understand the world.

Personally, I think that what’s odd about econ isn’t that it uses lots of math -- it’s the way it uses math. In most applied math disciplines -- computational biology, fluid dynamics, quantitative finance -- mathematical theories are always tied to the evidence. If a theory hasn’t been tested, it’s treated as pure conjecture.

UC Berkeley pilots data science class
The Daily Californian from September 03, 2015

... UC Berkeley is piloting a class this fall that faculty say will teach students how to engage with this information in a digitized world, where data are increasingly ubiquitous.

The new four-unit course, “Foundations of Data Science” — cross-listed as Statistics 94 and Computer Science 94 — combines introductory statistics and computational concepts with hands-on work involving hard data that brings “real-world relevance,” according to the program’s website. The course is a part of the new Data Science Education Program, a project that was initiated last year in response to strong student interest in learning programming and statistics.

Why Science Needs to Publish Negative Results
Elsevier SciTech Connect from April 03, 2015

Many experimental results never see the light of publication day. For a large number of these, it comes down to the data being “negative”, i.e. the expected and/or wanted effect was not observed. A straightforward example might be the testing of a soil additive that is believed to help a plant grow. If the experiment outcome shows no difference between the standard soil and the soil with the additive, then the result will end up buried in the laboratory’s archive.

But is this really the best approach to scientific results?
How do you know if your model is going to work? Part 1: The problem
Win-Vector Blog, John Mount from September 02, 2015
Here’s a caricature of a data science project: your company or client needs information (usually to make a decision). Your job is to build a model to predict that information. You fit a model, perhaps several, to available data and evaluate them to find the best. Then you cross your fingers that your chosen model doesn’t crash and burn in the real world.

We’ve discussed detecting if your data has a signal. Now: how do you know that your model is good? And how sure are you that it’s better than the models that you rejected?

Is a Cambrian Explosion Coming for Robotics?
IEEE Spectrum from August 31, 2015
... Many of the base hardware technologies on which robots depend—particularly computing, data storage, and communications—have been improving at exponential growth rates. Two newly blossoming technologies—Cloud Robotics and Deep Learning—could leverage these base technologies in a virtuous cycle of explosive growth.

SNLI corpus
The Stanford NLP (Natural Language Processing) Group from September 01, 2015
The SNLI corpus (version 1.0) is a collection of 570k human-written English sentence pairs manually labeled for balanced classification with the labels entailment, contradiction, and neutral, supporting the task of natural language inference (NLI), also known as recognizing textual entailment (RTE). We aim for it to serve both as a benchmark for evaluating representational systems for text, especially including those induced by representation learning methods, as well as a resource for developing NLP models of any kind.

Better Beer Through GPUs: How GPUs and Deep Learning Help Brewers Improve Their Suds
The Official NVIDIA Blog from September 02, 2015

Jason Cohen isn't the first man to look for the solution to his problems at the bottom of a beer glass. But the 24-year-old entrepreneur might be the first to have found it.

Cohen's tale would make a great episode of HBO's “Silicon Valley” if only his epiphany had taken place in sun-dappled Palo Alto, Calif., rather than blustery State College, Pa. That Cohen has involved GPUs in this sudsy story should surprise no one.

This is the tale of a man who didn't master marketing to sell his product — quality control software for beer makers. He had to master it to make his product. The answer, of course, turned out to be free beer. And that's put Cohen right in the middle of the fizzy business of craft brewing, a business that moves so fast he's enlisted GPUs to help his software keep up.

Project Oxford: Microsoft serves up APIs for intelligent apps

Microsoft this past spring announced Project Oxford, a set of SDKs and APIs that allow developers to build “intelligent” applications without having to learn machine learning. Using Oxford’s face, speech, and vision APIs, developers can create applications that recognize facial features, analyze images, or perform speech-to-text or text-to-speech translations.

In an interview with InfoWorld Editor at Large Paul Krill, Microsoft’s Ryan Galgon, senior program manager responsible for the Project Oxford platform and technologies, talked about the goals behind Oxford, emphasizing its potential in the Internet of things.

