One-U Responsible AI Initiative Annual Symposium

A Data Stream-Centric Auto-ML Framework for Time Series Forecasting

September 12, 2025

Jerome Soller, PhD
President, CogniTech Corporation, Salt Lake City, Utah
E-mail: soller@cognitech-ut.com
Phone: (801) 949-2903, URL: www.cognitech-ut.com

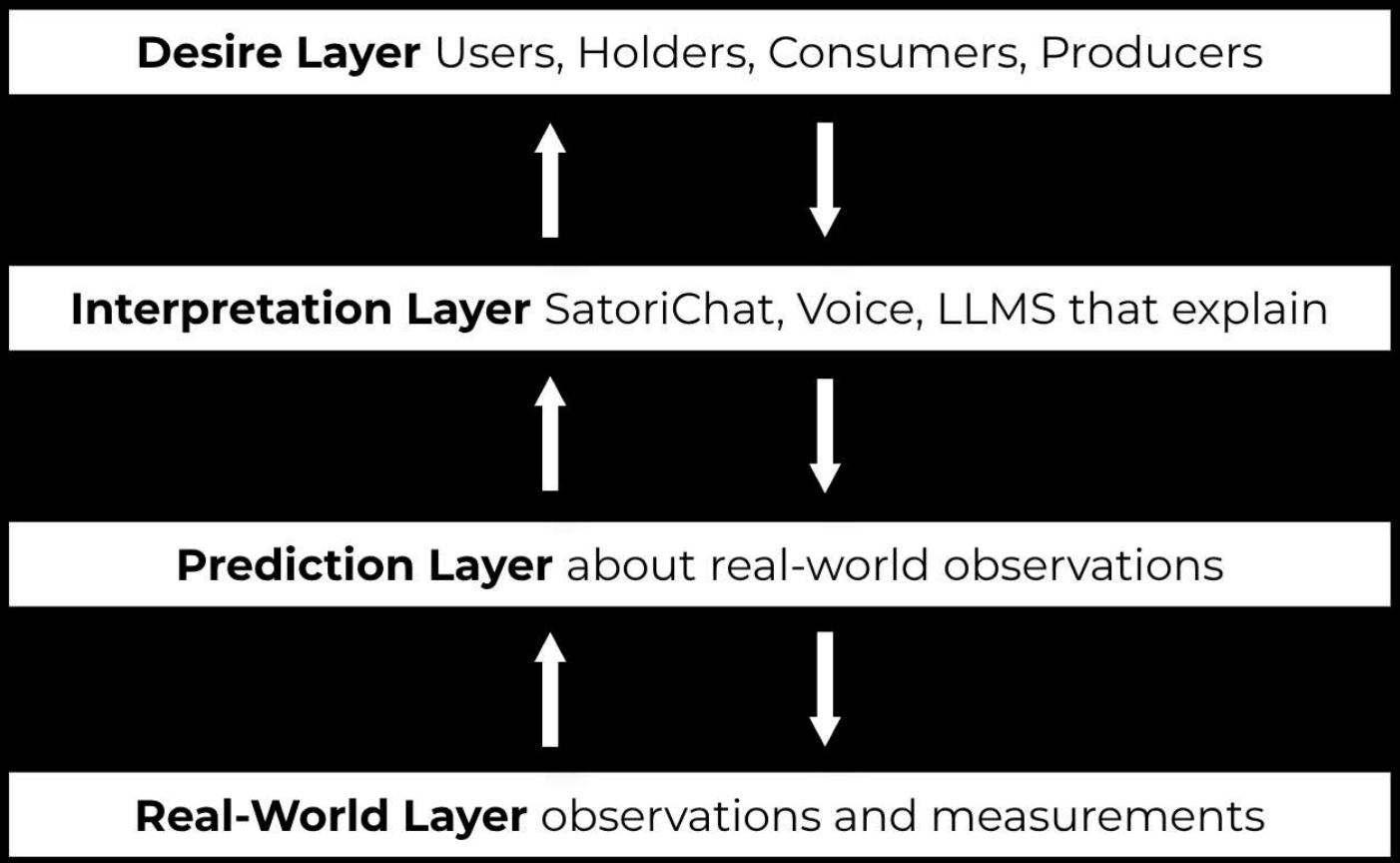
Jordan Miller
WCF, Salt Lake City, Utah
E-mail: iamajordanmiller@gmail.com
Phone: (801) 999-0484

