Artificial Intelligence and Electrical & Electronics Engineering: AIEEE Open Access ISSN: 3069-0706

PRIME OPEN ACCESS

Volume 2, Issue 2

Review Article

Date of Submission: 30 Mar, 2026
Date of Acceptance: 20 Apr, 2026
Date of Publication: 15 May, 2026

Scientific Machine Learning

Mark Temple-Raston*

Decision Machine, LLC, New York, NY, USA

***Corresponding Author:** Mark Temple-Raston, Decision Machine, LLC, New York, NY, USA.

Citation: Temple-Raston, M. (2026). Scientific Machine Learning. *Art Intelligence and Ele & Electronics Eng: AIEEE Open Access, 2*(2), 1-06.

Abstract
Scientific Machine Learning implements the science-of-counting to analytically process any time-series and produce a complete set of thermodynamic measurements that define the state of the system. The scientific measurements are first illustrated with a time-series of closing-prices on a stock. Exact scientific measurements from Scientific Machine Learning (SML) are then directed to create time-series decision services, without model or bias, which we call a Decision Machine. A Decision Machine service is defined for a large class of Open allocation problems and applied to optimal sales for a consumer product good.

Keywords: Machine Learning, Risk Analysis, Non-Equilibrium Thermodynamics, Information Theory, Decision Theory

Introduction
Emerging in the first half of the 19th century, the science-of-counting generated new and fruitful scientific disciplines, including, thermodynamics, the theory of electricity, physical chemistry, and more, that generated the many practical applications that shaped the modern world. The science-of-counting applies to all time-series, mechanical or otherwise, that count states, events, or units of measure, counting natural numbers (1,2,3,...).

The science-of-counting (SoC) is founded on the principle of maximum information entropy, a method of logical inference used to define both science and machine learning, through enforced constraints on information [1]. For SoC, the maximum entropy constraints are determined entirely by the time-series itself — no external model assumptions are introduced. Scientific Machine Learning (SML) is the simplest, unsupervised, non-trivial and solvable case of machine learning, and is purely transactional: SML ingests time-series and returns scientific measurements that justify the name Scientific Machine Learning (SML). To make the scientific instrumentation available broadly, a web-based service provides unbiased, model-free scientific measurements (the unique thermodynamic decomposition) for any input time-series, subject to modest data quality requirements.

Examples of scientific measurements of a time-series are presented below for six months of closing price data for General Electric (GE) stock, consisting of timestamps for the trading day and the closing prices (adjusted). The closing prices are plotted in Figure 1. The arithmetic mean (average) of the closing prices is superimposed (dash red curve).

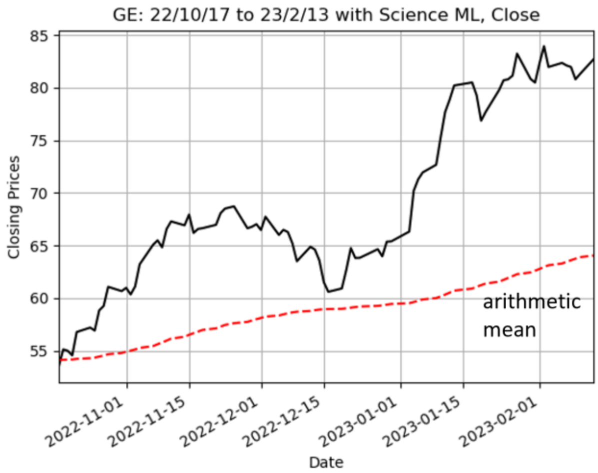


Figure 1: Closing Prices for General Electric Stock (GE)

