Best Practices for Scaling Modeling Platforms

Matt Greenwood, CIO, Two Sigma
Scott Clark, CEO, SigOpt

O'Reilly AI Conference, New York
AI Business Summit, Case Studies
Technology & Data

Investment Management

Research & the Scientific Method
$300B+
Assets Under Management*

*Current SigOpt trading customers represent over $300B in assets under management
Halite
Halite is our artificial intelligence programming challenge, organized annually around an original multi-agent game. The current competition runs through January 2019.
Learn More

Waiter
Waiter is an on-demand, web service platform that launches, manages, and automatically scales services running on Mesos.
Learn More

Cook
Cook is our dynamic fair job scheduler on Mesos for batch workloads and Spark.
Learn More

Flint
Flint is our project that provides time-series analysis in Spark.
Learn More

git-meta
Our project that lets you effectively manage many repositories with git submodules.
Learn More

BeakerX
BeakerX is a collection of kernels and extensions to the Jupyter interactive computing environment. It provides JVM support, interactive plots, tables, forms, publishing, and more.
Learn More

Marbles
Marbles is a Python unittest extension that allows test authors to write richer tests that expose more information on test failure.
Learn More

Webtau
Webtau (WEB Test AUtomation) is a concise and expressive way to create REST API and Web UI tests with Groovy or Java DSL.
Learn More
Data Management

Pre-Processing Pipelines
Feature Eng

Solutions: Evolving
Innovation: Incremental
Implication: Mixed

Modeling: Notebooks, Libraries, Frameworks
Solutions: Evolving  |  Innovation: Existential  |  Implication: Build

Experimentation: Tracking, Analysis, Optimization, Resource Mgmt
Solutions: Standard  |  Innovation: Incremental  |  Implication: Buy

Hardware: Scalable, Efficient Compute
Solutions: Evolving  |  Innovation: Existential  |  Implication: Build

Experimentation, Training, Evaluation

Serving Monitoring
Solutions: Evolving
Innovation: Existential
Implication: Build

Model Productionalization
Solution: Rapidly prototype models with SigOpt

**Optimization Engine**
Automate parameter tuning to maximize the impact of your models with our optimization ensemble

**Experiment Insights**
Track, analyze and reproduce any model to improve the productivity of your modeling teams

**Enterprise Platform**
Standardize experimentation across any combination of library, infrastructure, model, or task

---

**Graph Example**
- AUC for Method A: 0.768
- AUC for Method B: 0.567

**Experiment Insights Example**
- **Summary**: Analysis
- **Suggestions**: History
- **API Docs**: Properties
- **Experimental**: Admin
- **Client Admin**: User Admin

