

Trent Hauck

trent@trenthauck.com

Work Experience

Zymergen zymergen.com
Staff Data Scientist

Oct. 2016 - Present

- Technical lead managing a set of decision support tools used by biologists and process engineers, including:
 - Novel unsupervised deep learning technique for targeting gene edits.
 - Large scale Bayesian model to normalize batch effects and identify outperforming test subjects.
 - CNN-based tools to identify structural variants in sequencing data.
 - Building large scale information retrieval systems to search metagenomics data for relevant genes and proteins.
- Trained Data Science team members on advanced statistical and machine learning techniques like deep learning and probabilistic programming.

Metis thisismetis.com
Instructor

Dec. 2016 - Present

- Courses Taught: Introduction to Data Science, Live Online Data Science, Live Online SQL Fundamentals (2x).

Consultant

May 2015 - Sept. 2016

- Embedded with a team at an e-commerce company to develop their first recommendation system – the model based on LDA, produced item-to-item and item-to-user recommendations which was deployed as a microservice called by the main site.
- Built LTV scoring pipeline used for marketing strategy. Pipeline determined current value of incoming lead based on probabilities of purchasing certain policies given lead features, and likely duration of the policy being in-force.
- Consulted with a web analytics consulting firm to improve infrastructure for processing 3rd party APIs for their clients.

Zulily zulily.com
Data Scientist

Feb. 2014 - May 2015

- Built and deployed recommendation services to improve customer experience. Focusing on improving user/item recommendations and feature extraction/learning.
- Performed statistical analysis on business operations such as impact of customer repurchases as a function of product delivery duration.
- Provided internal consulting to marketing and accounting department to improve decision making through experimentation. For example, helping accounting to estimate total warehouse inventory given a sample count.

Alight Analytics alightanalytics.com
Data Scientist

Aug. 2012 - Feb. 2014

- Develop scalable architecture and ETL system to support adding additional data sources with little effort.
- Led reporting and analysis including content strategy and report implementation for several clients including: Monsanto, AMC, and Cargill Beef.

- Performed data analysis to support marketing and operational goals.
- Built the accounting database and surrounding processes for auditing and reconciling of policy income to sales.
- Managed online marketing and site development - including analytics agency and development in support of A/B testing.

Authored Books

- SciKit-Learn Cookbook, 2nd Edition through Packt Publishing, Published 2017
- SciKit-Learn Cookbook through Packt Publishing, Published 2014
- Instant Data Intensive Apps with pandas How-to through Packt Publishing, Published 2013

Talks

- Representation Learning for Graphs and Networks, KC Developer Conference, 2018
- Using an RNN to Make Song Recommendations, Seattle Data Science, 2016
- Low Friction NLP with gensim at PyData Seattle, 2015
- Analytics for Developers at KC Developer Conference, 2013
- Introduction to Pandas at Python KC User Group, 2013
- Python for Data Science at Data Science KC, 2013
- Maslow's Hierarchy of Analytics at Big Data KC, 2013

Education

University of Kansas

Winter 2011

Master of Science, Finance

- Performed research on Support Vector Machines for stock market prediction.
- Graduate Courses in Machine Learning, Data Mining, Econometrics, Probability, and Statistics.
- Teaching Assistant Math Department.

University of Kansas

Spring 2010

Bachelor of Science, Accountancy

- Elective courses in Linear Algebra, Vector Calculus, Statistics, Programming, and Digital Logic.
- Teaching Assistant Math Department.
- Research Assistant at E&Y CARAT, worked on the FRAANK project.

General Notes

- Broad exposure to ML and statistical based modeling approaches, including:
 - Frequentist and Bayesian Statistics
 - Machine Learning such as Clustering, Tree-based models, and Dimensionality Reduction
 - Deep Learning using RNNs, CNNs, and GCNs, especially for the generation of discrete structures
 - Sequential decision making via Bayesian Optimization and Evolutionary Algorithms
- Writing and maintaining applications:
 - Using various languages, most recently Python, Go, JavaScript, and Java
 - Productionizing systems via k8s-based deployments, or an approach native to the cloud provider
 - Documentation, communication with users, task prioritization
- Occasional contributor to OSS, including: Pandas, Scipy, docker-py, planout, and deepchem
- Certified ScrumMaster through Scrum Alliance