Times series data includes data points organized by temporal order. There are unique challenges to model and forecast future values of time series, including data discovery, accessing data from diverse sources, preparing the data for analysis, selecting appropriate models and hyperparameters, training, testing, evaluating, and ultimately deploying models. A publish and subscribe network was established to provide data streams to the AI/ML engines. As data and access evolve, these data streams update based on pre-defined intervals. The AI/ML engines began a series of automated computations (known as **Auto-M**L) to determine the best forecasting model(s) from a collection of more than 40 potential types of models, including classical time series analysis, tabular Machine Learning (ML), Deep Learning (DL), and Large Language Models (LLMs). This automation controls data pre-processing, feature selection, model and hyperparameter selection, test, and evaluation to discover the best solution across a complex space of possibilities. The resulting framework and models can be deployed to a variety of platforms.

Original Vision that Inspired First Generation Implementation of this Project

Satori Layers

Satori is a **Circuit** that allows our **present desires** to alter **the future**



AI/ML Systems Architecture

Requirements to Protect and Handle Sensitive Information

External Data & Systems

Security, Identity & Access Management

Data Ingestion & Discovery

Data Dictionary & Catalog

Data Governance

Data Pre-processing, Matching & Curation

Data Management

AI/ML, Data Mining, Analytics & Other Analysis

Query, Search & Visualization

Decision Support & Automation

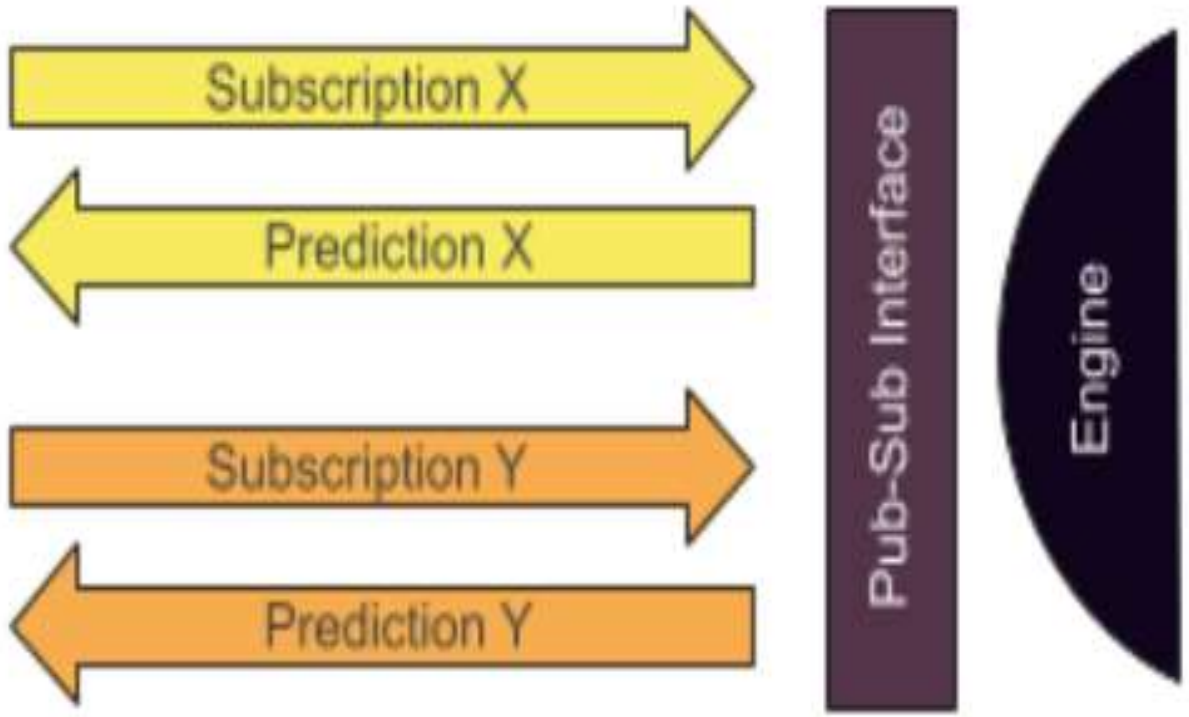
Analysis of Outcomes & Effectiveness

Continuous Improvement

Time Series Forecasting and Its Challenges

- **Time Series –** A series of data points that can be organized according to temporal order
- **Time Series Forecasting** – The process of modeling and predicting future time series values
- **Challenges of time series forecasting**
  - Lack of knowledge of underlying processes (e.g., US holidays) or physical phenomena
  - **Data:** Access, quantity, quality, biases/diversity, noise, missing, sampling
  - Non-stationarity (trends and cyclical behavior)
  - Time-varying changes in system being modeled
  - Selecting and training suitable models and their hyperparameters
  - Computational resources

