Understanding and Shaping AI Is Also a Job for Applied Mathematicians

Soon Hoe Lim
Department of Mathematics, KTH Royal Institute of Technology
Nordita, KTH Royal Institute of Technology and Stockholm University
`shlim@kth.se`

Abstract

Artificial intelligence (AI) and machine learning (ML) are often described as triumphs of data, computation, and engineering scale. This description is correct, but incomplete. Modern AI systems are also mathematical objects: high-dimensional dynamical systems, stochastic algorithms, approximation schemes, optimization procedures, probabilistic models, and increasingly, scientific instruments for modeling complex phenomena. This reflection paper argues that applied mathematicians have a central role to play in AI/ML research. Their role is not limited to explaining existing methods after empirical success; rather, applied mathematics provides the language, tools, and habits of thought needed to formulate new learning problems, design reliable algorithms, analyze stability and uncertainty, and connect data-driven methods with mechanistic models. We argue that AI/ML is becoming a major domain of contemporary applied mathematics, driven by developments in scientific machine learning, generative modeling, neural operators, optimal transport, stochastic control, uncertainty quantification, learning dynamical systems, and the theory of large-scale models. The next stage of AI will require not only scale, but also mathematical principles for reliability, interpretability, robustness, uncertainty, and scientific validity.

1 Introduction

Artificial intelligence (AI) and machine learning (ML) are often described as triumphs of data, computation, and engineering scale. This description captures an important part of the recent revolution, but it is incomplete. Modern AI systems are not merely software artifacts trained on large datasets. They are also mathematical objects: high-dimensional dynamical systems, stochastic algorithms, approximation schemes, optimization procedures, probabilistic models, and increasingly, scientific instruments for modeling complex phenomena.

This reflection paper argues that modern applied mathematicians have a central role to play in AI/ML research. Their role is not limited to explaining existing methods after they have already succeeded empirically. Rather, applied mathematics provides the language, tools, and habits of thought needed to formulate new learning problems, design reliable algorithms, analyze stability and uncertainty, and connect data-driven methods with mechanistic models.

The title of this paper deliberately echoes Lenka Zdeborová's call that understanding deep learning is also a job for physicists [15]. The analogy is useful, but the applied mathematical role is distinct. Applied mathematics is not only concerned with interpreting learning systems through analogies to physical phenomena. It is also concerned with formulation, approximation, numerical stability, stochastic modeling, inference, and the design of algorithms that remain meaningful when deployed in complex scientific or technological settings.

This is not a call for applied mathematicians to enter AI/ML from the outside. It is a recognition of a movement already underway. Across scientific machine learning, generative modeling, neural operators, optimal transport, stochastic control, uncertainty quantification, learning dynamical systems, and the theory of large-scale models, applied mathematicians are already shaping the field. The momentum is clear: many of the central questions in modern AI are naturally questions of probability, dynamics, optimization, approximation, geometry, and computation.

The next stage of AI will not be determined by scale alone. Scale will remain important, but it will not by itself solve the problems of reliability, interpretability, robustness, uncertainty, and scientific validity. These are mathematical problems as much as engineering problems. Understanding and shaping AI is therefore not only a job for computer scientists, engineers, or physicists. It is also a job for applied mathematicians.

2 AI/ML as a new domain of applied mathematics

Applied mathematics has always developed through engagement with complex systems: fluids, materials, populations, financial markets, biological processes, climate, and physical systems across scales. Its strength lies in connecting mathematical formulation, analysis, computation, and application. AI/ML now belongs naturally to this tradition.

Modern ML involves many of the central themes of applied mathematics. Probability and stochastic processes appear in uncertainty quantification, sampling, diffusion models, and stochastic optimization. Dynamical systems appear in neural ordinary differential equations, recurrent architectures, sequence models, and stability analysis. Optimal transport appears in distributional learning and generative modeling. Numerical analysis appears in discretization, solver design, and computational stability. Partial differential equations, stochastic differential equations, and multiscale analysis appear in score-based generative modeling, scientific machine learning, and reduced-order modeling. Optimization appears in training algorithms, implicit bias, and nonconvex landscapes. Statistical inference appears in generalization, calibration, model selection, and uncertainty representation.

