
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

AIM324

Maximize ML model accuracy with advanced hyperparameter tuning strategies

Valerio Perrone

Applied Science Manager
AWS

Mariano Kamp

Principal Solutions Architect
AWS

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Chalk talk

- What is a chalk talk?
  - Less of a presentation
  - Not a lecture
  - More of a demo and whiteboarding
  - Usually about 30-60 people
  - 60 minutes



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

- What is hyperparameter tuning and why is it difficult?
- How to find the optimal hyperparameters?
- Benefits of tuning your ML models with Amazon SageMaker
- Amazon SageMaker Automatic Model Tuning demo
- Q&A

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What is hyperparameter tuning?



Decision trees

Tree depth
Max leaf nodes
Gamma
Eta
Lambda
Alpha
…

Neural networks

Number of layers
Hidden layer width
Learning rate
Embedding dimensions
Dropout
…

Arbitrary task

(data prep, download, inference…)

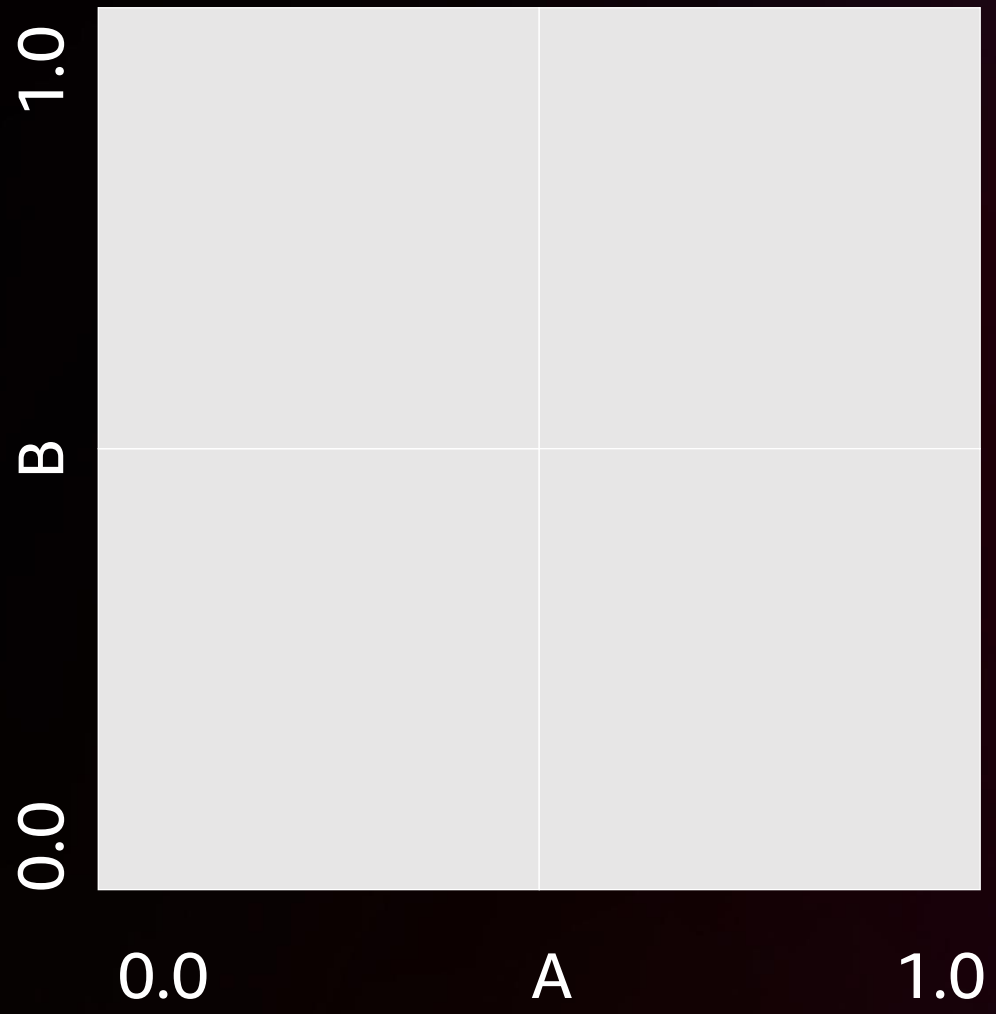
Parameter 1
Parameter 2
…

Hyperparameters

aws

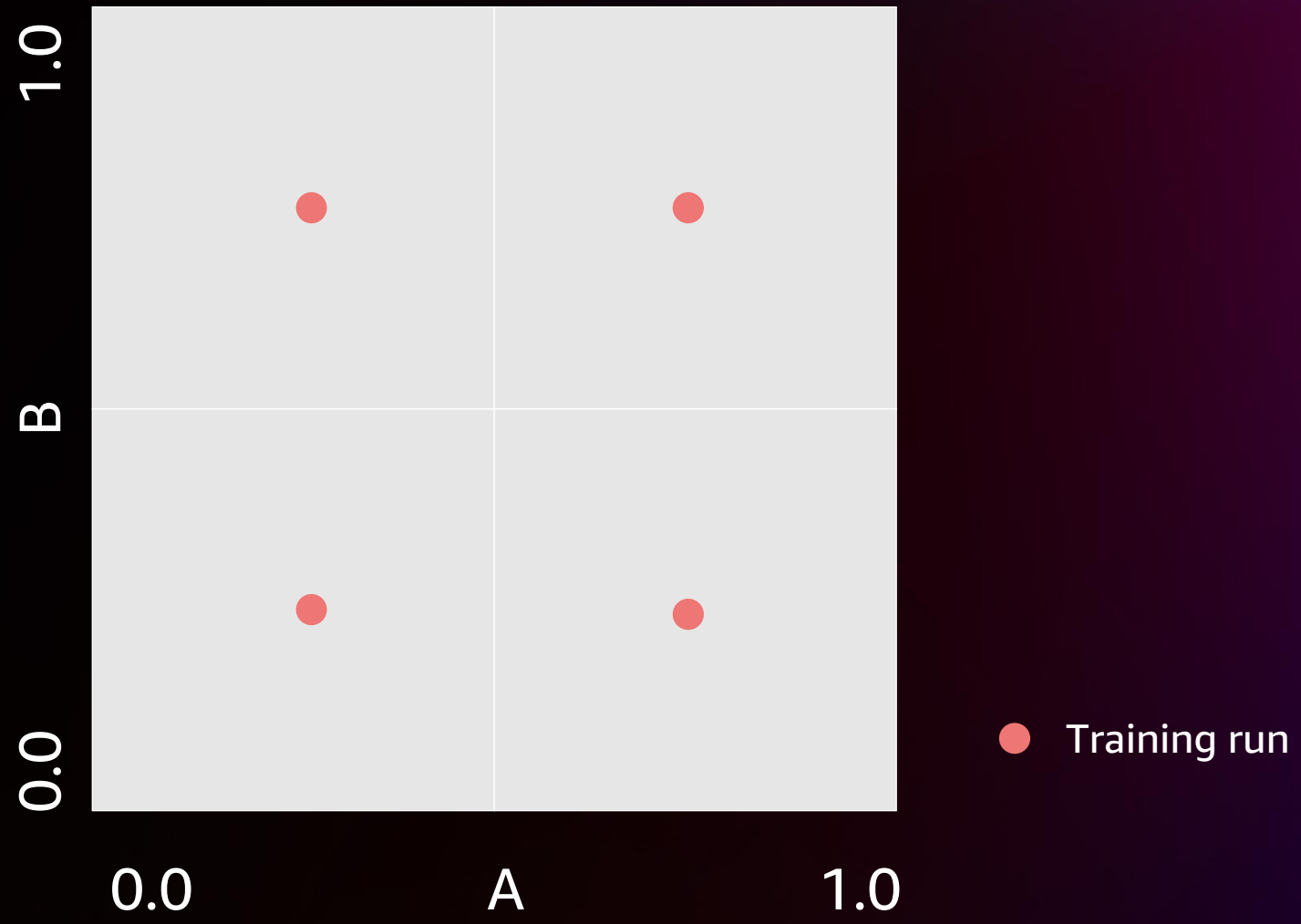
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How to find the optimal hyperparameters?



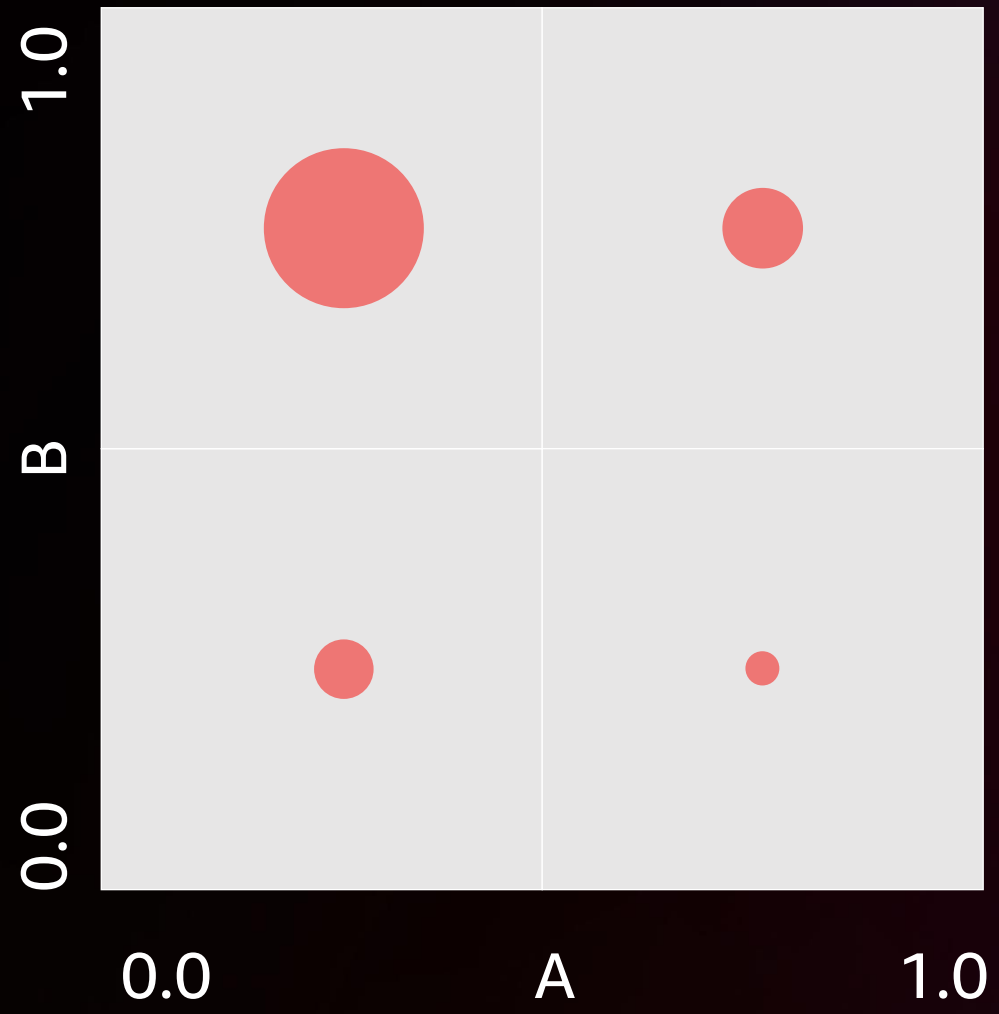
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Grid search