**On-Premise**
- **Hybrid/Multi**
Hyperparameter Optimization

- Model Tuning
- Deep Learning Architecture Search
- Training & Tuning
- Hyperparameter Search
- Evolutionary Algorithms
- Grid Search
- Random Search
- Bayesian Optimization
- Deep Learning Architecture Search
<table>
<thead>
<tr>
<th>Method</th>
<th>Pro</th>
<th>Con</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual Search</td>
<td>Leverages expertise</td>
<td>Not scalable, inconsistent</td>
</tr>
<tr>
<td>Grid Search</td>
<td>Simple to implement</td>
<td>Not scalable, often infeasible</td>
</tr>
<tr>
<td>Random Search</td>
<td>Scalable</td>
<td>Inefficient</td>
</tr>
<tr>
<td>Evolutionary Algorithms</td>
<td>Effective at architecture search</td>
<td>Very resource intensive</td>
</tr>
<tr>
<td>Bayesian Optimization</td>
<td>Efficient, effective</td>
<td>Can be tough to parallelize</td>
</tr>
<tr>
<td>Method</td>
<td>Pro</td>
<td>Con</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Manual Search</td>
<td>Leverages expertise</td>
<td>Not scalable, inconsistent</td>
</tr>
<tr>
<td>Grid Search</td>
<td>Simple to implement</td>
<td>Not scalable, often infeasible</td>
</tr>
<tr>
<td>Method</td>
<td>Pro</td>
<td>Con</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Manual Search</td>
<td>Leverages expertise</td>
<td>Not scalable, inconsistent</td>
</tr>
<tr>
<td>Grid Search</td>
<td>Simple to implement</td>
<td>Not scalable, often infeasible</td>
</tr>
<tr>
<td>Random Search</td>
<td>Scalable</td>
<td>Inefficient</td>
</tr>
<tr>
<td>Method</td>
<td>Pro</td>
<td>Con</td>
</tr>
<tr>
<td>---------------------</td>
<td>-------------------------------------------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>Manual Search</td>
<td>Leverages expertise</td>
<td>Not scalable, inconsistent</td>
</tr>
<tr>
<td>Grid Search</td>
<td>Simple to implement</td>
<td>Not scalable, often infeasible</td>
</tr>
<tr>
<td>Random Search</td>
<td>Scalable</td>
<td>Inefficient</td>
</tr>
<tr>
<td>Evolutionary Algorithms</td>
<td>Effective at architecture search</td>
<td>Very resource intensive</td>
</tr>
<tr>
<td>Method</td>
<td>Pro</td>
<td>Con</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Manual Search</td>
<td>Leverages expertise</td>
<td>Not scalable, inconsistent</td>
</tr>
<tr>
<td>Grid Search</td>
<td>Simple to implement</td>
<td>Not scalable, often infeasible</td>
</tr>
<tr>
<td>Random Search</td>
<td>Scalable</td>
<td>Inefficient</td>
</tr>
<tr>
<td>Evolutionary Algorithms</td>
<td>Effective at architecture search</td>
<td>Very resource intensive</td>
</tr>
<tr>
<td>Bayesian Optimization</td>
<td>Efficient, effective</td>
<td>Can be tough to parallelize</td>
</tr>
</tbody>
</table>
Data and models stay private.

Iterative, automated optimization

- Training Data
- AI/ML Model
- Testing Data
- Model Evaluation

New Configurations

Better Results

Objective Metric

Built specifically for scalable enterprise use cases.

EXPERIMENT INSIGHTS
Organize and introspect experiments

ENTERPRISE PLATFORM
Built to scale with your models in production

OPTIMIZATION ENSEMBLE
Explore and exploit with a variety of techniques

REST API
It there a **proprietary advantage** to DIY?
Does it benefit from **domain expertise**?
Is the **process the same** for each model?
Is the open source **well maintained**?
Is the open source **reliable**?
Are there **more performant** alternatives?
Is there a **low maintenance burden** to buy?
Can the product **scale** with our needs?
Can the product **evolve** with our needs?
Benefits of implementing SigOpt

Realize Performance Gains
Address entirely new problems
Maximize Resource Utilization
90% fewer training runs to optimize
https://devblogs.nvidia.com/sigopt-deep-learning-hyperparameter-optimization/

400x faster time to optimize

20x the cost efficiency to optimize
Better Results, 8x Faster

“We’ve integrated SigOpt’s optimization service and are now able to get better results faster and cheaper than any solution we’ve seen before.”

Matt Adereth
Managing Director
Two Sigma
Failed observations
Constraints on the model
Noise in the data
Competing metrics
Lengthy training cycles
Distributed training
Benefit: Address New Problems

Balancing Competing Metrics
Agnostic
Any Infrastructure
Any Model
Any Library
Any Language

Reliable
Requires no data access
99.9% SLA
24/7 on-call support
>1M API events per hour

Complete
Experiment reproducibility
Modeling insights
Differentiated algorithms
Easy to scale in parallel
Asynchronous parallelization is critical for resource utilization.
SigOpt Impact

Exploration
Identify valuable modeling use cases

Scale
Build repeatable process for scaling use cases

Performance
Fine-tune repeatable process to maximize gain

Modeling Scale
Thank you

https://www.twosigma.com/careers/
https://www.twosigma.com/insights/
https://opensource.twosigma.com/

https://sigopt.com/company/careers/
https://sigopt.com/blog/
https://sigopt.com/research/
https://sigopt.com/try-it