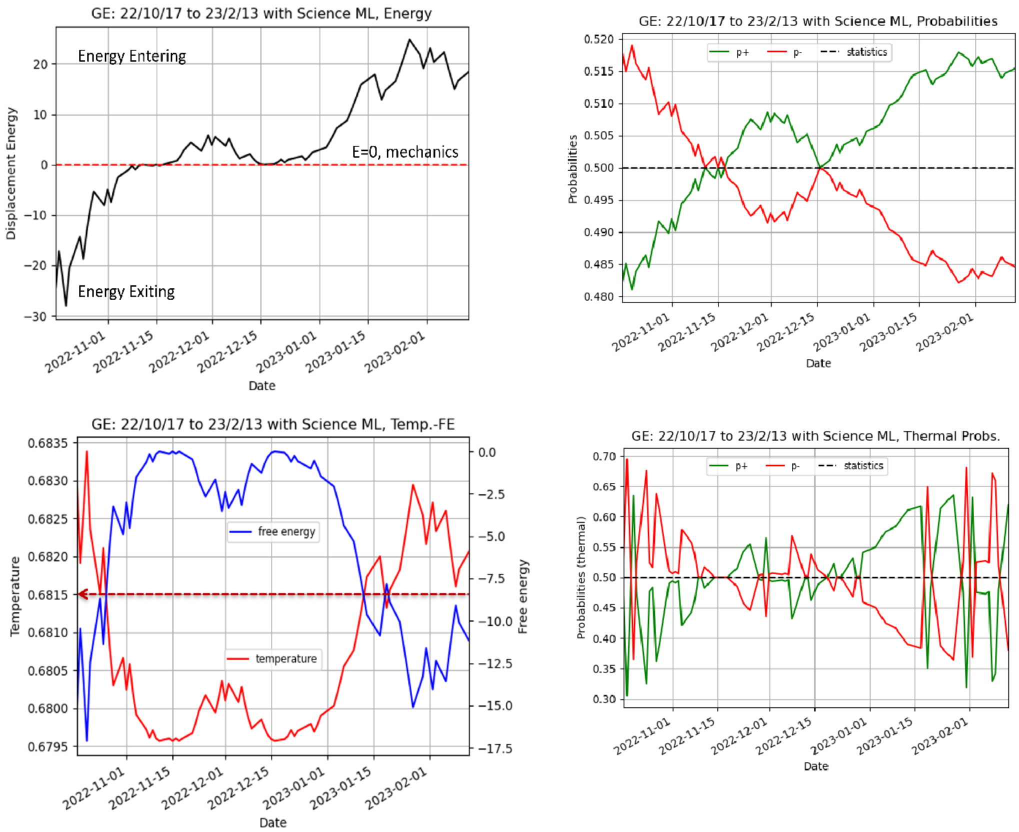


Figure 2: SML Measurements for GE (Binary/Up-Down States): A) Displacement Energy from Mechanics, B) Probabilities that the Stock Will Close Up/Down on the Next Day (Green/Red), C) Free Energy (Helmholtz) And Temperature, D) Thermal Probabilities for Closing Up/Down (Green/Red).

Several scientific measurements are plotted (Figure 2) from GE's historical closing prices (Figure 1). The measurement service for binary states counts the number of up or down states based on the close.

Figure 2a plots the displacement energy, the total energy minus the mechanical energy of the two-state system, equivalently, the energy entering (positive) or exiting (negative) the system. In practice, it is observed that real-world time-series deviate significantly from mechanics and statistical equilibrium, and that systems routinely store and release energy (dissipation of entropic heat). When the displacement energy, E=0, the system reduces to mechanics and the binary states behave statistically, $p_+=p_-=1/2$.

Figure 2b plots the probabilities, p_+ (green) and p_- (red), that GE stock will go up on the next trading day (green), or, close on the next trading day (red), respectively. When probabilities meet the dotted line, $p_+=p_-=1/2$, we observe that E=0 in Figure 2a.

Figure 2c presents a double-sided plot with the time-series measurements for the (entropic) temperature and free-energy (Helmholtz), the energy available to do price movement work. As the free-energy is used and observed to decrease, we see price movement work, and the temperature of the system increases. Observe the dissipative structure in the time-series in Figure 2c. When the internal temperature equals the body temperature, the system is in thermal equilibrium, and the measurement of body temperature is possible.

Finally in Figure 2d, we plot the elevated thermal probabilities for the two-state system for closing prices, using the same conventions as in Figure 2b. The thermal probabilities are calculated using the Crooks Fluctuation Equation in non-equilibrium thermodynamics [2].