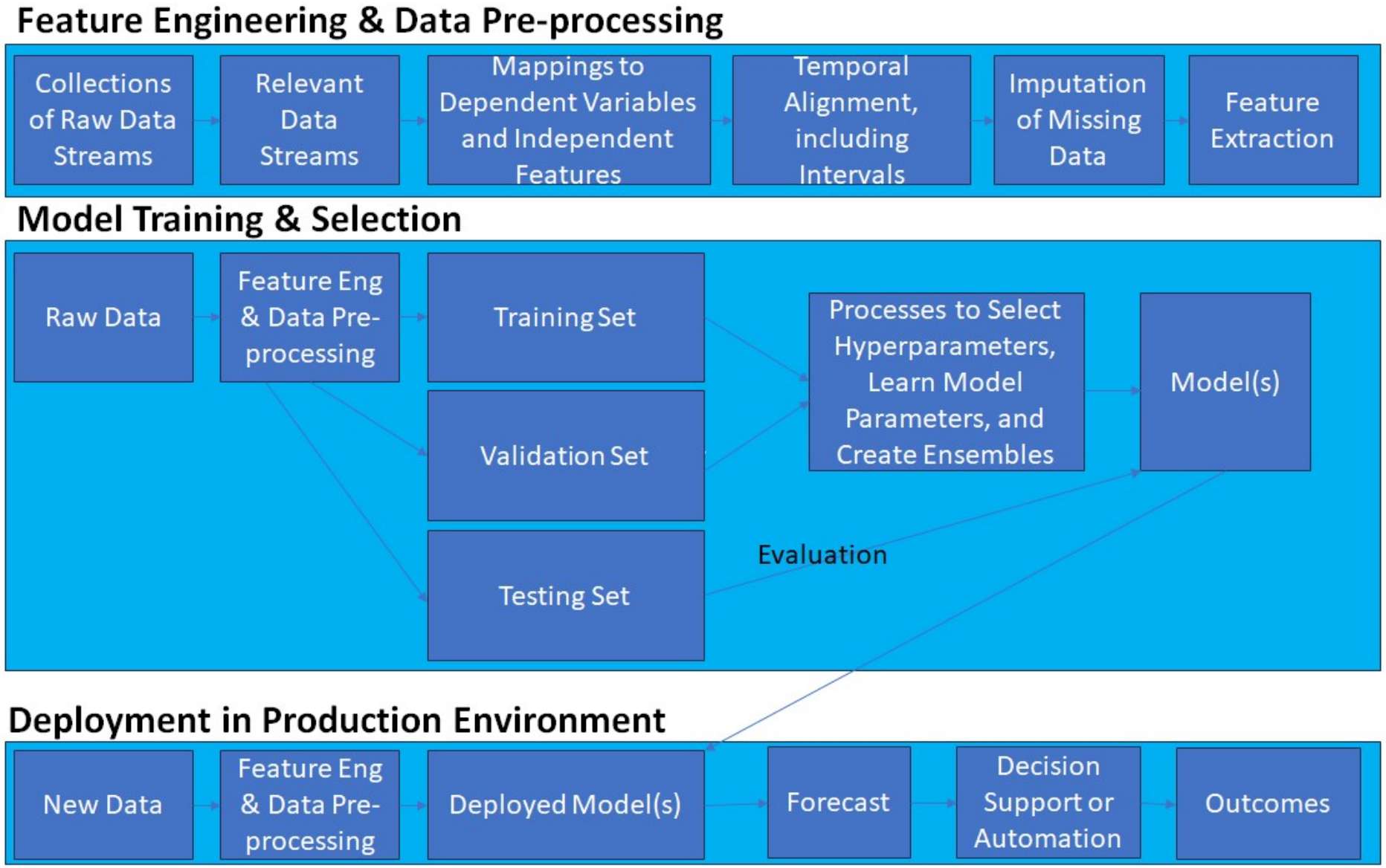
Publish & Subscribe Model



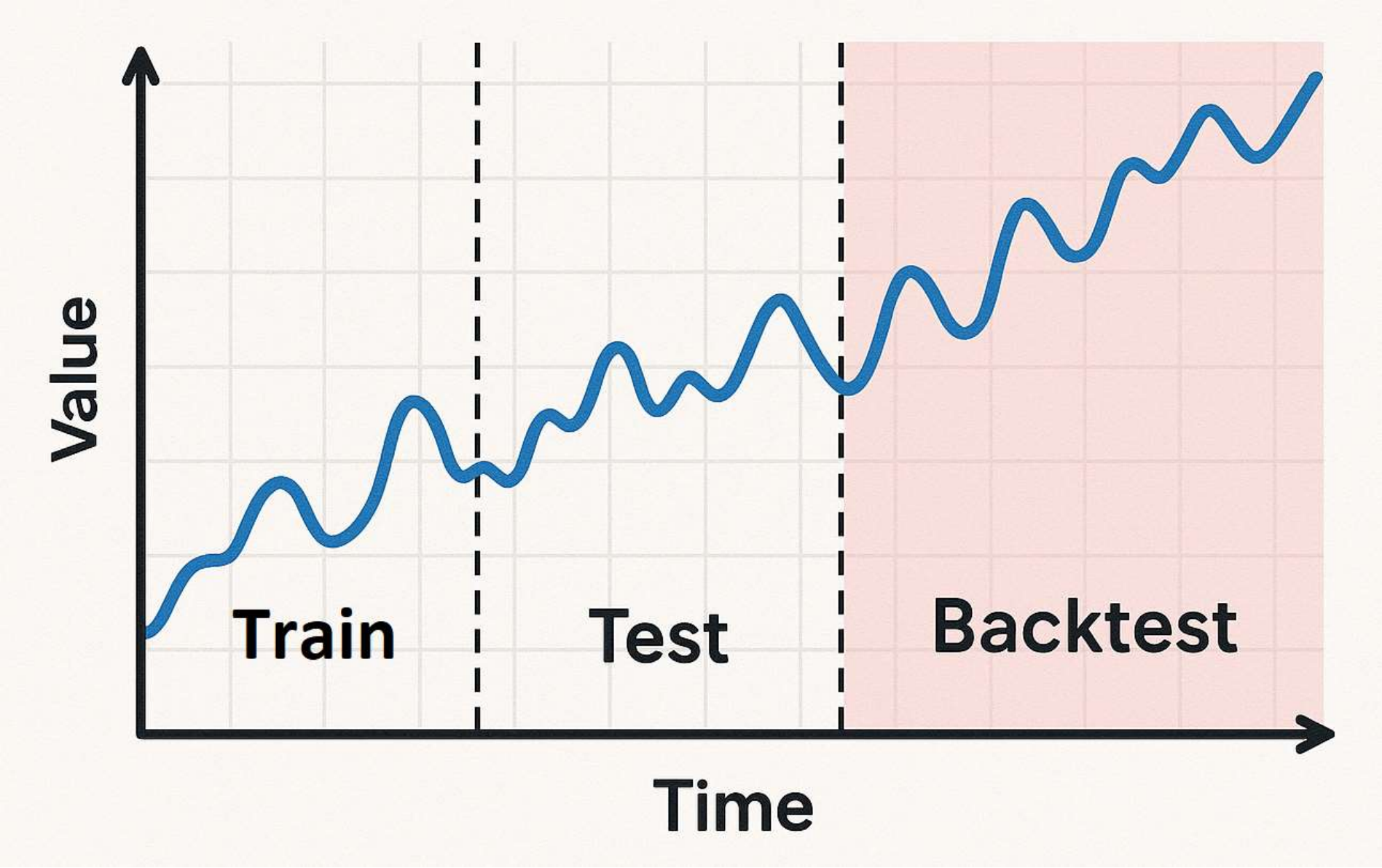
each datastream subscription corresponds to a published prediction datastream