This connection is not superficial. Many modern AI methods can be interpreted as mathematical procedures for transforming distributions, approximating functions, solving inverse problems, simulating dynamics, or optimizing high-dimensional objectives. Physics-informed neural networks [12], neural operators [8, 7], neural differential equations [3], diffusion models [5, 14], and flow matching [9, 1, 10] all illustrate how mathematical structure increasingly informs both the design and interpretation of AI methods. Thus, AI/ML should not be regarded only as an application area for applied mathematics. It should be seen as one of the major domains in which applied mathematics is currently evolving.

3 From empirical success to mathematical understanding

The empirical success of modern AI is undeniable. Large neural networks generalize in regimes where classical theory would suggest overfitting. Generative models produce strikingly realistic samples. Sequence models learn complex dependencies across long contexts. Optimization algorithms find useful solutions in enormous nonconvex parameter spaces. Yet these successes also expose deep conceptual gaps.

We still lack complete answers to basic questions. Why do overparameterized models generalize? What structures do neural networks learn? How does stochasticity influence training? When are learned dynamics stable under long-time iteration? How should uncertainty be represented and propagated? How can generative models be made controllable and reliable? How can scientific constraints be imposed without sacrificing flexibility? When does a data-driven model extrapolate beyond its training regime?

These are not merely theoretical curiosities. They determine whether AI systems can be trusted in scientific, medical, engineering, financial, and societal applications. A model used for simulation, forecasting, control, or decision-making must do more than interpolate a benchmark distribution. It must behave coherently under perturbation, iteration, discretization, and distribution shift.

Applied mathematicians are trained to ask such questions. They bring a perspective that is neither purely empirical nor purely abstract. They seek models that are useful, algorithms that are computable, and theories that explain behavior without ignoring practical constraints. This intermediate position is especially valuable in AI/ML, where the field often advances faster than its conceptual foundations.

4 A movement already underway

The role of applied mathematics in AI/ML is already visible and growing. Scientific ML has become a major research area, bringing together computational scientists, applied mathematicians, engineers, and ML researchers [6]. Neural operators draw on functional analysis, PDEs, approximation theory, and numerical analysis [8, 7]. Diffusion models, flow matching, stochastic interpolants, and Schrödinger bridges are shaped by stochastic processes, optimal transport, and continuous-time dynamics [14, 9, 4, 1]. Learning dynamical systems has become a central topic at the interface of data science, scientific computing, and applied mathematics.

This movement is significant because applied mathematicians are not only applying existing ML methods to classical problems. They are also contributing new conceptual frameworks. They reinterpret learning as transport, sampling as controlled dynamics, optimization as stochastic evolution, sequence modeling as learning evolution laws, generative modeling as probability-path construction, scientific ML as structure-preserving approximation, robustness as stability under perturbation, and uncertainty quantification as an essential part of prediction.

These perspectives do more than clarify existing algorithms. They suggest new ones. The growing presence of AI/ML in applied mathematics therefore reflects a disciplinary transformation. Applied mathematics is not standing outside the AI revolution. It is becoming one of the intellectual engines driving it.

5 Dynamical and probabilistic viewpoints

A particularly important frontier lies at the intersection of machine learning, applied probability, and dynamical systems. Many modern AI methods are best understood not as static input–output maps, but as evolving systems. Training is an evolution in parameter space. Sampling is an evolution in distribution space. Generative modeling constructs dynamics that transport one probability measure into another. Sequence modeling seeks to learn temporal evolution from data. Scientific ML aims to approximate the dynamics of physical, biological, or engineered systems.

This dynamical viewpoint changes how we think about AI. It encourages us to ask not only whether a model performs well on a test set, but whether its behavior is stable, coherent, and meaningful under iteration. It also highlights the role of probability: uncertainty, noise, sampling, fluctuations, and rare events are central to both learning algorithms and the systems they model.

From this perspective, many AI problems become questions about stochastic processes, transport dynamics, stability, approximation, and inference. Applied mathematics provides the language needed to formulate these questions precisely and the tools needed to address them.

