DAN VATTEROTT

St. Louis, MO
dvatterott@gmail.com
www.danvatterott.com | github.com/dvatterott
linkedin.com/in/dan-vatterott

PROFESSIONAL EXPERIENCE

Showtime Inc | New York, NY   **May 2017 – Present**
Data Scientist
- Architected and built data-pipeline (ETL) using Python and Apache Spark for transforming raw data from millions of customers into Hive tables describing business relevant insights
- Designed and manage Hive database housing the output of data-pipelines
- Built production Machine Learning models (Apache Spark) that predict future customer behaviors and relay these predictions throughout the company
- Built and manage Apache Airflow server (AWS) for coordinating scheduled, distributed-computing, production jobs (e.g., data-pipeline and machine learning models)
- Built testing platforms for verifying quality of data-pipeline algorithms and assumptions of raw data

Insight Data Science | New York, NY   **January 2017 – May 2017**
Data Science Fellow
- Built *Sifting the Overflow*, a chrome extension that highlights the most helpful parts of Stack Overflow answers
- Ingested 500,000 Stack Overflow answers into Postgres database
- Trained a recurrent neural network to identify helpful portions of Stack Overflow answers
- Built a web-app front-end (Flask, AWS) describing my chrome extension

Columbia University | New York, NY   **June 2015 – January 2017**
Postdoctoral Research Fellow | Department of Neuroscience
- Built data-pipeline which performs ETL on data from 100 electrodes sampling at 10,000 Hz
- Used Machine Learning algorithms to predict behavior using neural activity
- Used k-means clustering and PCA to extract eye-movement and electrophysiological signals from raw data
- Developed interactive, real-time tasks that react according to participants' eye-movements

University of Iowa | Iowa City, IA   **August 2010 – June 2015**
Graduate Researcher | Department of Psychology
- Built regression and ANOVA models that describe where people will direct their attention
- Taught 5 courses on topics such as Research Methods and Statistics (rated 5.8/6 by students)
- Published 8 research articles and presented at 10 international conferences

EDUCATION

University of Iowa | Iowa City, Iowa   **August 2010 – June 2015**
Doctor of Philosophy, Psychology, Concentration in Cognition and Perception

University of Wisconsin | Madison, Wisconsin   **August 2006 – May 2009**
Bachelor of Science, English and Psychology

SKILLS

Programming: Python (NumPy, Pandas, Sci-Kit Learn, Matplotlib), Spark, Bash, JavaScript (d3)
Data Analysis: Regression (GLM), Classification (Xgboost, neural networks), Hypothesis Testing (ANOVA, t-tests), Unsupervised Learning (k-means clustering, GMM)
Databases: SQL (PostgreSQL, HiveQL, MySQL)
Operating Systems: Linux (Ubuntu, Debian), OSX, Windows