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Grid search



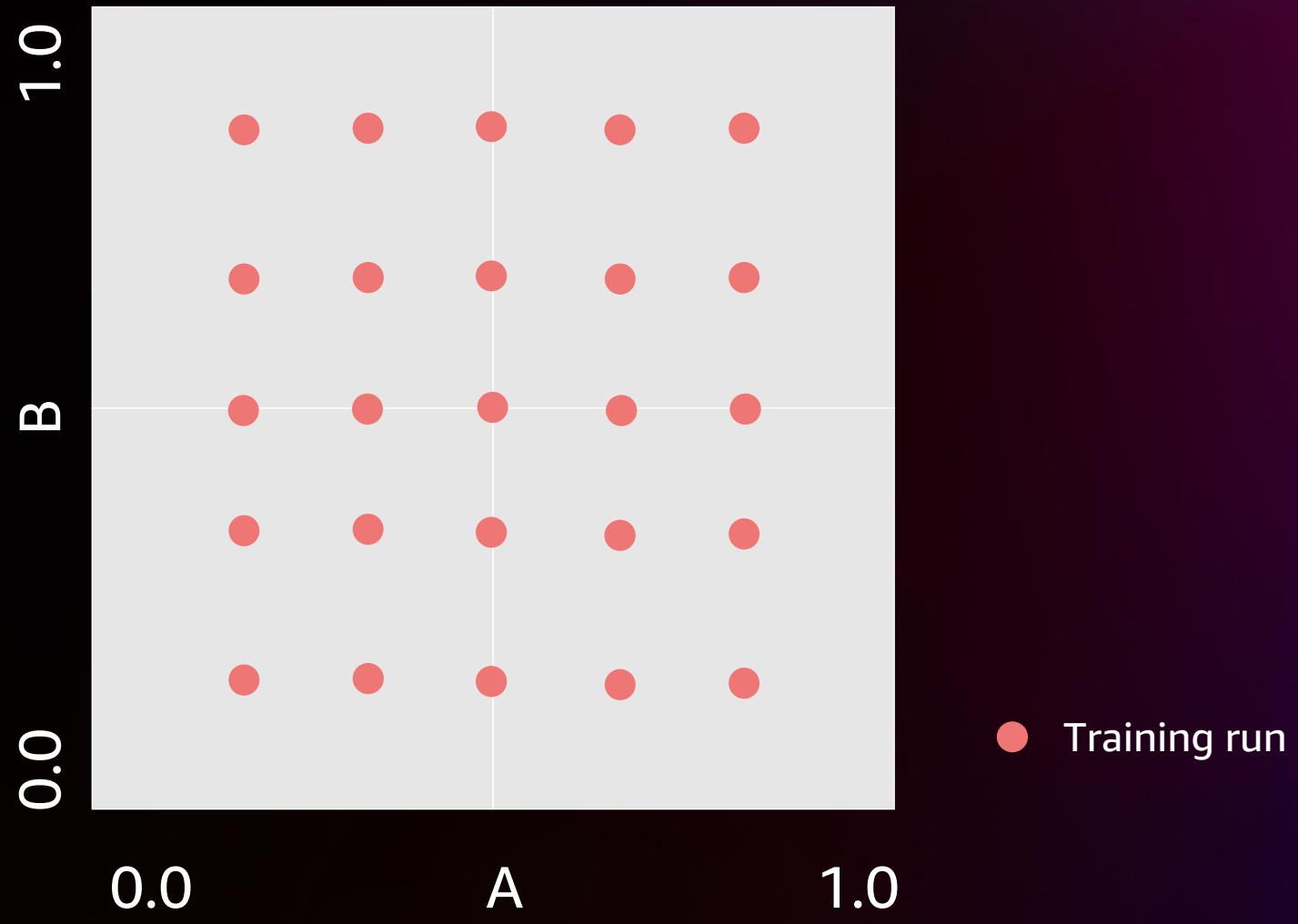
Loss function: measures how well your model predicts the expected outcome
(the lower, the better)

Example: validation error

Training run
(Diameter = Loss)

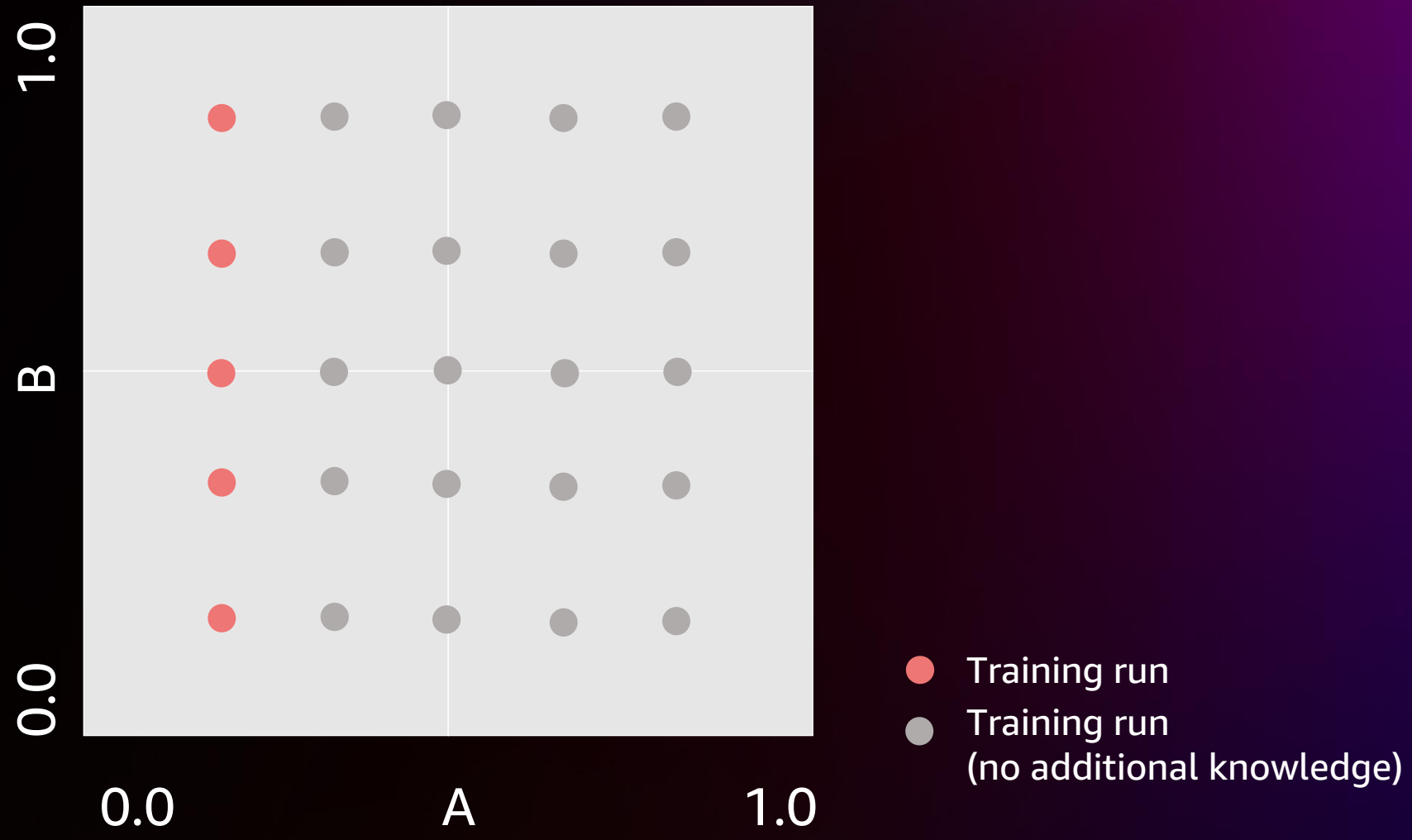
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Grid search



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Grid search



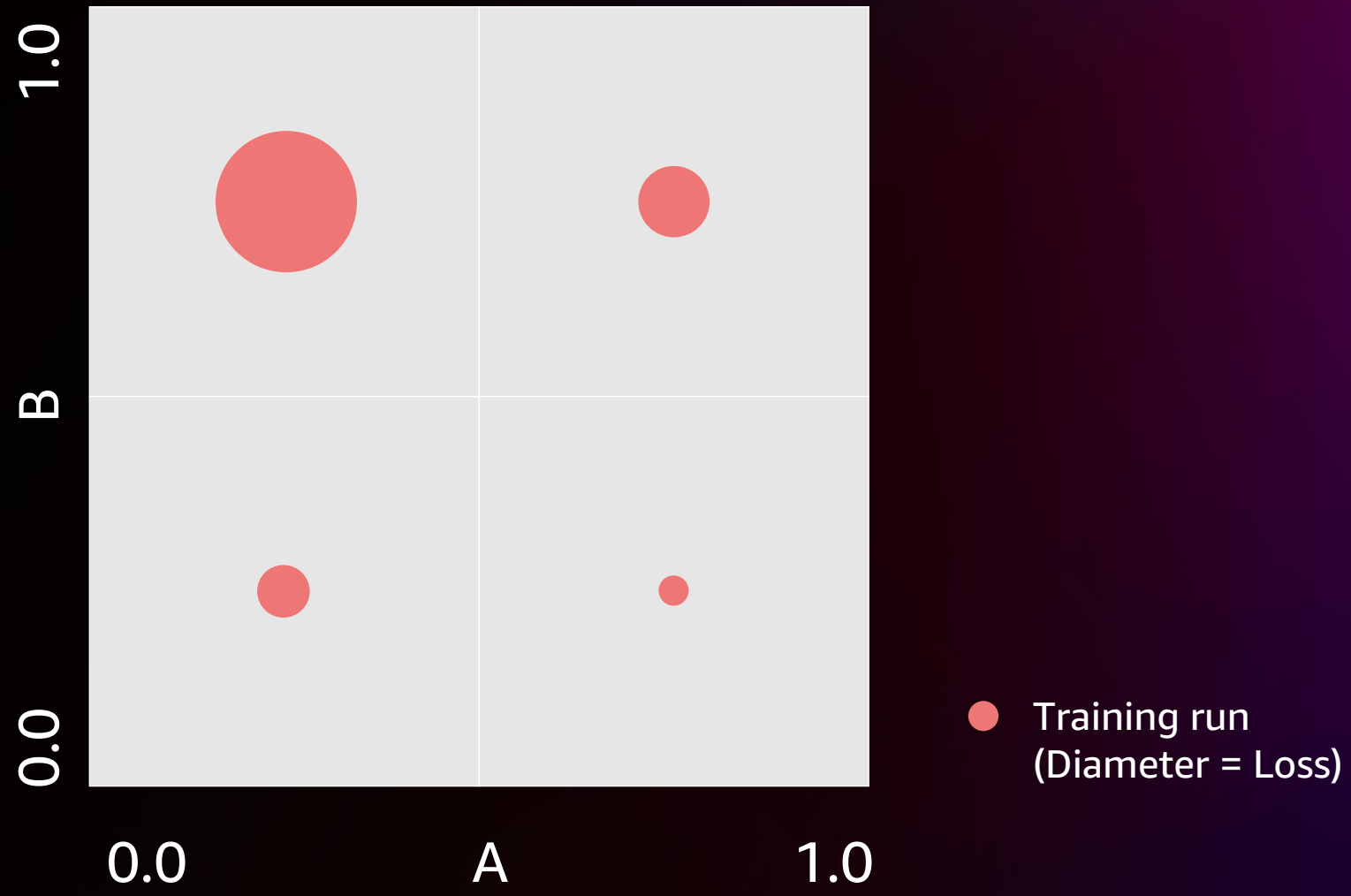
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Grid search



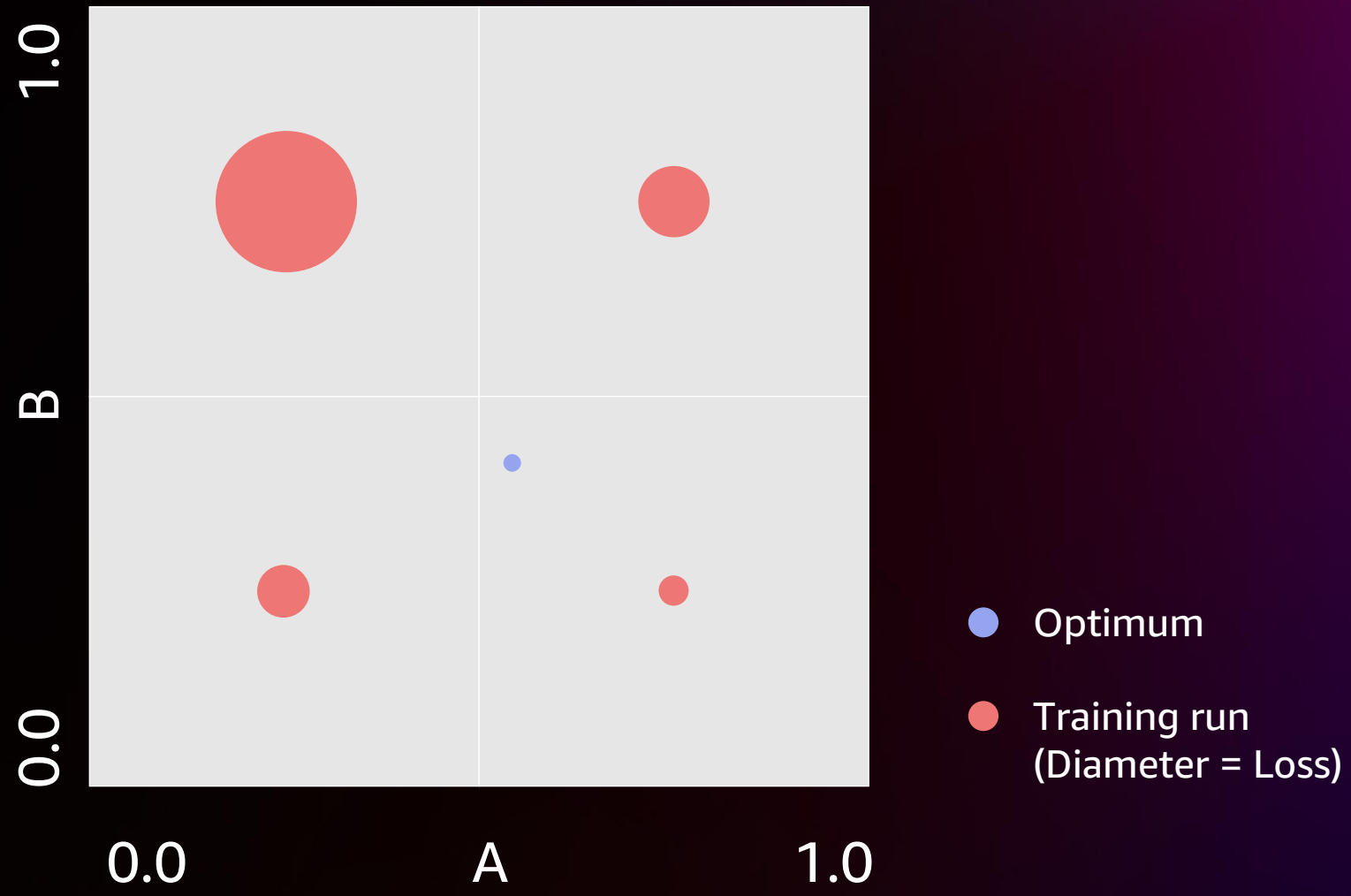
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Grid search



