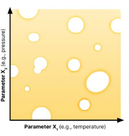
Atinary Emmental: ML Algorithm for Non-linear Constrained Optimization

Published the 24 may 2022



What is Atinary Emmental?

Emmental is a machine learning (ML) algorithm developed by Atinary Technologies that allows users to add a set of constraints in optimization problems. Solving non-linear constrained optimization is important across many real-world industry applications in pharma, biotech, materials science, chemistry and beyond.

Emmental uses the constraints to define the feasible regions in the parameter space, and avoid the non-feasible regions when solving the optimization problem. Emmental is compatible with Atinary's proprietary suite of ML algorithms available on Atinary's ML platform (SDLabs), described below.

Compared to popular Bayesian optimization libraries such as GPyOpt or BOtorch, Emmental can handle a broad variety of non-linear optimization problems with multiple parameters and constraints. Specifically, Emmental supports constrained optimization problems involving continuous, discrete and/or categorical parameters in combination with the following constraint types:

- Bounded constraints, which specify ranges of intermediate values that continuous/discrete parameters must avoid;
- Relational constraints, which define relationships between pairs of parameters that are to be avoided;
- Inequality and equality constraints, which indicate linear and non-linear relation between parameters.

What challenges does Emmental solve?

Real-world optimization problems are very often limited by multiple constraints. For example, the optimization of a chemical process may exhibit known physical and/or manufacturing constraints. These constraints have a direct impact on the parameter space of the optimization problem by effectively reducing its volume to only the subset of points that fulfill all the constraints (i.e. the feasible regions). Therefore, optimization techniques that lack the ability to incorporate this knowledge may choose parameter points that do not lie in the feasible regions.

What are typical use cases for Emmental?

Optimization problems with constraints are ubiquitous in nature, including in relevant applications in pharma, biotech, materials science, chemistry and beyond. For example:

- Mass balance constraints: In many chemical processes, it is required that the compositions of the input constituents add up to a specific total value. In certain cases, when optimizing the composition of A, B and C in a given chemical compound, it is necessary that the mass fractions mi add up to 1: $mA+mB+mC=1$.
- Thermodynamic constraints: Certain reactions and phase changes can only happen at particular thermodynamic conditions of pressure and temperature.
- Stoichiometric constraints: All the reactants of a chemical reaction should be present in specific proportions such that they are all consumed. This is particularly important in redox reactions to ensure that the correct reactants are consumed instead of producing undesired corrosion byproducts.

Atinary's proprietary ML algorithms

Atinary™ Falcon is a general-purpose optimization algorithm. It can solve optimization problems that include continuous, discrete and/or categorical variables with or without physicochemical descriptors, as well as batch-constrained optimization.

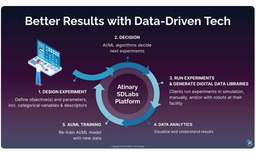
Atinary™ Falcon GPBO uses Gaussian Process Bayesian Optimization as surrogate model. Typically, GPBO is well suited for optimization problems that can potentially be solved with a relatively small number of experiments. However, GPBO scales cubically with the number of experiments. Thus, the computational cost can potentially be very high if used in complex simulation cases.

Atinary™ Falcon DNGO (Deep Network for Global Optimization) maintains desirable properties of the Gaussian Processes (e.g. management of uncertainty) while improving its scalability. Specifically, unlike a standard Gaussian process, DNGO scales linearly with the number of evaluations or experiments. Falcon DNGO creates a robust, scalable, and effective Bayesian optimization system that generalizes across many global optimization problems, for a suitable set of design choices.

About Atinary Technologies

Atinary is a Swiss/American deeptech startup based in Silicon Valley and Lausanne. Atinary accelerates optimization and discovery of breakthrough materials that are key to address climate change and sustainability challenges. Atinary enables the Self-driving Labs™ with its machine learning (ML) software platform (SDLabs) that minimizes waste and the R&D carbon footprint.

Atinary's SDLabs platform closes the loop in experimentation with ML. It enables users, customers and partners to deploy ML solutions in their existing workflows today. Atinary's solutions expand the space of what is currently possible in materials and molecules innovation. We solve complex optimization problems that are intractable with the current methods, including multi-objective and constrained-optimization with categorical variables and descriptors. Atinary accelerates R&D, reduces time to market and makes innovation more sustainable by minimizing waste and carbon footprint.



