

# Learning to optimize with a data-driven framework

Martin van der Schelling ([m.p.vanderschelling@tudelft.nl](mailto:m.p.vanderschelling@tudelft.nl))<sup>1</sup>, Deepesh Toshniwal ([D.Toshniwal@tudelft.nl](mailto:D.Toshniwal@tudelft.nl))<sup>2</sup>, and Miguel A. Bessa ([miguel\\_bessa@brown.edu](mailto:miguel_bessa@brown.edu))<sup>3</sup>

<sup>1</sup>Doctoral Researcher at Delft University of Technology, The Netherlands

<sup>2</sup>Assistant Professor at Delft University of Technology, The Netherlands

<sup>3</sup>Associate Professor at Brown University, USA

IFEEMCS520100 Fundamentals of Artificial Intelligence — August 29, 2022

**Objective:** Use machine learning to choose optimizers for an unseen problem – "learning to optimize".

**General Instructions:** Each group has to deliver **one PDF report** and a **Git repository**. The code should be easy to read, properly commented and it should be possible to replicate the results of the report.

## Introduction

Conventional materials design is based on experimental trial-and-error. However, just like solving an exam by trial-and-error requires many attempts, this is an impractical strategy because an overwhelming number of combinations become possible as the design space increases. One way to address this issue is to create much more data by using predictive physics-based models<sup>1</sup>, but simulating each design can be slow because these computer analysis need to be accurate. This remains true even considering the ever-increasing computational resources.

What if we had a more "intelligent" approach to design? In the past decade, machine learning techniques have soared to prominence in many applied fields of science and engineering. These techniques are able to discover patterns in input data, reducing the design space and informing the optimization process. The new generation of data-driven material design is here where data from predictive models and experiments can be used by machine learned models to discover new things. Once a model is trained, evaluating a material property using the model is significantly faster than measuring the property in an experiment or computing it in a simulation [1, 2].

## F3DASM framework

The framework for data-driven design and analysis of structures and materials (**F3DASM**) is an attempt to develop a systematic approach of inverting the material design process [3]. The framework integrates the following fields:

- Design of experiments (DoE), where input variables describing the microstructure, properties and external conditions of the system to be evaluated are sampled and where the search space is determined.
- Data generation, usually by computational analysis, where a material response database is created.
- Machine learning and optimization, where we either train a surrogate model to fit our experimental findings or iteratively improve the model to obtain a new design.

In this assignment the data generation part is very simple, as we will not be using any numerical method<sup>2</sup>. Instead, we will focus on the optimization part of the framework. The package is written in Python and it is still in the early days of development.

## Optimization

Optimization is a critical part of the data-driven process. The goal is to find the best parameters for some objective defined by the objective function  $f(\mathbf{x})$  that depends on an input vector  $\mathbf{x} = (x_1, \dots, x_d)$  with  $d$  variables and, in some cases, subjected to some equality or inequality constraints that define the feasible solution space  $\mathcal{X}$ . Each point in the solution space could, for instance, represent a micro-structure arrangement of a material.

Evaluating a sample of the objective function will result in an output vector  $\mathbf{y}$ . In analogy to the micro-structure input, this can represent some mechanical properties. For this project, we only consider single-objective optimization, i.e. where  $\mathbf{y}$  is just a scalar. Note that generally we don't know the underlying objective function. We can only get information by sampling from it.

<sup>1</sup>Such as finite element analyses, molecular dynamics, density functional theory, etc.

<sup>2</sup>You will not need to setup nor run computer simulations to create new data.

Trying all possibilities to find the optimum is not really feasible, as it requires a lot of (computational) resources and time. Instead, we try to iteratively improve the current solution with an optimization algorithm.

