Knowledge-Constrained Machine Learning:

A strategy for producing predictive process models in the absence of mechanistic understanding and large data sets

Daniel Griffin, Behnam Partopour, and Seth Huggins, Amgen Inc.

Abstract. This work explores the possibility of efficiently learning predictive process models from *small or sparse* data sets by blending domain knowledge, or understanding, into a machine learning strategy. We propose a strategy that we refer to as knowledge-constrained machine learning, in which process understanding is incorporated by way of mathematical constraints written into the learning algorithm. To make the proposed methodology concrete, we take a chemical reaction as our example process and explore the development of models that predict the reactant, product and impurity concentration profiles over time from select design variables. Using a fixed set of experimental data for a hydrogenation reaction in the synthetic route to AMG 986—a molecule in Amgen's portfolio—we compare the predictive accuracy of models resulting from the proposed methodology with those constructed by a pure machine learning strategy (no knowledge constraints) and with those produced by the more-commonly-applied approaches of semi-empirical rate law modeling and fully-empirical response surface modeling. The key result we take away from the analysis is: knowledge-constrained machine learning can be used to establish predictive models in both the absence of a detailed mechanistic understanding and the absence of a uniform and large data set that spans the full design space; the same cannot be said for the pure machine learning, semi-empirical rate law, or response surface approaches.

Scope: Modeling Objective, Example Process, and Development Task. There are at least two common objectives that motivate process modeling: to refine one's conceptual understanding, and (or) to build a *function*—mathematical or computational—that enables the quantitative prediction of process outcomes under different scenarios. The choice of modeling strategy should be informed by the dominant objective; in this work, we focus on process modeling with the objective of building a function that enables accurate predictions. And, more specifically, we explore strategies for creating domain-specific process models that predict process outcomes in a narrow domain—that is, over a low-dimensional and continuous design space. The example process we examine is a batch chemical reaction, and the development task in mind is reaction optimization. Mathematically expressed, the goal of the modeling task we consider is to identify a function, f, that accurately predicts the process outcome, $\boldsymbol{y}(t)$, for given values of the design variables, $\boldsymbol{x}$, over a fixed design space $\mathcal{X}$.

Goal. Identify a function, f, that accurately captures the mapping $\boldsymbol{x} \mapsto \boldsymbol{y}(t)$ across $\boldsymbol{x} \in \mathcal{X}$.

Example Reaction and Sources of Information. To aid in our exploration and discussion of modeling strategies for the above task, we use a hydrogenation reaction and consider two design variables: temperature and catalyst loading. With coded molecular structures, the reaction we consider is written as follows:

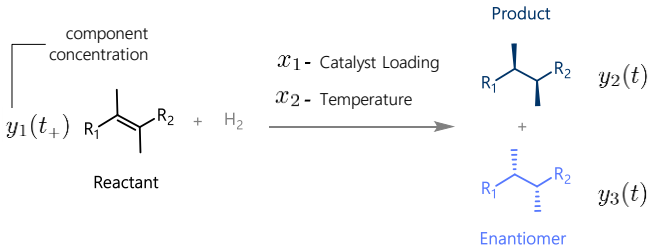

Figure 1. Example Reaction. *Coded representation of the hydrogenation reaction used as an example in this work.*

Taking this example, the modeling objective is to identify a function that predicts the concentration profiles for the reactant, $y_1(t_+)$, product, $y_2(t)$, and the undesired enantiomer, $y_3(t)$, as a function of the selected amount of catalyst, x_1, and the temperature, x_2. We explore creating this function applying different strategies that can pull from two sources of information: our understanding of the chemical reaction and experimental data collected under relevant conditions. The data we use consists of discrete time profiles for the reactant, product and enantiomer concentrations observed when the reaction was run at various points across the temperature-catalyst loading design space. For analysis, the data is broken into Training and Test sets.

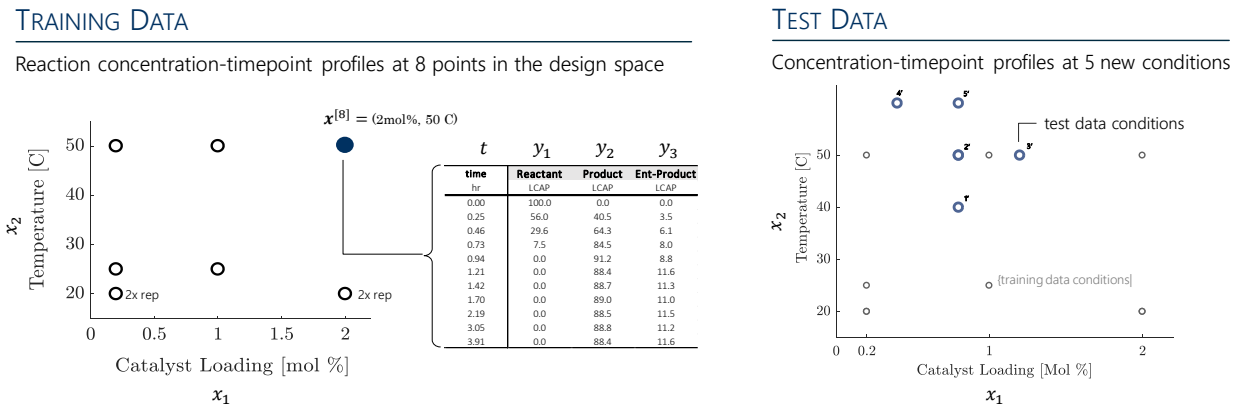

Figure 2. Training and Test data sets. *Training data consists of discrete concentration-time profiles collected at eight temperature-loading conditions; the Training data is not evenly distributed and does not span the full range of the possible design space $\mathcal{X}$. Test data also consists of discrete concentration-time profiles collected at a later time, at five new points in the design space; some of these points fall within the domain of the Training data, while others fall outside of that domain.*

