JONATHAN A. CHEUNG

Los Angeles, CA 90025 • jacheung415@gmail.com • github.com/jacheung • (415) 519-2029

Summary

- **6 years of experience using machine learning tools for data-driven scientific discovery**
- **4 publications demonstrating high productivity, clear communication, and solution-focused nature**
- **2 contracts leveraging data to deliver business solutions**

Experience

Doctoral Researcher – Hires Laboratory, University of Southern California Sept. 2014 – Apr. 2020

- Produced key findings for 2 major grants, increasing funding from $1M to $4.5M over 4 years, by collaborating with colleagues to author 4 manuscripts in high-impact peer reviewed journals
- Established protocols to answer "How does the brain represent touch?" by crafting data pipelines to collect and synchronize 30 million time points from automated trial testing, high-speed video, and neural recordings
- Generated more accurate predictive behavioral and neural models through creative feature engineering via applying domain expertise, time-series filtering and physics models
- Resolved a decade-long debate regarding touch search strategies using predictive behavioral modeling on 16 uniquely extracted touch features
- Discovered a neural representation of touch location and hypothesized a circuit model for this sensorimotor transformation using interpretable machine learning models such as generalized linear models (GLM)
- Promoted a legacy of collaboration by recruiting, mentoring, and training new hires in the scientific method

Data Science Contractor Aug. 2019 – Present

- Sene Studio
  - Ensured data quality and improved accessibility to data by replacing disorganized spreadsheets with an automated extract, transform, load (ETL) workflow with data warehousing in AWS S3 and postgreSQL
  - Reduced order returns and costs by 38% ($58k), leading to a 30% increase in positive customer review, using a boosted regression model to optimize fit for made-to-measure suits and jeans
- Structure Research
  - Standardized metrics, improved data quality, and saved days of manual calculations for a premier data center research group by building an automated analyses pipeline in Python
  - Improved customer experience and web traffic by 20% through application of an interactive data visualization monthly newsletter

Skills

- Tools: Python | Matlab | SQL | Spark | AWS S3
- Python libraries: pandas | NumPy | scikit-learn | SciPy | TensorFlow | PyTorch | PySpark | matplotlib | plotly
- Supervised learning: linear and logistic regression | generalized linear models (GLM) | gradient boosted machines (GBM) | XGboost | time-series forecasting | convolutional neural networks (CNNs) | SVM |
- Unsupervised learning: clustering (k-means, DBSCAN, OPTICS), dimensionality reduction (PCA, t-SNE)

Notable Publications, Awards, and Achievements

Cheung JA, Maire P, Kim J, Lee K, Flynn G, Hires SA (2020) – Active touch remaps barrel cortex output from a representation of self-motion to object location. *(in review)*

Top Research Proposal in Systems Neuroscience 2018 (*USC Annual Research Symposium*)

National Institute of Health Intramural Research Training Award 2013-2014 (*National Institute on Aging)*

Education

Ph.D. Neuroscience
University of Southern California
Los Angeles, CA / 2020

B.S. Human Biology
University of California, San Diego
San Diego, CA / 2013