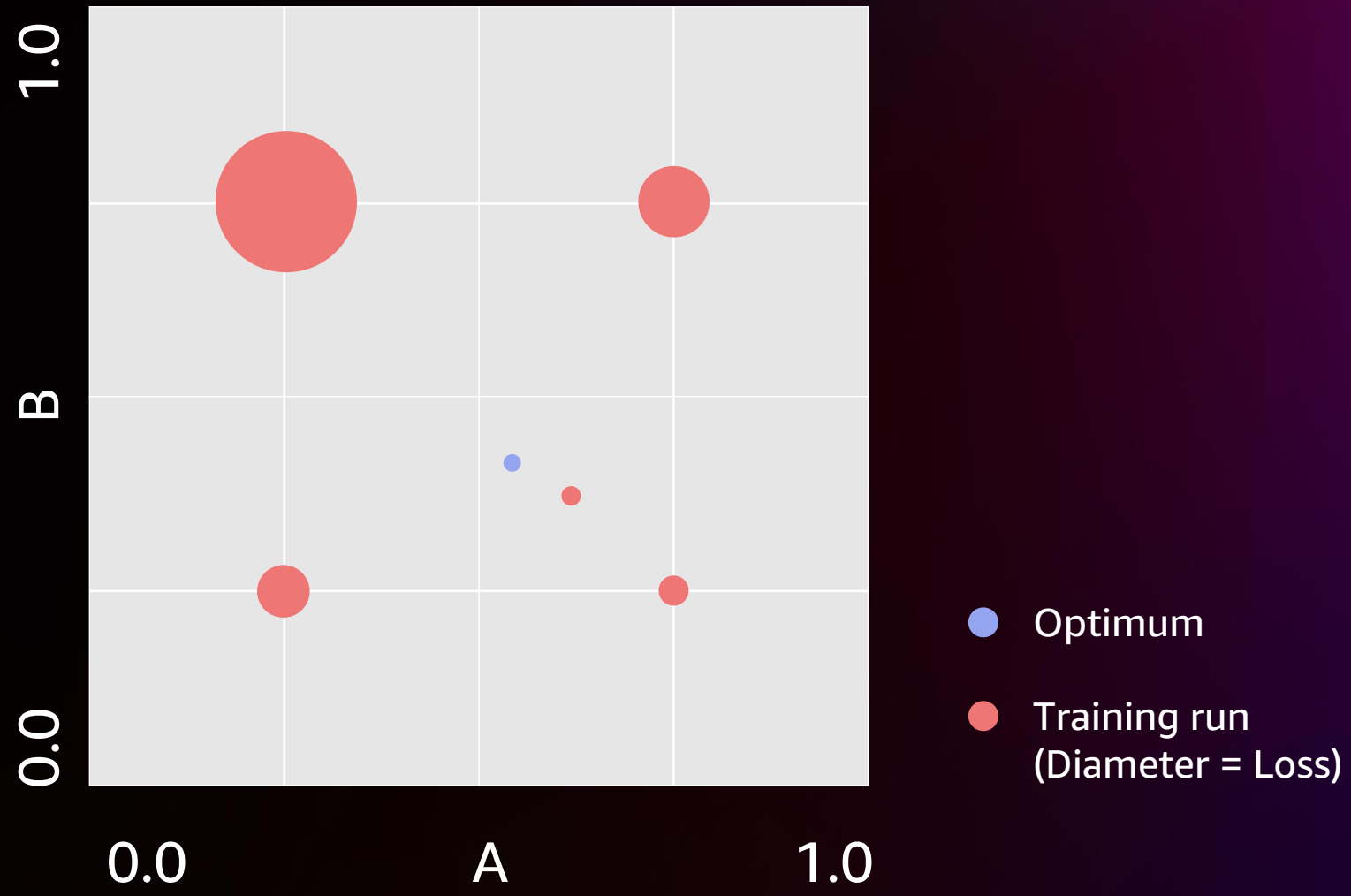
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Grid search



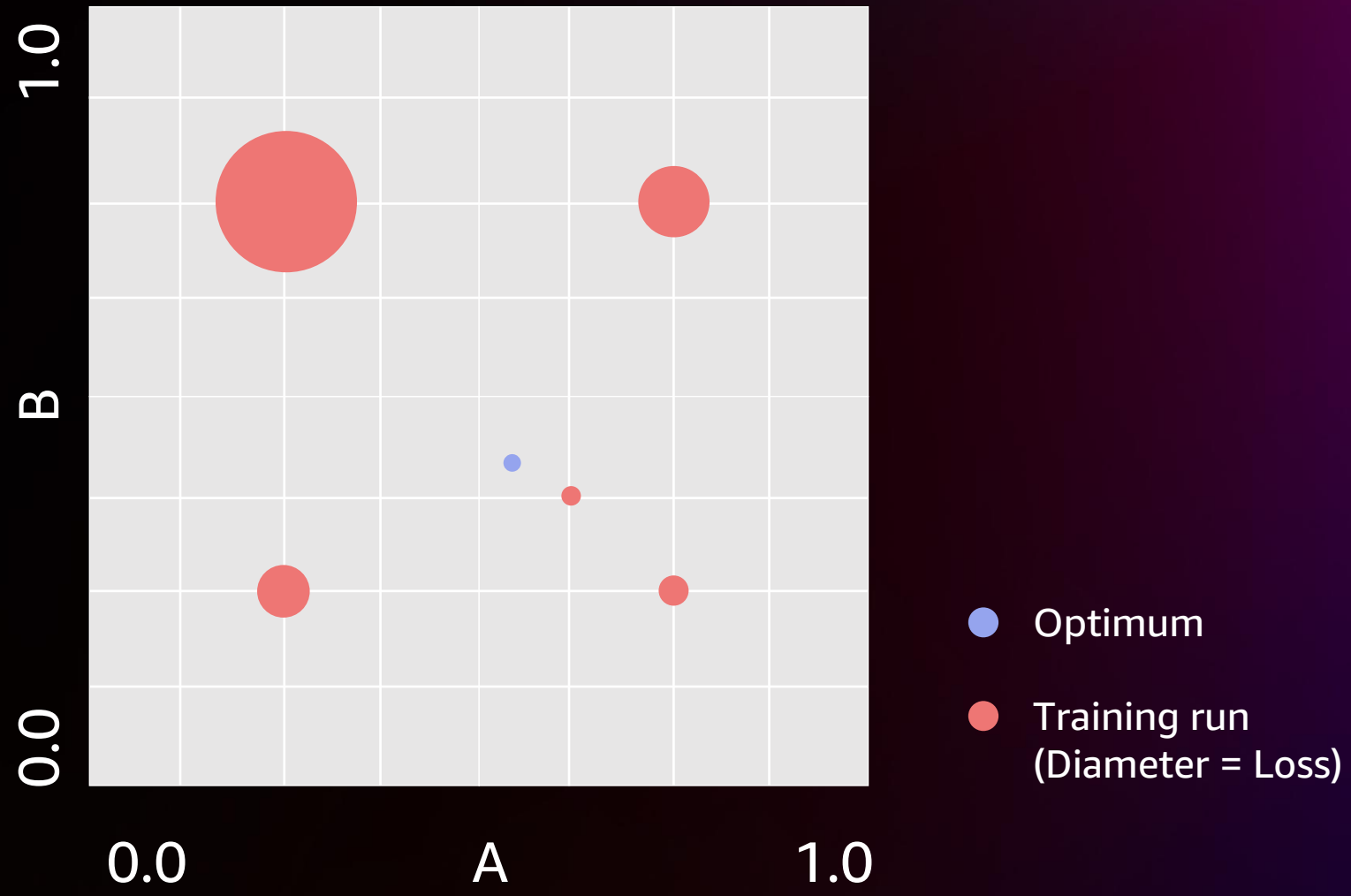
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Grid search

