# aws re: Invent

#### **AIM307**

# Amazon SageMaker – A Deep Dive

#### **Shyam Srinivasan**

Product Marketing Lead, ML Amazon Web Services

#### **Julien Simon**

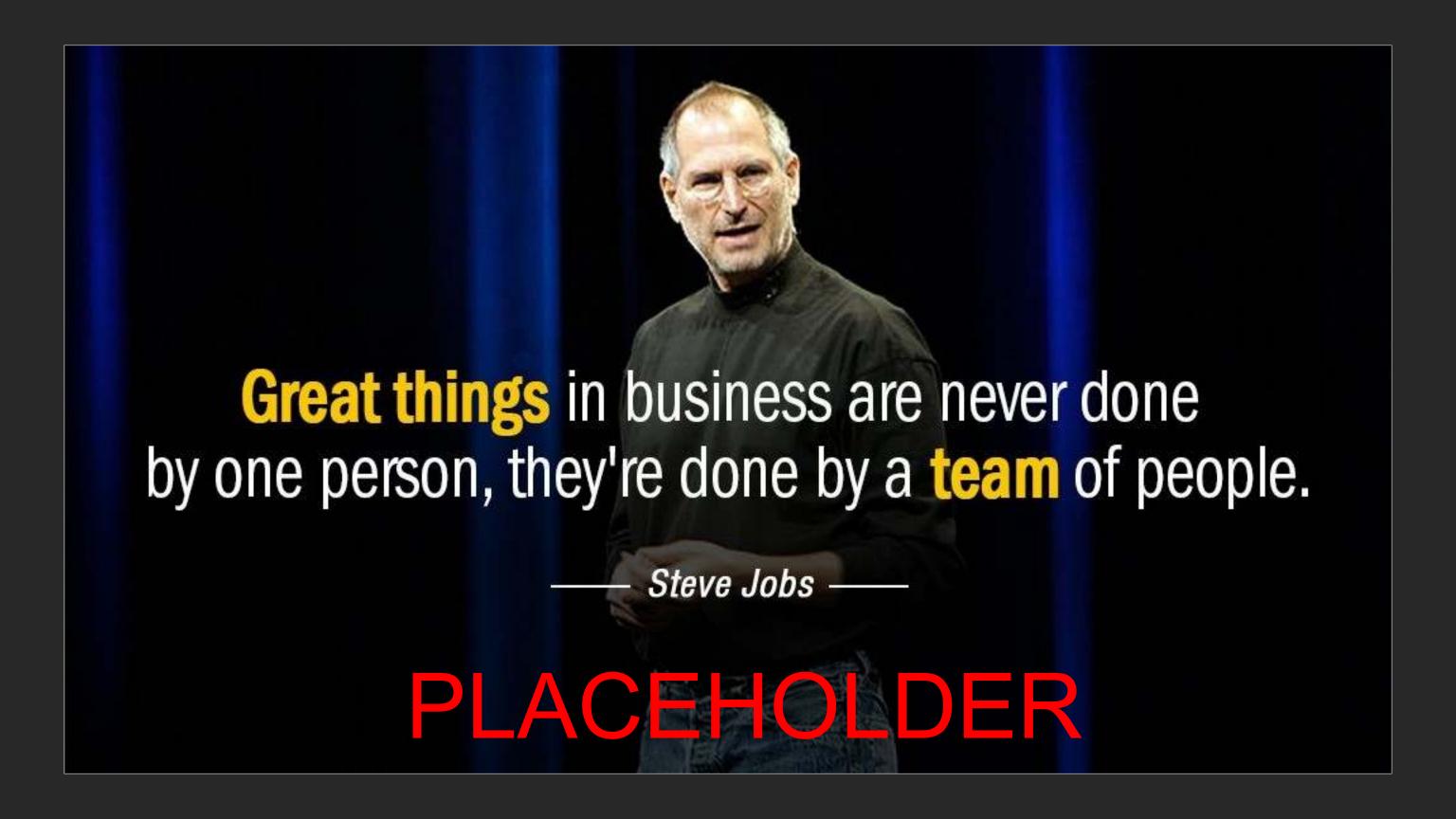
Global Evangelist, AI/ML Amazon Web Services