Methodology: Knowledge-Constrained Machine Learning. Strategies for creating reaction kinetic models tend to rely on either a detailed understanding of the reaction mechanism (first-principles and semi-empirical rate law modeling), or on having uniformly-distributed and accurate data that spans the design space of interest (response surface modeling). A core motivation behind the present work is to create a strategy that does not require either of these—that is, a modeling strategy that can work in *the absence* of a detailed mechanistic understanding and in *the absence* of uniformly-distributed and accurate data that spans the full design space. For this, we are proposing a knowledge-constrained machine learning methodology, that operates as follows. Starting with a set of training data we construct a "blank-slate" supervised learning algorithm for selecting the function f that *best fits* the mapping $\boldsymbol{x} \mapsto \boldsymbol{y}(t)$ across the data available. The blank-slate supervised learning algorithm we have chosen uses least-squares regression to first interpolate between time points, fitting a 10^{th}-order orthogonal polynomial bases set to the discrete concentration-time data. And then uses non-parametric, weighted least-squares regression to interpolate and extrapolate across the $\boldsymbol{x}$-domain, at any fixed time. This blank-slate supervised learning algorithm can be seen as fitting a very flexible hypothesis set of functions (could

reasonably represent *any* continuous function of $\boldsymbol{x}$ and $\boldsymbol{t}$). As such, it is highly susceptible to overfitting, and unlikely to produce a predictive model when applied to the limited training data set that we have in our example. To go from such strategy to one with a better chance of success, we take advantage of the general knowledge or expectations for reaction kinetics to impose constraints on the model *a priori*. For example, in the given hydrogenation, we expect the reaction to be irreversible—and the reactant concentration to decrease over time, while the product and enantiomer concentrations increase over time. We also expect the reaction to *end* eventually—that is, the concentrations should reach a steady state. Each of these expectations, along with any physical constraint, can be written as an inequality imposed on the function or an inequality imposed on a derivative of the function. That set of inequalities can then be added to the blank-slate supervised learning algorithm to create a knowledge-constrained fit. By this approach, the algorithm now searches for the best fit among functions *that also meet each of the knowledge constraints imposed*. The global learning algorithm can be written as a convex optimization problem and solved. In this work we use CVX, software for convex optimization, in Matlab.

Key Results. Applying the knowledge-constrained machine learning algorithm described to the given example, we find that the methodology produces a model that predicts accurately over the Test data set and could be used for accurate reaction optimization over the design space. Comparatively: the pure or blank-slate machine learning strategy overfits the training data and predicts poorly; the semi-empirical rate law model we construct is a simplification of the true underlying mechanism (which is not fully known) and fits the training data well, but shows a notable bias in predicting the reaction outcome at conditions outside the training data set; and, finally, the response surface model that we use fits the data reasonably well and is easy to construct, but also results biased predictions that are both inaccurate and, in some instances, non-physical.

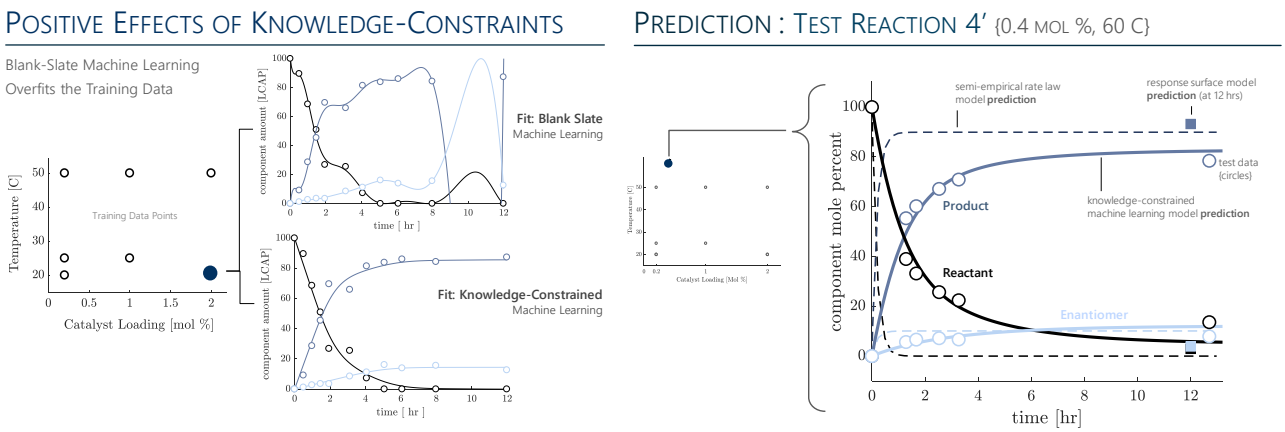

Figure 3. Select Results. *Applying a blank-slate machine learning algorithm results in a function that overfits the training data; the proposed knowledge-constraints are effective in mitigating this type of overfitting. In predictions across the Test set, the model produced by knowledge-constrained machine learning performs well; comparatively, clear biases are seen in the predictions made by the semi-empirical rate law and response surface models.*

Conclusion. The idea of knowledge-constrained machine learning holds promise for producing predictive process models in the absence of detailed mechanistic understanding and in the absence of large and accurate experimental data sets that span the full design space. We therefore expect the strategy to be of interest to broader community of Process Analytics and Machine Learning researchers. Our poster will expand on the strategy and detail the modeling results for the given example reaction.