MATLAB EXPO 2016
KOREA

4월 28일 (목)

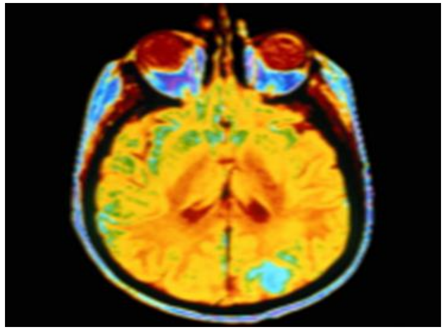
등록 하기 matlabexpo.co.kr

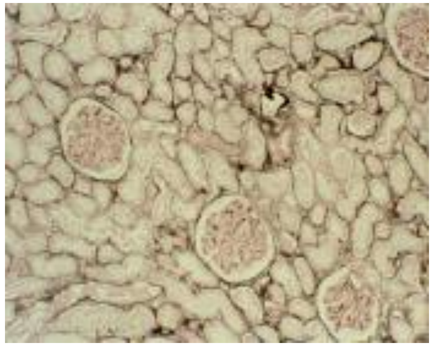
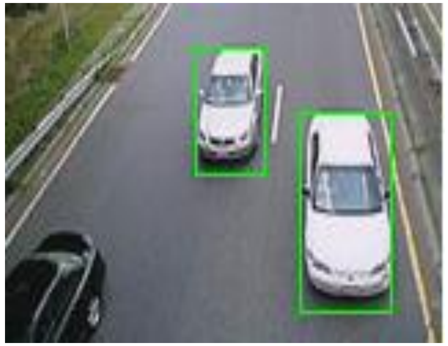
데이터 애널리틱스를 위한 머신 러닝 기법

Application Engineer
엄 준 상 과장

Machine Learning is Everywhere

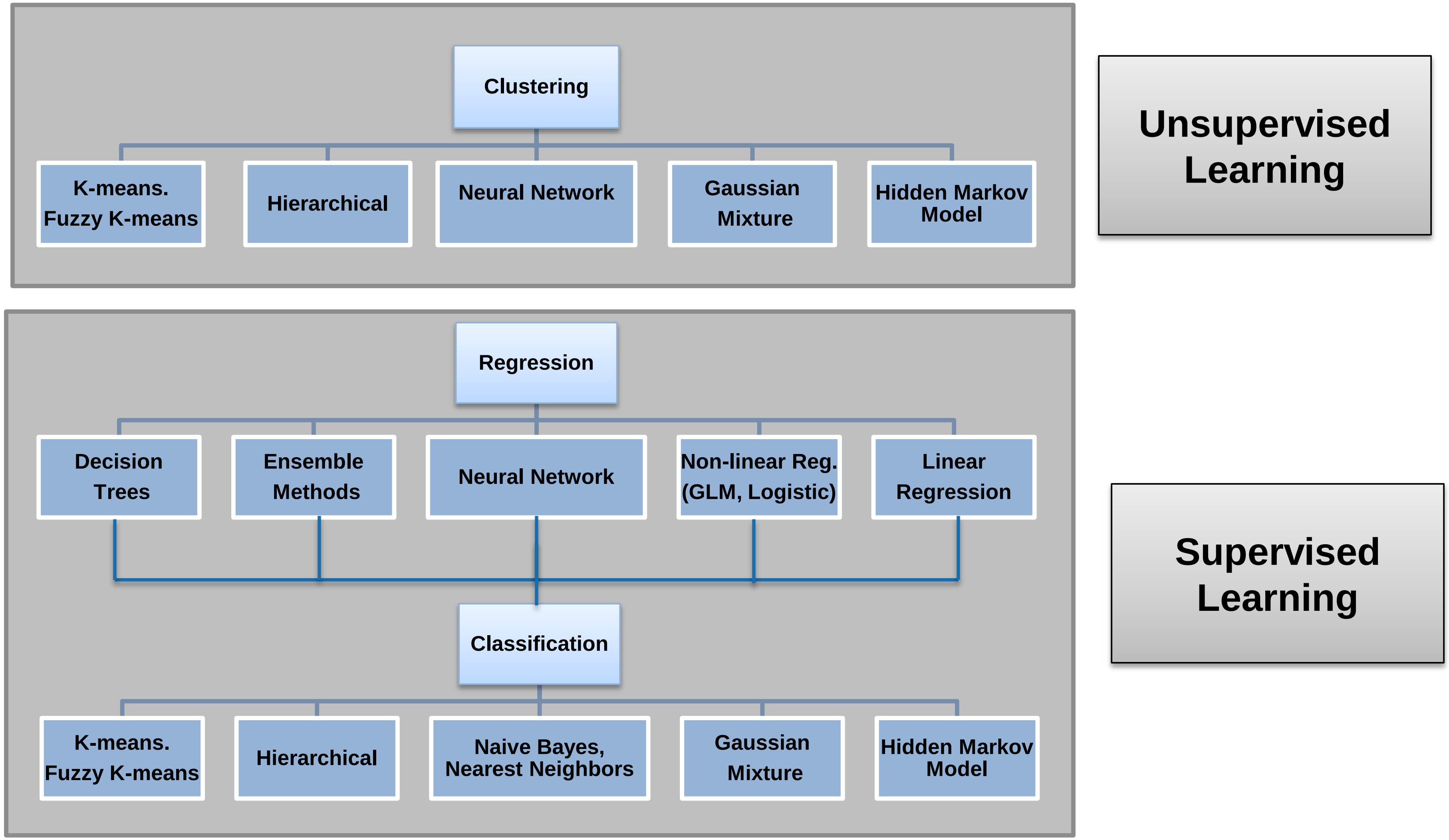
- Image Recognition
- Speech Recognition
- Stock Prediction
- Medical Diagnosis
- Data Analytics
- Robotics
- and more…