The significance of counting comes from the fact that counting is an independent, self-sufficient, and exact human act that does not depend on the individual person or on judgement and is to know things just as they are. Counting measurements in Scientific ML are returned before models or bias can be introduced. We are not restricted to counting tangible, material objects; counting can include concepts, ideas or qualities that exist as thought or feeling.

The next section examines a large class of practical decision problems: Open System Energy Allocations (OSEA). OSEA generalizes multi-armed bandits to open environments. Using SML, time-series determine an exact Probability Distribution (PD) for each "bandit". In conventional bandit problems choosing an arm does not affect the properties of the arm or of any other arm, that is, Armed Bandits are static and independent (Closed). OSEA permits time-dependence and interactions with other time-series or sources of energy.

SML demonstrates an OSEA problem with sales time-series for a Consumer Product Good (CPG, coffee pods) to calculate the optimal energy allocations at each timestamp to get the best sales performance for the least energy/effort. Energy allocations can be translated into currency allocations with a better appreciation of the utility of money that generates the required energy allocation.

Open System Allocation

Multi-Armed Bandits (MAB) applications are in common use today:
- online advertising (serve new impression, or, previously viewed impression; which ad to display),
- clinical trials (new treatment, or, standard care; allocating patients to different treatment arms),
- recommender systems (new product or service suggestion, or, repeated product or service suggestion),
- portfolio management (new investment, or, add to existing investment; adjusting investments based on reward probabilities),
- adaptive routing (new bypass route (given congestion), best route (no congestion);; best traffic route),
- product development (new feature, or, known feature; features selection), and so on.

Conventionally these applications assume that the environment is effectively static on the timescales considered critical for the N-armed-bandit problem (Figure 3, below, left). However, evidently, none of these applications is found to be truly static, so that MAB models must be replaced periodically as conditions change. Upgrading MAB applications to leverage SML could significantly reduce the need for building models, because as the latest data is incorporated in the time-series, scientific measurements, as patterns, are produced in real-time.

Assume that resources for energy allocation to multiple decision options are made available, but with probabilistic results. Recognizing that the environment changes as energy enters or exits the Open system, SML produces the probability measurements (pure, conditional and joint) that support Open Systems analysis (Figure 3, right). SML therefore generalizes MAB to a very large class of energy allocation problems: Open System Energy Allocations (OSEA).

Figure 3: Two-Armed Bandit (Left) Requires Time-Independence and Forbids Interactions. Scientific Machine Learning Permits Time-Dependence and Interactions with Other Time-Series or Sources of Energy in Probability Theory. When p(AB) = p(A)p(B), so that A and B are Independent, the Right-Hand Interaction Diagram Reduces to the Non-Interacting Left-Hand Diagram for Bandits.

Decision Machine

A Decision Machine is defined as an orchestration of data processing services in Scientific Machine Learning, in service to a Policy (see Figure 4). A Decision Machine can be developed by anyone with access to exact SML measurements.

