Running MATLAB Machine Learning Jobs in Amazon SageMaker

Dr. Shun Mao, AWS

Gandharv Kashinath, MathWorks

MATLAB EXPO

Agenda

- Engineering applications and Machine Learning
- Designing a Pump Health Monitoring System
- Machine Learning using MATLAB in Amazon SageMaker (demo)
- Amazon Sagemaker Highlights

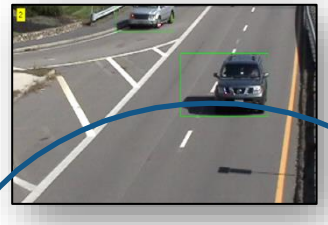
Agenda

- Engineering applications and Machine Learning
- Designing a Pump Health Monitoring System
- Machine Learning using MATLAB in Amazon SageMaker (demo)
- Amazon Sagemaker Highlights

Popular Machine Learning applications with MATLAB and Simulink

Solution is too complex for hand written rules or equations

Speech Recognition

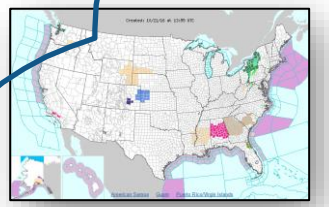


Object Recognition

Engine Health Monitoring

learn complex non-linear relationships

Solution needs to adapt with changing data



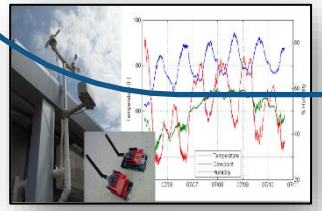
Weather Forecasting

Energy Load Forecasting

Stock Market Prediction

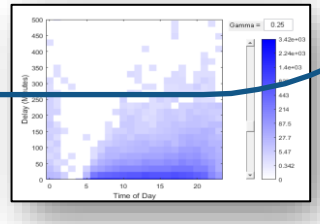
update as more data becomes available

Solution needs to scale



IoT Analytics

Taxi Availability



Airline Flight Delays

learn efficiently from very large data sets

Typical Machine Learning Workflow

Access Input Data

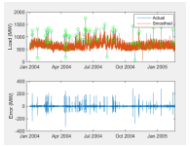
Files

Databases

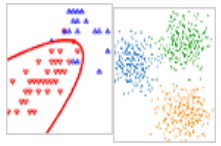
Sensors

Pre-Processing

Working with Messy Data



Data Reduction/ Transformation



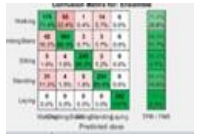
Feature Extraction

Model Training

Model Creation e.g. Machine Learning

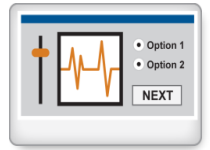
Parameter Optimization

Model Validation



Model deploy and testing

Desktop Apps



Enterprise Scale Systems

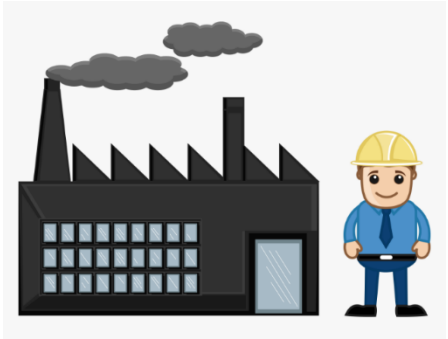
MATLAB Excel .NET .exe C/C++ Java .dll

Embedded Devices and Hardware

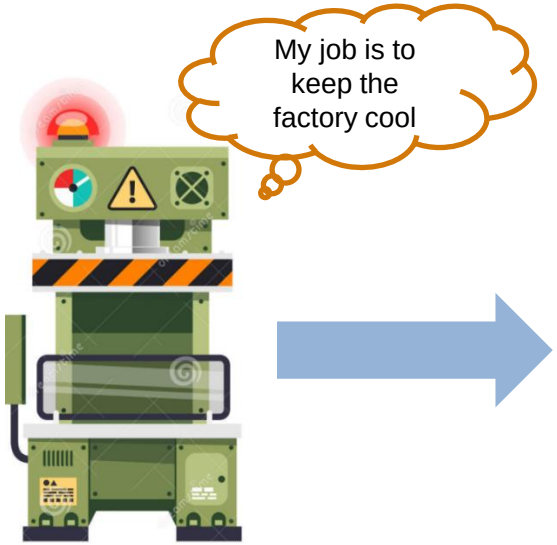
Agenda

- Engineering applications and Machine Learning
- Designing a Pump Health Monitoring System
- Machine Learning using MATLAB in Amazon SageMaker (demo)
- Amazon Sagemaker Highlights

