



# INNOVATE2018

## ONLINE CONFERENCE

DEVELOPER EDITION

# From Notebook to Production with Amazon SageMaker

October 2018

**Jason Barto, Solutions Architect, AWS**

@awscloud



#AWSInnovate

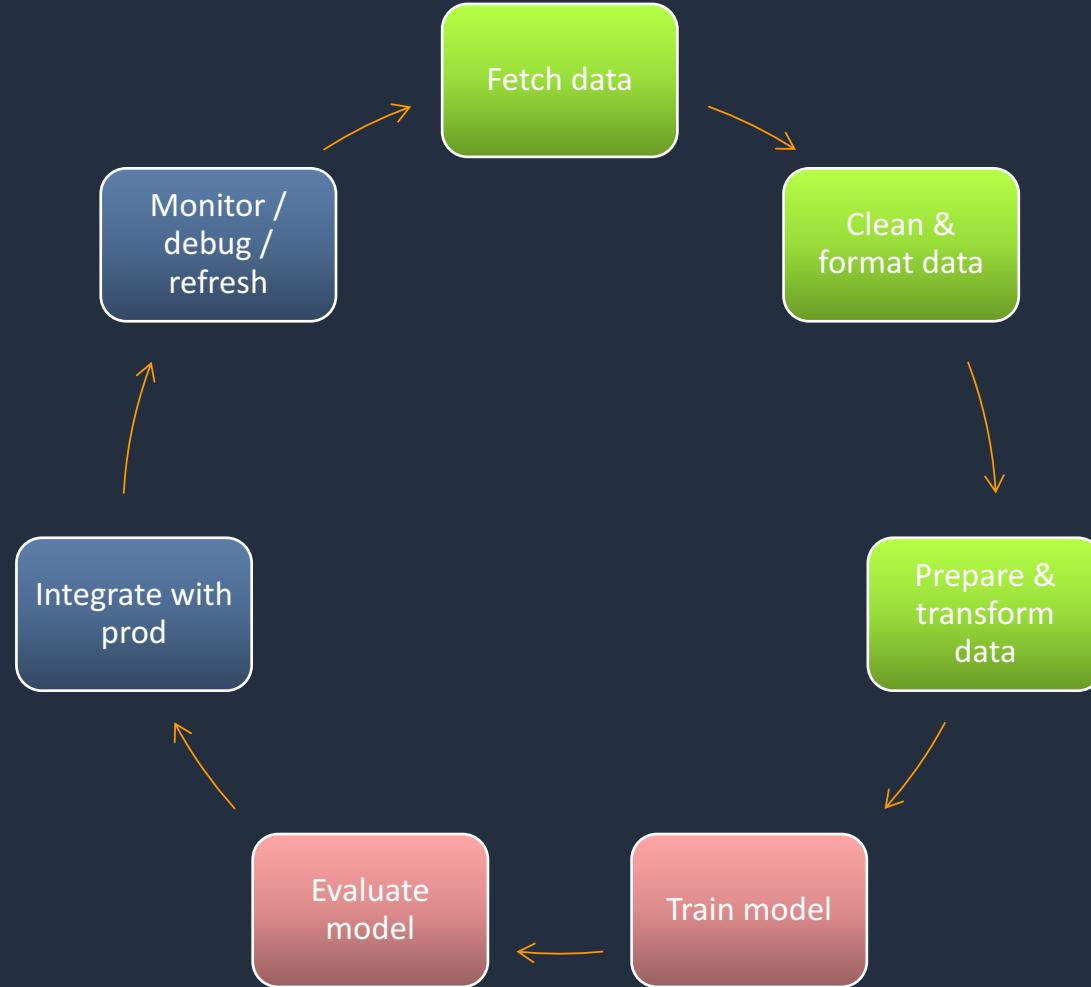
# Topics in this talk

- Amazon SageMaker features
- Customer stories
- Creating Jupyter Notebooks
- Performing hyperparameter tuning
- Hosting production models securely