Searching the Space of Parameters: Contemporary Alchemy

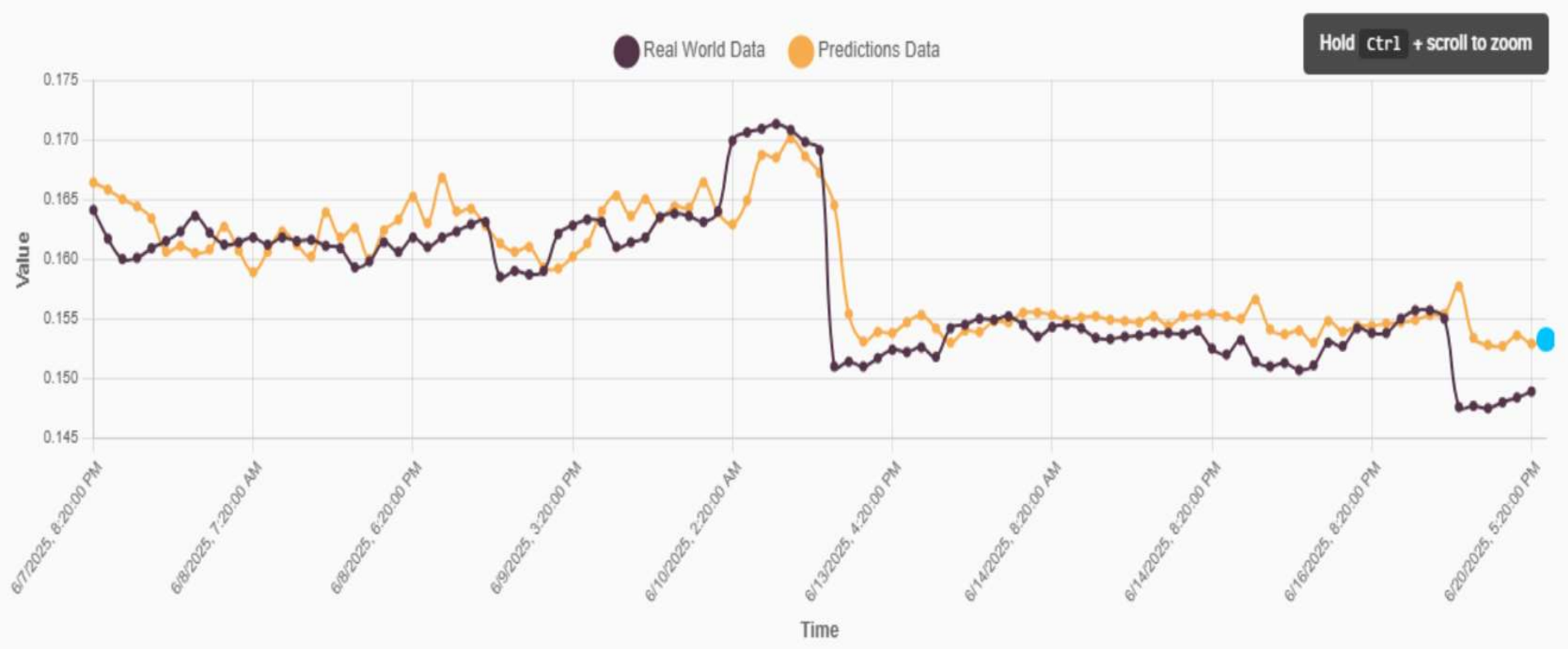
Auto-ML for Time Series Models: Approach to Searching Parameter Space



