JONATHAN T. FARLAND

TimeGPT Launch Event ◇ Gen AI Training Course ◇ Silicon Angle ◇ Spark Summit
3491 Rogers Road◇ Concord, California, 94519 ◇ (508) · 237 · 8192
jonfarland@gmail.com ◇ LinkedIn ◇ github

EXPERIENCE

H2O.ai Dec 2021 - Present
Director of AI Solutions Engineering, Senior Data Scientist *San Francisco Bay Area, CA*

- **Leadership**
  - Executive staff member leading AI solutions engineering in North America and globally.
  - Lead Gen AI Masterclasses for C-suite, data science leaders and executives at the Fortune 100.
  - Hired roles ranging from Engineer, to Director, to Senior Principal Architect for an interdisciplinary team of data scientists, SMEs and software engineers.
- **Technical**
  - Deployed and fine-tuned a wide range of LLMs and Embedding models optimized for a range of tasks including Doc QA, Summarization, and Entity Extraction
  - Developed Gen AI Applications, including 'ProposalGPT' which can use advanced Retrieval-Augmented Generation to analyze RFPs for competiveness and automatically compose winning proposal.
  - Built a SOTA forecasting product using Foundational Models for Time Series and other best-of-breed approaches.
- **Business Development**
  - Proven track record of personally driving revenue growth, sometimes exceeding more than 2x the quarterly Net ARR of all of North America sales, combined.
  - Reduced sales cycle by 25 percent and increased product, industry visibility through the development of a system to match technologists to indidivual accounts and opportunities
  - Responsible for analyst relations (e.g. Gartner, Forrester) and historically achieved fist place placements portions of both CAIDS and DSML Magic Quadrants

TROVE Predictive Data Science (Acquired by E Source) Oct 2017 - Dec 2021
Senior Data Scientist, Manager *San Francisco Bay Area, CA*

- Developed operational AI Models still used in production across North American Utilities
- Lead a team of data scientists and software engineers to deploy large scale, real time predictive modeling systems
- Customer-facing data science lead for main product line, focusing on forecasting and customer segmentation.

KEMA Consulting (Acquired by DNV Energy) Aug 2012 - Oct 2017
Senior Technical Consultant, Data Scientist, Project Manager *Boston, MA*

- Proposed and managed technical studies related to the evaluation of energy programs both across the US and abroad. These include demand response, behavioral programs, distributed generation, renewables, and electric vehicle penetration.

- Responsible for stakeholder reporting and advisory services as lead technical consultant for the predictive analytics team.
- Developed hierarchical energy forecasting approaches to address growth in emerging technologies using machine learning and statistical techniques.

EDUCATION

University of Massachussetts, Amherst, USA *Aug 2012*
Masters of Science - Applied Econometrics: Deans List, Cum Laude
Bachelors of Business Admin - Operations Research & Finance
Minor in Resource Economics

University of Naples Federico, Portici, Italy *Aug 2011*
Certificate of Course Completion - Advanced Micro-Econometrics

TECHNICAL STRENGTHS

Generative AI	HuggingFace, NVIDIA Nemo Guardrails, Langchain, Enterprise h2oGPT, TimeGPT
Languages	Python, R, SQL, SAS, Matlab, Mathematica, VBA, AMPL
Computing	Unix-based Systems, Spark, Docker, K8S, AWS, GCP, Azure, Databricks, Digital Ocea
Database	Postgresql, MySQL, MongoDB, Cassandra, DBeaver, QuestDB
Top Python Packages	pytorch, pandas, numpy, sklearn, sktime, tensorflow
Top R Packages	xgboost, shiny, caret, neuralnet, hts, h2o, sparklyr
Applications	Tableau, Advanced Microsoft Office

SELECTED PROJECT WORK

AI for Genomics Study
Senior Data Scientst *Boston, Massachusetts*

Built a classifier to correctly predict cancer diagnosis using genomics data with 50k features

Improved True Positive Rate by about 30 percent by preprocessing genomics data with autoencoder and postprocessing with a metalearner

Predictive Analytics for Utility DR Operations
Project Manager *San Francisco, California*

Managed day to day client interactions and support while also tracking and developing the longer term product roadmap.

Designed and built a large-scale Demand Response forecasting system across all 7+ million meters of the Pacific Gas and Electric service territory and across the demand response portfolio, including residential, commercial and industrial sectors. These programs cover various designs ranging from behavioural to direct load control and dispatch.

Research and Development of novel hierarchal load impact forecasting algorithms using machine learning and statistical models.

Advanced Electric Distribution Grid Forecasting
Data Scientist *San Francisco, California*

Developed advanced forecasting system to provide short-term hourly forecasts of energy demand across two large utility operating areas in California. Forecasts accounted for emerging technologies as well distributed generation such as solar and storage.

Smart meter and other telemetry data (e.g., SCADA) were paired with weather data and ingested into our propietary data model on an ongoing basis across millions of sensors in near real-time.

Linking down-stream sensors across the complex electrical distribution network enabled the software to forecast by aggregation for grid nodes that did not have any historical telemetry. This provided significant cost savings to the customer from not having to install expensive sensors across all grid nodes.

Transformer Asset Risk Modeling
Data Scientist *Kansas City, Kansas*

Worked with in-house subject matter experts to codify institutional knowledge and include as a compliment to machine learning and AI-based predictive models

Used data from Dissolved Gas Analysis (DGA) testing to predict the risk of various failure modes as well as the anticipated time-to-failure across electrical transformers.

Time series of key gases such as Acetylene, Hydrogen, Ethane, Methane, Ethylene and Nitrogen were used to classify transformers based on their perceived risk level.

