Christopher Fleisch

New York, NY 914-886-2300
chris@chrisfleisch.com .:. www.linkedin.com/in/chris-fleisch .:. www.chrisfleisch.com

Profile

Technology professional transitioning into data science. Experience with machine learning, statistics, and data infrastructure. Interested in helping organizations learn from their data and implementing data science projects.

Education

UC Berkeley School of Information May 2018
Master of Information and Data Science

Stanford Online April 2015
Statistical Learning Course from Stanford.edu, Statement of Accomplishment

The Flatiron School February 2015
Introduction to Data Science Course: Analyzing a dataset using the Pandas library and Python

Rensselaer Polytechnic Institute May 2003
Bachelor of Science in Information Technology

Data Science Projects

Causality and linear regression: Built models in R looking for causality using infant data.
Hypothesis testing: Tested hypotheses in R using political survey data.
Puffin Really Matters: Designed a case study that attempts to use images of Puffins to identify individuals in populations.
Exploratory data analysis: EDA of CEO data using R, associations between performance and salary.
Best Hospitals: Created a data lake in HDFS for government hospital data. Using Hive and Spark transformed the data for analysis. Discovered the best performing hospitals across a wide variety of measures.
Tweet Storm: Analysis of live tweets using Apache Storm and the Twitter API. Tweets were parsed using python and stored in Postgresql.
Whiskey Business: A comparison of whiskey prices and reviews from a liquor controlled state and free market states to find the best deals outside the state controlled prices. Whiskey prices and reviews were taken from multiple web sources, data was cleaned with python, stored in HDFS, combined into one table, exported to Google sheets, and connected to Data Studio for interactive visualizations.
Predicting Tree Coverage: Based on a Kaggle competition, Sklearn was used to predict the type of tree coverage from forest service data and Random Forest models.
Machine Learning at Scale: Implemented embarrassingly parallel algorithms such as Naive Bayes with MapReduce, KMeans with MapReduce and Spark, Cosine similarity with MRJob, Linear Regression using gradient descent with MRJob, Single Source Shortest Path with MRJob, PageRank with MRJob and Spark, Logistic Regression using gradient descent with Spark, pairwise data mining in MapReduce.
Disaster Perplexity: Detected perplexity and topic changes on social media when natural and man-made disasters happened in local communities.
Email Response Experiment: Ran an experiment to see if women receive slower responses to emails than men by sending out emails to businesses with randomly assigned male/female names and measuring response times.
Taughtfully.com: A data collection and insight engine designed to turn a teacher's observations into credible data that can inform classroom and school level decision making.
Rodents in NYC: A visualization created in Tableau using data from the NYC Open Data site. This project visually and interactively explores signs of rodents in NYC buildings and restaurants.

Skills and Tools

Machine Learning: classification, regression, clustering, natural language processing (Kneser-Ney, LSTM, part-of-speech, phrase structure), feature engineering
Statistical Methods: regression, hypothesis testing, survival analysis, confidence intervals, principal component analysis
Software and Programming Languages: Python (scikit-learn, TensorFlow, numpy, scipy, pandas, jupyter), R, Hadoop (HDFS, MapReduce), Apache Spark, SQL, D3, Tableau, Linux, LaTeX

Work Experience

Simons Foundation, Senior Systems Architect November 2016 - Present

- Manage a small team of Linux administrators.
- Plan and implement cloud infrastructure for Informatics group.
- Collect, archive, and distribute genomic data to hundreds of researchers.
- Build AWS Lambda services for backup, scheduled tasks, and a small website.
- S3 log aggregation and analysis with Athena and EMR used to troubleshoot access issues.
- Launch foundation websites and load test for scalability.
- Oversee cloud infrastructure costs and invoicing.

Simons Foundation, Technology Director January 2016 - November 2016
Simons Foundation, Technology Manager July 2008 - January 2016
Simons Foundation, Programmer/Analyst July 2007 - July 2008

- Managed internal IT team and external IT consultants that supported the foundation as it expanded from 12 to over 150 employees.
- Created, analyzed, and implemented an annual technology plan for the growing foundation.
- Oversaw desktop support for an OS X based organization with 200+ clients and 150+ users on and off site.
- Managed IT infrastructure (network, servers, backup) locally and for externally hosted servers.
- Purchased hardware, software, and IT services based on foundation requirements.
- Implemented client management software (JSS) for managing OS X clients and pushing custom software.
- Administered Linux servers hosted at Rackspace and AWS using Ansible.
- Setup large science data sharing service with Amazon S3, Aspera, and Globus.
- Initiated annual security audit on the foundation's network, servers, web applications, and performed a social engineering test with foundation staff.
- Implemented a computer security training program for all staff with monthly social engineering tests.
- Created internal Django applications for IT inventory, science journal requests and Python scripts for administrative tasks like email reminders, AWS functions, and a web crawler for software updates.
- Planned and implemented a disaster recovery plan for the foundation, set up offsite alternate office location, conduct bi-annual testing with the finance team.
- Migrated email hosting from single Linux IMAP host to Exchange 2007, Exchange 2010, Exchange 2013, and Google Apps. Implemented DMARC and DKIM to secure email.
- Planned, executed, and seamlessly moved all IT services from one building to another over a weekend during foundation expansion.
- Created foundation websites using the Symfony framework, MySQL, HTML, and CSS for simonsfoundation.org and sfari.org (2007-2009). Implemented single sign on using OpenID.

Christopher Fleisch

New York, NY
chris@chrisfleisch.com .:. www.linkedin.com/in/chris-fleisch/ .:. www.chrisfleisch.com