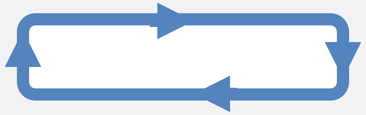


Overview of Machine Learning in MATLAB



Challenges in Machine Learning

Hard to get started

Steps	Challenge
Access, explore and analyze data	**Data diversity** Numeric, Images, Signals, Text – not always tabular
Preprocess data	**Lack of domain tools** Filtering and feature extraction Feature selection and transformation
Train models	**Time consuming** Train many models to find the "best"
Assess model performance	**Avoid pitfalls** Over Fitting Speed-Accuracy-Complexity tradeoffs
Iterate	

Faulty braking system leads to windmill disaster

Why perform predictive maintenance?

- Example: faulty braking system leads to windmill disaster
  - https://youtu.be/-YJuFvjtM0s?t=39s

- Wind turbines cost millions of dollars

- Failures can be dangerous

- Maintenance also very expensive and dangerous

Types of Maintenance

- Reactive – Do maintenance once there’s a problem
  - Example: replace car battery when it has a problem
  - Problem: unexpected failures can be expensive and potentially dangerous

- Scheduled – Do maintenance at a regular rate
  - Example: change car’s oil every 5,000 miles
  - Problem: unnecessary maintenance can be wasteful; may not eliminate all failures