Designing a pump health monitoring system at a manufacturing facility



- Alex is an automation engineer working at car manufacturing facility
- He is tasked with producing a ML solution to monitor the health of the cooling system at the facility
- One of the critical systems is a coolant pump, which has measurements of Pressure, Flow Rate and Motor RPM



Cooling system

Pressure

Flow Rate

Motor RPM

Sensor outputs streaming over time

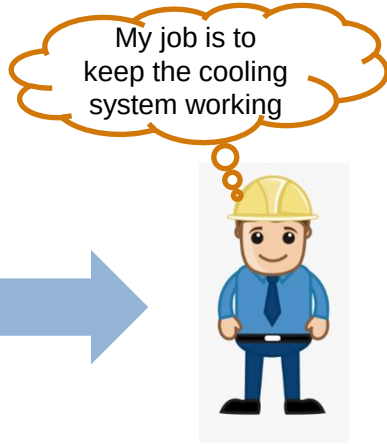
Normal (No Fault)

Leak Fault

Blocking Fault

Bearing Fault

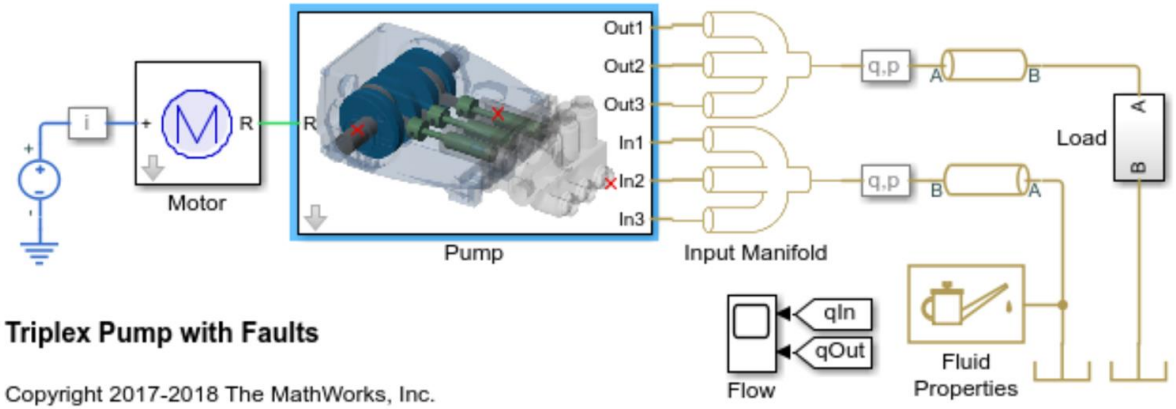
Health status determined from outputs



Monitor, take action!

Coolant pump modeled in Simulink to create synthetic data for training a machine learning model to predicts fault states

Simulink pump model



Numerical features extracted from time-series data

	fPeak	pLow	pMid	pHigh		Fault Code
1	43.9087	0.8468	117.7258	18.8553	27(	0
2	43.9087	0.4622	125.9801	18.9564	12.4	0
3	43.9087	1.1679	138.0097	17.5404	11.589	0
4	14.7785	235.2714	193.4912	26.7282	197.0199	1
5	14.7785	287.4136	198.7871	25.3212	487.5828	1
6	43.8481	4.3805	137.3085	19.1752	110.9272	6
7	14.8391	303.7370	176.3257	23.6653	392.3841	7
8	44.1509	133.9878	159.0942	26.9734	434.602	1

- Access Input Data
- Pre-Processing
- Model Training
- Model deploy and testing

Pressure

Flow Rate

Motor RPM

Normal (No Fault)

Leak Fault

Blocking Fault

Bearing Fault

Machine Learning models trained and tuned to predict fault codes using sensor measurements

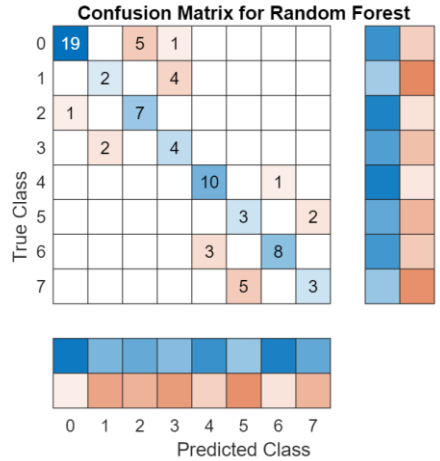
Numerical features extracted from time-series data

	fPeak	pLow	pMid	pHigh	
1	43.9087	0.8468	117.7258	18.8553	27(
2	43.9087	0.4622	125.9801	18.9564	12.4
3	43.9087	1.1679	138.0097	17.5404	11.589
4	14.7785	235.2714	193.4912	26.7282	197.0199
5	14.7785	287.4136	198.7871	25.3212	487.5828
6	43.8481	4.3805	137.3085	19.1752	110.9272
7	14.8391	303.7370	176.3257	23.6653	392.3841
8	44.1509	133.9878	159.0942	26.9734	434.602(