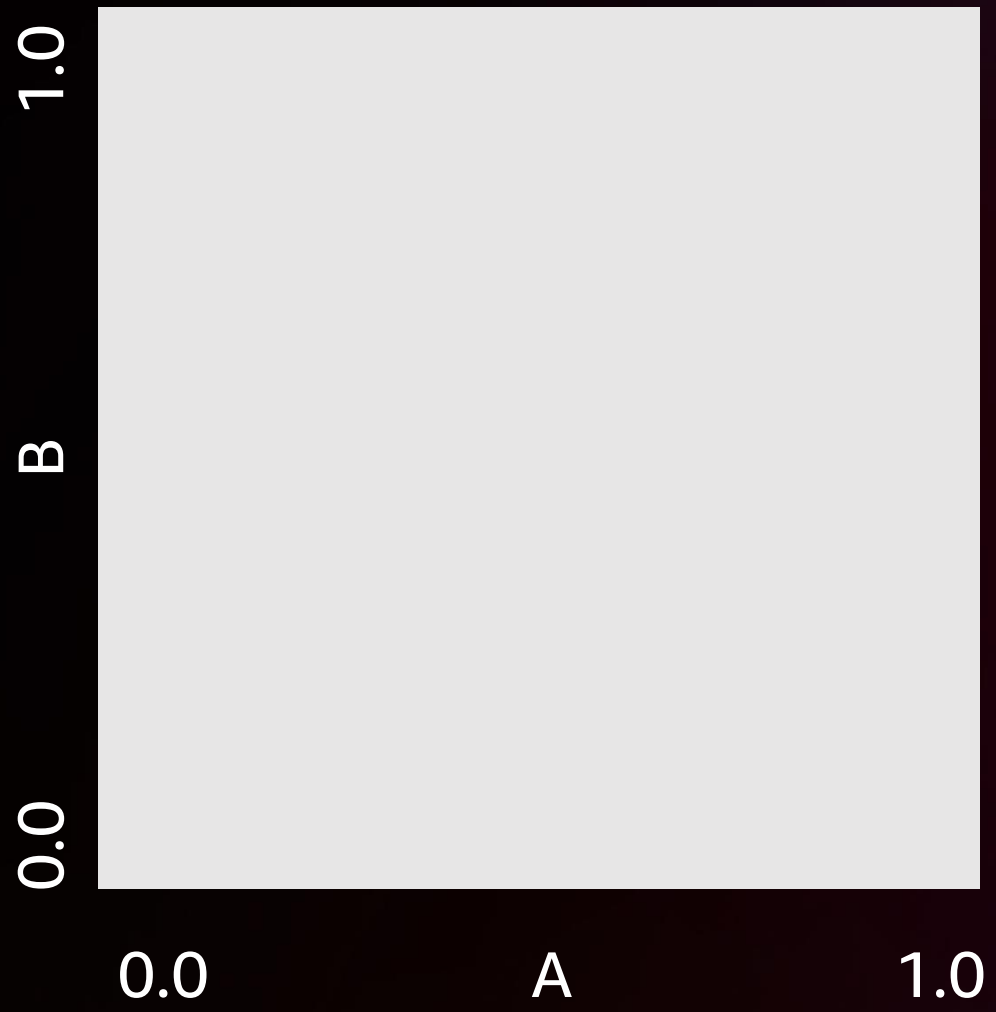


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Grid search

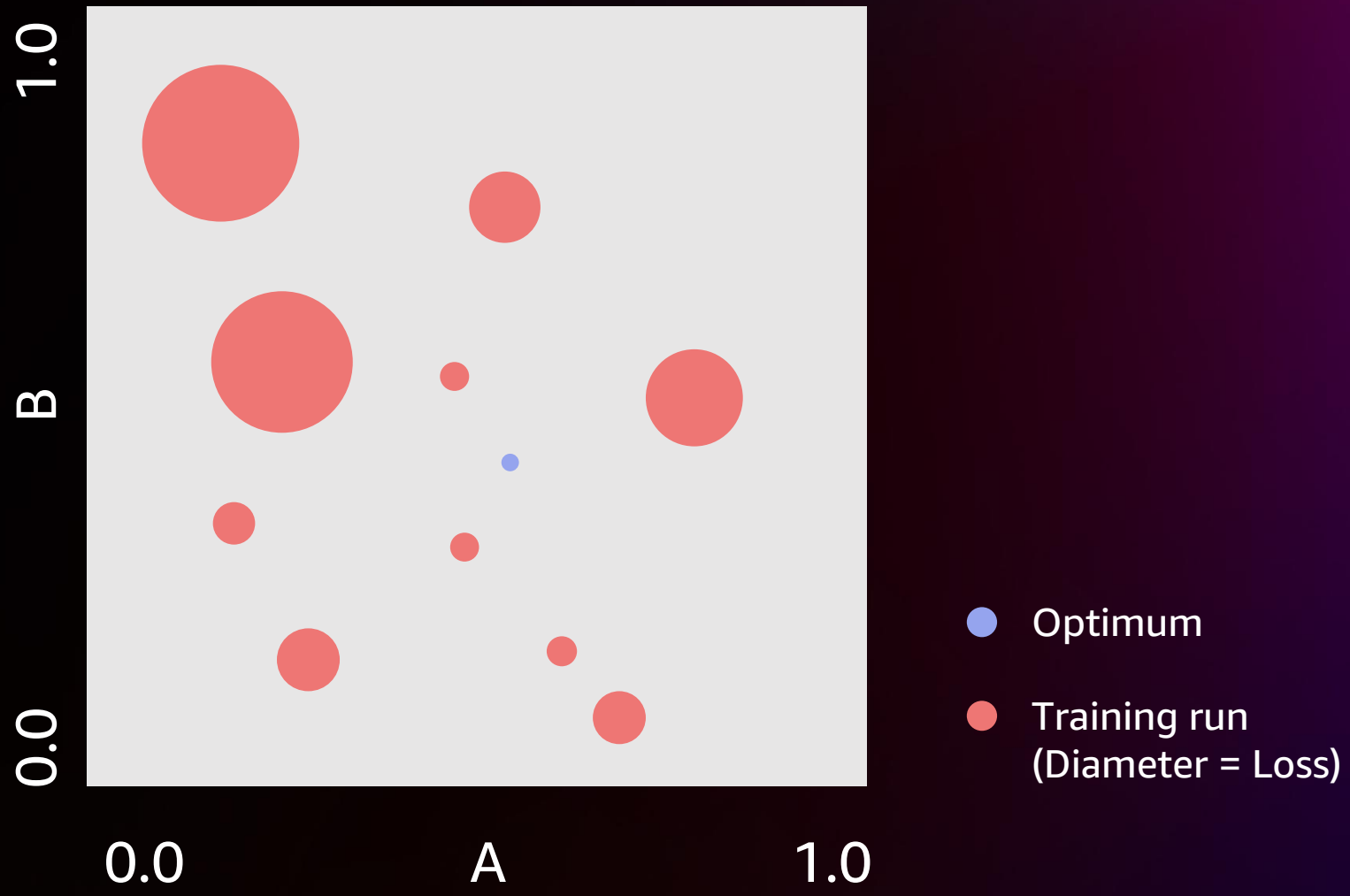


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

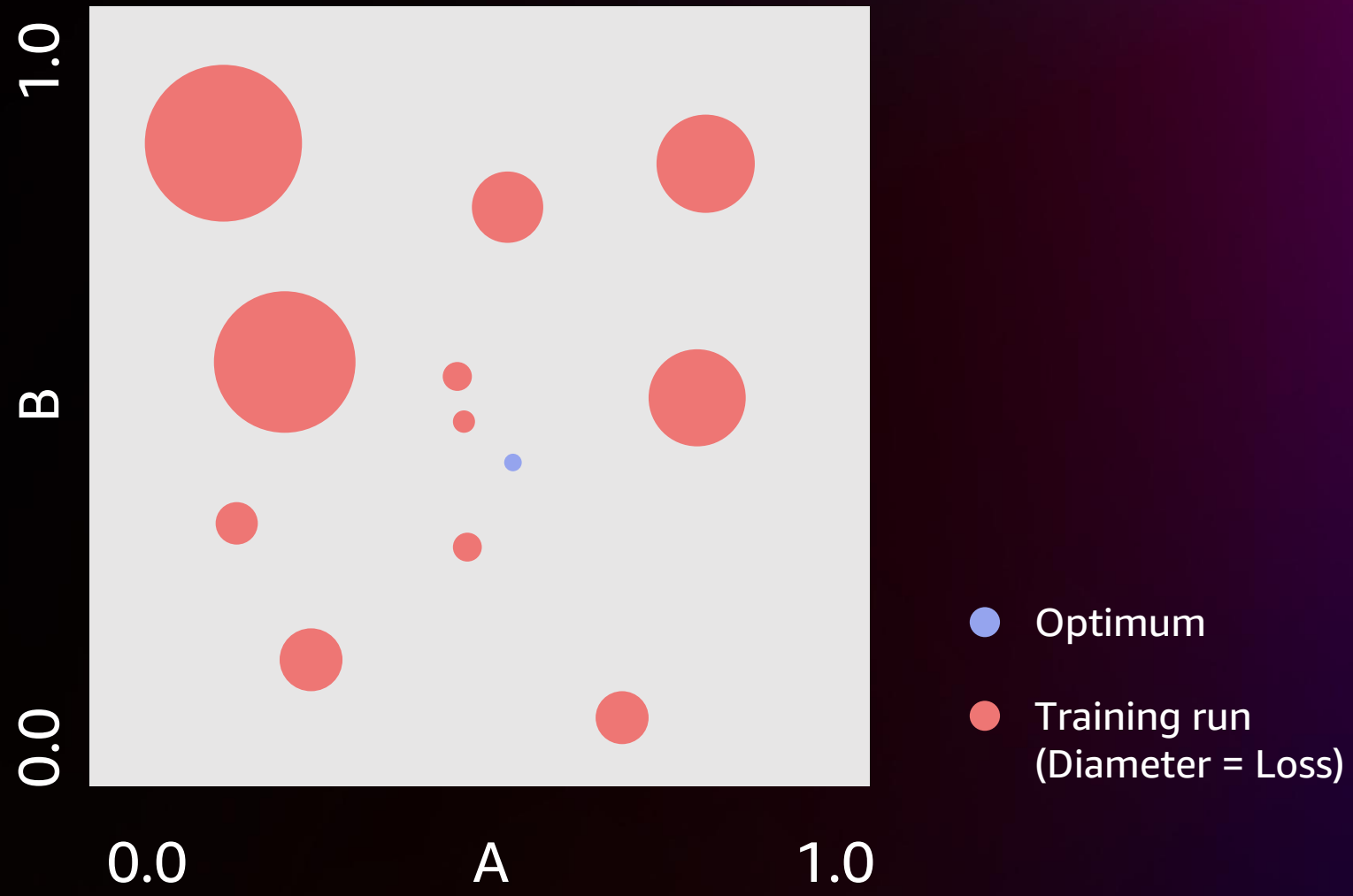
Random search



aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Bayesian search

EXPLORE VS. EXPLOIT



Illustrative only

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon SageMaker Automatic Model Tuning