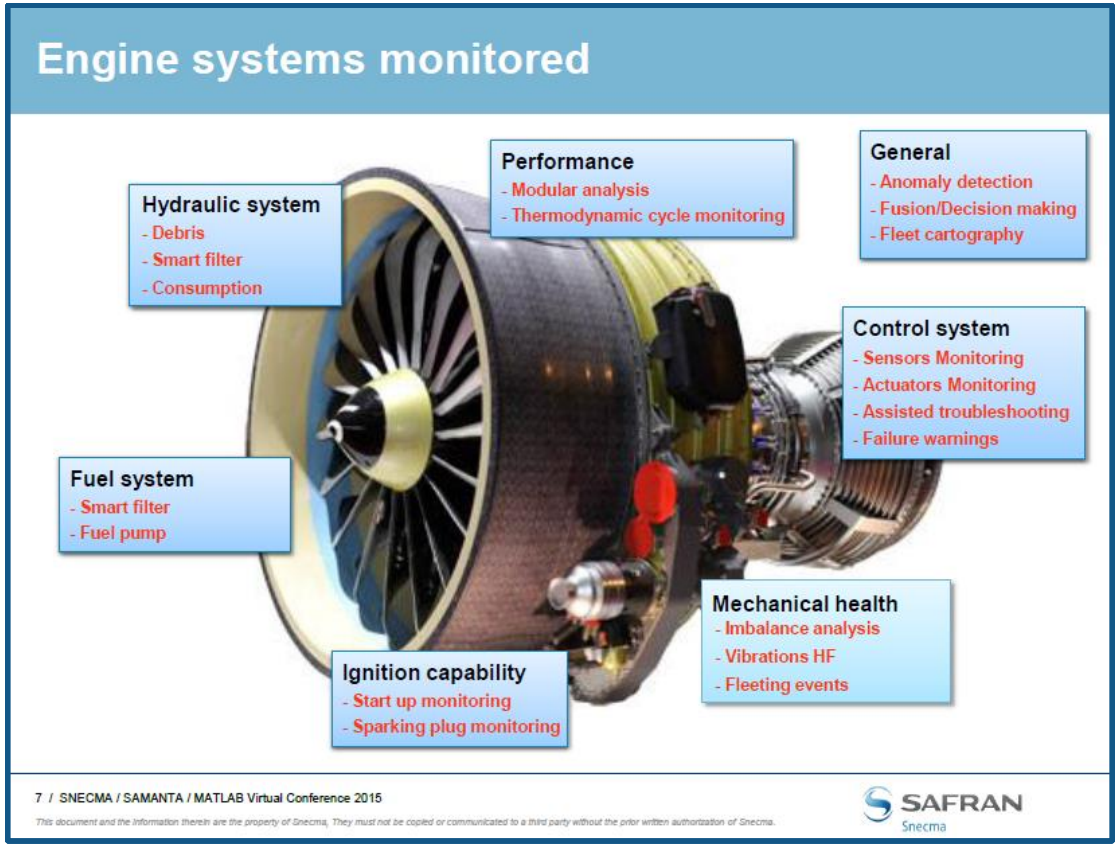
- Predictive – Forecast when problems will arise
  - Example: certain GM car models forecast problems with the battery, fuel pump, and starter motor
  - Problem: difficult to make accurate forecasts for complex equipment

Benefits of Predictive Maintenance

Increase “up time” and safety	**Reliability**
Minimize maintenance costs	**Cost of Ownership**
Optimize supply chain	**Reputation**

What Does Success Look Like?

Safran Engine Health Monitoring Solution



- Monitor Systems
  - Detect failure indicators
  - Predict time to maintenance
  - Identify components
- Improve Aircraft Availability
  - On time departures and arrivals
  - Plan and optimize maintenance
  - Reduce engine out-of-service time
- Reduce Maintenance Costs
  - Troubleshooting assistance
  - Limit secondary damage

Desktop
- Ad-hoc data analysis
- Analytics to predict failure

Compiled Shared
- Suite of MATLAB Analytics
- Shared with other teams
- Proof of readiness

Enterprise Integration
- Real-time analytics
- Integrated with maintenance and service systems