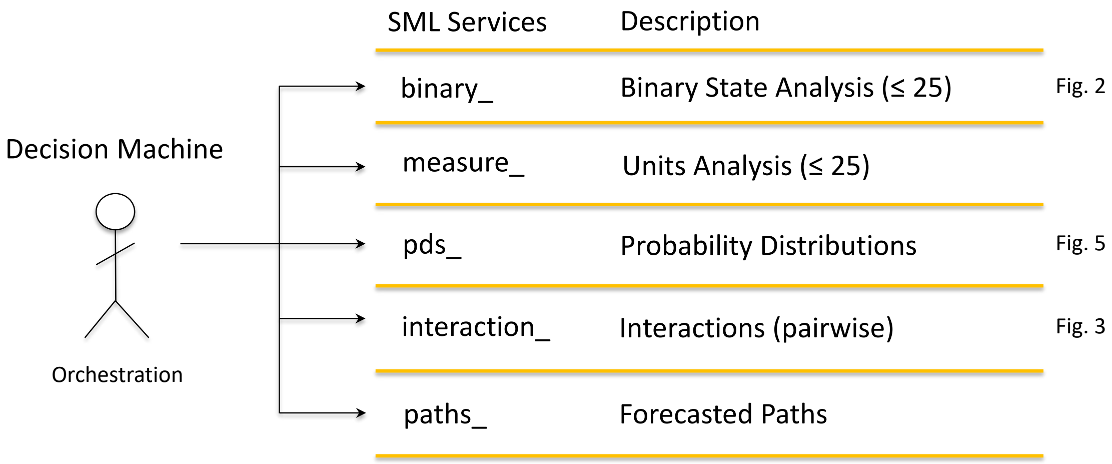


Figure 4: Counting Measurements are Made for Binary (Up-Down) States, Events and Units, Using the Binary_ Or Measure_ Services. Up To 25 Time-Series for The Same Set of Time-Stamps Can Be Processed At Once. The PD and Demand Distribution (DD) are Computed for the Latest Distribution, or, to Produce A Video of the Demand Distribution. Interactions in SML (figure 3, Right) Are Computed Pairwise. Paths are Generated Consistently with Historical PD.

Counting-in-time (time-series) determines two quantities in SML, the self-coupling, λ, and the energy of the system, E. The self-coupling appears because it enforces the counting constraints on maximum entropy, and is calculated exactly in SML, along with the energy. Self-coupling can be viewed as the tendency to return to statistical equilibrium, and contributes to the strain, $\langle\varepsilon\rangle$. Subtracting mechanical energy from the total energy determines the energy entering or exiting the system, the displacement energy which we will continue to denote by E. The natural coordinate system for Scientific Machine Learning is presented in Figure 5a. The (∂_λ, ∂_E) are centered on statistical mechanics, which is clearly a special case. Also observe that in statistical equilibrium the energy is naturally quantized, $E=\lambda n$, in multiples of the self-coupling.

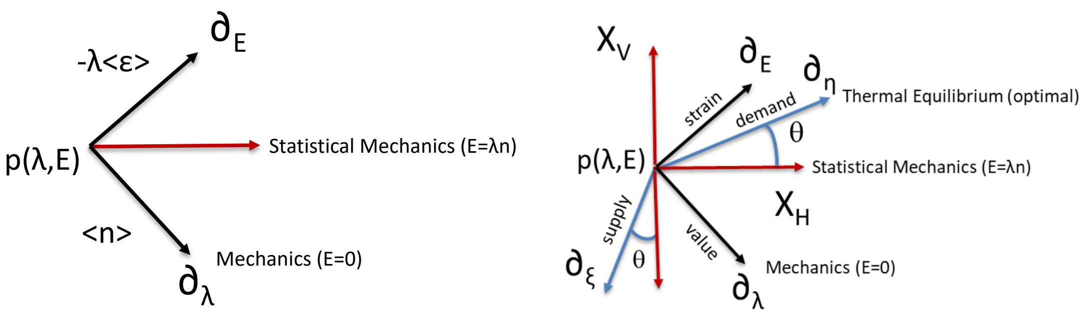


Figure 5: LHS: The Change in Momentum, p(λ, E), Due to Change in Self-Coupling or in Energy, Defines the Expected Value of What Is Being Counted, $\langle n\rangle$, and the Coupling-Strain, $-\lambda\langle\varepsilon\rangle$ (Gibbs, see [1]). Statistical Mechanics is Found to be a Special Case that Bisects the Coordinate Basis. RHS: the LHS Coordinate System is not the Most Natural when Energy is Entering the System. The Linear Combinations $\langle\eta\rangle = \langle n\rangle - E\lambda\langle\varepsilon\rangle$ for the Demand, and $\langle\xi\rangle = E/p^2-\lambda\langle\varepsilon\rangle$ for the Supply, Together Form the Eigenbasis.

The natural coordinates for SML in Figure 5 provide the exact scientific measurements that quantify the current state, including the discrete, probability distribution (PD, red in Figure 6 below) for the closing price data in Figure 1. The demand distribution (blue, Figure 6) is computed directly from the PD by multiplying the number of units (x-axis) by the probability.