Atinary combines a multi-disciplinary team of chemists, software engineers, AI/ML experts and data scientists with technology and business leaders. Together with its rich network of academic experts, leaders, and strategic partners, Atinary deploys its vision of the term we coined as the Self-driving Lab(™).

References

Schonlau, M., Welch, W.J., & Jones, D.R. (1998). Global versus local search in constrained optimization of computer models.

Gardner, J.R., Kusner, M.J., Xu, Z.E., Weinberger, K.Q., & Cunningham, J.P. (2014). Bayesian Optimization with Inequality Constraints. *ICML*.

Snoek, J., Rippel, O., Swersky, K., Kiros, R., Satish, N., Sundaram, N., Patwary, M.M., Prabhat, & Adams, R.P. (2015). Scalable Bayesian Optimization Using Deep Neural Networks. *ICML*.

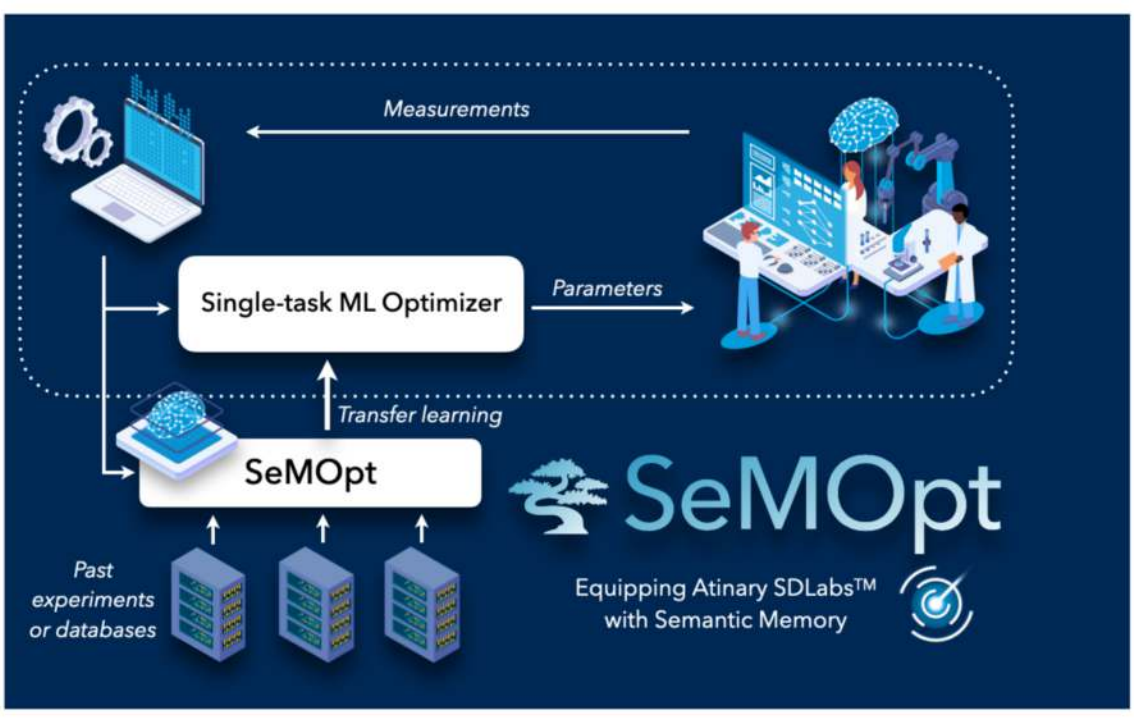
The GPyOpt authors. (2016). GPyOpt: A Bayesian Optimization framework in Python.

Liaw, R., Liang, E., Nishihara, R., Moritz, P., Gonzalez, J.E., & Stoica, I. (2018). Tune: A Research Platform for Distributed Model Selection and Training. *ArXiv, abs/1807.05118*.

Balandat, M., Karrer, B., Jiang, D.R., Daulton, S., Letham, B., Wilson, A.G., & Bakshy, E. (2020). BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization. *NeurIPS*.