Fault Code
0
0
0
1
1
6
7
1

Random Forest classification model

Random Forest classification model



Access Input Data

Pre-Processing

Model Training

Model deploy and testing

Pressure

Flow Rate

Motor RPM

Normal (No Fault)

Leak Fault

Blocking Fault

Bearing Fault

Using MATLAB in the Cloud via Amazon Sagemaker makes it easy to monitor the health of the cooling system

Access Input Data

Pre-Processing

Model Training

Model deploy and testing

Amazon SageMaker

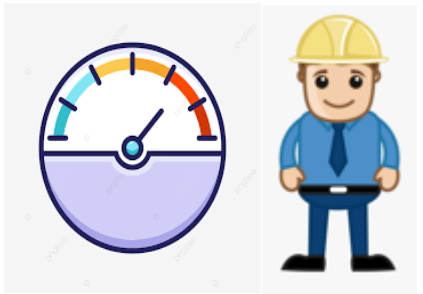
Model is trained in **SageMaker**

Model is deployed to **SageMaker** end-point

Amazon S3

Outputs cooling system health predictions to dashboard

MATLAB WebApp Server



Apache kafka

Human in the loop annotates labels showing fault states, uploads latest labeled dataset

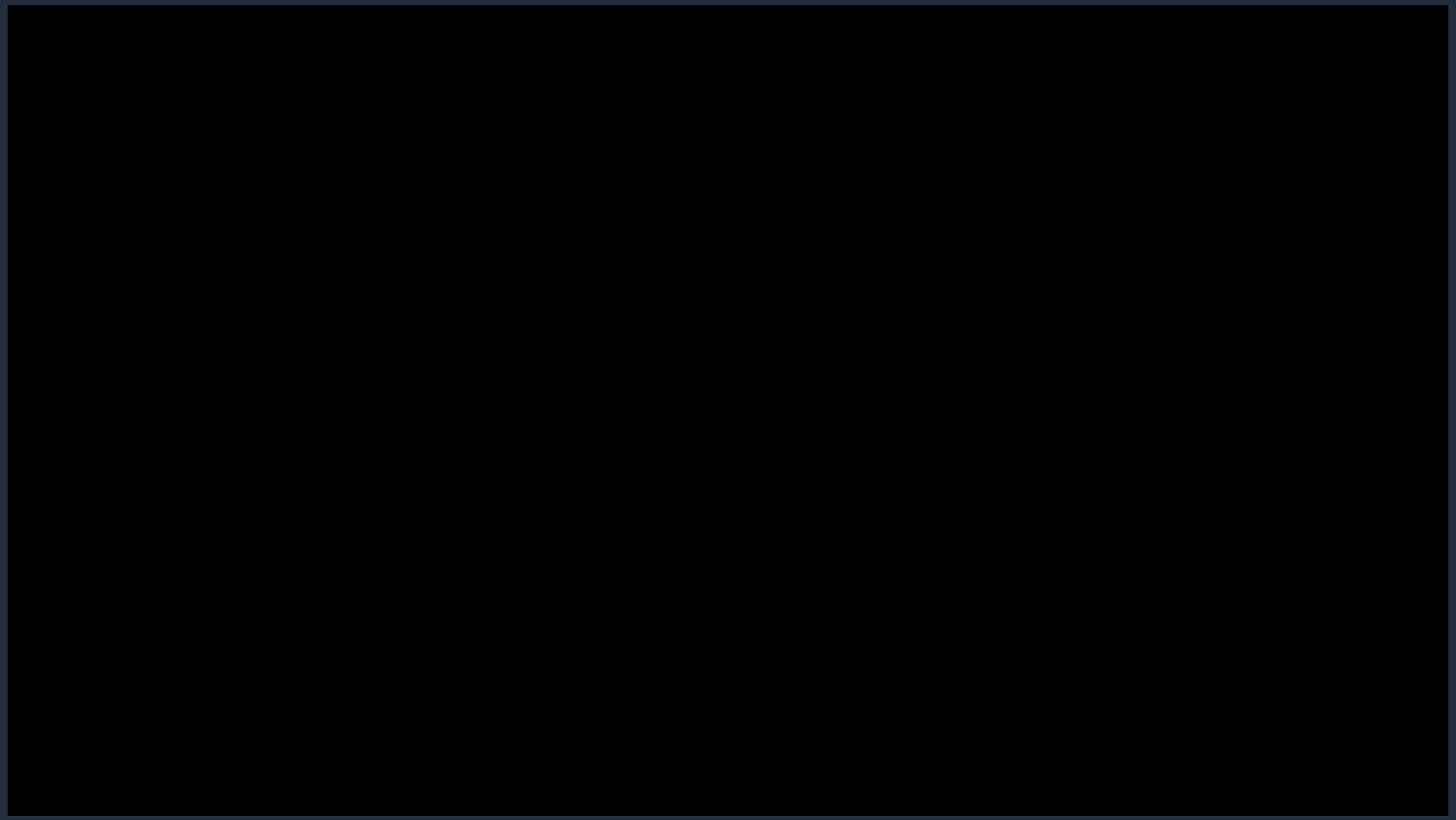
Sensor readings are uploaded to the cloud using kafka framework