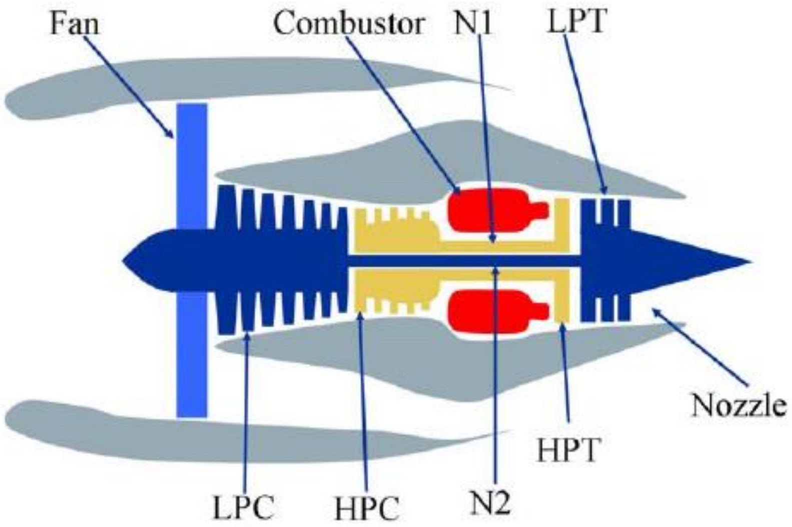
http://www.mathworks.com/company/events/conferences/matlab-virtual-conference/

Predictive Maintenance of Turbofan Engine

Sensor data from 100 engines of the same model

Predict and fix failures before they arise

- Import and analyze historical sensor data
- Train model to predict when failures will occur
- Deploy model to run on live sensor data
- Predict failures in real time



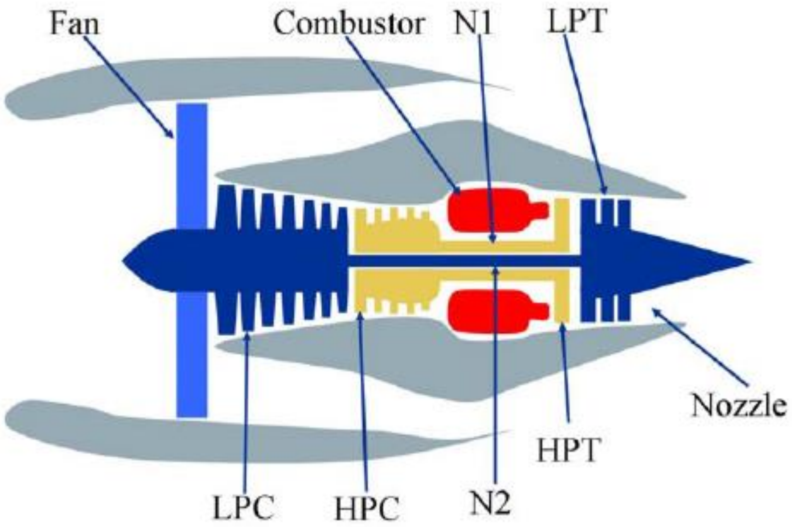
Data provided by NASA PCoE
http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/