Atinary Technologies Releases New Transfer Learning Algorithm for Bayesian Optimization (SeMOpt)

Published the 10 may 2022



Atinary Technologies Inc.
542 followers

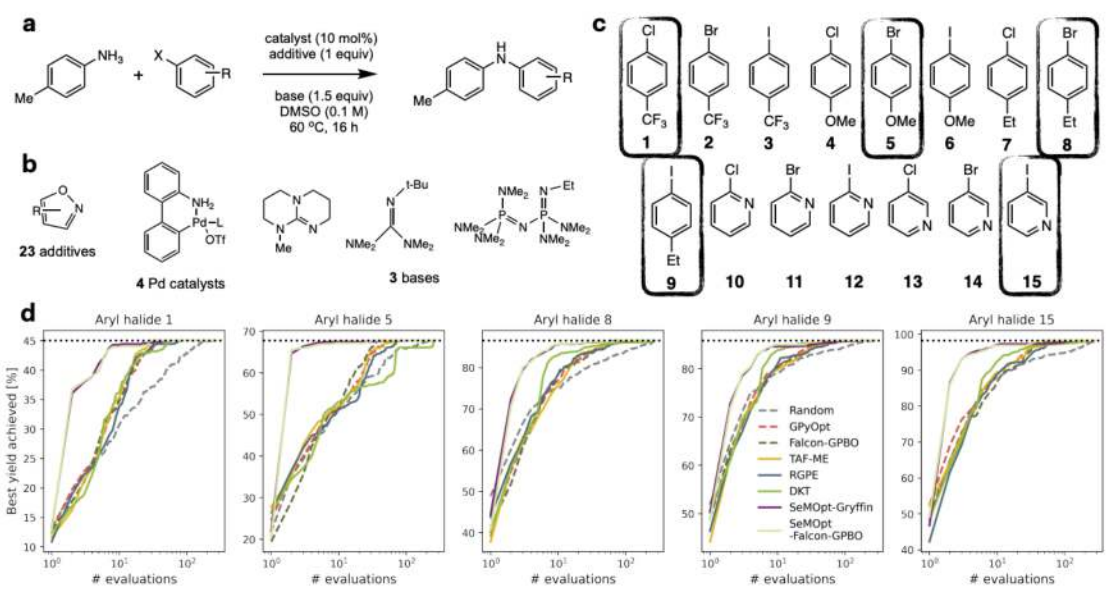
✓ Following

We are excited to announce Atinary's new ML algorithm and research paper on Semantic Memory Enhanced Optimization (SeMOpt). SeMOpt is Atinary's transfer-learning algorithm to accelerate optimization processes by leveraging existing knowledge learned from past experiments and existing databases. It automatically identifies key information and guides the optimization algorithms to the most promising parameter region from the start. SeMOpt works hand-in-hand with any general-purpose Bayesian optimization ML algorithms.

SeMOpt strengthens Atinary offering of data-driven ML solutions to further accelerate R&D of new molecules and materials. SeMOpt overcomes limitations of ML optimization algorithms by using state-of-the-art techniques from meta-learning and few-shot learning. These techniques implicitly learn inductive biases that strongly resemble concepts in experimental science. This learned intuition is then used to strengthen and guide optimization processes on novel tasks.

We expect SeMOpt to provide opportunities to accelerate R&D across the experimental sciences. Sample applications of SeMOpt accelerated include the following optimization challenges:

- parameters for accurate dispensing of liquids and solids by robotic platforms;
- chemical reactions with novel (unstudied) catalysts and/or ligands and/or substrates with access to yield measurements from related reactions;
- design of drug delivery systems for novel pharmaceutical compounds.



The paper presents SeMOpt's performance on the optimization of five simulated cross-coupling reactions, as well as on the palladium-catalyzed Buchwald-Hartwig cross-coupling of aryl halides with 4-methylaniline. These chemical reactions are widely used in organic synthesis, with application across industrial sectors, including pharma and biotech.

SeMOpt accelerates the optimization rate by a factor of 10 or more compared to standard single-task ML optimizers (those without transfer learning capabilities). Moreover, the case studies presented show that SeMOpt outperforms several existing ML Bayesian optimization strategies that leverage historical data.

Our new algorithm is a valuable technical contribution for general-purpose optimization and provides further support of the benefits from replacing the traditional trial-and-error experimentation process with Atinary Self-driving Labs™ technology, now powered with semantic memory.