Automatically tune hyperparameters in your algorithms



Tuning at scale

Adjust thousands of different combinations of algorithm parameters



Intelligent

Uses ML to find the best parameters



Faster

Eliminate days or weeks of tedious manual work



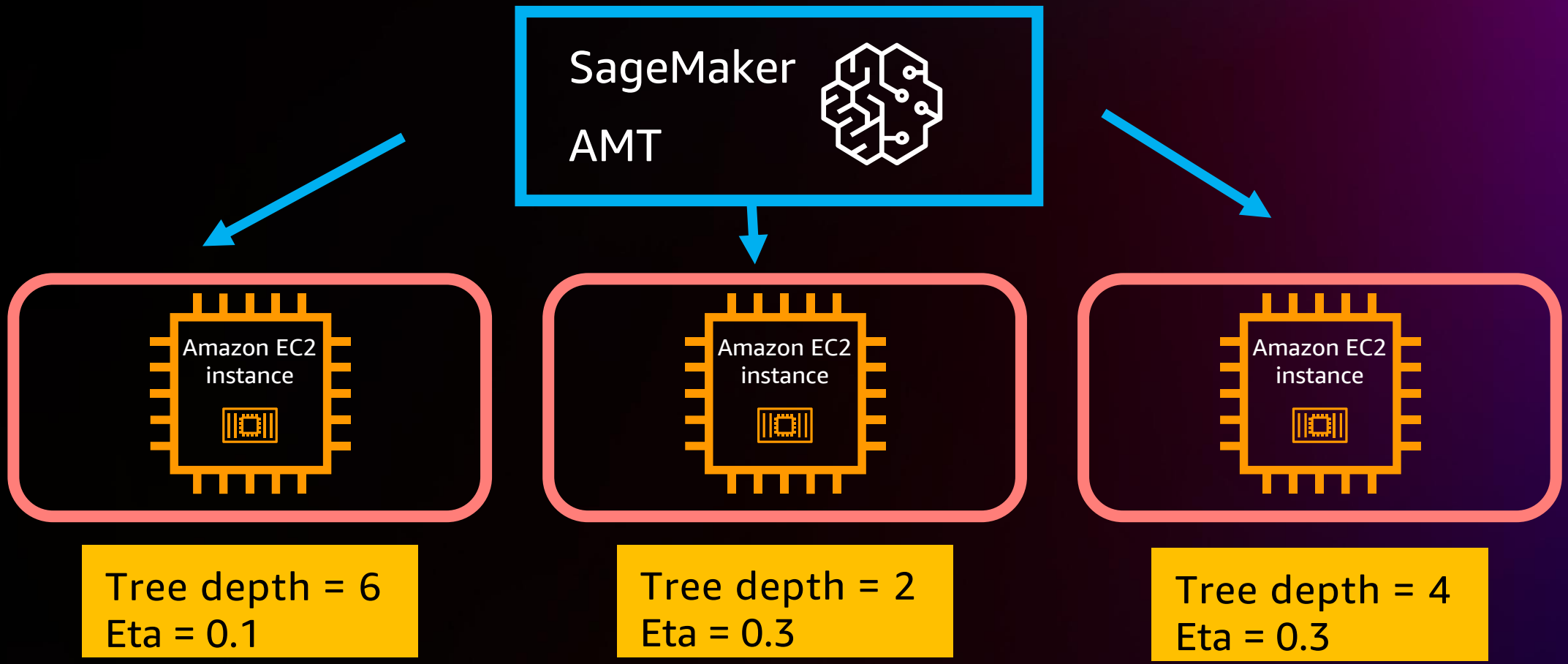
Production-Ready

Serverless managed online service – not just a library

aws

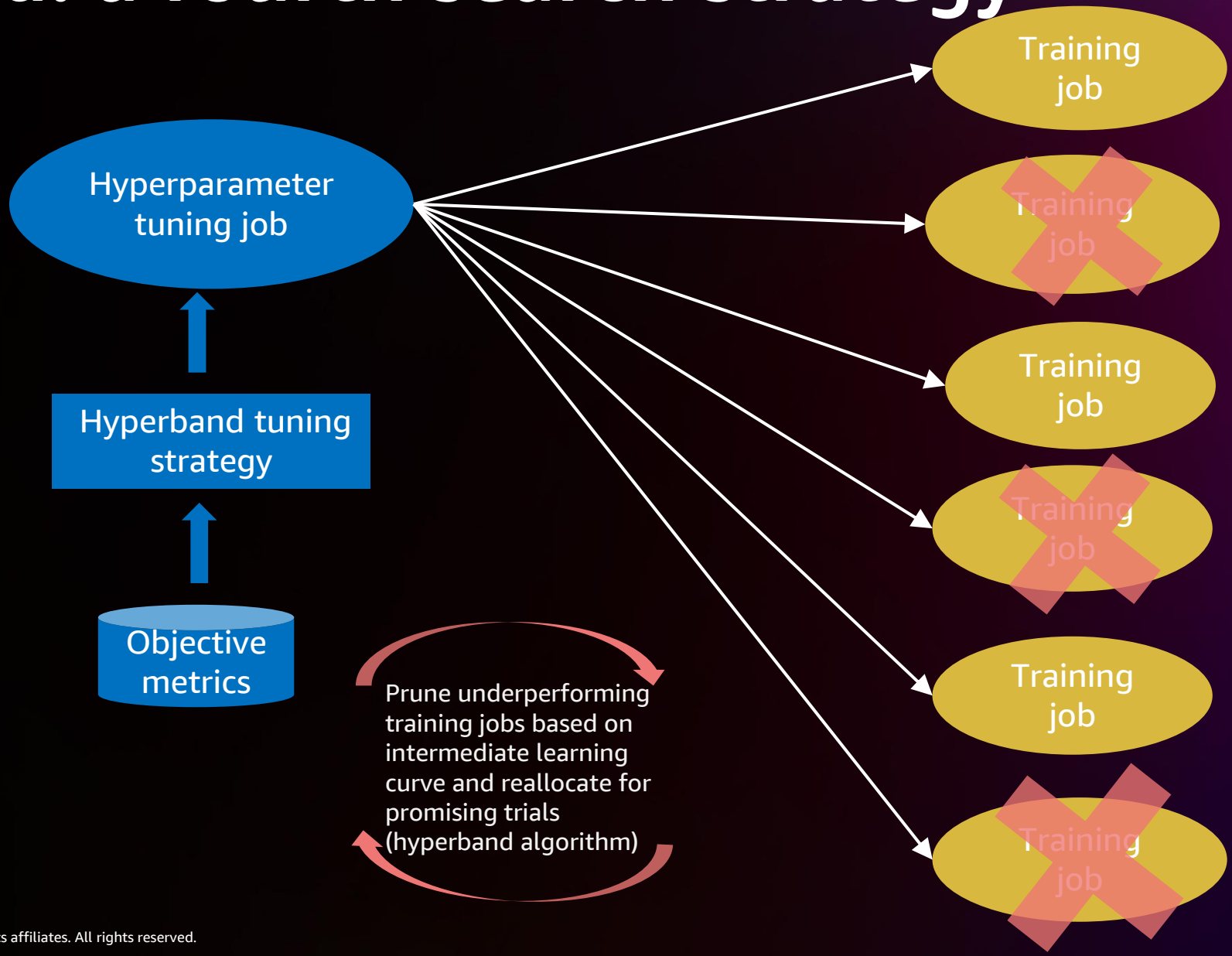
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Hyperparameter tuning jobs



aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Hyperband: a fourth search strategy



aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Which search strategy should I pick?

Grid search

Evaluate the impact of hyperparameters systematically

Random search

Introduce sampling to improve efficiency and scale up to larger search spaces

Bayesian optimization

Sample new data points by incorporating past experience

Hyperband

Maximize diversity of starting points and progress the most promising trials

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

We’d like to know about you

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Where do I start?



Amazon SageMaker Automatic Model Tuning web page
Resources and tools to help you tune your ML models



GitHub Repository (**amazon-sagemaker-amt-visualize**)
GitHub demo notebook to help you onboard to AMT



Optimize hyperparameters with Amazon SageMaker AMT
Blog post associated to the demo notebook



Tune your ML models 3x faster
Amazon SageMaker Automatic Model Tuning now supports Hyperband, a new multi-fidelity search method

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Valerio Perrone
https://www.linkedin.com/in/valerio-perrone/

Mariano Kamp
https://www.linkedin.com/in/mariano-kamp/



Please complete the session survey in the **mobile app**

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

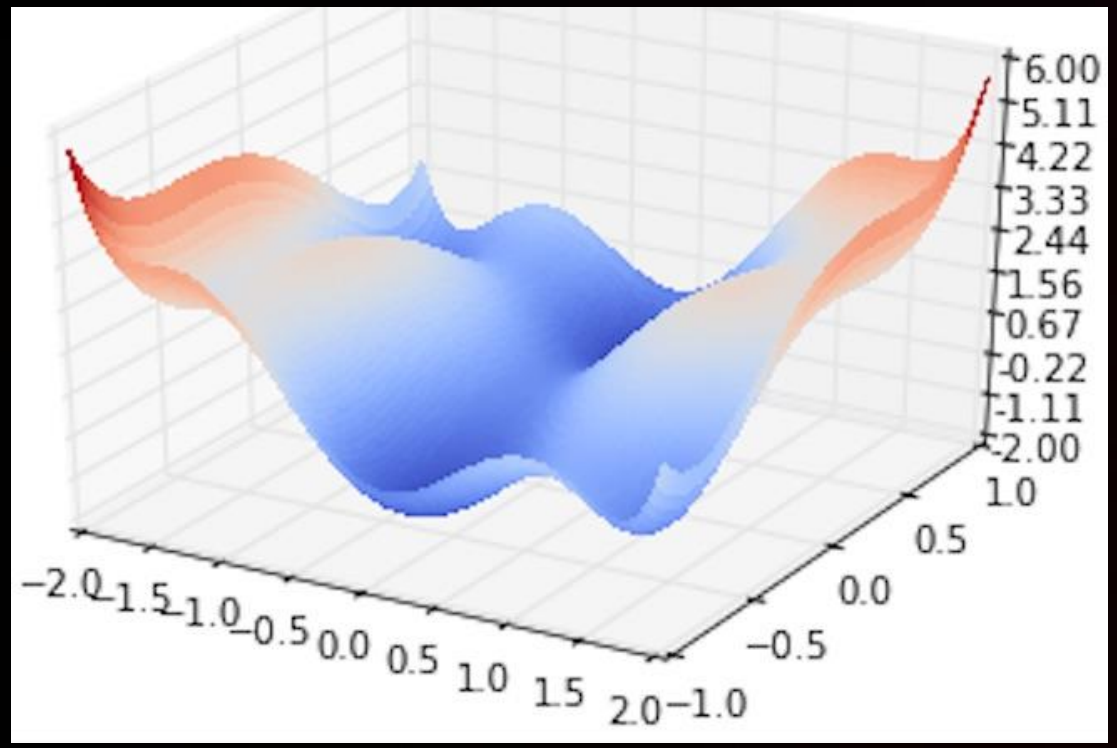
Appendix

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Bayesian optimization

Goal: find the best hyperparameter configuration x^* optimizing a function f: $x^* = argmin_{x \in X} f(x)$

f(x)



Challenges:

- Evaluating f(x) is expensive
- No analytical form or gradient
  → Can't use SGD
- Evaluations may be noisy

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Bayesian optimization: two ingredients

Ingredient 1: model (Gaussian process)

- Use ML to learn the relationship between hyperparameters and your objective

Ingredient 2: acquisition function

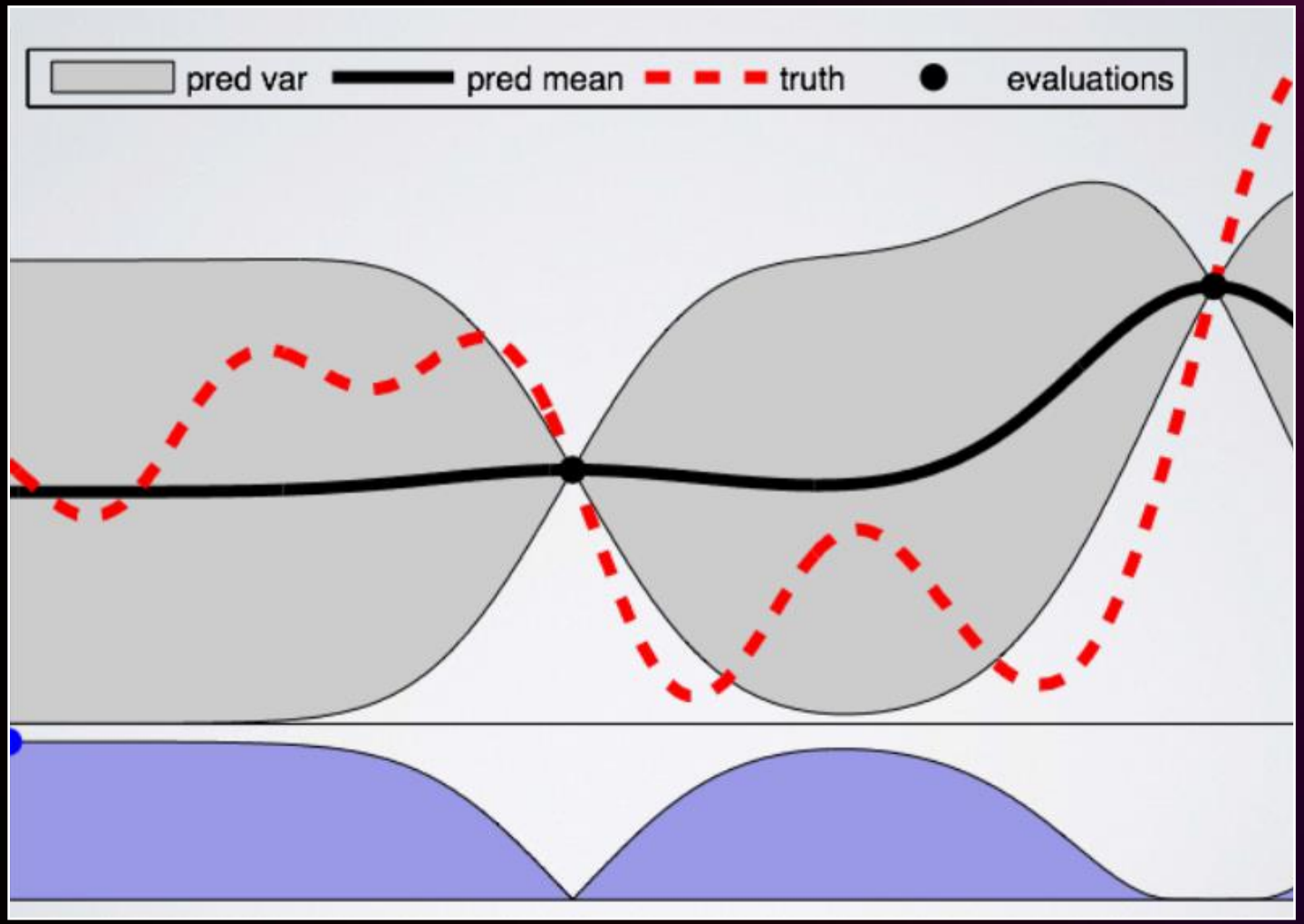
- Rank all possible hyperparameter configurations and pick the best one
- Trade off exploitation and exploration

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Bayesian optimization: a cartoon

