Justin A. Le, Ph.D.

Predictive Analytics, Machine Learning & Data Science

Orange, CA 92869
☎ *please inquire*
✉ *justin@jle.im*
https://cv.jle.im

Objectives

I'm looking for a place I can apply my passion for finding the **story** that data has to tell us! My experience is in a wide collection of tools from **statistics**, **math modeling**, **machine learning**, and **data science** — tools that I have proven effective in taming data in climate, energy, finance, and condensed matter physics.

As a *scientist and mathematician*, I focus not only on prediction, but also in developing full mathematical and data-driven **frameworks** for fully **understanding** and *exploring* processes and systems. As a *programmer*, I focus on developing **performant** and **robust** systems that are statically **verifiable** and prioritized for long-term **maintainability** and extensibility.

I am excited to **build systems** for conquering data and expanding the horizons of what your data can do for you, equipped with both the **cutting edge** and the **tried and tested** tools of the trade.

Education

2016 - 2019 **Ph.D., Computational and Data Science**, *Chapman University*, Orange, CA.
Application of Differentiable Programming and Machine Learning to Predictive Weather and Climate Analysis and Forecasting

2014 - 2016 **M.S., Computational and Data Science**, *Chapman University*, Orange, CA.

2010 - 2014 **B.S. in Physics w/ Specialization in Computational Physics; Minor in Computer Science and Engineering**, *University of California, San Diego*, La Jolla, CA.

Skills

Computer Science: Machine Learning (clustering, classification, artificial neural networks), Large-scale data analytics, Numerical algorithms (FEM, stochastic methods), Digital signal processing, Functional programming, Static analysis, DSL design

Languages: C++, Haskell, Python, Matlab, R, Ruby, Fortran

Mathematics: Multivariate statistics, Numerical analysis, Real/Complex analysis, Stochastic processes, Dynamical systems, Abstract algebra, Differential equations, Wavelet analysis, Applied Category Theory

Selected Work and Research Experience

2016 - 2017 **Machine Learning and Data Science Specialist**, *Schmid College of Science and Technology*, Orange, CA.
Developed an ensemble-based Machine Learning system for forecasting and predicting frequency and intensity of power outages for major Energy and Utility company providing for over 3 million people. Developed mathematical models based on stochastic principles for analysis and pre-processing of data. Worked with Neural Network, Self-Organizing Map, Stochastic models, and ARIMA models to provide a ensemble forecast. Worked also on developing an on-line platform to manage updating models and generating predictions as weather data was submitted.

2015 - 2017 **Climate Analysis with Recurrent Neural Networks**, *El-Askari Lab, Schmid College of Science and Technology*, Orange, CA.
Developed statically verified neural network technology in Haskell for the purpose of performant modeling and analysis of climate trends in relation to the mid-2010's California drought and the 2015 — 2016 season El Niño. Implemented high-performance recurrent neural networks and training algorithms, and integrated modern highly parallelized cluster training techniques with an automated operational back-end to make confident and validated projections about future climate trends. Research paper published in *Atmospheric Research*.

2015 - 2018 **Machine Learning Specialist / Educational Supervisor**, *Intela Solutions*, Irvine, CA.
Involved in the development of the technology, underlying mathematics, and user interface for MathDB, an abastracted data store used for real-time streaming data analysis. Assisted in the promotion and integration of MathDB technology in different capacities. Directed the planning of educational programs in Machine Learning and Data Science aimed for university students and industry professionals in Ukraine.

2014 **Condensed Matter Modeling and Simulation**, *Dynes Lab, UCSD Physics Department*, La Jolla, CA.
Modeling complex topologies of superconducting quantum interference devices for magnetoscopic applications, and implementing efficient, parallel numerical simulations under those models for calibration and experimentation.

Selected Projects

Machine Learning **Differentiable Programming (Backpropagation) and Optimization Platform**, *Numerical Computation / Computational Science.*
Authored and maintained open-source *[backprop][]*, *backprop-learn* platform/library for the Haskell language, providing automatic differentiation in support of differentiable programming and machine learning based projects. Currently used by many in the Haskell open source community to build richer data science platforms. Additionally, authored the *opto* platform for efficient extensible numerical optimization.

Physics / Programming **Path Integral Monte Carlo Simulation**, *Numerical Computation / Parallel Programming.*
Applied principles of the Feynman Path Integral Formulation of Quantum Mechanics to create real-time high-performance, parallelizable numeric simulations in multiple languages, including C++ and Fortran, for live analysis and exploration of ground state quantum systems.

Education / Writing **Functional Programming and Haskell Blog**, *Machine Learning / Computer Science.*
Maintaining a top Functional Programming and Haskell blog with 50,000 pageviews per year, appearing multiple times on the front page of high-visibility platforms such as *[Hacker News][]*. Topics include mathematical models, functional programming, and dependently-typed and type-safe programming.

Selected Publications & Presentations

Geoscience & Machine Learning **J. A. Le**, *H. M. El-Askary, D. C. Struppa (President, Chapman University)*, "Long-term drought impact on the El Niño-driven precipitation over Southern California using recurrent neural networks".
Atmospheric Research `https://www.sciencedirect.com/science/article/pii/S0169809517300157` (January 2017)

Machine Learning **J. A. Le**, "A Purely Functional Approach to Trainable Models".
`https://blog.jle.im/entries/series/+functional-models.html` (May 2018)

Algebra & Comp. Sci. **J. A. Le**, "Applicative Regular Expressions using the Free Alternative".
Compose Conf 2019, New York, New York `http://talks.jle.im/composeconf-2019/` (May 2019)

Geoscience & Machine Learning **H. M. El-Askary**, *J. A. Le*, "Forecasting Interactions Between ENSO and Extreme Drought Conditions with Recurrent Neural Networks".
AOGS 13th Annual Meeting, Beijing, China `http://talks.jle.im/aogs-2016/` (August 2016)

Selected Coursework

CS 611 **Time Series Analysis**, *Chapman University.*
Study of statistical time series analysis and statitiscal models for studying and analyzing time series data with mathematical rigor. Applied to financial time series analysis data, comparing the efficacy of different statistical models.

CS 533 **Computational Methods in Financial Markets**, *Chapman University.*
The computational study of various mathematical models and simulation techniques in historical financial data, specializing in comparative market analysis and currency exchange.

Phys 520 **Principles of Remote Sensing**, *Chapman University.*
Survey of remote sensing techniques, including the acquisition, aggregation, processing, analysis, and physical considerations of geophysical satellite data. In-depth look at a wide range phenomenology including meteorological anomalies, dust, fire, and anthropological impacts.

CS 540 **High-Performance Computing**, *Chapman University.*
Study of the modern state of high performance computing and big data. In-depth look at parallel and concurrent computing through various approaches, architectures, and network topologies. Applying cluster and grid computing algorithms to compute-intensive tasks.