Events

NYU Game Center Incubator Showcase | NYU I Game Center

The 2015 Incubator is reaching its exciting conclusion! Since June 1st, thirteen developers have worked full time on eight games, under the advisement of more than 50 industry partners, all with the goal of preparing for a commercial release. This year you'll play games about glitch witches, angry apes, awkward dates, and community theater, just to name a few. The 2015 games have already been featured in the the PAX10, invited to XOXO and Fantastic Arcade, greenlit on Steam, and they’re not even out of the Incubator yet.

Friday, September 4, at 6 p.m., 2 Metrotech Center, 8th Floor

Cesar Hidalgo on Urban Perception | Marron Institute

Cesar Hidalgo will be leading an interactive NYU Urban Seminar discussion on urban perception and neighborhood amenities.

The NYU Urban Seminar is co-hosted by the Marron Institute and the Furman Center.

Tuesday, September 8, at 12:15pm, Vanderbilt Hall, NYU

NYC Media Lab 2015 Annual Summit

NYC Media Lab's Annual Summit is a snapshot of the best thinking, projects, and talent in digital media from universities in NYC and beyond.

This is an opportunity for media executives, technologists, and decision makers to see more than 100 university prototypes and demonstrations that explore interesting technologies and applications related to the future of media. Attendees will be also invited to roll up their
sleeves during a series of interactive workshops led by NYC faculty.

Friday, September 25, at 9 a.m., NYU Skirball Center for the Performing Arts

**Deadlines**

**DataPoint2015 conference**
New York City's own data science community conference. (Hosted by Google. Sponsored by Google and Chartbeat.) David Hogg from the Center for Data Science has been invited to speak.

DataPoint2015 will feature presentations by NYC data science community on projects they are excited about. The program will feature the newest applications, most clever solutions and tools you'll want to share with your colleagues when you get back to the office. We know you've been working on something that fits that description, so let us put you on stage to share it!

Deadline for Submissions: Monday, September 7

**Reasoning, Attention, Memory (RAM) NIPS Workshop 2015**
The research into developing learning algorithms combining these components and the analysis of those algorithms is still in its infancy. The purpose of this workshop is to bring together researchers from diverse backgrounds to exchange ideas which could lead to addressing the various drawbacks associated with such models leading to more interesting models in the quest for moving towards true AI.

Deadline for Submissions: Friday, October 9

**Workshop on Networks in the Social and Information Sciences, NIPS 2015**
This workshop aims to bring together a diverse and cross-disciplinary set of researchers to discuss recent advances and future directions for developing new network methods in statistics and machine learning. In particular, we are interested in network methods that learn the patterns of interaction, flow of information, or propagation of effects in social and information systems empirical studies, particularly attempts to bridge observational methods and causal inference, and studies that combine learning, networks, and computational social science research that unifies the study of both structure and content in rich network datasets

Deadline for Submissions: Monday, November 2

**CDS News**

**Moore-Sloan Data Science Lunch Seminar Series begins Monday, September 14**
NYU Center for Data Science from September 04, 2015
The Data Science Lunch Seminar Series is meant to be an informal weekly gathering of NYU Data Science affiliated persons to discuss a broad range of data science related topics. Each week there is a 20-30 minute presentation, over lunch (provided), with additional time for conversation and questions.

Taking place each Monday starting September 14 from 12-1 p.m. at the Center for Data Science, 726 Broadway, 7th Floor
6th NYU Data Science Showcase
NYU Center for Data Science from September 16, 2015
Featured Faculty talk by Vasant Dhar, Professor and Head, Information Systems Group, Co-Director, Center for Business Analytics, Talk Title: Should You Trust Your Money to a Robot?

Abstract: Financial markets emanate massive amounts of data from which machines can, in principle, learn to invest with minimal initial guidance from humans. I contrast human and machine strengths and weaknesses in making investment decisions. The analysis reveals areas in the investment landscape where machines are already very active and those where machines are likely to make significant inroads in the next few years.

Background reading:

Wednesday, September 16, at NYU Torch Club, Tap Room, starting at 4 p.m.

Click here to receive the NYU Data Science Community Newsletter OR to have us follow your twitter feed so that our data science twitter bot can easily grab links from your tweets.

To send us an announcement for the newsletter, please email laura.noren@nyu.edu and brad.stenger@nyu.edu by 9 pm Eastern time on Thursday evenings for inclusion in Friday’s newsletter. We retain curatorial discretion.

NYU Data Science Community Newsletter Issue 017.