Eric Cotner, PhD

Principal data scientist

Email: 2.71828cotner@gmail.com
Website: ecotner.com
Phone: +1 (530) 748-9935
GitHub: github.com/ecotner
LinkedIn: eric-cotner
Location: San Diego, CA

Data scientist with experience in machine learning, statistics, optimization, automated decision-making, experiment design and analysis, causal inference, visualization, and end-to-end product development in partnership with engineering/product teams. Have worked on real-time recommendation/optimization systems within the grocery delivery business, and optimization of supply chain processes, vehicle routing, and time series forecasting within the automotive industry. Background in theoretical physics, lending to strong conceptual/logical/mathematical aptitude. Plenty of experience communicating analytical results to various non-technical audiences at all levels.

Technical skills/tools

- Languages: Expert: Python | intermediate: C/C++, JavaScript, bash/sh | beginner: Rust, Java, Julia
- Data analysis/manipulation/visualization: pandas, numpy, scipy, jupyter, matplotlib/seaborn, plotly, folium/leaflet
- Statistics and machine learning: statsmodels, scikit-learn, (num)pyro, pytorch, TensorFlow/Keras, mxnet/gluonTS, JAX, XGBoost, LightGBM
- Databases: Snowflake, BigQuery, Oracle, SQL Server, MySQL, PostgreSQL, SQLite, DuckDB
- Cloud computing: AWS EC2/EBS, S3, ECS, Google Cloud services like Compute Engine, Cloud SQL, BigQuery, GCS, Cloud Run, Container Registry, Data Studio
- Web dev: Flask, FastAPI/pydantic, bottle, WebAssembly/emscripten, WebGL, jQuery, HTML/CSS/JS/TypeScript
- HPC: GPGPU, CUDA, numba, GPU.js, Cython
- Operations research: OR-Tools, OSRM, Valhalla, MILP (SCIP, COIN-OR, MIPCL, cvxpy), MDP's, dynamic programming
- Other: Docker (compose), Kubernetes, Kafka, Redis, MLFlow/comet.ml, git/GitHub, airflow, Drone

Experience

Principal Data Scientist, Shipt (subsidiary of Target), 2022-Present

Senior Data Scientist, Shipt (subsidiary of Target), 2021-2022

- **Experimentation and causal inference**: Designed and executed multiple experiments on the marketplace (switchbacks, synthetic control, A/B tests). Introduced multi-armed treatments and researched appropriate design/analysis techniques like randomized block design, ANOVA, and clustered standard errors. Helped teammates set up automated experiment analysis/results dashboards. Created simulation framework for performing power analyses to explore alternative experiment design/analysis strategies, and developed new techniques to speed power analyses up. Led adoption of usage of causal DAGs and "double ML" regression techniques to infer causal information from observational studies while minimizing confounder/selection bias. Utilized hierarchical Bayesian models to extract information from longitudinal panel data and determine impact of customer loyalty programs on revenue.
- **Real-time order recommendation**: Created algorithms to match shoppers/drivers with orders (in a potentially supply/demand unbalanced marketplace) in ways that are equitable, enjoyable for drivers, and result in more successful deliveries. System pairs delivery drivers with orders by using ML models to predict probability of success and "score" potential matches, maximize match scores under various constraints using mixed integer programming (OR-Tools + SCIP), then finally send offers/recommendations to drivers. Offers were generated in real time via multiple k8s-managed web services, serving close to a million recommendations each day on a nationwide scale, which necessitated careful balancing of model complexity/accuracy with inference and solver time. Worked very closely with partner engineering team to set up API contracts, data storage and retrieval plans, design and execution of experiments, etc.
- **Optimal decision-making**: When orders go unclaimed, we forego individual offer recommendations and make them visible to all drivers by sending them to "open metro" in an effort to advertise to more people in exchange for potentially lower fulfillment quality. Built a stochastic simulation of our marketplace that was used to simulate outcomes of various policies that decided when to send an order to open metro, based on time left until scheduled delivery and supply of drivers that have/haven't been offered this order. State transitions for mechanics where time to event was uncertain were modeled using survival analysis techniques to produce probability distributions that could be sampled from to provide stochastic (yet realistic) evolution of the simulation state. Parametrized policy function was optimized utilizing evolution strategies to maximize simulation reward signal without access to gradient information.

