

Amazon SageMaker

Build, Train, and Deploy TensorFlow models with Amazon SageMaker

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark



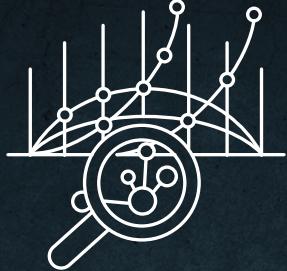
A g e n d a

- Amazon SageMaker Overview
- Q & A
- Setup for Workshop
- Hands-on Workshop Modules



Customer-focused

90%+ of our ML roadmap is defined by customers



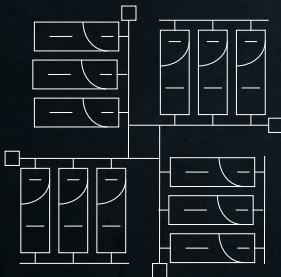
Pace of innovation

200+ new ML launches and major feature updates last year



Breadth and depth

A wide range of AI and ML services



Multi-framework

Support for the most popular frameworks



Security and analytics

Deep set of security with robust encryption and analytics



Embedded R&D

Customer-centric approach



More machine learning happens on AWS than anywhere else

10,000+ customers | 2x the customer references | 85% of TensorFlow projects in the cloud happen on AWS



© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark



THE AWS ML STACK

Broadest and deepest set of capabilities

ML Frameworks + Infrastructure

| FRAMEWORKS | INTERFACES | INFRASTRUCTURE | | | | | | | | |
|---|--|---|---|--|--|---|--|---|---|--|
|  TensorFlow  PYTORCH |  mxnet  Keras |  GLUON |  EC2 P3 & P3DN |  EC2 G4 |  EC2 C5 |  FPGAs |  DL CONTAINERS & AMIS |  GREENGASS |  ELASTIC INFERENCE |  INFERENTIA |



© 2019 Amazon Web Services, Inc. or its affiliates. All rights reserved |

5

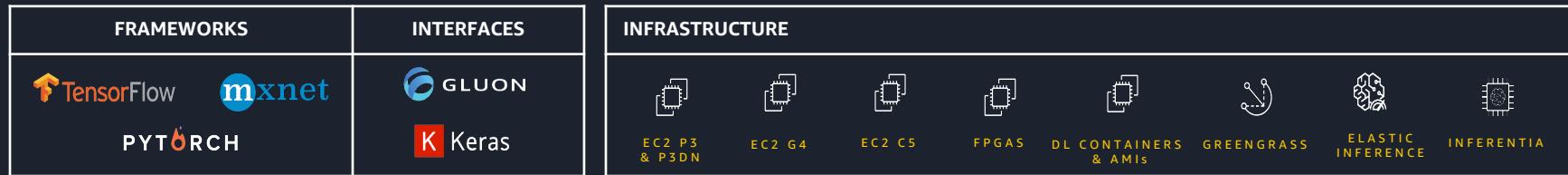
THE AWS ML STACK

Broadest and deepest set of capabilities

ML Services



ML Frameworks + Infrastructure



THE AWS ML STACK

Broadest and deepest set of capabilities

AI Services

| VISION | SPEECH | LANGUAGE | CHATBOTS | FORECASTING | RECOMMENDATIONS | | | |
|---|---|---|---|--|---|---|--|---|
|  REKOGNITION IMAGE |  REKOGNITION VIDEO |  TTEXTRACT |  POLLY |  TRANSCRIBE |  TRANSLATE COMPREHEND & COMPREHEND MEDICAL |  LEX |  FORECAST |  PERSONALIZE |

ML Services



ML Frameworks + Infrastructure

| FRAMEWORKS | INTERFACES | INFRASTRUCTURE | | | | | | | | |
|---|--|---|---|--|--|---|--|---|---|--|
|  TensorFlow  PYTORCH |  mxnet  Keras |  GLUON |  EC2 P3 & P3DN |  EC2 G4 |  EC2 C5 |  FPGAS |  DL CONTAINERS & AMIS |  GREENGASS |  ELASTIC INFERENCE |  INFERENTIA |

Amazon SageMaker:

Build, Train, and Deploy ML Models at Scale



Collect and prepare
training data



Choose and
optimize your
ML algorithm



Set up and
manage
environments
for training



Train and
Tune ML Models



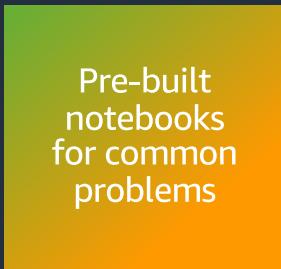
Deploy models
in production



Scale and manage
the production
environment

Amazon SageMaker:

Build, Train, and Deploy ML Models at Scale



Collect and prepare training data



Choose and optimize your ML algorithm



Set up and manage environments for training



Train and Tune ML Models



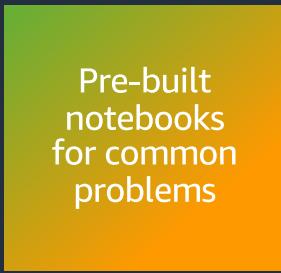
Deploy models in production



Scale and manage the production environment

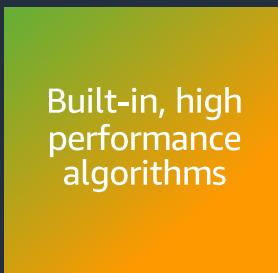
Amazon SageMaker:

Build, Train, and Deploy ML Models at Scale



Pre-built
notebooks
for common
problems

Collect and prepare
training data



Built-in, high
performance
algorithms

Choose and
optimize your
ML algorithm



Set up and
manage
environments
for training



Train and
Tune ML Models

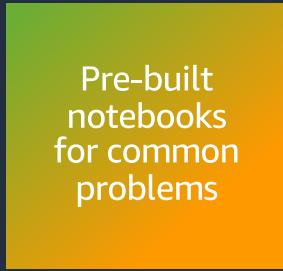


Deploy models
in production



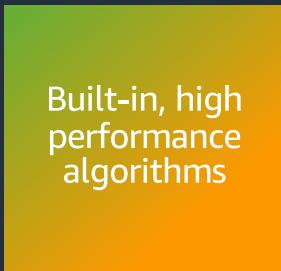
Scale and manage
the production
environment

Amazon SageMaker: Build, Train, and Deploy ML Models at Scale



Pre-built
notebooks
for common
problems

Collect and prepare
training data



Built-in, high
performance
algorithms

Choose and
optimize your
ML algorithm



One-click
training on the
highest
performing
infrastructure

Set up and
manage
environments
for training



Train and
Tune ML Models



Deploy models
in production



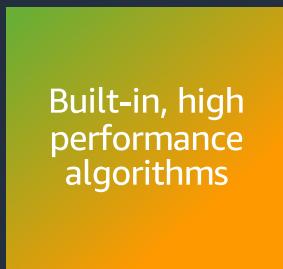
Scale and manage
the production
environment

Amazon SageMaker: Build, Train, and Deploy ML Models at Scale



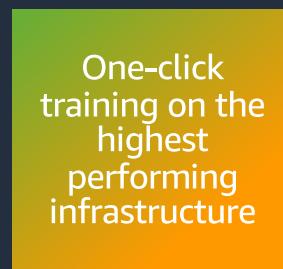
Pre-built
notebooks
for common
problems

Collect and prepare
training data



Built-in, high
performance
algorithms

Choose and
optimize your
ML algorithm



One-click
training on the
highest
performing
infrastructure

Set up and
manage
environments
for training



Model
Optimization

Train and
Tune ML Models



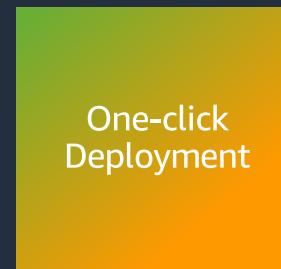
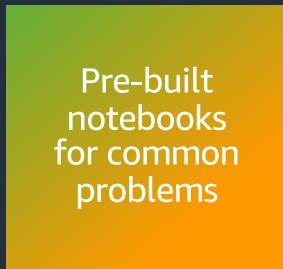
Deploy models
in production



Scale and manage
the production
environment

Amazon SageMaker:

Build, Train, and Deploy ML Models at Scale



Pre-built notebooks for common problems

Collect and prepare training data

Built-in, high performance algorithms

Choose and optimize your ML algorithm

One-click training on the highest performing infrastructure

Set up and manage environments for training

Model Optimization

Train and Tune ML Models

One-click Deployment

Deploy models in production

Scale and manage the production environment

Amazon SageMaker: Build, Train, and Deploy ML Models at Scale

Pre-built notebooks for common problems

Collect and prepare training data

Built-in, high performance algorithms

Choose and optimize your ML algorithm

One-click training on the highest performing infrastructure

Set up and manage environments for training

Model Optimization

Train and Tune ML Models

One-click Deployment

Deploy models in production

Fully managed with auto-scaling for 75% less

Scale and manage the production environment



DOW JONES





Machine Learning is a highly collaborative
process every step of the way

SageMaker Workflows

Experiment Management

Organize, track and evaluate
model training experiments
with SageMaker Search

SageMaker Workflows

Experiment Management

Organize, track, and evaluate
model training experiments
with SageMaker Search

Automation

Use AWS Step Functions to
automate end-to-end
workflows

Integrate with Apache Airflow

SageMaker Workflows

Experiment Management

Organize, track, and evaluate model training experiments with SageMaker Search

Automation

Use AWS Step Functions to automate end-to-end workflows

Integrate with Apache Airflow

Collaboration

Link GitHub, AWS CodeCommit and self-hosted Git repositories to notebooks

Clone public and private repositories

Secure information with IAM and AWS Secrets Manager



Q & A

© 2019, Amazon Web Services, Inc. or its Affiliates.



Getting Started on the Workshop

© 2019, Amazon Web Services, Inc. or its Affiliates.



Workshop Options

Lab 1 - NLP Use Case – Sentiment Analysis: train a textCNN to classify movie reviews as positive or negative, and perform batch inference on a set of reviews.

- Local Mode Training
- Hosted Training Mode
- Batch Prediction

Workshop Options, Cont'd...

Lab 2 - Structured Data Use Case – Boston Housing: while training a regression MLP to predict house prices, use Automatic Model Tuning to find the best hyperparameters.

- Local Mode Training
- Local Mode Endpoint
- Hosted Training Mode with model extraction for running it elsewhere
- Automatic Model Tuning

Workshop Options, Cont'd...

Lab 3 - Computer Vision Use Case – Image Classification: train a CNN from scratch to classify images. This notebook showcases two distributed training approaches.

- Distributed Training using Parameter Servers
- Distributed Training using Horovod

Thank You!