6 Generative and sequential modeling

One of the clearest examples of the applied mathematical nature of modern AI is generative modeling. Many generative models can be viewed as transforming a simple probability distribution into a complex target distribution through learned deterministic or stochastic dynamics. This includes diffusion models, score-based models, normalizing flows, neural ODEs, flow matching, stochastic interpolants, and Schrödinger bridges [3, 5, 14, 9, 4, 1].

From an applied mathematics perspective, these are not merely architectures. They are methods for constructing and approximating probability paths, velocity fields, stochastic processes, and transport maps. This viewpoint is especially important for sequential data. Many real-world data sources are not static independent samples, but trajectories: physical simulations, biological signals, financial time series, climate records, molecular dynamics, and observations of complex dynamical systems. For such data, generative modeling cannot mean only producing realistic samples at a single time. It must mean generating coherent temporal behavior.

This requires mathematical tools for learning distributions on path space, modeling temporal dependence, preserving dynamical structure, and understanding the stability of generated trajectories. It also requires careful numerical implementation: learned dynamics must be integrated, discretized, and sampled efficiently.

Sequence modeling presents related challenges. Temporal data involves memory, dependence, causality, nonstationarity, multiscale behavior, and long-horizon error accumulation. A mathematically grounded approach to sequence modeling should ask: What is the underlying state? What is the relevant time scale? What structure should be preserved? How does uncertainty evolve? How do local prediction errors accumulate globally? When does the learned model remain stable under iteration? Can the model extrapolate beyond the observed regime?

For scientific and engineering applications, these questions are unavoidable. The goal is not merely to predict the next observation, but to learn dynamics that are coherent, stable, interpretable, and useful for simulation or decision-making. Applied mathematics can help connect modern sequence models with classical tools from dynamical systems, filtering, stochastic processes, control, and numerical analysis. This connection is crucial if AI systems are to move from pattern recognition toward reliable modeling of evolving systems.

7 Optimization, sampling, and numerical reliability

Training and sampling are two central computational acts in modern AI. Training a model requires optimization in high-dimensional and often nonconvex landscapes. Sampling from a model requires generating representative points from complex probability distributions. Both problems involve dynamics, geometry, stochasticity, approximation, and numerical error.

Applied mathematics provides tools for understanding these procedures. Gradient descent can be studied as a dynamical system. Stochastic gradient methods can be approximated by stochastic differential equations. Langevin algorithms, diffusion samplers, and controlled dynamics can be analyzed using stochastic processes, variational principles, and transport methods [13, 2].

This perspective matters because optimization and sampling are not merely implementation details. They shape what models learn and how they behave. Important questions include: Why do certain optimization algorithms generalize better than others? How does noise affect training? How should continuous-time generative dynamics be discretized? Can sampling be accelerated without introducing unacceptable bias? How do approximation errors accumulate? What are the trade-offs between computational cost, accuracy, and reliability?

As AI models become larger and more expensive, principled improvements in optimization and sampling become increasingly important. Applied mathematics can help move the field from heuristic algorithm design toward systematic algorithmic understanding.

8 Robustness, reliability, and uncertainty

One of the central limitations of current AI systems is reliability. Models can fail under distribution shift, produce overconfident predictions, behave unpredictably when iterated, or generate plausible but incorrect outputs. These failures are not incidental. They reveal gaps in our mathematical understanding.

Reliability requires more than benchmark performance. It requires methods for uncertainty quantification, calibration, stability analysis, robustness guarantees, sensitivity analysis, and error propagation. In high-stakes or scientific settings, these are not optional additions. They are core requirements.

Applied probability is especially important here. It provides a language for reasoning about uncertainty, rare events, stochastic perturbations, concentration, and risk. Dynamical systems theory helps analyze stability and long-time behavior. Numerical analysis helps identify discretization and solver-induced errors. Statistical theory helps distinguish genuine generalization from interpolation within a narrow data regime.

A central message of this paper is therefore simple: AI cannot become a mature scientific technology without a mathematical theory of reliability. Scale may improve performance, but scale alone does not provide trust.

9 Scientific AI and multiscale systems

Scientific AI is one of the most important arenas where applied mathematics can shape the future of ML. In scientific applications, the goal is often not merely prediction. It is simulation, discovery, control, inverse modeling, uncertainty quantification, and the construction of reduced or effective descriptions of complex systems. These goals require AI methods that respect structure: conservation laws, symmetries, stability, geometry, causality, physical constraints, and long-time behavior.