Predictive Maintenance of Turbofan Engine

Sensor data from 100 engines of the same model

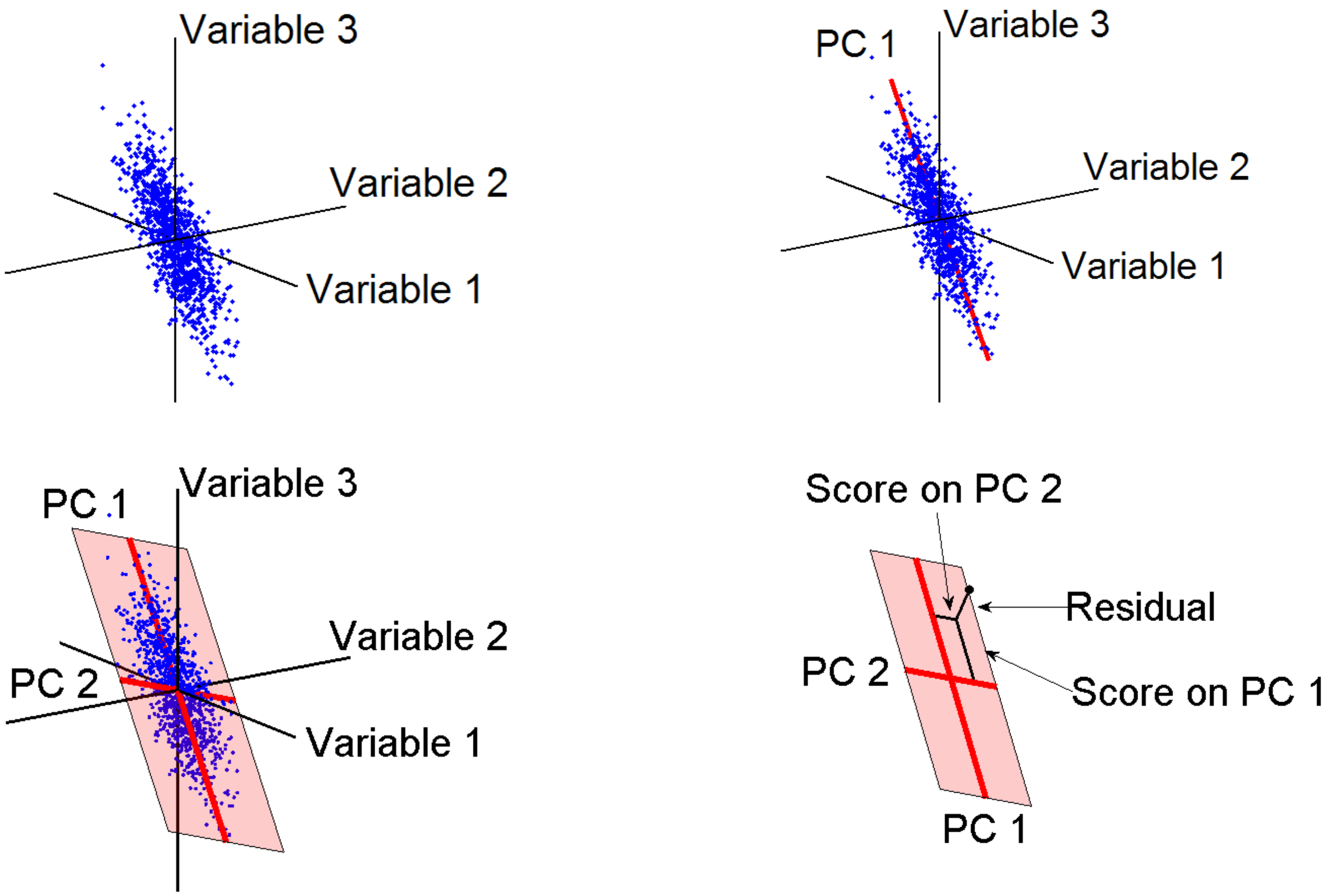
Scenario 1: No data from failures

- Performing scheduled maintenance
- No failures have occurred
- Maintenance crews tell us most engines could run for longer
- Can we be smarter about how to schedule maintenance **without** knowing what failure looks like?