# The machine learning process

## Deployment

- Set up and manage inference clusters
- Manage and auto scale inference APIs
- Testing, versioning, and monitoring



## Data wrangling

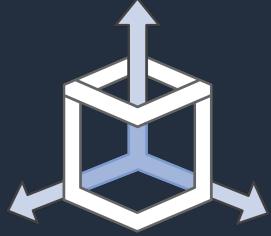
- Set up and manage data workbench
- Get data to workbench securely

## Experimentation

- Set up and manage clusters
- Scale / distribute ML algorithms



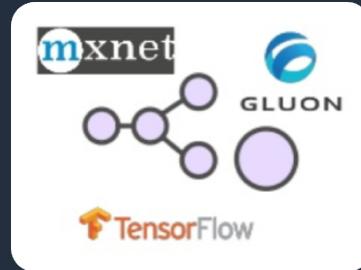
# Introducing Amazon SageMaker



End-to-end  
machine  
learning  
platform



Zero  
setup



Flexible  
model  
training



Pay by  
the  
second

# Real-time Fraud Detection with Amazon SageMaker

AI and ML at **Intuit** have three areas of focus.

- ✓ Smart Products
- ✓ Fraud detection and prevention
- ✓ Customer Care and Expert advice



In order to keep fraudsters out of their systems and data, Intuit always stays several moves ahead by leveraging AI/ML-generated insights from data that can determine **real-time fraud detection** in TurboTax: Specifically

- Account take-over detection at login
- Identity theft detection at filing



# Improve customer service with Amazon SageMaker

Zendesk builds software for better customer relationships. It empowers organizations to improve customer engagement and better understand their customers. More than 94,000 paid customer accounts in over 150 countries and territories use Zendesk products.



**“Amazon SageMaker will lower our costs and increase velocity for our use of machine learning. With Amazon SageMaker, we can transition from our existing self-managed TensorFlow deployment to a fully-managed service. Amazon SageMaker also gives us easier access to other popular deep-learning frameworks, while managing the infrastructure for authoring, training and serving our models.”**

- David Bernstein, Director of Strategic Technology, Zendesk

# Customize your travel using Amazon SageMaker

Hotels.com is a leading global lodging brand operating 90 localized websites in 41 languages.



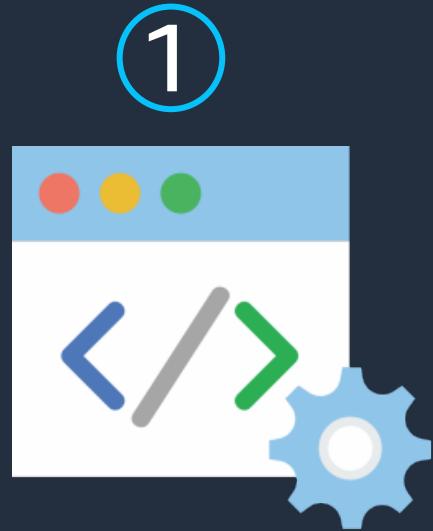
- Hotels.com is always interested in ways to move faster, to leverage the latest technologies and stay innovative.

**“With Amazon SageMaker, the distributed training, optimized algorithms, and built-in hyperparameter features should allow my team to quickly build more accurate models on our largest data sets, reducing the considerable time it takes us to move a model to production.”**

- Matt Fryer, VP and Chief Data Science Officer, Hotels.com & Expedia Affiliate Network



# Amazon SageMaker's four components



Notebook  
Instances



Algorithms



ML Training  
Service



ML Hosting  
Service

# Zero setup for data exploration

1



Resizable  
as you  
need



Common  
tools pre-  
installed



Easy  
access to  
your data  
sources



No  
servers to  
manage

# Demo

# Algorithms designed for scale ②



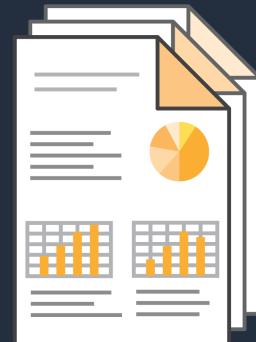
Streaming  
datasets,  
for  
cheaper  
training



