Patrick McCarthy

Data Scientist and Technologist

resume@patrickmccarthy.cc
http://patrickmccarthy.cc

Skills

- Python; scikit-learn, pandas, jupyter
- R; dplyr, data.table, sparklyr, Rstudio
- Cloud; Spark, Hadoop, Hive, AWS

- Visualization: shiny, ggplot2, d3.js
- UNIX CLI; bash, grep, ssh, vim
- Misc; SQL, git, SVN, JIRA

Experience

Dstillery Inc., New York, NY   Sr. Data Scientist, Machine Learning Eng. …..2018-Present
(Data Scientist, Machine Learning Eng. …..... 2016-2018)

- Designed and built Python API for existing C++-based modeling tools, providing the flexibility and interactivity of modern data science environments.
- Implemented and pioneered the use of Spark for production use among data scientists and engineers, developed workflows and best practices.
- Researched and produced models for real-time bidding infrastructure, predicting ad viewability and likelihood of customer interaction.
- Partner closely with engineers, building and maintaining machine learning systems.

eXelate Inc., New York, NY   Data Scientist…….2013-2016 (*acquired by Nielsen, 2015)*

- Designed and built end-to-end pipelines against 1bn+ records for use on personally identifiable information in secure remote environments.
- Implemented high-performance concurrent ETL system with Python, EC2 and S3.
- Researched novel applications of statistical practice to business problems including latent class analysis, bayesian modeling.

Witness LA, Los Angeles, CA   Consultant, Investigative Print Journalism………..….2012

- Discovered a high likelihood of "pay for play" practices in Sheriff's Department records, in which political donations correlated strongly with promotions.
Best Database-Driven Journalism award, LA Press Club. http://goo.gl/GYvI7R.

Acumen LLC, Burlingame, CA   Quantitative Analyst Intern………………….…..…..2012

- Enhanced client relationships by developing a general use JavaScript visualization tool to intuitively display hierarchical datasets. http://goo.gl/iO6BcY.
- Developed new healthcare policy insights through exploration and analysis of a multi-gigabyte dataset, including extensive use of SAS macros and ODS.

Boston University Medical Campus, Boston MA Technical Specialist…………..2008-2011

- Managed teams through multiple phases of enterprise infrastructure deployment including research, initial implementation and ongoing maintenance
- Took initiative in streamlining department operations, both in procedure and technology

Education

M.S., Statistics, 2013
University of California Los Angeles, Los Angeles, CA
B.A. Economics and Mathematics, 2008
Boston University, Boston, MA