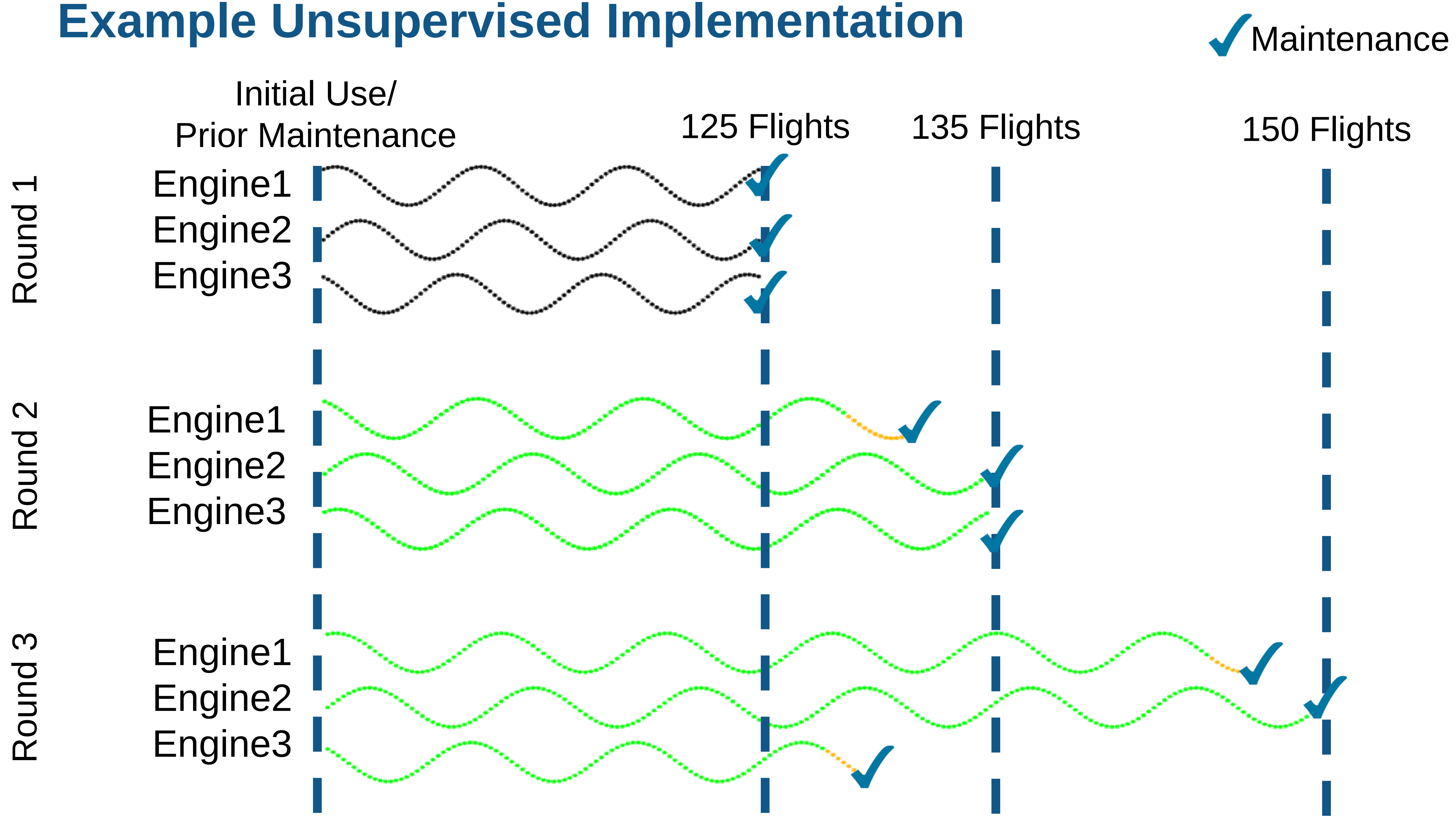


Data provided by NASA PCoE
http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/