Model

Acquisition function

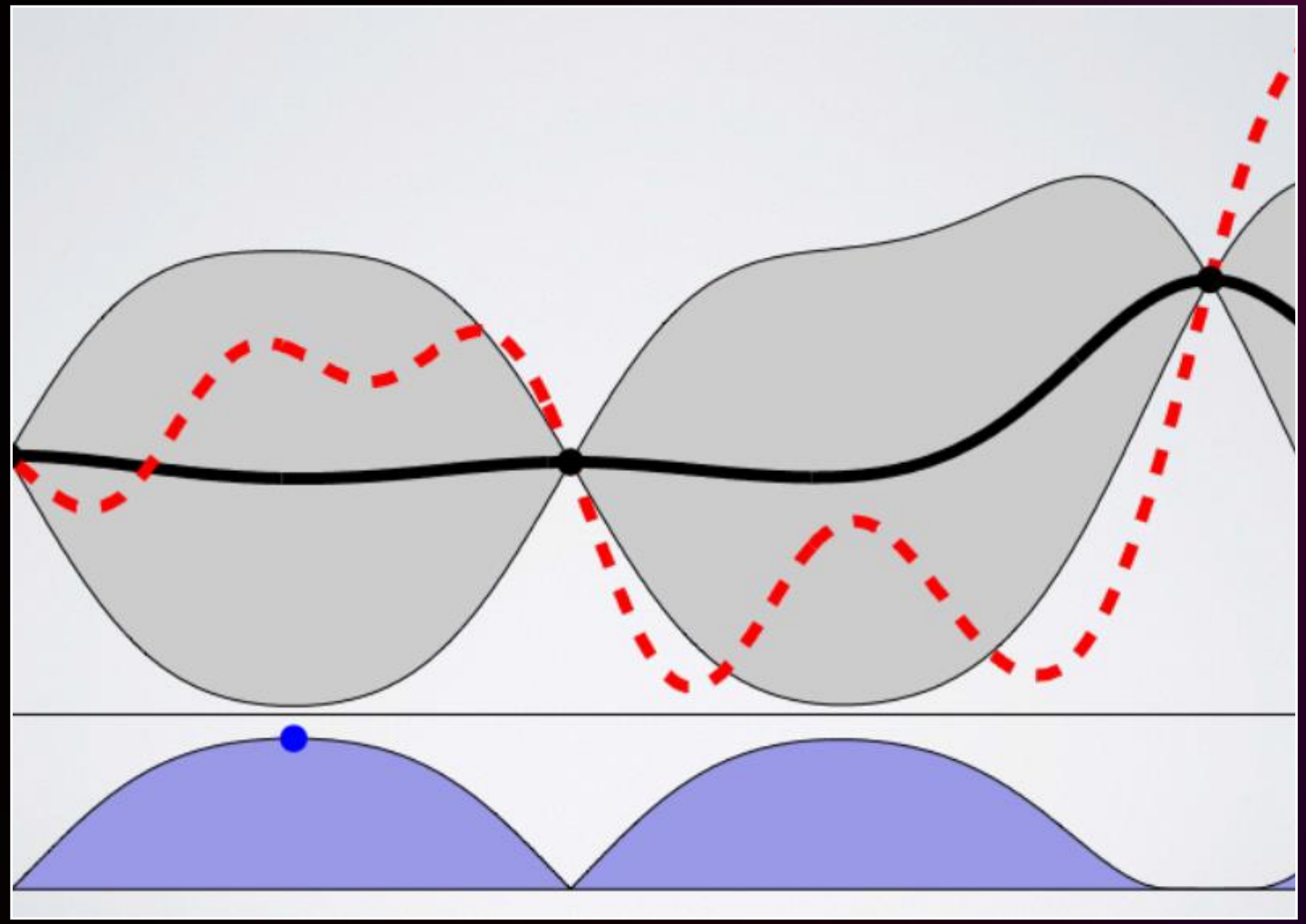


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Bayesian optimization: a cartoon

Model

Acquisition function



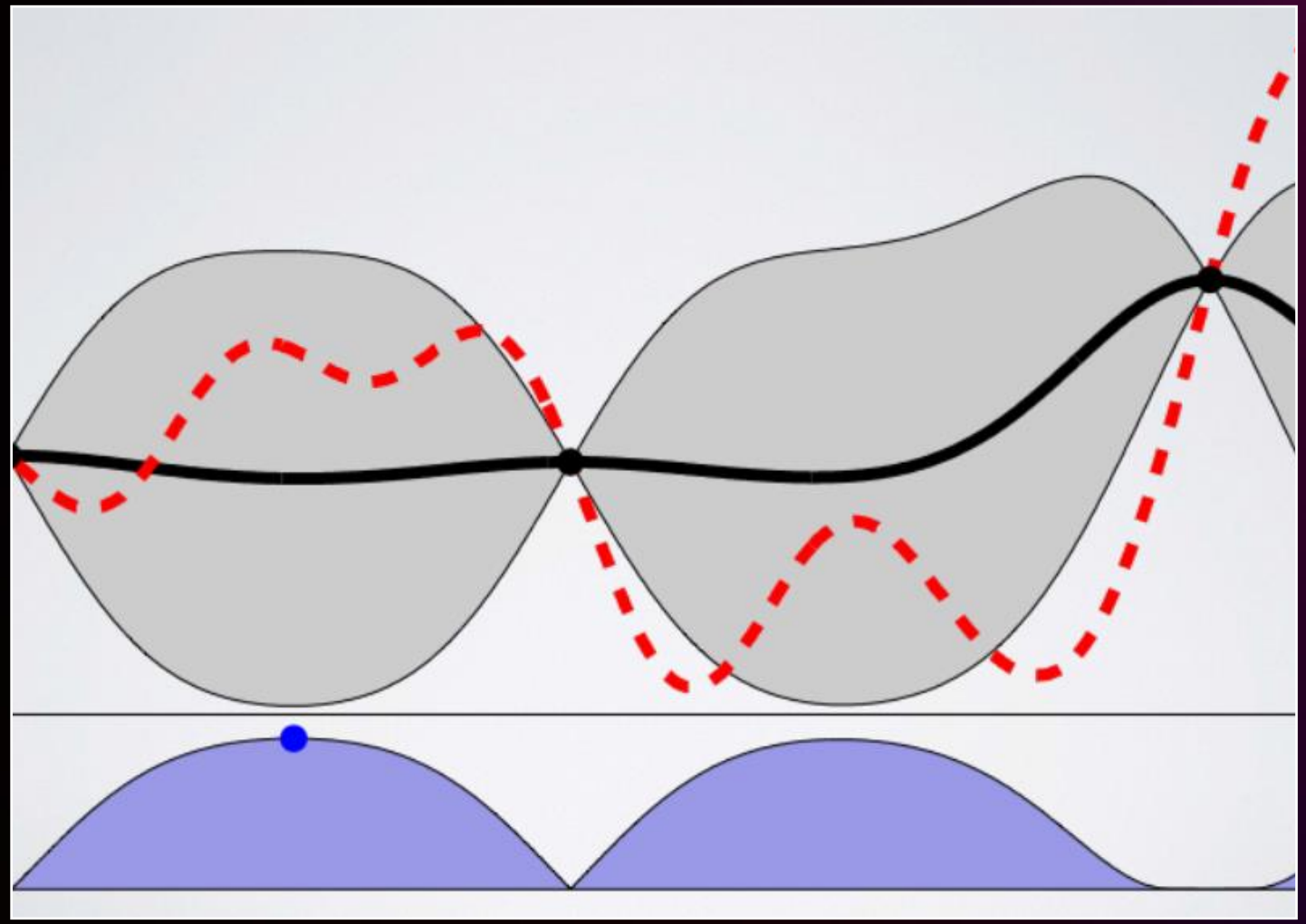
aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Bayesian optimization: a cartoon

Model

Acquisition function



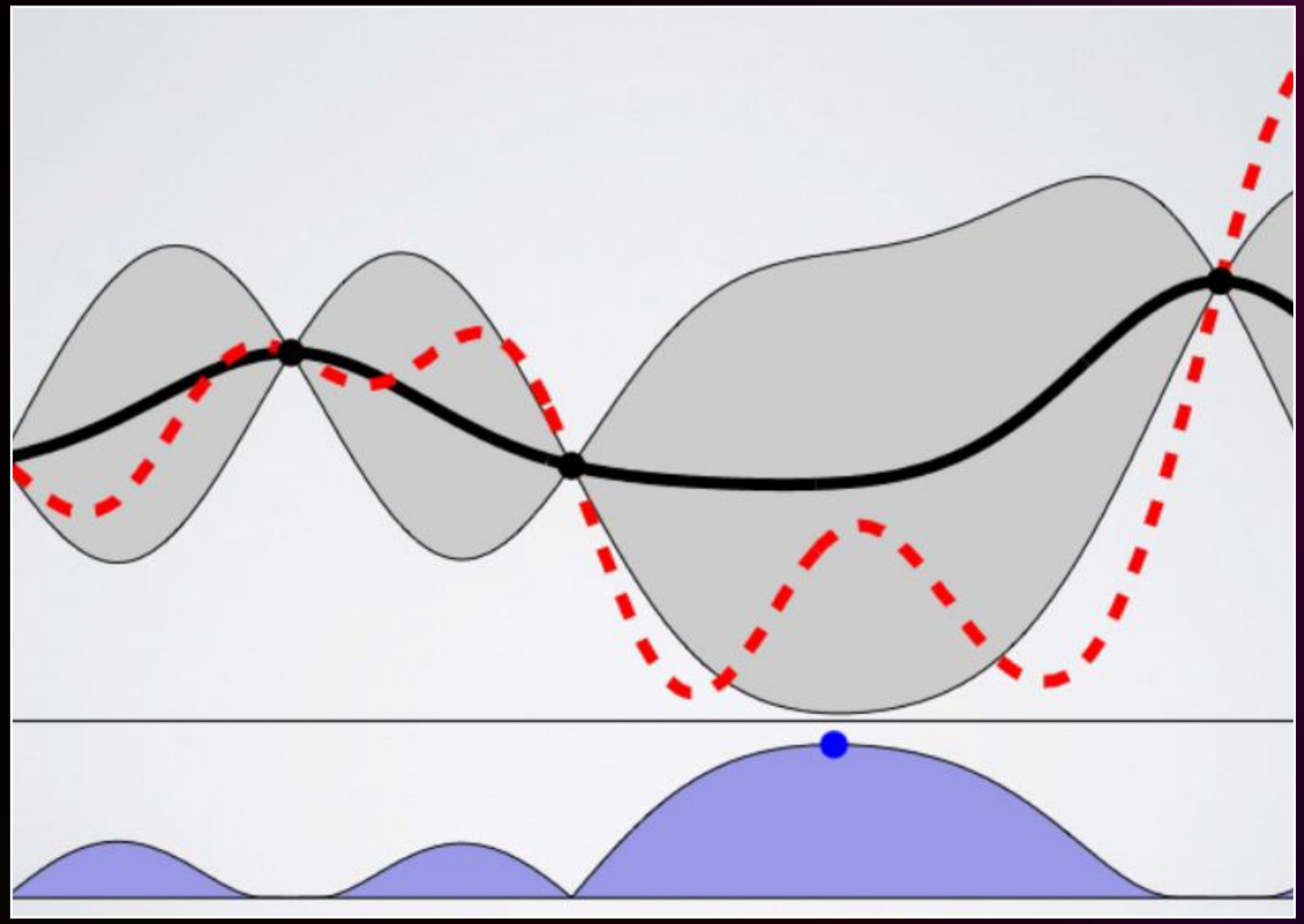
aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Bayesian optimization: a cartoon

Model

Acquisition function

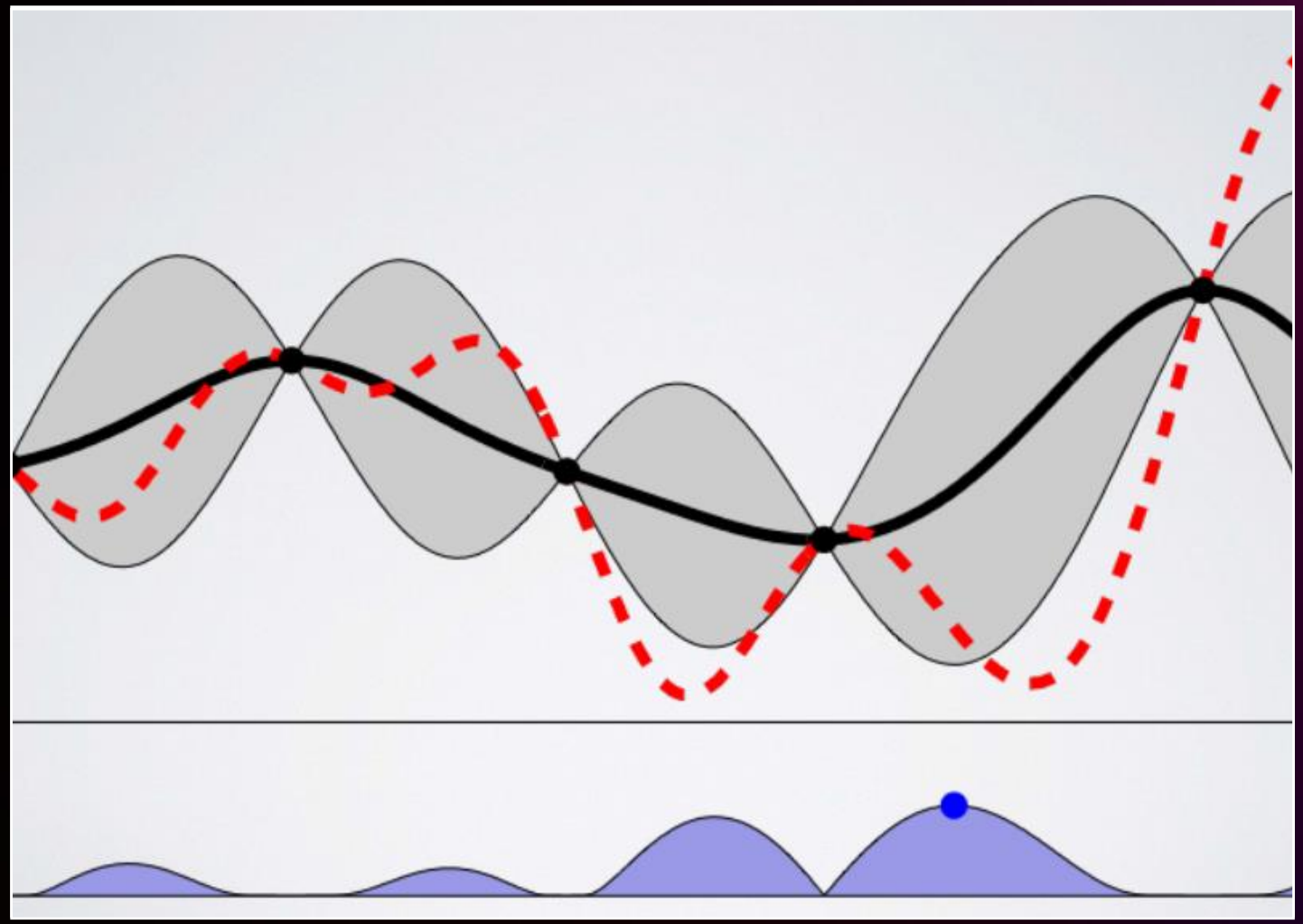


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Bayesian optimization: a cartoon