#### **Nils Mohr**

Flight Data Recorder Programming Analyst British Airways







### Agenda

What is Amazon SageMaker

Why Amazon SageMaker

The ML Workflow

Demo

Seeing is believing – Amazon SageMaker at British Airways

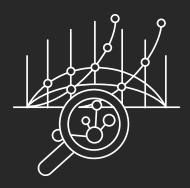


### Amazon SageMaker

Machine Learning for every developer & data scientist Build, Train, Deploy ML models at scale



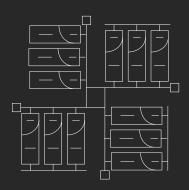
Customer-focused
90%+ of our ML roadmap is
defined by customers



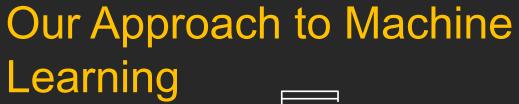
Pace of innovation
Hundreds of ML launches
and major feature updates last year

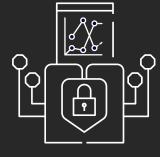


Breadth and depth
A wide range of Al and ML services



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

### Machine Learning is Hard



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 Getting models from concept to production

Pre-built notebooks

AWS Marketplace for ML

Amazon EC2 P3dn Instances Amazon Auto-scaling SageMaker Neo infrastructure

Amazon Elastic Inference

Amazon SageMaker Ground Truth Amazon SageMaker RL Managed Spot Training

Open source runtime

Multi-model endpoints

Operators for Kubernetes

Hundreds of notebook examples

Collect and prepare training data

Built-in, highperformance algorithms

Choose and optimize your ML algorithm

One-click training

Train and tune models

Model Optimization

Optimize models for the cloud and the edge

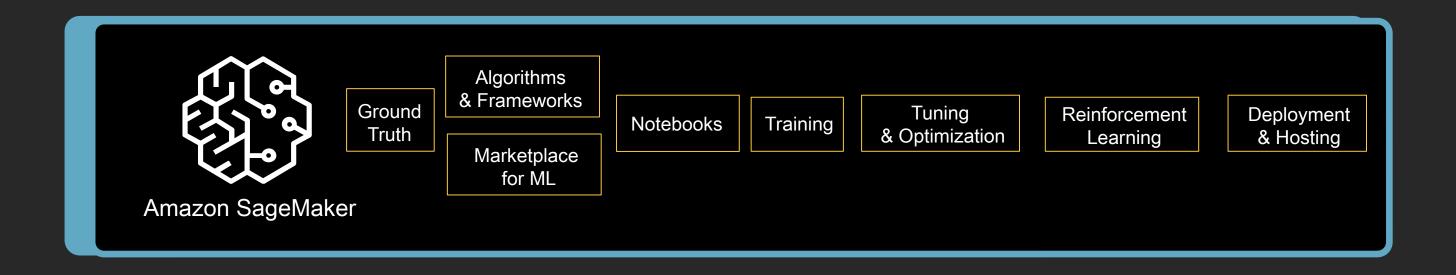
One-click deployment

Deploy models in production

Inference and inter-operability

Scale and manage the production environment

### Build, Train, Deploy Machine Learning Models Quickly at Scale





# Machine learning is iterative involving dozens of tools and hundreds of iterations

Multiple tools needed for different phases of ML process

+

Lack of an integrated experience

+

Large number of iterations

Ξ

Cumbersome, lengthy processes,, resulting in loss of productivity to get ML solutions to address business requirements



### Introducing Amazon SageMaker Studio

The first fully integrated development environment (IDE) for machine learning



Collaboration at scale

Share scalable notebooks without tracking code dependencies



Easy experiment management

Organize, track, and compare thousands of experiments



Automatic model generation

Get accurate models for with full visibility & control without writing code



Higher quality ML models



Increased productivity

Automatically debug errors, Code, build, train, deploy, & monitor models, & maintain monitor in a unified visual high quality interface