For each iteration  $t$ , the current solution  $\mathbf{x}_t$  is altered to acquire a new solution  $\mathbf{x}_{t+1}$ . Ultimately, we want to find an optimal set of values  $\mathbf{x}^*$  that will minimize or maximize the objective function  $f(\mathbf{x})$ . We can express a minimization optimization problem with the following formulation:

$$f(\mathbf{x}^*) \leq f(\mathbf{x}) \quad \forall \mathbf{x} \in \mathcal{X} \quad (1)$$

A minimization in  $f(\mathbf{x})$  is the same as a maximization in  $-f(\mathbf{x})$ . For the sake of consistency, we will view the optimization as a minimization problem for  $f$ .

## Learning to optimize

When solving a new optimization problem, we do not know what is the best optimization algorithm and its parameters that would quickly get us to the optimum. Unfortunately, the "one optimizer to rule them all" does not exist. According to the famous paper of Wolpert et al. [4], algorithms tend to exploit certain problem-specific characteristics, and if this particular feature is not present in the optimization problem then the performance of the optimizer will be worse than random search, i.e. random sampling (random trial-and-error!).

However, what if we could train a machine learning model on many different optimization problems considering many different optimizers? The machine learning algorithm would start to learn when an optimizer should be used or not. If it is trained properly, then the next time we consider a new problem this algorithm could make an informed decision on which optimizer to use at a given iteration! This is a new field of research called "learning to optimize".

Adequate training so that a machine learning algorithm can "learn to optimize" is challenging because being prepared for any kind of optimization problem requires a diverse set of algorithms to mix and match and a diverse set of optimization problems to learn from.

In the masters thesis of van der Schelling [5], different possibilities of implementing a general optimizer that can adapt to problem-specific features of the input data have been investigated. The thesis proposed a data-driven framework for general optimization problems to adapt the algorithm during optimization. With an adaptive data-driven optimization strategy, we eliminate the need to hand-design a machine learning model. By integrating this idea into the F3DASM framework, the proposed software package does not require any machine learning architectural choices beforehand. This will increase accessibility of the framework to pure experimentalists.

## Aim of the project

The aim of the project is to design a general optimization strategy with machine learning that competes with the conventional hand-engineered optimizers. You will investigate how different optimization algorithms behave on different problem-specific characteristics.

You will use the F3DASM framework to set-up your search space, sampling, objective functions and optimizers. You will assess the optimizers on various benchmarks, and come up with fair performance metrics.

At the end of the project you will gain:

1. Knowledge on different optimizers and why they do or do not work.
2. Understand the key challenges of selecting appropriate machine learning models for applications.

## Preparations

- In order to assist you effectively during the project, please install the following:
  - Create a [GitHub](#) account
  - Install GitHub Desktop for [Windows/Mac](#) or [Linux Ubuntu](#).
  - Follow the F3DASM installation instructions from the documentation found [here](#).
  - Create a new GitHub repository; this will be the place where you store your code.
  - Add Martin van der Schelling as a collaborator (Github username is `mpvanderschelling`).

- Please read the following material:
  - To understand the problem and ambiguity of optimization, read the literature review of this thesis [5], which can be downloaded [here](#).<sup>3</sup>
  - To understand the idea of the framework: read the original article on the F3DASM framework [3].
  - To see an implementation of machine learning and optimization into materials design, read the 'fragile becomes supercompressible' article [6]. You can also check out this short video about it [here](#)!

## Getting started

To get familiar with the F3DASM package, try to replicate the following experiments:

1. We start of by creating and sampling from a design space:
  - 1.1. Create a 2D dimensional continuous design space with a single-objective output. Set the boundaries for all of the parameters from -1.0 to 1.0.
  - 1.2. Sample 50 points from this design space with the `RandomUniformSampling` sampler. Plot these samples for visualization.
  - 1.3. Repeat 2. for the `SobolSequenceSampling` and `LatinHyperCubeSampling`. What is the difference?
  - 1.4. Sampling is important for initialization of optimization algorithms. Can you explain why?
2. Next we instantiate several objective functions:
  - 2.1. Choose a function that is compatible with a 2D input space from each of the following categories found in the [objective function list](#) at the documentation:
    - Convex functions
    - Separable function
    - Differentiable functions
    - Multimodal functions
  - 2.2. Plot a 3D visualization of the loss-landscape. Use the same boundary values from 1.1.
  - 2.3. Create a contour plot (a top-down view of the loss-landscapes with its contours) for each of the functions. Indicate the global minimum on the plots.
  - 2.4. Create the same heatmap of the loss-landscape but with the sampling points from one of the samplers used in exercise 1. Tip: you may want to use the `plot_data()` function!
3. Now we are ready to optimize our objective functions:
  - 3.1. Select two of the optimizers from the [optimizer list](#) in the documentation and shortly explain for each of them:
    - How does the update step of the optimizer work?
    - What are the prerequisites of the optimizer? In other words; are there any constraints on how the optimizer can be used?
    - When is this optimizer effective and when is it not?

Support your answers with scientific references.

- 3.2. Create an instance for each of the optimizers picked in 3.1 with the default hyper-parameters and the data from one of the samplers. Set the seed so that the experiment can be replicated later.
- 3.3. Run the optimization with the `run_optimization()` function considering 500 iterations for each of the chosen functions.
- 3.4. Create a plot showing the objective function value on the vertical axes and the iteration number on the horizontal axes for each of the optimization trials.

---

<sup>3</sup>You can skip the part about bio-based composites, although it could help you understand the relevance of machine learning for materials design.

4. The initial conditions of the optimization process can heavily influence the performance. Therefore it is important to redo the experiment with different initial conditions. These are called realizations.
  - 4.1. Run the same optimization of 3.3 with the `run_multiple_realizations()` function for 500 iterations with 10 different realizations.
  - 4.2. Create a plot showing the mean objective function value on the vertical axes and the iteration number on the horizontal axes for each of the different optimizer trials.
  - 4.3. Explain the results of your optimization:
    - Which optimizer is better on each of the objective functions and why do you think this is?
    - Do the results comply with your findings in scientific literature?
5. A big thing in improving the optimizer is to tune the hyper-parameters. We have been using the default settings but maybe we can do better by changing them!
  - 5.1. Choose one of the objective functions of the results created in 4.2 and select the worst-performing optimizer.
  - 5.2. Choose one of the hyper-parameters of this optimizer and explain what it controls.
  - 5.3. Create different instances of the selected optimizer with a different value of the selected hyperparameter
  - 5.4. Run the same optimization process as in 4.2 for the different hyperparameter settings.
  - 5.5. Plot your results and try to improve the optimizer!

## Creating a performance metric

stub

## Create your own a smart optimization strategy

stub

## References

- [1] Sanket Kadulkar, Zachary M. Sherman, Venkat Ganesan, and Thomas M. Truskett. Machine Learning-Assisted Design of Material Properties. *Annual Review of Chemical and Biomolecular Engineering*, 13(1), mar 2022.
- [2] Kai Guo, Zhenze Yang, Chi-Hua Yu, and Markus J. Buehler. Artificial intelligence and machine learning in design of mechanical materials. *Materials Horizons*, 8(4):1153–1172, 2021.
- [3] M. A. Bessa, R. Bostanabad, Z. Liu, A. Hu, Daniel W. Apley, C. Brinson, W. Chen, and Wing Kam Liu. A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality. *Computer Methods in Applied Mechanics and Engineering*, 320(April):633–667, jun 2017.
- [4] David H. Wolpert and William G. Macready. No free lunch theorems for optimization. *IEEE transactions on evolutionary computation*, 1(1):67–82, 1997.
- [5] Martin van der Schelling. A data-driven heuristic decision strategy for data-scarce optimization: with an application towards bio-based composites. mathesis, Delft University of Technology, March 2021.
- [6] Miguel A. Bessa, Piotr Glowacki, and Michael Houlder. Bayesian machine learning in metamaterial design: Fragile becomes supercompressible. *Advanced Materials*, 31(48):1904845, 2019.