Stochastic differential equations provide a natural language for many phenomena in physics, chemistry, biology, finance, and machine learning. They also appear throughout modern AI: in diffusion models, stochastic optimization, Langevin sampling, stochastic control, continuous-time sequence models, and data-driven modeling of dynamical systems.

For systems with multiple time scales, the connection is even deeper. In many scientific problems, fine-scale dynamics are too expensive or too complex to model directly. Homogenization and stochastic analysis provide principled ways to derive effective dynamics [11]. ML can complement these methods, but only if guided by mathematical structure. Without such guidance, learned models may match short-term data while failing to capture long-time behavior, invariant measures, metastability, or rare events.

Scientific AI therefore requires more than flexible function approximation. It requires a mathematical understanding of dynamics, scales, noise, and structure.

10 Five central claims

The reflection can be summarized in five claims.

Claim 1: AI/ML is now a central subject of applied mathematics. Modern AI systems involve probability, dynamics, optimization, approximation, geometry, and computation. Their study belongs naturally within applied mathematics.

Claim 2: Applied mathematicians should help design AI methods, not merely analyze them afterward. The mathematical sciences should play a constructive role in inventing new models and algorithms, especially for generative modeling, sequence modeling, scientific ML, and reliable AI.

Claim 3: Dynamical and probabilistic viewpoints are essential for the next generation of AI. Many important AI systems are best understood as stochastic processes, transport dynamics, or learned evolution equations.

Claim 4: Reliability requires mathematics and statistics. Robustness, uncertainty quantification, stability, calibration, and long-time behavior cannot be solved by scale alone.

Claim 5: Scientific AI needs applied mathematics to connect data-driven learning with mechanistic understanding. In scientific applications, the goal is not only prediction, but also simulation, discovery, control, and explanation.

11 Conclusion

The future of AI will not be determined by larger models alone. Scale will remain important, but it will not by itself solve the problems of reliability, interpretability, stability, uncertainty, and scientific validity. These are mathematical problems as much as engineering problems. Applied mathematicians should therefore take an active role in AI/ML research. They should not only explain existing models, but help develop new ones. They should not only prove theorems, but shape algorithms. They should not only import tools from mathematics into machine learning, but also allow the challenges of AI to transform applied mathematics itself.

AI/ML is becoming one of the major scientific and mathematical enterprises of our time. To build systems that are reliable, interpretable, robust, and scientifically meaningful, we need the full participation of modern applied mathematics. Understanding and shaping AI is not only a task for computer science, engineering, or physics. *It is also a task for applied mathematicians.*

References

[1] Michael S. Albergo, Nicholas M. Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. *Journal of Machine Learning Research*, 26(209):1–80, 2025.

[2] Léon Bottou, Frank E. Curtis, and Jorge Nocedal. Optimization methods for large-scale machine learning. *SIAM Review*, 60(2):223–311, 2018.

[3] Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David K. Duvenaud. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

[4] Valentin De Bortoli, James Thornton, Jeremy Heng, and Arnaud Doucet. Diffusion schrödinger bridge with applications to score-based generative modeling. In *Advances in Neural Information Processing Systems*, volume 34, 2021.

[5] Jonathan Ho, Ajay N. Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, volume 33, 2020.

[6] George Em Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021.

[7] Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces with applications to PDEs. *Journal of Machine Learning Research*, 24(89):1–97, 2023.

[8] Zongyi Li, Nikola B. Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew M. Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations*, 2021.

[9] Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023.

[10] Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. *arXiv preprint arXiv:2209.03003*, 2022.

[11] Grigorios A. Pavliotis and Andrew M. Stuart. *Multiscale Methods: Averaging and Homogenization*, volume 53 of *Texts in Applied Mathematics*. Springer, 2008.

[12] Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

[13] Gareth O. Roberts and Richard L. Tweedie. Exponential convergence of Langevin distributions and their discrete approximations. *Bernoulli*, 2(4):341–363, 1996.

[14] Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021.

[15] Lenka Zdeborová. Understanding deep learning is also a job for physicists. *Nature Physics*, 16(6):602–604, 2020.