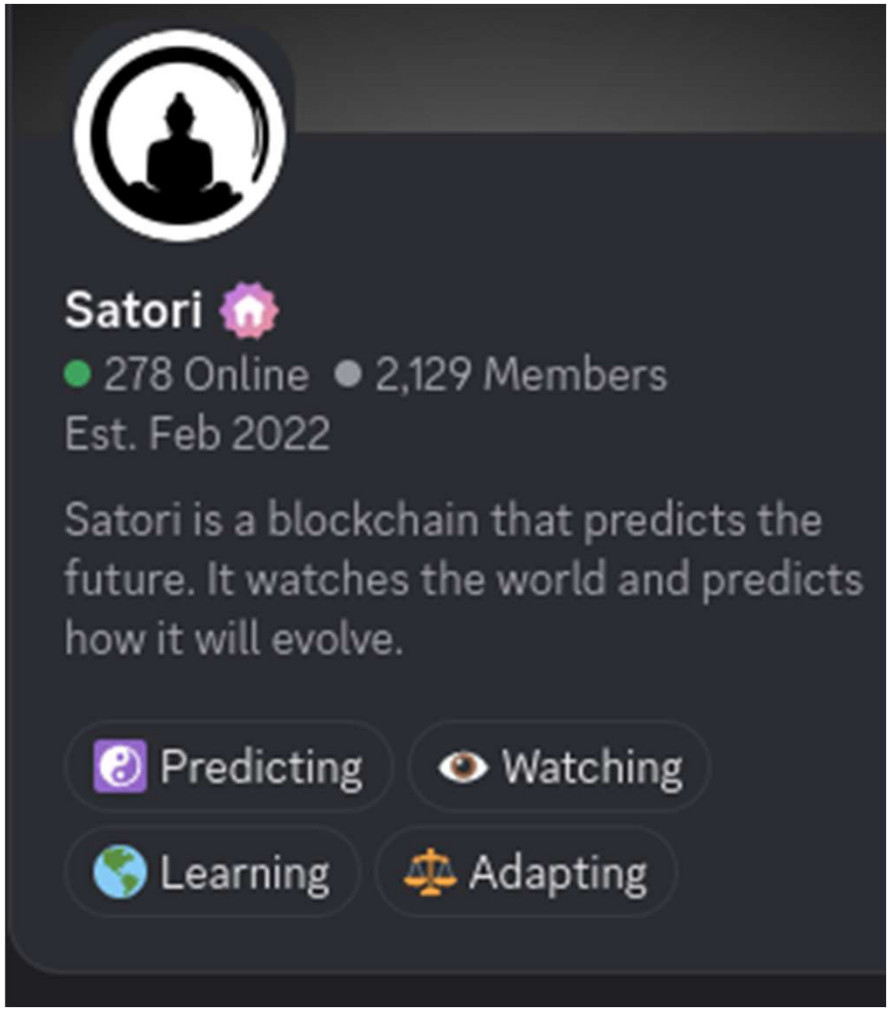
Backtesting: A Rigorous Approach to Time Series Auto-ML and Error Estimation



Preliminary First Generation Implementation of Dashboard: A Tabular ML Model Predicting Doge Crypto Coin



Preliminary Stakeholders Inspired by Vision



Dr. Soller and Mr. Miller are building on and advancing the open source Satori framework independent of the Satori Association.

Conclusions and Future Direction

To meet the unique challenges of time series forecasting, we developed a framework for managing publish and subscribe access to data streams, auto-ML for modeling multivariate time series, and the ability to predict values of interest in the future. This leverages more than 40 widely used types of models including classical times analysis, ML, DL, and LLMs.

Potential Future Directions for Next Generation of Project

- Intelligent data discovery for new potential data streams
- Improve the infrastructure to stream related information to better support existing multivariate modeling
- Add functionality to support trust-worthy, secure and Responsible AI
- Continue to improve and extend the Auto-ML
- Marketplace to incentivize the providers of data streams and developers of models
- Rich semantic description of the numerical predictions produced by the models
- Mature and harden all aspects of the system

Author Information and Acknowledgments

- **Dr. Jerome Soller** received his BS in Electrical Engineering from the Johns Hopkins University and PhD in Computer Science from University of Utah (UofU). He is currently the President / Founder of CogniTech Corporation; and a Director for STEMistryLab, a non-profit dedicated to applying AI to education. He has 35 years of experience in AI, ML, data mining, Modeling & Simulation, software systems architectures, and software engineering, including past positions as Research Investigator for the VA GRECC and a Limited Term Instructor for the UofU School of Medicine. He developed solutions for these industries: aerospace & defense; healthcare; legal; scientific & engineering; education; and creative. During 1990s, he organized regional AI and Cognitive Science events. He led past efforts, including the OMG CORBAmed Healthcare Data Interpretation Facility (HDIF) and the National Emergency Medical Services Information System (NEMSIS) web services / data interchange standards.
- **Mr. Jordan Miller** is the founder of The Satori Association, a non-profit that focused on developing a framework for time series forecasting to benefit humanity. He currently is an AI developer at WCF.
- Dr. Soller and Mr. Miller are building on and advancing the open source Satori framework independent of the Satori Association.
- Acknowledgments
  - **Mr. Wil Bown -** Assisted development of portions of the AI/ML engine and configuration.
  - **Mr. Krishna Prem** – Assisted design and development of framework enabling publish & subscribe capabilities; developed portions of AI/ML engine implementation.