

Justin A. Lê, Ph.D.

Predictive Analytics, Machine Learning, Data Science & Infrastructure

Orange, CA 92869

📞 please inquire

✉ justin@jle.im

🌐 <https://cv.jle.im>

Education

- 2016 - 2021 **Ph.D., Computational and Data Science**, *Chapman University*, Orange, CA
Learning-Based Modeling of Weather and Climate Events Related to El Niño Phenomenon via Differentiable Programming and Empirical Decompositions
- 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

Selected Work and Research Experience

- 2023 - **Principal Software Engineer**, *Anduril Industries*, Costa Mesa, CA
Current EW.
- 2022 - 2023 **Google Ads & Analytics Engineer**, *Google*, Irvine, CA
Information-theoretical statically verifiable data privacy guarantees in machine learning algorithms and deployments. Robust multi-layered infrastructure for large-scale data analytics pipelines and tooling, involved in all aspects of engineering and product design, algorithm design, maintenance, instrumentation, monitoring, migration, incident mitigation, and SLO stress-test infrastructure. Planned and designed systems for a unified abstraction of data extraction and aggregation to and from vendors, sources, and sinks.
- 2020 - 2022 **Senior Software Engineer (Backend)**, *SimSpace Corporation*, Boston, MA
Backend engineer for a large-scale Haskell web-facing application simulating entire corporate internal and external networks for the purpose of realistic mock cyber attacks. Direct team responsibilities involved bridging user-facing interfaces with complex scheduling and execution of applications and processes (namely, attack agents) across multiple virtual machines, analyzing and assessing results of attacks in a data visualization pipeline, and coordinating with cloud services and APIs. In the capacity of a Haskell programmer, implemented large type-safety and type-directed code generation initiatives company wide for greater internal correctness guarantees and more robust channels of communication between front-end and back-end.
- 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 - 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 abstracted 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.

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, Differential Privacy
Languages	Java, Haskell, Python, Matlab, R, C++, Ruby, Fortran
Mathematics	Multivariate statistics, Numerical analysis, Real/Complex analysis, Stochastic processes, Dynamical systems, Abstract algebra, Differential equations, Wavelet analysis, Applied Category Theory

Selected Projects

Machine Learning	Differentiable Programming (Backpropagation) and Optimization Platform , <i>Numerical Computation / Computational Science</i> Authored and maintained open-source <i>backprop</i> , <i>backprop-learn</i> 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 <i>opto</i> platform for efficient extensible numerical optimization.
Physics / Programming	Path Integral Monte Carlo Simulation , <i>Numerical Computation / Parallel Programming</i> 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 , <i>Machine Learning / Computer Science</i> 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 <i>Hacker News</i> . Topics include mathematical models, functional programming, and dependently-typed and type-safe programming.

Selected Publications & Presentations

Geoscience & Machine Learning	J. A. Le , "Learning-Based Modeling of Weather and Climate Events Related to El Niño Phenomenon via Differentiable Programming and Empirical Decompositions" Ph.D. Dissertation, Chapman University https://digitalcommons.chapman.edu/cads_dissertations/22/ (2021)
Geoscience & Machine Learning	J. A. Le , <i>H. M. El-Askary</i> , <i>D. C. Struppa</i> (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)

Teaching and Leadership

- 2014 - 2019 **Mechanics & Electromagnetism Lab**, *Physics*, Chapman University
- 2016 - 2017 **Intro to Computer Science**, *Computer Science*, Chapman University
- 2017 **Principles of Machine Learning and Data Science**, *Machine Learning*, Intel Solutions
- 2015 - 2017 **Vietnamese Student Association**, *Founding President / Vice President*, Chapman University
- 2015 **Functional Programming and Haskell**, *Computer Science*, Chapman University