Sharing data science and the burden of setting up and managing resources

### Setup and manage resources

+

Code dependencies for collaboration

+

Lack of IDE for large data science teams

=

Reproducing and building on existing work is a significant time sink



### Introducing a new Amazon SageMaker notebook experience Setup free notebooks for collaboration at scale



Easy access with single sign-on

Access your notebooks in seconds with SSO



No setup of resources

Start your notebooks without spinning up compute



Full managed and secure

Administrators manage access and permissions



Collaboration at scale

with a single click



Scale on the go

Share with peers as a URL Scale up your notebooks without restart



# Debugging and profiling deep learning is painful

Large neural networks with many layers

₽

Many connections

+

Additional tooling for analysis and debug

Extraordinarily difficult to inspect, debug, and profile the 'black box'

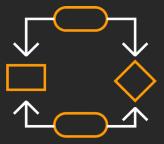


# Introducing Amazon SageMaker Debugger Analysis & debugging, explainability, and alert generation



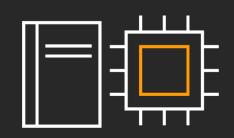
Relevant data capture

Data is automatically captured for analysis



Automatic data analysis

Debug data with no code changes



Automatic error detection

Errors are automatically detected based on rules



Improved productivity with alerts

Alerts are sent with violations, to take action



Visual analysis and debug

Visually analyze & debug from SageMaker Studio



# Managing trials and experiments is cumbersome

Hundreds of experiments

+

Hundreds of parameters per experiment

+

Compare and contrast

 $\equiv$ 

Very cumbersome and error prone



# Introducing Amazon SageMaker Experiments A system to organize, track, and evaluate training experiments



Experiment tracking at scale

Track parameters & metrics across experiments & users



Organize experiments

Organize by teams, goals, & hypotheses



Visualization for best results

Visualize & compare between experiments



Flexibility with Python SDK & APIs

Log custom metrics & track models using APIs



Iterate quickly

Iterate & develop highquality models



# Deploying a model is not the end, you need to continuously monitor it in production and iterate

Concept drift due to divergence of data

+

Model performance can change due to unknown factors

+

Continuous monitoring of model performance and data involves a lot of effort and expense

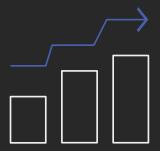
Model monitoring is cumbersome but critical



### Introducing Amazon SageMaker Model Monitor Continuous monitoring for deployed models



Automatic data collection



Visual data analysis



Higher model quality



CloudWatch Integration



**Flexibility** with rules

from endpoints

ata is automatically collected Data is analyzed based on Model quality is monitored and built-in rules or custom rules alerts are sent for deviations in quality

Act on emitted metrics on Amazon CloudWatch

Use built-in rules or custom rules for monitorin



# Successful ML requires complex, hard to discover combinations

of algorithms, data, parameters

Largely explorative & iterative

+

Requires broad and complete knowledge of ML domain

Lack of visibility

Time consuming, error prone process even for ML experts



# Introducing Amazon SageMaker Autopilot Automatic model creation with full visibility & control



Quick to start

Provide your data in a tabular form & specify target prediction



Automatic model creation

Feature engineering & automatic model tuning are automatically done



Transparency

Get the model notebook with source code



Recommendations & Optimization

Get recommendations on a leaderboard & continue to better your model

### Build, Train, Deploy Machine Learning Models Quickly at Scale



## TO BE UPDATED

## Demo





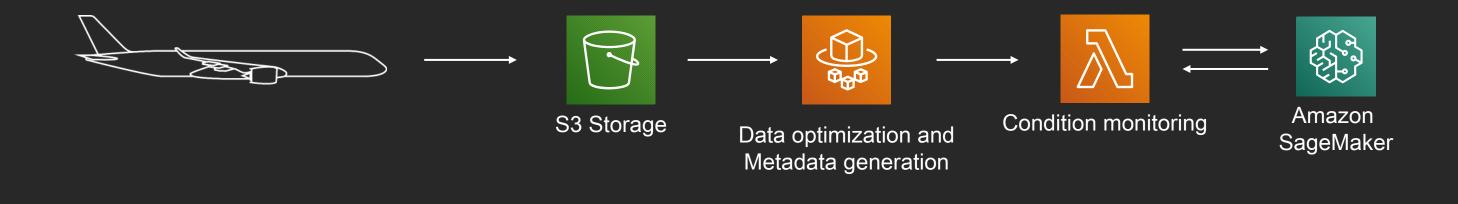
# Machine learning at British Airways Engineering