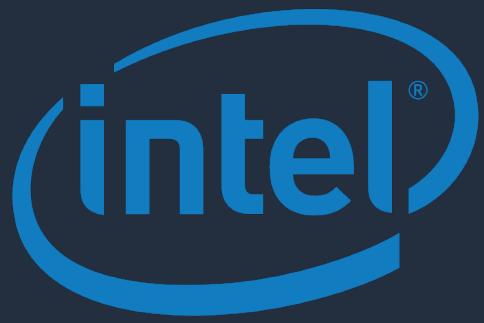
Train  
faster, in  
a single  
pass



Greater  
reliability  
on  
extremely  
large  
datasets



Choice of  
ML  
algorithms



Tuned for  
the Intel  
Chipset



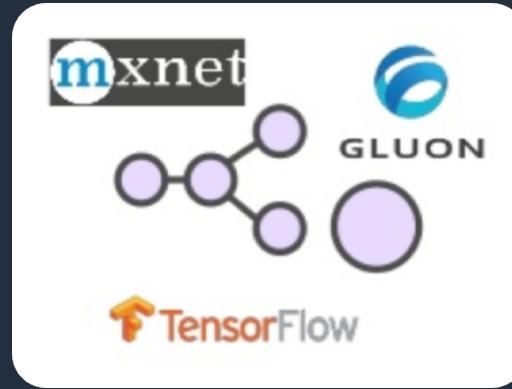
# Distributed training that works for you ③



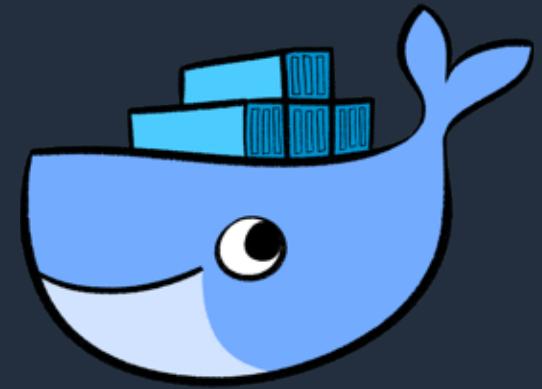
Amazon-optimized algorithms using the AWS SDK



Apache  
SparkML  
SageMaker  
Estimators



Bring your own deep learning script



Custom algorithm Docker images



# Hyperparameter optimization

- Bayesian optimization
- Tunes hyperparameters as a linear regression problem
- Test with 1 permutation at a time or hundreds

*Learning Rate,  
Regularization  
x Network Depth  
x Network Width  
x Dropout*

Points / Axis	Searching	2 HP	3 HP	4 HP	5 HP
2	Corners	4	8	16	32
3	Corners and Middle	9	27	81	243
5	Small variation	25	125	625	3,125
10	Reasonable range	100	1000	10,000	100,000



# Demo

# Quickly deploy to production ④

Launch



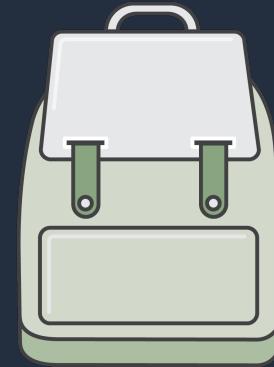
One step deployment



Low latency,  
high  
throughput,  
and high  
reliability



A/B  
testing



Use your  
own  
model

# Demo

# Amazon SageMaker

- End-to-end support for machine learning projects
- APIs for exploration, training, and hosting
- Automatically scaled, automatically secured



# INNOVATE2018

ONLINE CONFERENCE

DEVELOPER EDITION

# Thank you!

@awscloud



#AWSInnovate