

Keir Havel

Machine Learning Engineer / Data Scientist

[linkedin.com/in/keir-havel-3ab67993](https://www.linkedin.com/in/keir-havel-3ab67993) github.com/LumenPallidium

Applied mathematician (M.S., SDSU) with 5+ years in production data science and research in deep learning for more efficient training and weather forecasting. Skilled in Python, PyTorch, LLM integration, and time-series forecasting; experience with LLM coding assistants like Claude Code.

Experience

San Diego State University Research Foundation	San Diego, CA
<i>Machine Learning Research Intern, Climate Informatics Lab</i>	June 2025 -- Sep 2025

- Built diagnostic tools to measure conservation-law violations in neural weather models (GraphCast, FourCastNet)
- Prototyped physics-constrained training methods for atmospheric conservation; showed available data was insufficient to measure conservation reliably, a negative result that redirected the lab's approach

Center for Sustainable Energy	San Diego, CA
<i>Data Scientist</i>	June 2023 -- Aug 2024
<i>Junior Data Scientist</i>	Oct 2021 -- June 2023
<i>Research Analyst Assistant</i>	Feb 2019 -- Oct 2021

- Integrated large language models (LLMs; OpenAI, Anthropic APIs) into document review workflows, cutting processing time from 30 hours to about 1 hour while improving accuracy
- Automated survey report generation, including statistical hypothesis testing and figure creation, reducing manual effort from 80 to 10 hours per cycle
- Developed machine learning models for geographic vehicle traffic prediction using ensemble methods, with Shapley Additive Explanations (SHAP) for interpretability
- Built automated data ingestion (ETL) pipelines with data cleaning, enrichment, and anomaly detection
- Improved budget forecasting by evaluating time-series models (ARIMA, additive models)

Education

San Diego State University	August 2026
<i>M.S. Applied Mathematics</i>	Coursework: Optimization, Bayesian Statistics

- Thesis: *Compressibility and Multifractal Properties of Self-Compressing Neural Networks*. Developed a regularization technique that substantially accelerated training in small-scale settings, and characterized how the speedup diminishes at scale

University of California, Santa Barbara	June 2018
<i>B.S. Mathematical Sciences</i>	Coursework: Data Mining (graduate level)

Selected Projects

JEPA	45 stars, 12 forks
Unofficial PyTorch implementation of Meta's I-JEPA for self-supervised image representation learning; extended with saccade-based and predictive coding architectures.	

Backprop Alternatives	22 stars, 7 forks
Implementations of alternatives to backpropagation, showing significant sample-efficiency gains in small-scale experiments.	

Energy Transformer	14 stars, 2 forks
Energy-based transformer implementation incorporating modern Hopfield networks.	

Technical Skills

Languages: Python, R, SQL

Machine Learning / Deep Learning: PyTorch, JAX, scikit-learn, LLMs, self-supervised learning

Statistics & Modeling: Time-series forecasting, Bayesian statistics, hypothesis testing, optimization, SHAP

Data & Visualization: Spark, Pandas, NumPy, SciPy, Vaex, Matplotlib, Plotly, Dash

Tools: Git, Linux, AWS (boto3), Claude Code, OpenAI Codex