Karl Lundquist, PhD

klundquist@gmail.com
+1 (765) 237-9267
Data Science Blog
Google Scholar (400+ citations)
LinkedIn
Github

PROFESSIONAL EXPERIENCE

Calyxt

Data Scientist

May 2022 - Dec 2022

- Developed ML models to predict rearrangement of genetic elements that maximizes production of rare compounds in plants
- Utilized Jira API along with Azure Functions and Tables to extract employee worklog and project data to create Power BI dashboards for management and finance teams
- Created a Dash web app to help researchers discover pathways to compounds of interest. Extracted and parsed data from plantcyc.org and deployed app on heroku and Azure App Service

NYC Data Science Academy

Data Science Fellow

Sept 2021 - Dec 2021

- Completed over 400 hours of intensive training in Python, R and SQL with focus on data analysis, statistics, and machine learning. [curriculum]
- Project: Worked with a leading real-estate data analytics company to make Zillow Observed Rent Index (ZORI) predictions with data obtained from Google Cloud and IRS. Used VIF, PCA, and k-means and reduced 252 variable dataset to 20. Created models with linear regression, random forest, and gradient boosting to forecast rent prices in California to RMSE < $100. [blog] [github]
- Project: Predicted housing prices in Ames, Iowa. Starting with a dataset of 81 variables, reduced features to 19. Tuned and trained linear regression, random forest, SVM, and gradient boosting models to produce a final R^2 of 0.89. Built app to showcase analytics and provide homeowners with an interface for estimating added value of renovations based on ML model. [blog] [github] [app]
- Project: Fatal Encounters with Police. Performed EDA in R addressing high rate of deaths at the hands of law enforcement in the United States. [blog] [github]
- Project: Heart Disease Prediction. Carried out python EDA and trained random forest classifier determining indicators of heart disease risk. [blog] [github]

Purdue University

Postdoctoral Research Associate

Sept 2019 - Sept 2021

- Designed experiments, performed biochemical assays and gathered cryo-EM imaging data to determine outer-membrane protein assembly model.
- Caputured first structure of an outer-membrane protein as it is being assembled with Wu et al. Nat Comm (2021)

Georgia Institute of Technology

Graduate Research Associate

Aug 2012 - Sept 2019

- Carried out statistical modeling and data analysis of time-series molecular dynamics simulations executed on high-performance computing systems to characterize outer-membrane protein assembly. Developed regression models to characterize molecular features.
- Witnessed a key assembly event and resolved associated free energy landscape with ß-barrel assembly machinery Lundquist et al. PNAS (2018)
- Observed delivery process for a crucial bacterial membrane component Lundquist et al. BBA (2020)
- Worked with a team that discovered a new immune disease and characterized its molecular mechanisms Fernandez et al. J Exp Med (2019) [media coverage]

EDUCATION

PhD, Physics

Georgia Institute of Technology
Chemistry minor
Aug 2012 - May 2019
Atlanta, Georgia

BS, Physics

University of Michigan
Mathematics minor
Aug 2007 - May 2012
Ann Arbor, Michigan

PUBLICATIONS

- Author or co-author of 16 peer-reviewed publications
- h-index: 10
- 400+ citations

SKILLS

- Python (NumPy, Pandas, Scikit-learn, Dash)
- Machine Learning (linear models, ridge, lasso, elastic net, random forest, gradient boosting)
- Azure
- R, RStudio, RShiny
- SQL
- Data Analysis and Visualization
- Git, GitHub
- Linux, Bash scripting
- Statistical modeling
- High-performance computing systems