### Architecture & Intentions



Aeroplan	Dat	Flight	Rout	Condition	Condition	Condition
е	е	No	е	a	b	С
G-	12 <sup>th</sup>	BA	$LH \longrightarrow LA$		•	
ZBWW	Nay 13 <sup>H</sup>	B12	B → Sh		•	
ZBWW	Nay	B13	Sh → JF	•	•	
ZBWW	Nay	132 BA	<b>晁→</b> 仏	•		
ZBWW	Nov	133	K R			



## Train, Test, and Deploy





### Using built in algorithms

### Unsupervised

K-Means
 Clustering of single sectors

Random Cut Forest
 Anomaly detection on data arrays

### Supervised

XGBoost

Further classification/root cause detection

DeepAR

Timeseries forecasting



### Our dataset

**ACARS** 

QX QUKMEBA LONKFBA

.LONAXBA 221134

OFF01

BA414 G-NEOV 221134

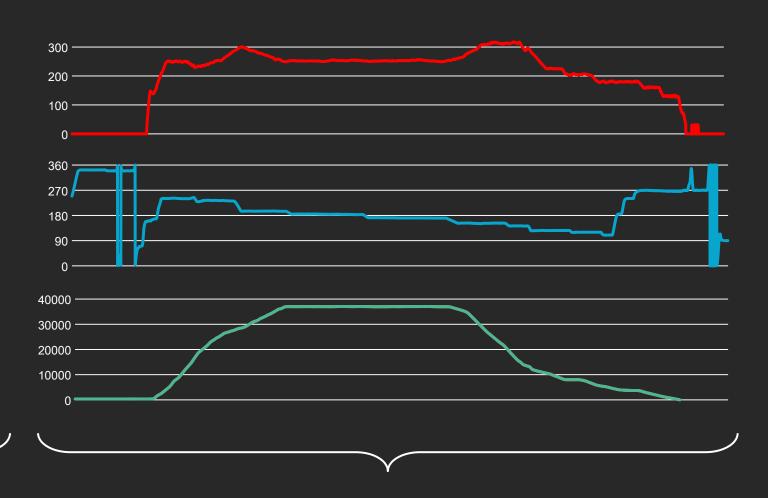
1134DIA02EGLLEDĎT154 15.4t 85.3t take-off