Model

Acquisition function

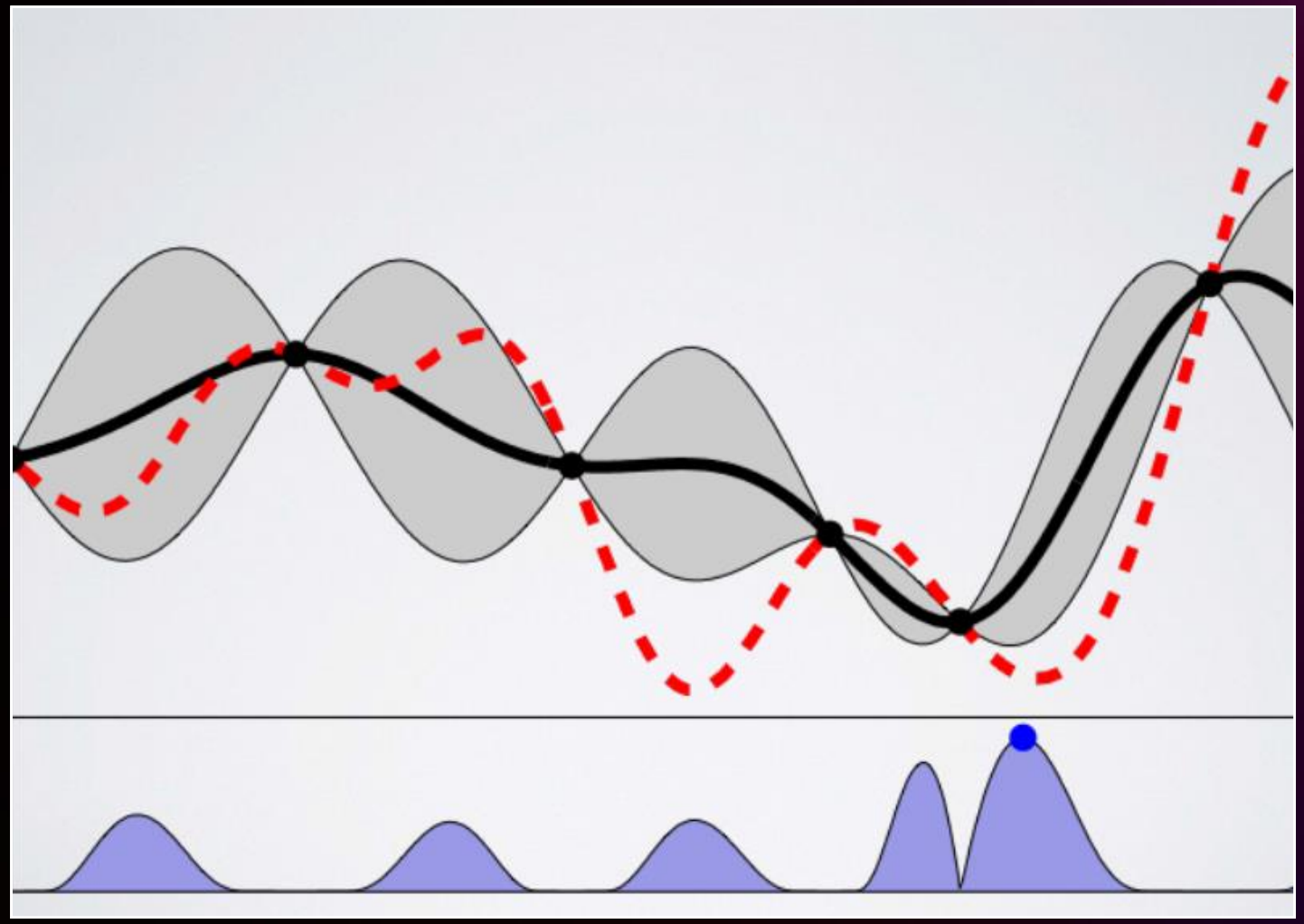


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Bayesian optimization: a cartoon

Model

Acquisition function

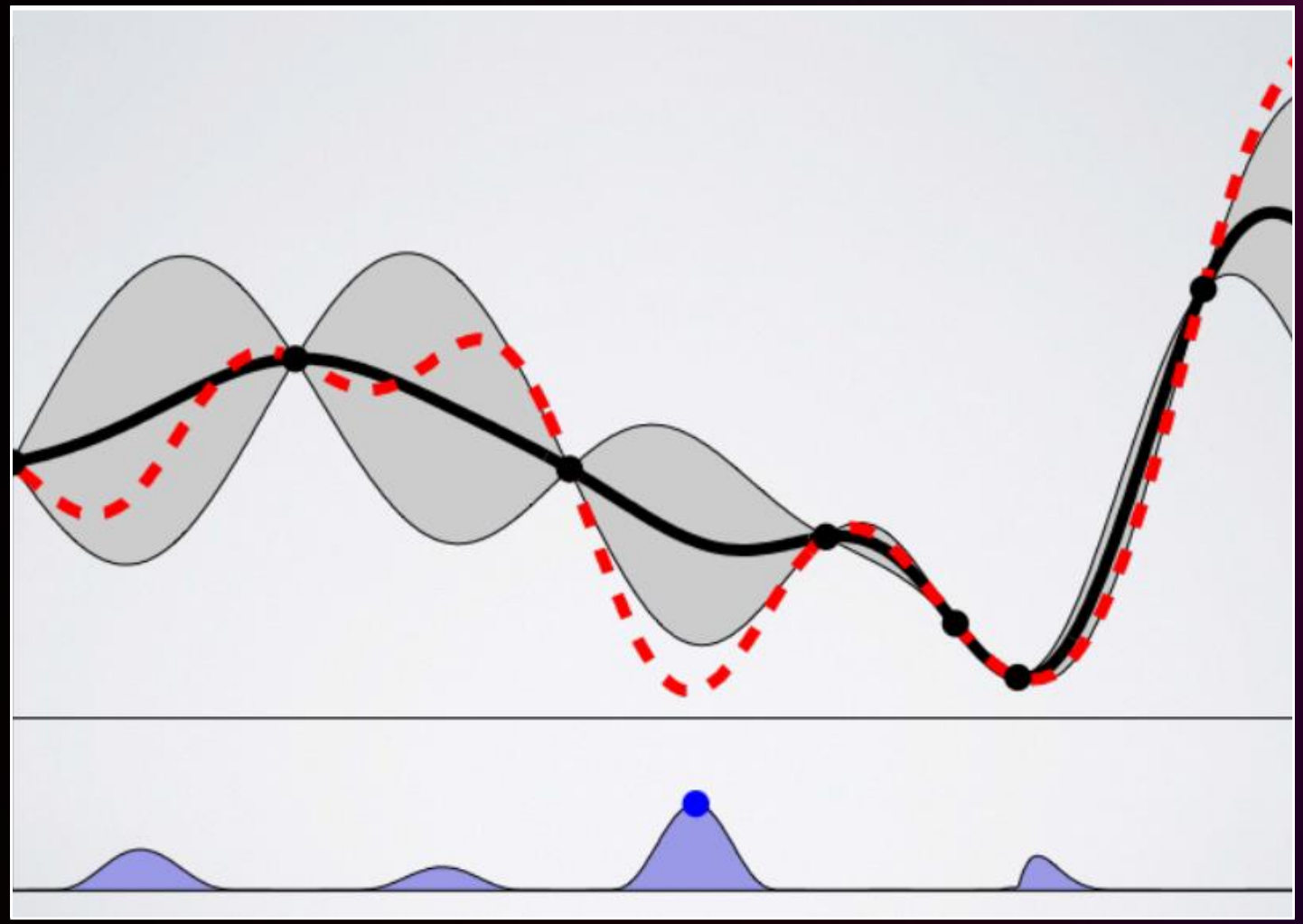


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Bayesian optimization: a cartoon

Model

Acquisition function

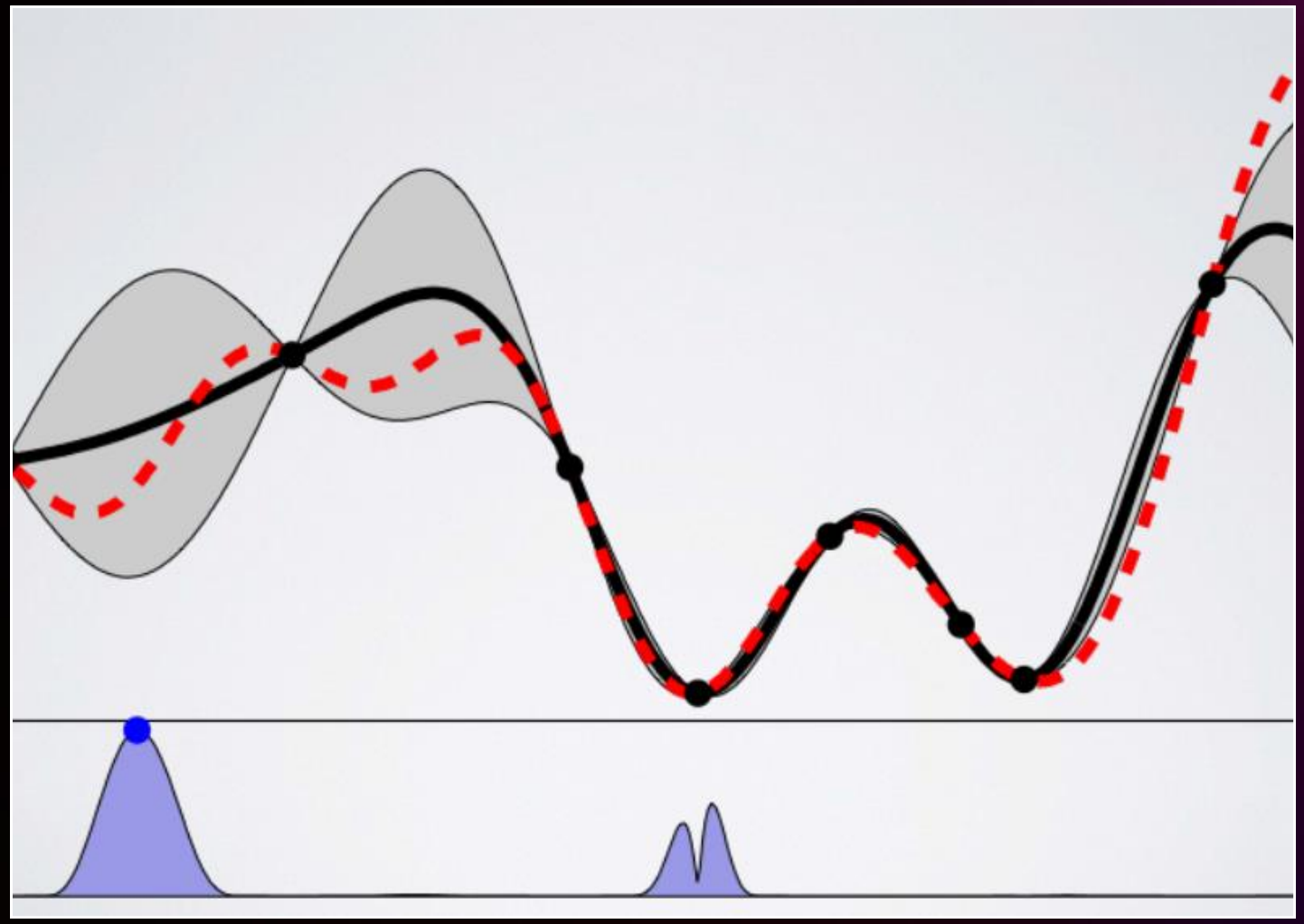


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Bayesian optimization: a cartoon