Real-time Customer Baseline and Financial Settlement Modeling
Data Scientist *Portland, Oregon*

Consulted for a medium sized utility company in the Pacific Northwest to launch their first Peak Time Rebates program across the service territory.

Counterfactual customer-level baselines were generated for each customer using an ensemble of approaches including similar-day, regression, control groups and gradient boosting machine approaches.

Method selection was the result of an optimization model based on performance data in the most recent cross-validation period.

Hierarchical Forecasting of Energy and Peak Demand in the Kingdom of Saudi Arabia
Senior Data Scientist *Riyadh, Saudia Arabia*

Data analytics and modeling for the largest end-use metering project in the world. Developed and delivered a three-day course on predictive analytics to subsequently train analytical staff at client site in Riyadh, Saudi Arabia. Seminar participants included staff from Saudi Aramco, Saudi Electricity Company, and the Electricity and Cogeneration Regulatory Authority.

Global Energy Transition Outlook
Technical Advisor *Oslo, Norway*

Technical advisor for DNV GL's Energy Transition Outlook (ETO). This annual report seeks to identify and measure the major industry implications of the ongoing global energy transition for each of the OECD's regions. Developed bottom-up and top-down predictions of energy demand for each region of the globe until 2040.

Caltrack Beta Test
Lead Data Scientist *San Francisco, California*

Primary code base developer for rapid measurement of site-level, weather normalized energy savings. Process and predictive results benchmarked across open-source implementations from Open EE Meter. Algorithms implemented using R, Python and the Spark distributed computing framework on compute-optimized instances in the Amazon cloud.

Behavioural Demand Response Evaluation
Project Manager *Ottawa, Canada*

Project Manager and Lead Data Scienstist for an impact evaluation pertaining to a hybrid energy program targeting both ongoing behavioural impacts as well as event-based hourly demand reductions. Analytics and reporting were generated using the Spark (1.5) distributed computing framework, Amazon Web Service S3 and EC2 instances, and the Databricks browser based platform.

Electric Vehicle Pilot Evaluation
Lead Data Scientist *State of Virginia*

Lead Data Scientist for a pilot involving time-of-use charging rates for electric vehicle owner's in Virginia. The project focused on whole house impacts for EV owners, as well as vehicle-charging only impacts. Statistical methods were used to enumerate differences between control and treatment average load shapes for the vehicle-charging only members. A synthetic control group was generated for the whole house treatment members.

Macroeconomic Modeling of State Commercial and Industrial Energy Sectors
Lead Data Scientist *State of Massachusetts*

Lead Data Scientist for a macroeconomic consumption modeling project of Massachusetts' Commercial and Industrial sectors. Our team developed a database of billing data pertaining to commercial and industrial premises for all energy efficiency program administrators in Massachusetts for three years. Along with NOAA weather data and economic data from the US Census Bureau and other sources, macroeconomic consumption models were used to estimate the impact of energy efficiency programs at the county and town level.

Mathematical Programming Approach Towards Risk Mitigation in Sports Betting
Graduate Researcher *University of Massachusetts*

Developed a modeling technique to maximize profit subject to a zero probability of loss from sports betting. Solved using the Simplex Algorithm, with the CPLEX solver in AMPL.

PUBLICATIONS AND RELEVANT PRESENTATIONS

Utility Load Research: The Future of Load Research is Now, C. Puckett, C. Williamson, C. Godin, W. Gifford, J. Farland, T. Laing, T. Hong, IEEE Power and Energy Magazine, May/June, 2020

Model Based Matching and Other Benefits of High Frequency Interval Data, L. Getachew, J. Farland, K. Agnew, P. Franzese, V. Richardson, G. Sadhasivan, International Energy Program Evaluation Conference, Baltimore, USA, 2017

Electricity End Use Forecasting Using Non-Intrusive Load Metering Technology, J. Farland, C. Puckett, F. Coito, International Symposium on Forecasting, Cairns, Australia, 2017

High Resolution Energy Modeling that Scales with Apache Spark 2.0, J. Farland, Spark Summit Boston, USA, 2017

Load Forecasing with Distributed Energy Resources, J.Farland, F. Farzan, R.J. Hyndman, International Symposium on Forecasting, Santander, Spain, 2016

Breaking Down Analytical and Computational Barriers in Energy Data Analytics, J. Farland, Spark Summit San Francisco, USA, 2016

Zonal and Regional Load Forecasting in The New England Wholesale Electricity Market: Semi-parametric Regression Approach, Masters Thesis, University of Massachusetts, 2012

PROFESSIONAL ORGANIZATIONS AND DISTINCTIONS

Board of Directors, Peak Load Management Alliance, 2020-2021

Member, International Institute of Forecasting

Invited Speaker, 41st International Symposium on Forecasing, Charlottesville, VA, 2023

Speaker, 39th International Symposium on Forecasting, Thessaloniki, Greece, 2019

Invited Speaker, 38th International Symposium on Forecasting, Boulder, Colorado, 2018

Speaker, 37th International Symposium on Forecasting, Cairns, Australia, 2017

Speaker, Spark Summit East, Boston, USA, 2017

Session Chair, 36th International Symposium on Forecasting, Santander, Spain, 2016

Speaker, AEIC Advanced Load Research Applications, Nashville, Tennessee, USA, 2016

Speaker, Spark Advisory Forum, San Francisco, USA, 2016

Speaker, Spark Summit West, San Francisco, USA, 2016

Speaker, 34th International Symposium on Forecasting, Rotterdam, Netherlands, 2014

Vijay Bhagavan Distinquished Teaching Award, University of Massachusetts Amherst, 2012

International Advanced Econometrics Scholarship from Italian Ministry of Agriculture, 2011