Uncertainty and How it Impacts Modeling, Optimization, and Metrics

There are a wide variety of machine learning optimization challenges that are well-researched in academia. By deeply exploring each particular challenge, this academic research creates a robust foundation for applied research situations. But applying this research in the real world also typically involves the consideration of tradeoffs between each of these challenges that cannot be assumed away. As a team responsible for helping customers solve real-world optimization problems, SigOpt Research spends considerable time weighing these tradeoffs. We hope this work serves to build upon much of the leading academic research in optimization. There are a number of examples of this type of research:

- **Discrete variables**: Flexibility in features, models, and parameters
- **Failure regions:** Identifying and avoiding configurations that fail to yield outcomes
- **Constraints:** A priori limits on the configuration space
- **Multimetric:** Solving for competing objectives
- **Uncertainty:** Addressing noise and its impact on models

This research note focuses on uncertainty. In a three-part blog series, we share some of the more compelling research on uncertainty and how it impacts applied machine learning optimization techniques. Given how frequently noise interferes with a real-world machine learning problem, effectively addressing it can be the difference between whether a situation can or cannot be effectively modeled.

1: Modeling with Uncertainty

Uncertainty can appear intrinsically in a given situation or as a byproduct of measuring that situation. In these scenarios, the model must learn from the data without being tricked or confused by the noise. Modeling in the presence of uncertainty and noise is hard. A good model respects the data but does not overfit, producing useful predictions. In recognition of this challenge, the ML community has developed a variety of techniques, such as the regularization of models and prevention of overfitting, that serve as useful solutions.

2: Bayesian Optimization with Uncertainty

If modeling with noisy observations is complicated, than optimization in the same conditions is even more challenging in that it introduces an additional "search" component. An efficient, model-based optimization technique must both evaluate the objective through this noise and take noise into consideration when efficiently searching for the optima. In fact, Random Search is often severely crippled by nontrivial noise. Bayesian optimization experts, however, have spent a decade developing strategies that perform well in this noisy context.

3: Balancing Multiple Metrics with Uncertainty

Solving for multiple metrics adds a third component. Multimetric optimization efficiently searches a space in which improvements in one metric necessitate tradeoffs in the other. In the presence of uncertainty, understanding this tradeoff is complicated, because the tradeoff itself becomes obscured by noise. Strategies for dealing with this, such as a data smoothing process that creates a more accurate sense of and probabilistic estimate for efficiency, will improve outcomes.

SIGOPT

sigopt.com
sales@sigopt.com

More on SigOpt

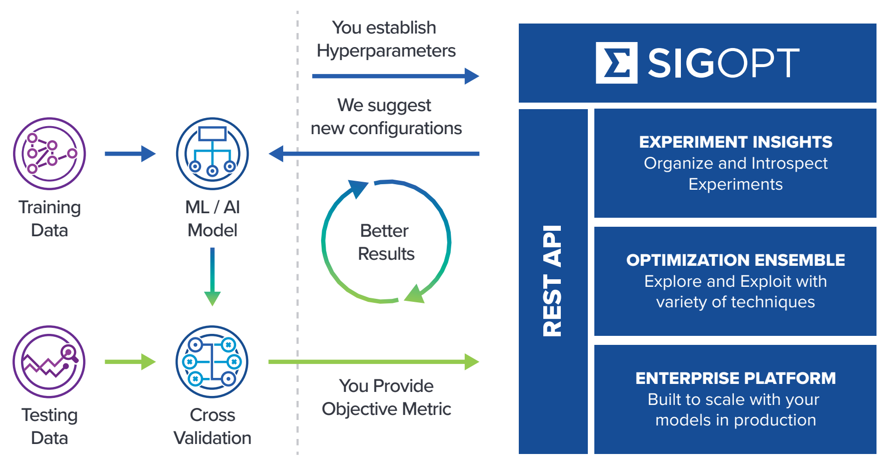
Overview of the solution that supports Multitask Optimization

SigOpt automates tuning of any model through a process described in more detail below. This drives significant value for our customers in the form of:

1. **Expert productivity:** Automated optimization reduces the time experts spend on tasks like hyperparameter tuning that do not benefit from their expertise
2. **Wall-clock time savings:** Bayesian techniques minimize the time it takes to train and tune any model, compressing the timeline from development to production
3. **Compute efficiency:** Efficient exploration and learning of the problem space maximizes the use of expensive computing, which directly contributes to bottom line savings
4. **Sustained model performance:** SigOpt tunes models that would otherwise be too expensive to optimize, and ensures they are consistently re-tuned in production

This combination empowers teams to apply optimize earlier in the development process, and more frequently retune models in production. The result is a greater return on any machine learning, deep learning, or other data science investment.

How SigOpt's optimization solution works



SIGOPT

sigopt.com
sales@sigopt.com