853

fuel

weight

#### Timeseries data

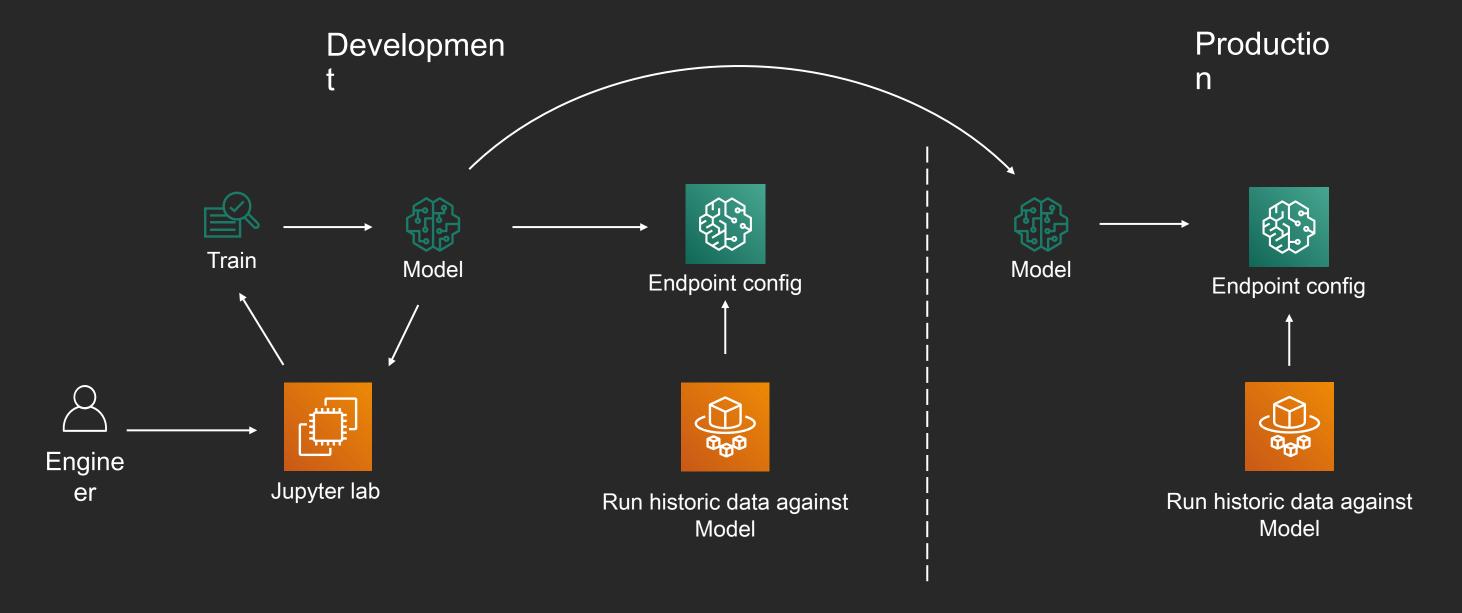


Single Values per flight: K-Means

Array of datapoints: RCF



### Build, train, and test in Amazon SageMaker



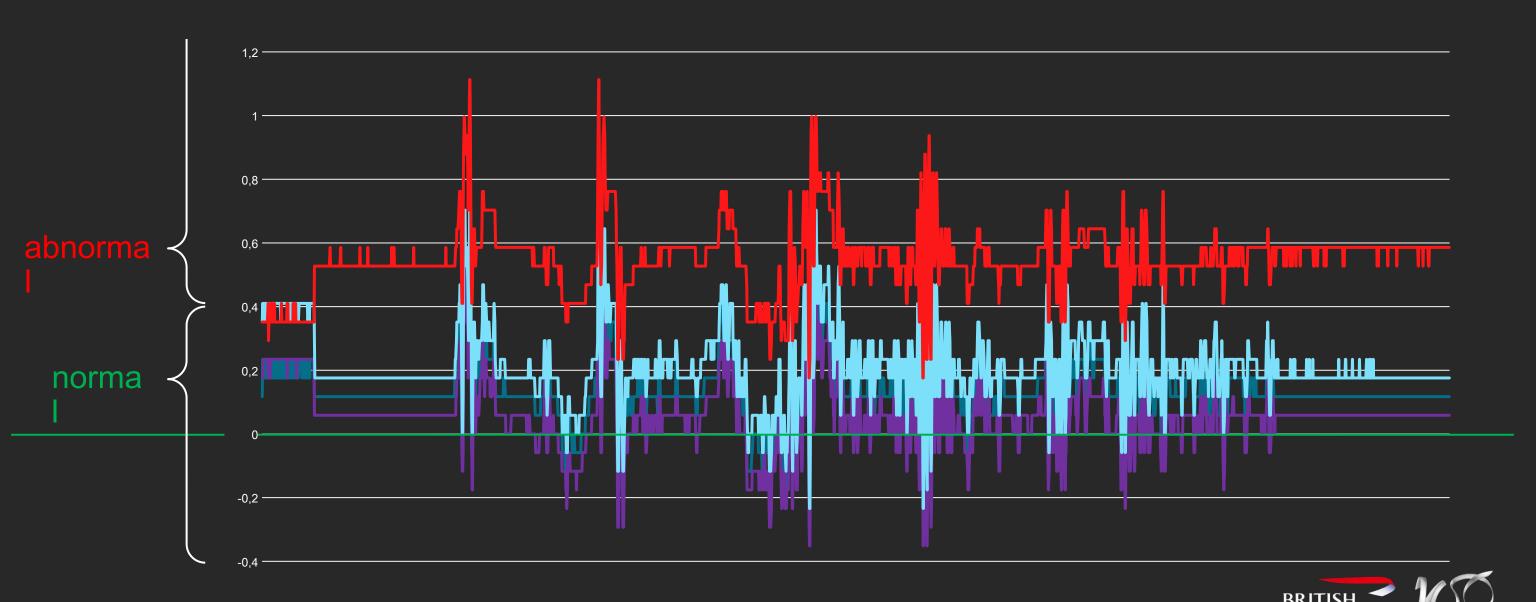


# Sample use case

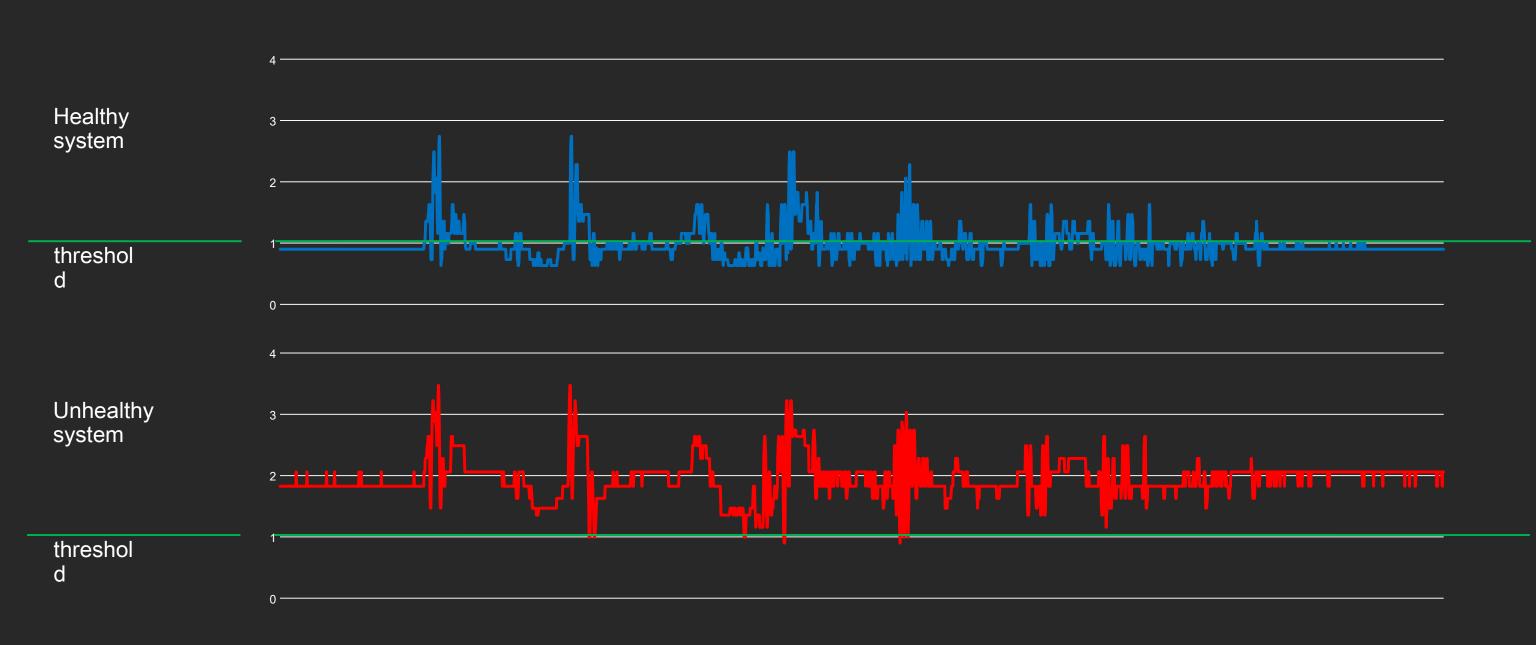




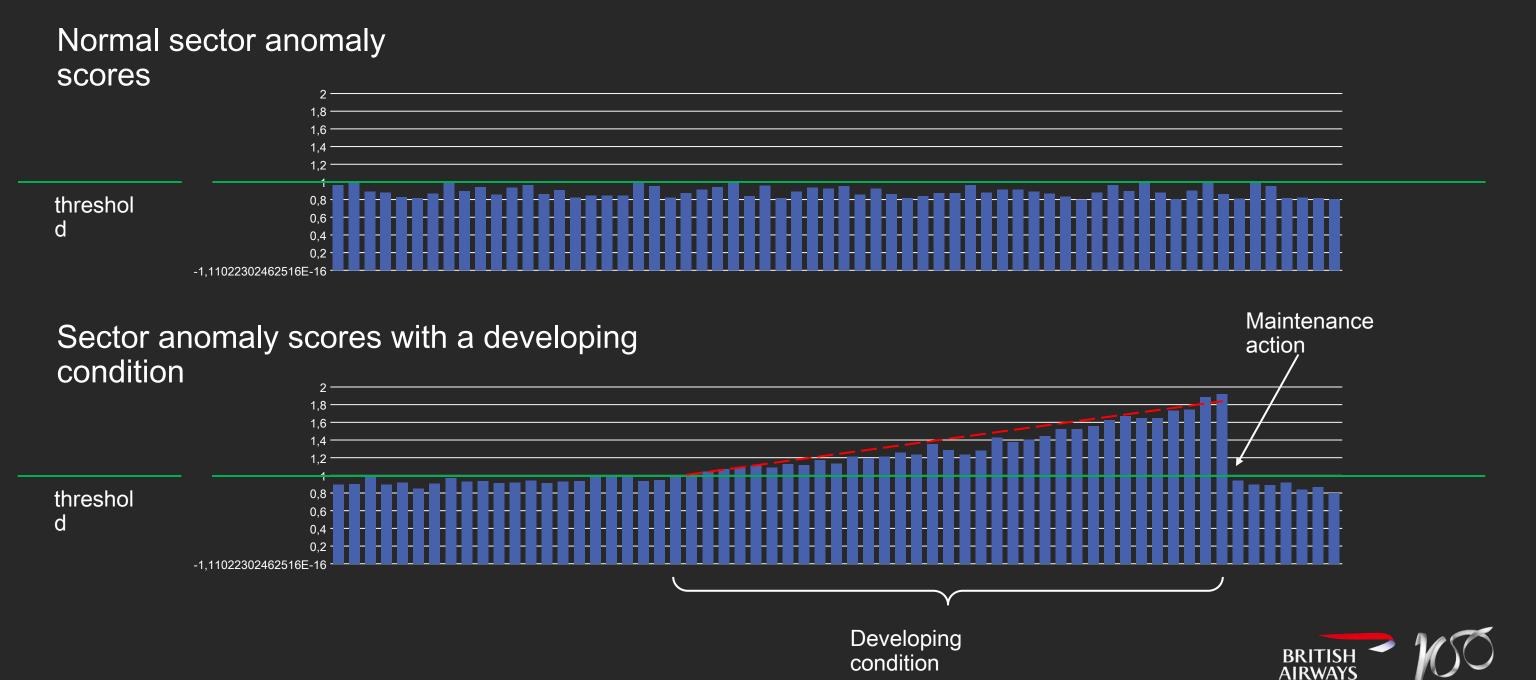