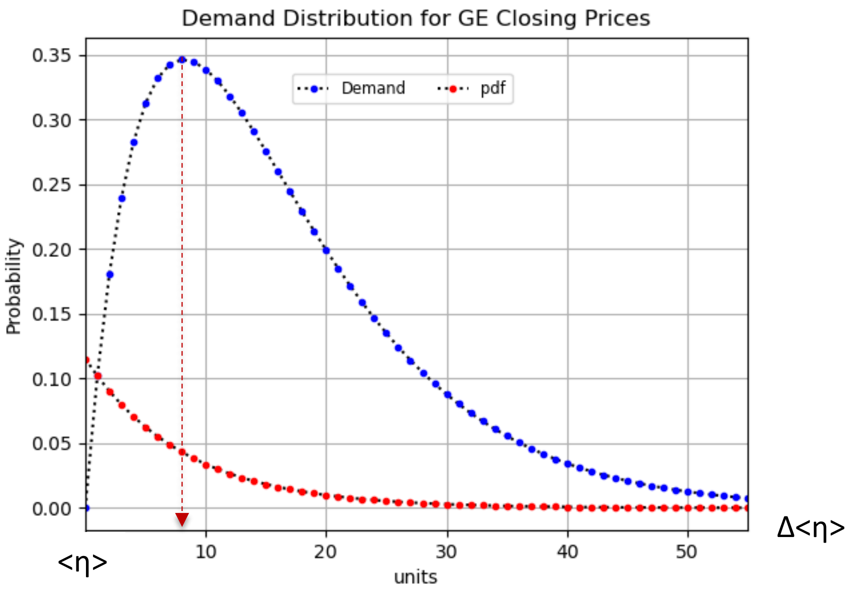


Figure 6: Exact PD and DD for GE Closing Prices. SML Measures the Exact Probability Distribution (PD) for a Time-Series (Equivalently, A Sampling Distribution), and from the PD Directly Computes the Demand Distribution (DD).

Decision frameworks also demand forecasting. To forecast using Figure 6, users download the exact PD into a Monte Carlo environment to locally generate paths for a Monte Carlo simulation, replacing a subjective probability distribution with the exact PD. In addition, SML can generate a set of paths consistent with time-series history and the geometry of the science-of-counting.

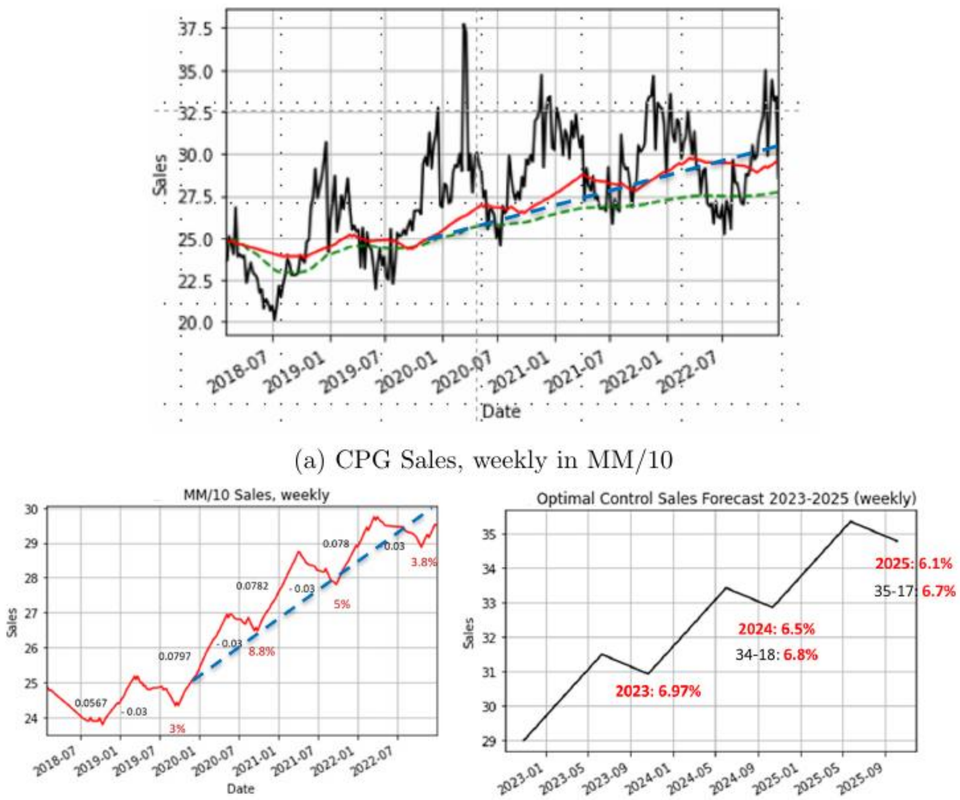
Finally, two time-series may be found to affect each other, that is, to interact. The interaction diagram in Figure 3 (RHS) follows from the product rule in probability theory. SML, in addition to the pure PD and DD above, calculates joint and conditional probabilities and demands.