Eric Cotner, PhD

Principal data scientist

Email: 2.71828cotner@gmail.com
Website: ecotner.com
Phone: +1 (530) 748-9935
GitHub: github.com/ecotner
LinkedIn: eric-cotner
Location: San Diego, CA

Experience (continued)

Shipt (continued)

- **Education/communication**: Spread awareness of machine learning and experiment design techniques through presentations, and regularly consulted with various other data scientists (from my team and others) to help them understand concepts like causal inference, Bayesian models, recommendation algorithms, optimization techniques, etc. Managed a team book club that met bi-weekly to discuss subjects like python performance tuning, infrastructure for handling big data, principles of microeconomics and markets, causal inference, and reinforcement learning.

Lead Data Scientist, Torqata Data and Analytics (subsidiary of ATD), 2020-2021
Data Scientist, American Tire Distributors, 2019-2020

- **Route optimization**: Created algorithm to optimize flexible order delivery by using a Markov decision process and Monte Carlo simulation of probabilistic route forecast to calculate the expected risk of delaying the decision to deliver, in the context of uncertain future orders. Constructed production system to pull live order data from Oracle database, generate delivery recommendations, and serve to route planners in the warehouse. Generates decisions for orders at 130 warehouses on an hourly basis and has been shown to save roughly $0.50/order in transportation/labor costs.
- **Dynamic routing**: Led an initiative to research the application of dynamic routing optimization to our daily delivery routes using constraint satisfaction programming (OR-Tools). Ran simulations on historical data to demonstrate proof-of-concept, considerable potential savings (roughly 15% of variable costs) and effects of establishing concrete delivery service levels, leading to adoption of the idea by leadership. Implementation is currently underway.
- **Forecasting**: constructed multiple time-series forecasting pipelines to predict product flow in/out of warehouses (to anticipate labor requirements). Reduced forecasting error (uMAPE) by over 50% in some cases relative to pre-existing models. Built an interactive web application for inventory planners to consume forecast information.
- **Mentorship**: Mentored several interns and employees as part of our "Data Science Accelerator" program.

Academic background

PhD, MSc, Theoretical Physics, UCLA, 2013-2018
Graduate student researcher, UCLA, 2014-2018

- Studied exotic dark matter candidates, black holes, cosmology, and particle physics
- Performed numerical simulations of boson star collisions
- Modeled statistical fluctuations of Q-ball density, leading to formation of primordial black holes
- Publications in multiple top physics journals, and invited speaker to multiple conferences

Teaching Fellow, UCLA 2013-2018

- Led undergraduate discussion/lecture/laboratory sections for a large variety of subjects (quantum, electromagnetism, mechanics, thermodynamics, nuclear/particle, cosmology, etc.)

BSc, Physics, UCSD, 2009-2013

- Majored in physics with a minor in mathematics

Hobby/volunteer projects

- **Route optimization**: built a web application to optimize delivery routes with an interactive map using OR-Tools, Valhalla, and GCP.
- **Open-source work**: occasional contributor to several open-source python packages with scientific and machine learning applications such as pyro, matplotlib, pytorch, and jax.
- **Meet-ups/hackathons**: gave presentation on Markov decision processes to Charlotte, NC PyData chapter; mentor/judge for HackNC annual hackathon.
- **CoderDojo**: Taught elementary/middle school children computer science concepts using Scratch. Recipient of *Volunteer of the Month award*.
- **Science communication**: Current volunteer at the San Diego Fleet Science Center. Graduate student coordinator for *Exploring Your Universe* expo at UCLA for 4 years.

Eric Cotner, PhD

Principal data scientist

Email: 2.71828cotner@gmail.com
Website: ecotner.com
Phone: +1 (530) 748-9935
GitHub: github.com/ecotner
LinkedIn: eric-cotner
Location: San Diego, CA

Hobby/volunteer projects (continued)

- **Eviction aid**: worked with lawyers at Legal Aid Chicago and urban development researchers at UC Berkeley to analyze evictions in Cook County using historical eviction data from the Sheriff's office combined with ACS data from the US Census Bureau to create maps highlighting racial segregation and unfairness in the eviction process.
- **Video games**: maintain a Valheim server for playing games with friends. Server is hosted on GCP and has a front end web app for managing server state, including integration with Steam, discord status bot, and auto-shutdown to save costs.