Agenda

- Engineering applications and Machine Learning
- Designing a Pump Health Monitoring System
- Machine Learning using MATLAB in Amazon SageMaker (demo)
- Amazon Sagemaker Highlights

DEMO

Run MATLAB Machine Learning Jobs in Amazon Sagemaker



Agenda

- Engineering applications and Machine Learning
- Designing a Pump Health Monitoring System
- Machine Learning using MATLAB in Amazon SageMaker (demo)
- Amazon Sagemaker Highlights

Amazon SageMaker

A COMPREHENSIVE MACHINE LEARNING PLATFORM FOR VIRTUALLY ANY DATA AND ML USE CASES

Predictive maintenance

Manufacturing, automotive, IoT

Demand forecasting

Retail, consumer goods, manufacturing

Fraud detection

Financial services, online retail

Credit risk prediction

Financial services, retail

Extract and analyze data from documents

Healthcare, legal, media/ent, education

Computer vision

Healthcare, pharma, manufacturing

Autonomous driving

Automotive, transportation

Personalized recommendations

Media and entertainment, retail, education

Churn prediction

Retail, education, software and internet

Tens of thousands of customers use Amazon SageMaker

Amazon SageMaker Key feature tour

PREPARE DATA AND BUILD, TRAIN, AND DEPLOY ML MODEL FOR ANY USE CASE

Prepare Input Data

Geospatial
Visualize geospatial data

Ground Truth
Create high quality datasets for ML

Data Wrangler
Aggregate and prepare data for ML

Feature Store
Store, catalog, search, and reuse features

Processing

Studio Notebooks & Notebook Instances
Fully managed Jupyter notebooks with elastic compute

Studio Lab
Free ML development environment

Processing
Built-in Python, BYO R/Spark

JumpStart
UI based discovery, training, and deployment of models, solutions, and examples

Model Training

Fully Managed Training
Broad hardware options, easy to setup and scale

Distributed Training Libraries
High performance training for large datasets and models

Training Compiler
Faster deep learning model training

Automatic Model Tuning
Hyperparameter optimization

Managed Spot Training
Reduce training cost by up to 90%

Bring Your Own
Bring your own container and algorithms

Model deploy and testing

Fully Managed Deployment
Ultra low latency, high throughput inference

Real-Time Inference
For steady traffic patterns

Serverless Inference
For intermittent traffic patterns

Asynchronous Inference
For large payloads or long processing times

Batch Transform
For offline inference on batches of large datasets

Multi-Model Endpoints
Reduce cost by hosting multiple models per instance

Multi-Container Endpoints
Reduce cost by hosting multiple containers per instance

Shadow Testing
Validate model performance in production

Inference Recommender
Automatically select compute instance and configuration

Model Monitor
Maintain accuracy of deployed models

MLOps: Pipelines | Projects | Model Registry
Workflow automation, CI/CD for ML, central model catalog

Canvas
Generate accurate machine learning predictions—no code required

Studio | RStudio
Integrated development environment (IDE) for ML

Governance
Model Cards | Dashboard | Permissions

Summary

- MATLAB is the best-in-class ML tool for engineering systems
- Amazon Sagemaker platform can be used for any ML and data use cases
- Leverage MATLAB in Amazon Sagemaker for your data and ML use cases

Reach out to us if you want to run MATLAB in Amazon SageMaker

cloud@mathworks.com

Dr. Shun Mao, AWS

Gandharv Kashinath, MathWorks

MathWorks®

MATLAB EXPO

Thank you