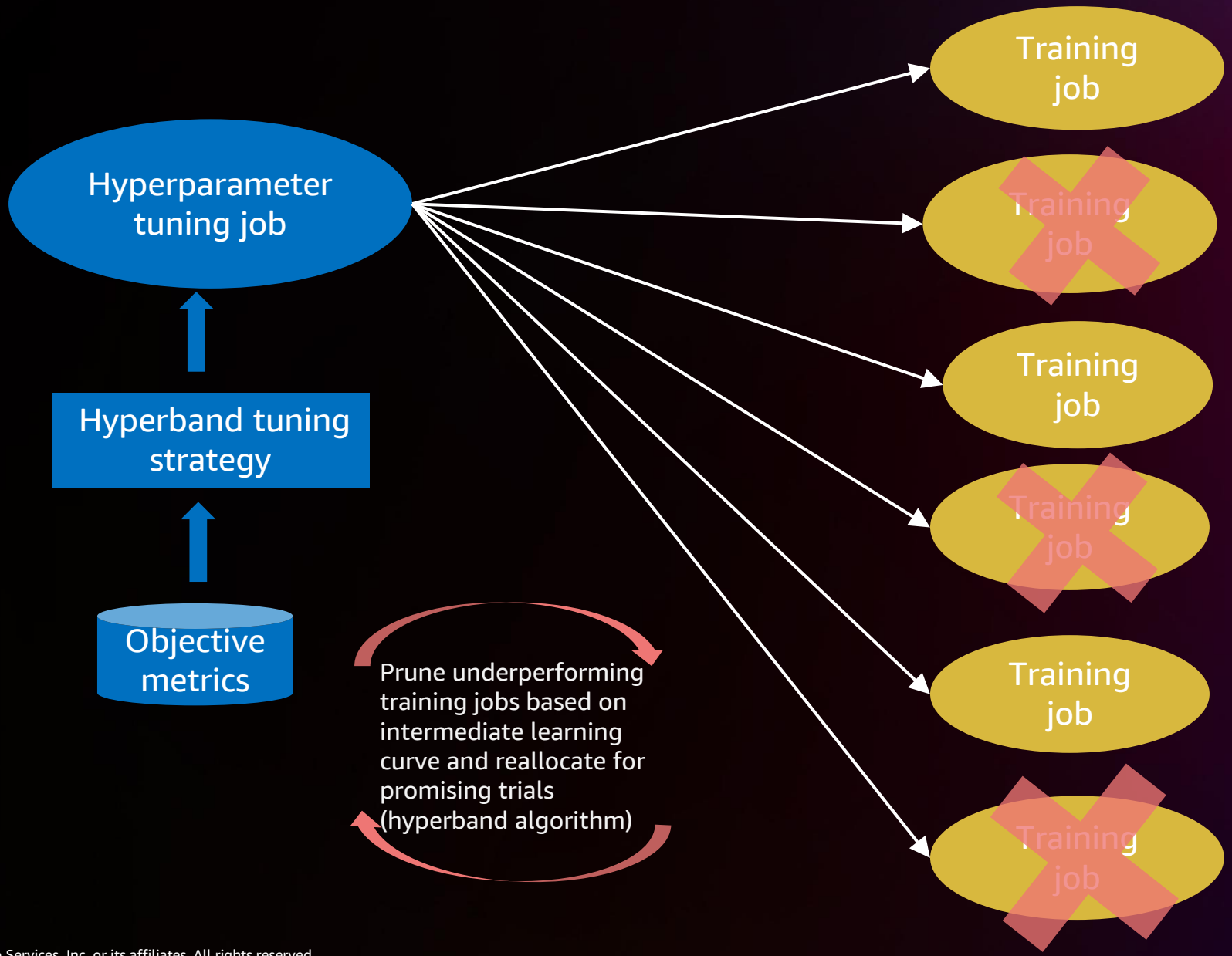
Model

Acquisition function



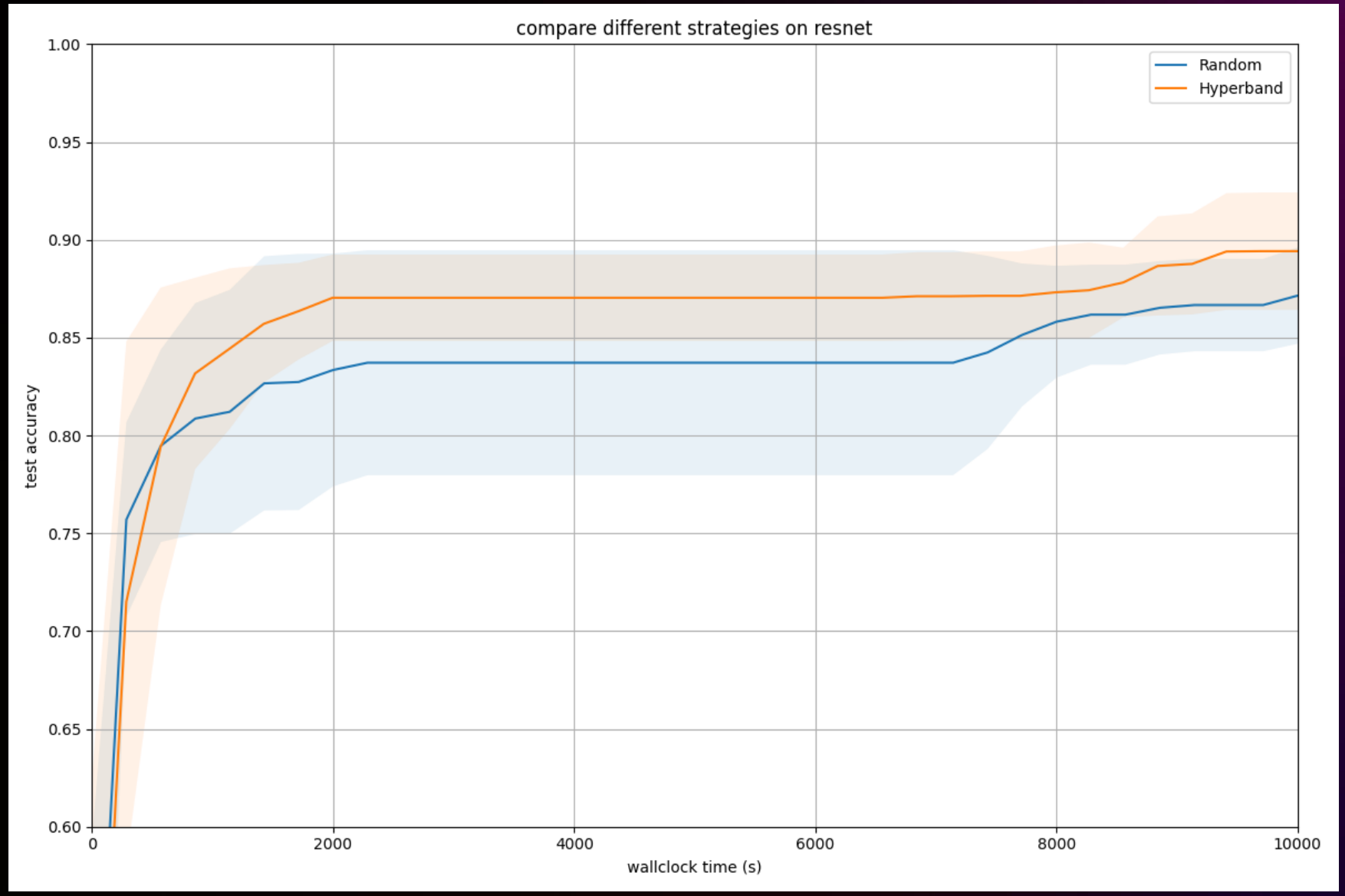
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Hyperband



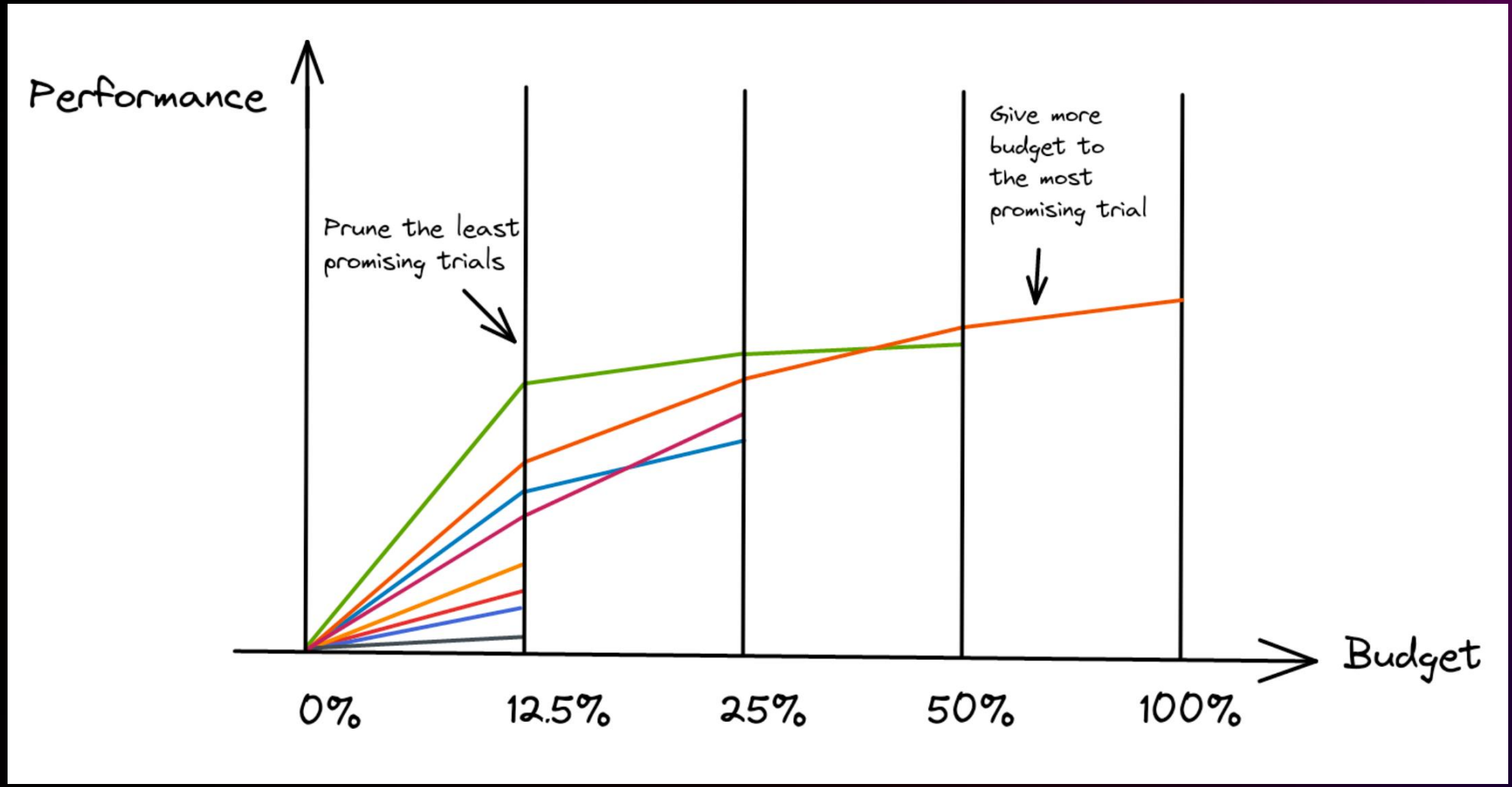
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Hyperband – 5x tuning on ResNet-20



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Successive Halving idea



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.