

Data Science News

February 2019

Data Science News

Despite excellent performance on stationary test sets, deep neural networks (DNNs) can fail to generalize to out-of-distribution (OoD) inputs, including natural, nonadversarial ones, which are common in real-world settings.

<https://arxiv.org/pdf/1811.11553.pdf>

A team of scientists led by researchers at the University of Georgia Center for Food Safety in Griffin has developed a machine-learning approach that could lead to quicker identification of the animal source of certain Salmonella outbreaks.

<https://www.sciencedaily.com/releases/2019/02/190211163952.htm>

Cambridge and Pfizer collaborated on an algorithm which looks at both molecules known to be active and molecules known to be inactive, and learns to recognise which parts of the molecules are important for drug action and which parts are not for drug discovery.

<https://phys.org/news/2019-02-machine-algorithm-drugs.html>

Purdue and USC developed an ML algorithm to quickly correct computer-aided design models and produce parts with improved geometric accuracy for assembly and manufacturing

<https://www.industryweek.com/technology-and-iiot/universities-new-machine-learning-tool-improves-additive-manufacturing>

Data Science News

We must recognize that those patterns are merely correlations amongst vast realms of data, rather than causative truths or natural laws governing our world.

<https://www.forbes.com/sites/kalevleetaru/2019/01/15/a-reminder-that-machine-learning-is-about-correlations-not-causation/#587ccaf46161>

Uber recently presented Manifold, a model-agnostic visual debugging tool for machine learning similar to LIME

<https://eng.uber.com/manifold/?amp>

Airbnb recently presented Bighead, their approach to an end-to-end machine learning platform (built on Python, Spark, and Kubernetes)

<https://conferences.oreilly.com/strata/strata-ny-2018/public/schedule/detail/69383>

Databricks recently open-sourced Flint, a time-series library for Apache Spark

<https://databricks.com/blog/2018/09/11/introducing-flint-a-time-series-library-for-apache-spark.html>

<https://github.com/twosigma/flint>

Data Science News

Uber also recently open-sourced Ludwig, a deep learning toolbox built on top of TensorFlow that allows users to train and test deep learning models without writing code.

<https://eng.uber.com/introducing-ludwig/>

<https://uber.github.io/ludwig/>

White House recently signed an executive order which calls for a variety of actions in the federal government to make AI research and education a priority.

<https://www.purdue.edu/newsroom/releases/2019/Q1/artificial-intelligence-sparks-new-arms-race.html>

The Vatican says it is teaming up with Microsoft on an academic prize to promote ethics in artificial intelligence.

<https://www.seattletimes.com/business/vatican-microsoft-team-up-on-artificial-intelligence-ethics/>

Data Science News

DeepMind is now beating competitive StarCraft II players

<https://9to5google.com/2019/01/24/deepmind-starcraft-ii-livestream/>

Alphabet starting to acknowledge the risks of AI:

“New products and services, including those that incorporate or utilize artificial intelligence and machine learning, can raise new or exacerbate existing ethical, technological, legal, and other challenges, which may negatively affect our brands and demand for our products and services and adversely affect our revenues and operating results.”

<https://www.wired.com/story/google-microsoft-warn-ai-may-do-dumb-things/>