Build ML models using SageMaker Studio Notebooks

Kunal Jha, Senior Product Manager
Sean Morgan, Senior Machine Learning
Solutions Architect
Learning Objectives

• See how to launch and use SageMaker Studio Notebooks
• Learn how to install open-source extensions
• Learn how to track and manage training and data processing jobs and test machine learning model performance
Amazon SageMaker Studio is a web-based, integrated development environment (IDE) to prepare data, build, train, deploy, and monitor your machine learning models.
Prepare Data

Easily prepare data, directly in notebooks

- **Prepare data in a few clicks**
  Use little to no code with SageMaker Data Wrangler

- **Manage and interact with Amazon EMR clusters and debug Spark Jobs directly in Studio Notebooks**
  Create, browse and connect to Amazon EMR clusters

- **SageMaker Processing**
  Run data processing workloads reliably and at scale

- **SageMaker Feature Store, a fully managed, purpose-built repository**
  Store, update, share, and retrieve ML features in both real-time and in batch

- **SageMaker Clarify**
  Detect and limit potential bias during data preparation, after model training, and in your deployed model
Build ML models

Fully managed shareable notebooks on Amazon EC2

- **Fully managed one-click Jupyter notebooks**
  Run notebooks on elastic compute resources

- **Built-in algorithms**
  15 built-in algorithms available in prebuilt container images

- **SageMaker JumpStart**
  Get started quickly with pre-built solutions and over 300 popular open-source models

- **AutoML**
  Automatically create ML models with full visibility

- **Support for major frameworks and toolkits**
  Optimized for popular deep learning (DL) frameworks such as TensorFlow, PyTorch, Apache MXNet, and Hugging Face
Train ML models

Fast and cost-effective ML model training

- **Experiment management and model tuning**: Save weeks of effort by automatically tracking training runs and tuning hyperparameters.
- **Debug and profile training runs**: Use real-time metrics to correct performance problems.
- **Distributed training**: Complete distributed training up to 40% faster.
- **Training compiler**: Accelerate training times by up to 50% through more efficient use of GPUs.
- **Managed spot training**: Reduce the costs of training by up to 90%.
Deploy ML models

Fully managed deployment for inference at scale

- **Wide selection of infrastructure**: 70+ instance types with varying levels of compute and memory to meet the needs of every use case
- **Single-digit millisecond overhead latency**: For use cases requiring real-time responses
- **Asynchronous inference**: Supports large models with long-running processing times
- **Cost-effective deployment**: Multi-model/multi-container endpoints, serverless inference, and elastic scaling
- **Built-in integration for MLOps**: ML workflows, CI/CD, lineage tracking, and catalog
- **Automatic deployment recommendations**: Optimal instance type/count and container parameters, and fully managed load testing
SageMaker Studio Notebooks

Fully managed shareable notebooks on Amazon EC2

**Quick start**
Launch in few clicks with pre-configured data science and ML frameworks such as TensorFlow and PyTorch and over 200 sample notebooks

**Manage and interact with Amazon EMR clusters and debug Spark Jobs directly**
Create, browse and connect to Amazon EMR clusters

**Elastic compute**
Scale underlying compute resources up or down with a simple click

**Collaborative, built for teams**
Set up team access with a single click and share notebooks easily with shareable links

**Customizable**
Bring your own notebook environment to SageMaker Studio using a custom docker image
Product Demo