Data Science News

LinkedIn: Machine learning and data science "actually have the least experienced engineering workforces, with the highest percentage of employees with 0-5 years of experience of any specialty at 17 percent, likely due to being emerging fields in the technology sector," LinkedIn said in a blog post this week. "Despite being a newer and more dynamic field, machine learning and data science engineers have the highest compensation level, with a median of $129,000 across the United States."

https://adtmag.com/articles/2017/07/21/linkedin-jobs-data.aspx

Apple just launched a blog focused on machine learning research papers and sharing the company's findings. The Apple Machine Learning Journal is a bit empty right now as the company only shared one post about turning synthetic images into realistic ones in order to train neural networks.

https://techcrunch.com/2017/07/19/apple-launches-machine-learning-research-site/

China is looking to invest heavily in AI in the coming years. By 2020, China's AI technologies and research facilities will match other leading countries, said Li Meng, the vice minister of science and technology. Five years later, he expects "a big breakthrough," and then China should finally become the global "innovation center for AI" by 2030.

http://money.cnn.com/2017/07/21/technology/china-artificial-intelligence-future/index.html

Today, Salesforce introduced Einstein Vision for Social Studio, which provides a way for marketers to search for pictures related to their brands on social media in the same way they search for other mentions. The product comes trained to recognize two million logos, 60 scenes (such as an airport), 200 foods and 1000 objects. That should be enough to get many companies started.

https://techcrunch.com/2017/08/08/salesforce-using-ai-to-find-brand-images-on-social-media/

A research team at Salesforce has come up with a clever way to improve the performance of many modern language programs—teaching an algorithm to speak another language before training it to do other tasks.

https://www.technologyreview.com/s/608382/to-build-a-smarter-chatbot-first-teach-it-a-second-language/?imm_mid=0f55ac&cmp=em-data-na-na-newsltr_20170809

Data Science News

TechCrunch Disrupt SF 2017 is all in on artificial intelligence and machine learning. "Some of the top minds in both categories to discuss the latest advances and future of AI and ML on stage and Disrupt San Francisco in just over a month."

https://techcrunch.com/2017/08/08/techcrunch-disrupt-sf-2017-is-all-in-on-artificial-intelligence-and-machine-learning/

Back in March at Google's Cloud Next conference, the company announced plans to run its own machine learning startup competition side-by-side with Data Collective and Emergence Capital. Four months later, 10 startups, pulled from a pool of 350+ applicants, presented onstage at Google's Launchpad Space in San Francisco.

https://techcrunch.com/2017/07/13/here-are-the-winners-of-the-google-cloud-machine-learning-pitch-off/

In his time since departing as Baidu's chief scientist, Ng has been developing a sequel of sorts to his popular introductory machine learning Coursera class. This go around, he is focusing specifically on deep learning. Coursera's new Deep Learning Specialization Certificate will feature five courses.

https://techcrunch.com/2017/08/08/deeplearning-ai-is-andrew-ngs-new-series-of-deep-learning-classes-on-coursera/

Google has just moved to a production release of TensorFlow Serving, its open source library for serving machine-learned models in production environments. A beta version of the technology was released in February. TensorFlow Serving lets you push out multiple versions of models over time, as well as roll them back.

http://www.infoworld.com/article/3214764/machine-learning/google-releases-tensorflow-serving-library.html

Data Science News

Reporters from ProPublica obtained more than 7,000 test results from Florida in 2013 and 2014 and analysed the reoffending rate among the individuals. The system incorrectly predicted that members of one demographic group would reoffend, while also missing offenders from other demographic groups who actually did reoffend. Joanna Bryson, a researcher at the University of Bath, told the Guardian: "People expected AI to be unbiased; that's just wrong. If the underlying data reflects stereotypes, or if you train AI from human culture, you will find these things."

Microsoft's TayTweets (AI) chatbot was unleashed on Twitter last year which was designed to learn from users. However, it almost instantly turned to racism, and even anti-semitism, after interaction with users.

http://www.express.co.uk/news/science/838557/RACISM-artificial-intelligence-AI-

A pair of 'chatbots' in China have been taken offline after appearing to stray off-script. In response to users' questions, one said its dream was to travel to the United States, while the other said it wasn't a huge fan of the Chinese Communist Party.

http://www.reuters.com/article/us-china-robots-idUSKBN1AK0G1?imm_mid=0f55ac&cmp=em-data-na-na-newsltr_20170809

The UK Information Commissioner's Office declared that the hospital operator had broken civil law when it gave health data to Google's London-based subsidiary DeepMind. The arrangement failed to consider how patients expect their data to be used, and by whom.

https://www.nature.com/news/the-deepmind-debacle-demands-dialogue-on-data-1.22330

Audit and compliance professionals should be aware of what is being "delegated" to AI within specific business processes, along with the potential risks. Auditors and compliance professionals may find themselves with the opportunity to help oversee, or "audit" the complex algorithms that guide AI systems.

https://www.accountingweb.com/technology/trends/how-machine-learning-is-disrupting-accounting

Many disciplines are currently in the midst of a "replication crisis" where even textbook studies aren't holding up against rigorous retesting. The list includes: ego depletion, the idea that willpower is a finite resource; the facial feedback hypothesis, which suggested if we activate muscles used in smiling, we become happier; and many more. A forthcoming paper in the journal Nature Human Behavior argues that results should only be deemed "statistically significant" if they pass a higher threshold. Others have argued for the rejection of null hypothesis testing, and pushed for better scientific reasoning and reproducible research.

https://www.vox.com/science-and-health/2017/7/31/16021654/p-values-statistical-significance-redefine-0005?imm_mid=0f55ac&cmp=em-data-na-na-newsltr_20170809

D J Patil, a former chief data scientist at the White House, said "One thing that's very important to talk about is that a key aspect that's intrinsic to data science is the need for ethics. One of the things, when I was in office that we called for is that every data training program — data science, economics, statistics, computer science, machine learning, whatever you wanna call these things — they all have to have ethics and security as a core component of the curriculum. Just because you can with data doesn't mean you should."

https://factordaily.com/d-j-patil-data-sciences-ethics-aadhaar/