

Joseph M. Hahn, Ph.D.

Data Scientist

jhahn@spacescience.org

<http://gemelli.spacescience.org/~hahnjm>

Qualifications

Skilled data scientist, author and educator, with expertise in the analysis and visualization of big data in linux-based cloud-compute environments. An expert with many years experience in analytics and machine learning using python's data-science stack (numpy pandas scikit-learn tensorflow matplotlib jupyter), big-data tools like Spark and Hadoop, executing in the Oracle and AWS clouds. With some experience in GPU-accelerated deep learning and neural networks. Other skills include mathematics, numerical modeling, public speaking, and technical writing. By day I'm a Data Scientist that uses Spark + machine learning to predict the future, but I also moonlight as a Research Scientist in Planetary Dynamics.

Goals

To work on challenging, high-impact projects using machine learning or neural network models trained on large datasets to predict future events. Especially interested in the use of algorithms to reduce manufacturing mishaps, to optimize business decisions, investment portfolios, and logistics

Education

PhD in physics, 1997, University of Notre Dame

BS in physics and BA in astronomy, 1990, University of Texas at Austin

Professional Experience

Data Scientist

March 2018—present

Oracle:

- I deliver proof-of-concepts (POCs) for Oracle customers in the cloud, using python, machine learning (ML), spark, tensorflow, and Oracle analytics tools. Use cases include:
 - using ML to estimate current and past market value of \$20B of residential properties
 - ML to anticipate and avert production mishaps, for a major steel producer
 - training ML models to automate insurance claim processing
 - simulating a predictive maintenance strategy for upstream oil & gas
 - ML to forecast city-wide utility expenses
 - ML to forecast crimes rates across a major metro area
 - ML models to detect high-value users of a loyalty card program

Consulting Data Scientist

August 2017—Jan 2018

The Judge Group:

- Onsite data scientist for client Charles Schwab, where I developed Spark + ML code on website traffic, to identify customers having greatest lead potential

Associate Partner Technical Architect

February 2014—August 2017

DXC (formerly CSC formerly Computer Sciences Corporation formerly Infochimps):

- Delivered solutions on CSC's Analytics Platform as-a-Service (CAPS) for customers in the Amazon AWS cloud, for clients in aviation, nuclear power, information management, and fleet management. Projects delivered include:
 - using machine learning models to predict interpurchase intervals for customers at a global car-rental firm, as well as an ancillary-purchase recommendation engine
 - built machine learning models to anticipate project costs and completion times, for the many thousands of concurrent engineering projects in progress at an aviation firm.
 - demo'ed NLP on Hadoop to automatically parse a large volume of medicinal-product-specification texts, to extract key facts (product name, dosage, biochemical interactions, warnings, etc) from millions of human-generated documents, to help drug manufactures in EU meet new product documentation rules
 - demo'ed how to use the Data Alignment/Map (ADAM) suite of Spark applications to pre-process a large volume of genetic data, and the trained a classifier to predict a person's population membership using that person's genetic variations
 - demo'ed predictive maintenance on Hadoop; this demo simulated the time-history of a large fleet of vehicles that experience maintenance issues with usage. The demo used ML to predict those vehicles in greatest need of repair, and illustrated how machine learning can optimize maintenance schedules and reduce vehicle downtime and expenses,
 - ML to classify telemetry emitted by a large fleet of military drones.

Research Fellow

February 2012—July 2016

University of Texas Center for Space Research (CSR)

- taught a graduate-level course on planetary dynamics, and wrote textbook on planetary dynamics that is now 80% complete and to be published by Wiley

Research Scientist

July 2006—present

Space Science Institute (headquartered in Boulder CO)

- prior to February 2014 I was a full-time research scientist with SSI, and since then I have continued to moonlight as a scientist at ~5 hours/week, with that work supported by my National Science Foundation (NSF) research grant in astrophysics
- research includes modeling of data acquired by the Cassini orbiter at Saturn, the Messenger spacecraft at Mercury, and Hubble observations of circumstellar planet-forming disks
- co-advisor to PhD student in astrophysics, graduated 2020 at the University of Maryland
- author or co-author of 31 peer-reviewed scientific publications, available at <http://gemelli.spacescience.org/~hahnjm/pubs.html>

Associate Professor of Astronomy

July 2003—July 2006

Saint Mary's University in Halifax, NS, Canada

- Tier II Canada Research Chair (CRC) in Astrophysics, member of the Institute for Computational Astrophysics (ICA), principle investigator (PI) on grant-supported astrophysics research, taught physics and astrophysics, mentored student research, organized an international astrophysics conference, advised thesis research by Masters student that graduated 2006.

Research Postdoctoral Fellow

1996—2003

Lunar and Planetary Institute in Houston, TX

- pursued astronomy research, PI on NASA research grant, organized astrophysics conference, and mentored student research

Links

LinkedIn: <https://www.linkedin.com/in/hahnjoe>

Github: <https://github.com/joehahn>

JMH Datasciences LLC: <http://jmh-datasciences.com>

Academic homepage: <http://gemelli.space-science.org/~hahnjm>

Blog post:

- [Predictive maintenance for upstream oil & gas](#)

Recent Publications:

- [Unusual one-armed density waves in the Cassini Division of Saturn's rings](#), 2020, French, French, Nicholson, Hedman, Rappaport, Marouf, Longaretti, & Hahn, *Icarus*, **339**, 11360.
- [The Shape of Saturn's Huygens Ringlet Viewed by Cassini ISS](#), 2016, J. Spitale & J. Hahn, *Icarus*, **279**, 141.
- [Deciphering the Embedded Wave in Saturn's Maxwell Ringlet](#), 2016, by R. French, P. Nicholson, M. Hedman, J. Hahn, C. McGhee-French, J. Colwell, E. Marouf, & N. Rappaport, *Icarus*, **279**, 62.
- [Impact vaporization as a possible source of Mercury's calcium exosphere](#), 2015, by R. Killen & J. Hahn, *Icarus*, **250**, 230, see also this [NASA press release on this work](#).

Links to selected chapters from the [textbook I am writing](#), on planetary dynamics.

Source code for my numerical models: <http://gemelli.space-science.org/~hahnjm/software.html>, see <http://gemelli.space-science.org/~hahnjm/animations.html> for animations of those models.