## Sample dataset



### Resulting SageMaker scores



### Evaluation of the data set in a test environment

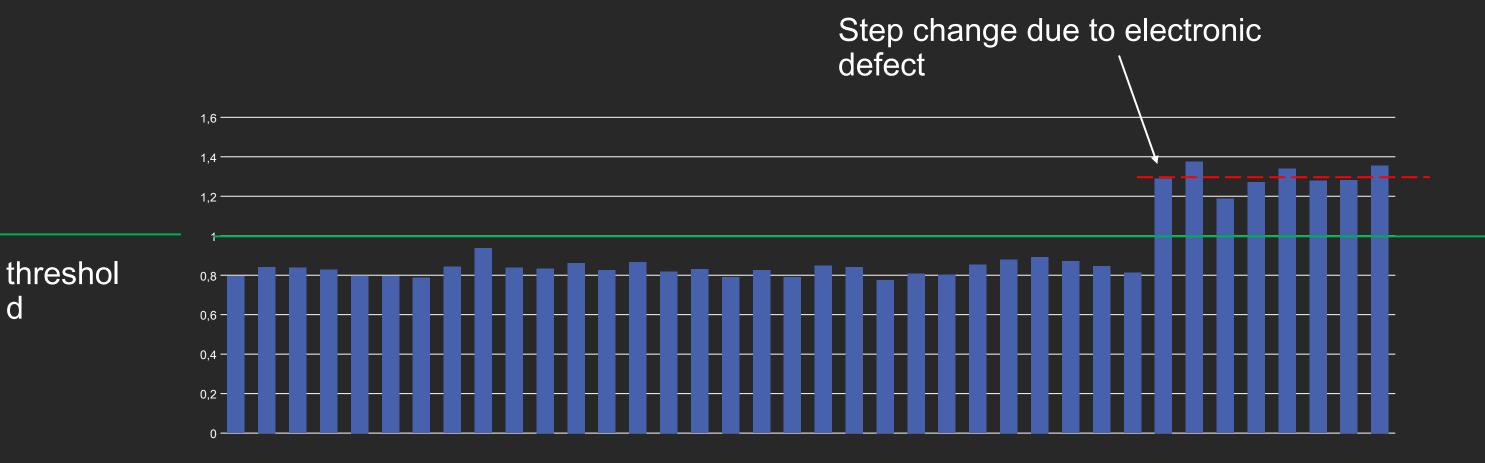


### Moving SageMaker into production

- About 4 flights earlier detection than a "traditional" detection mechanism
- SageMaker built in algorithms allow us to focus on our domain expertise