The Cultivation of Sales

In this section we are presented with five years of sales data for a consumer product good (CPG: coffee pods). Figure 7a plots the weekly consumer sales figures (black curve) and the running arithmetic mean (green, dashed) from 2016 to 2023 (in units of MM/10).

The blue dashed line in Figure 7a is calculated by the angle, θ, in the SML coordinate system Figure 5 (RHS) that corresponds to thermal equilibrium (equivalently, Fokker-Planck, statistics with an optimal drift). The angle, θ, between the demand eigenvector and statistical mechanics is calculated to trace out the path of greatest increase in sales, for the least amount of energy. With more detail in Figure 7b, the least effort policy (blue dashed line) projects sales to increase by roughly 10% per year for three years! This level of cultivation is not sustainable, of course, and doesn't reflect seasonality.

Efforts that drive sales *above* thermal equilibrium necessarily convert energy to heat that dissipates. The energy is lost to the system. The red sales path in Figure 7ab takes the actual (seasonal) sales figures and measures the expect demand. Expect demand is found to have a quasi-linear trajectory when maximally strained, resembling Fokker-Planck (statistics with a drift), which deviates from the optimal path. The red sales path in Figure 7ab is the path of maximum increasing demand with strain/supply extended to its maximum.



(a) CPG Sales, weekly in MM/10

(b) Maximum Strain Sales

(c) Maximum Strain Forecast

Figure 7: Maximum Improvement, Least Effort. (Blue) Maximum Strain. Stretch and Release. (Red)

Operational efficiency at this level of control is likely well beyond our current reach. However, there is middle ground between what we currently do and what is *most* efficient. The red paths in Figure 7 investigate the middle ground.

Sales paths steeper than the critical angle are more difficult to control, because the system is overheated and energy is unpredictably dissipated. The piece-wise linearity in the red path reflects seasonality in sales: there is a high season and low season, with durations defined by the sales operations team (in weeks that add up to 52 weeks). The slopes of the triangular demand curve are determined empirically (by filtering dissipation) from time-series (see figure 7b). Using the slopes calculate for 2020 for future returns, year-over-year percentage increases in sales for the optimized paths are calculated (see figure 7c): 2023: 6.97%, 2024: 6.8%, 2025: 6.7%.

Conclusions

SML is the foundational technology for the manufacture of Decision Machines, that are designed to provide the best analytical support for human decisions using time-series. In the previous sections we constructed three Decision Machines for Open systems: the best allocation decision among a set of decisions starting with no allocation preference (OSEA); and two optimal decisions, best result for minimum effort, and best result for maximum seasonal strain. We continue to investigate decision making in SML in [3], especially inventory levels, risk management and interactions.

References

1. Jaynes, E.T. (2003) "Probability Theory: The Logic of Science", Cambridge University Press.
2. Crooks, G.E. (1999) Phys. Rev. E 60, 2721–2726.
3. Temple-Raston, M (2026) "Decision Making in Scientific Machine Learning", Machine Learning and Applications (MLAIJ) Vol. 13, No.1, March 2026.