   generating flight data
- BUT: we found some false positives which turned out to be genuine events of a different type



### Expect the unexpected



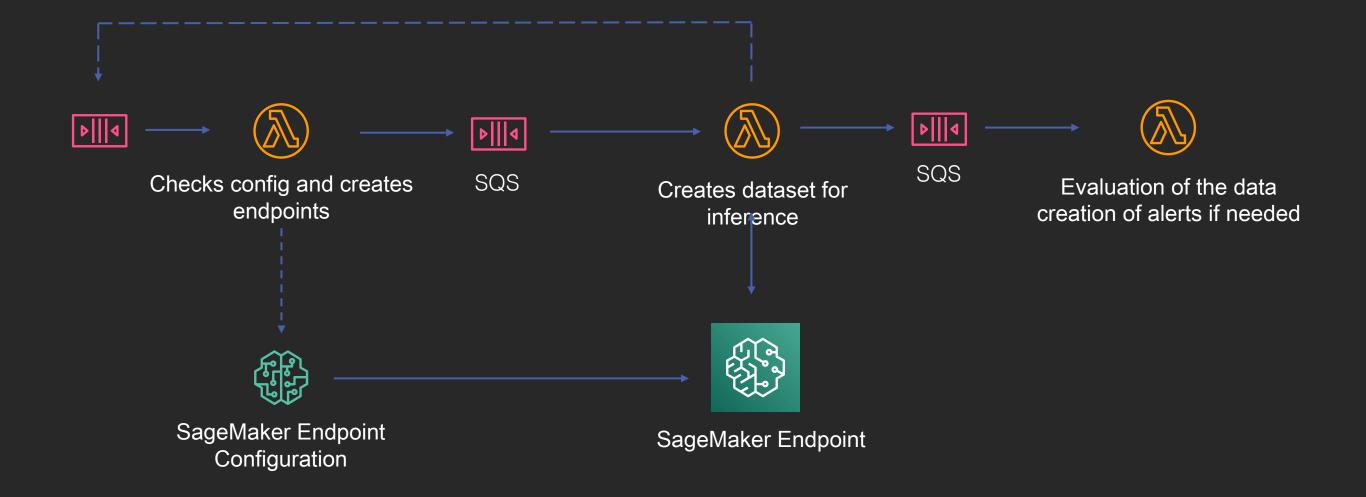


## Moving SageMaker to Production





### Detailed architecture





### Build, Train, Deploy Machine Learning Models Quickly at Scale



## TO BE UPDATED

### Related breakouts

[Session ID] [Session Title]

## TO BE UPDATED

# Thank you!

Shyam Srinivasan
Product Marketing Lead, ML
Amazon Web Services
srishyam@amazon.com

Julien Simon
Principal Evangelist, AI/ML
Amazon Web Services
julsimon@amazon.com







Please complete the session survey in the mobile app.