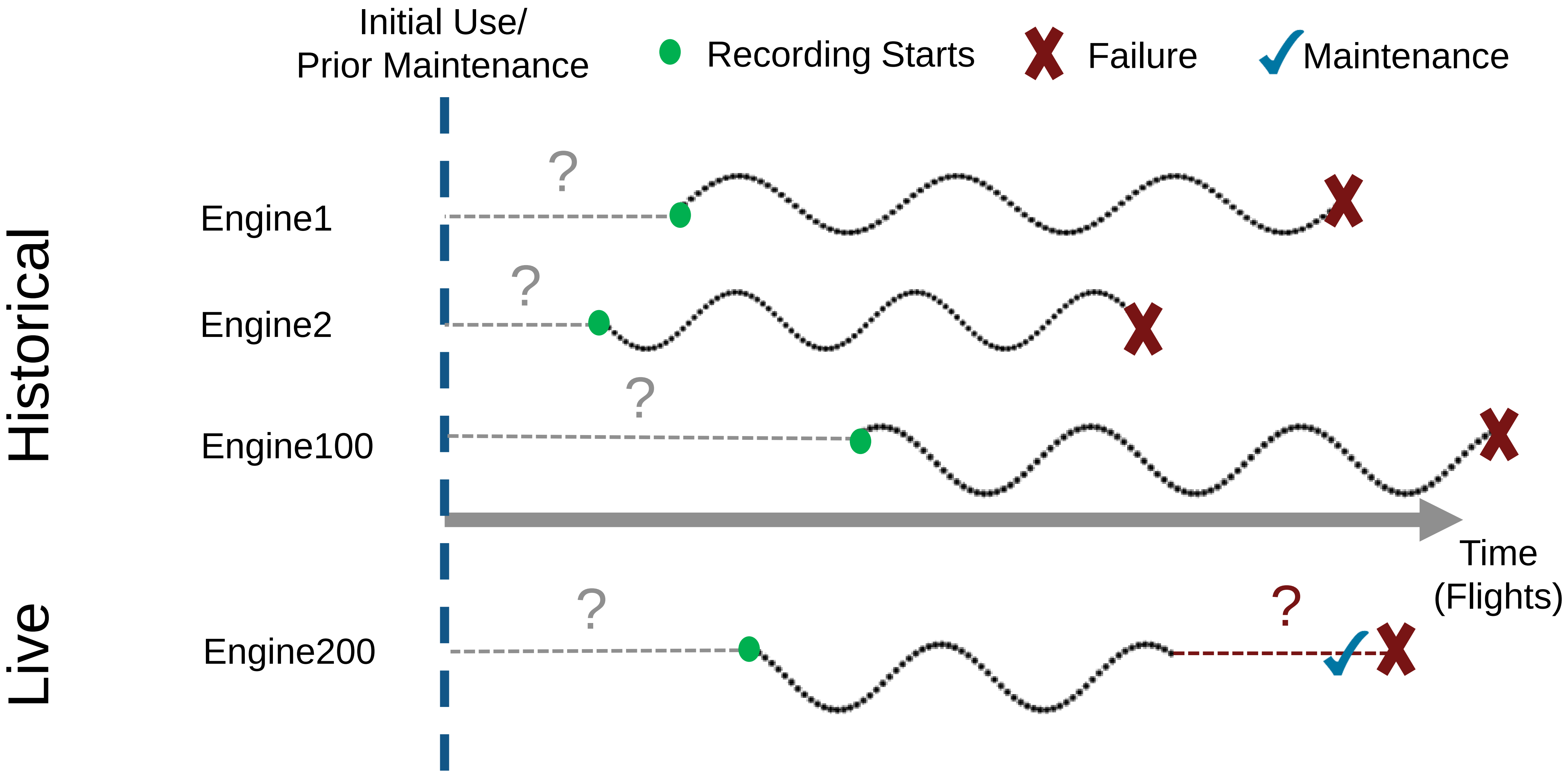
Principal Components Analysis – what is it doing?



Example Unsupervised Implementation



How Data was Recorded

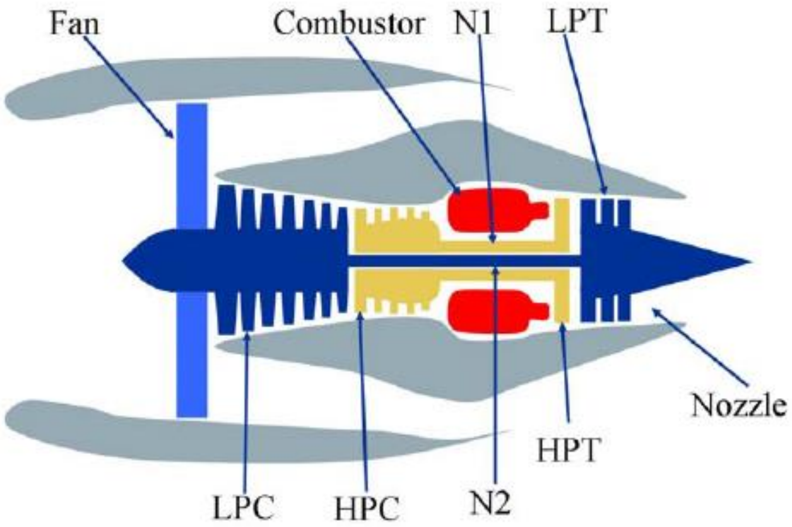


Predictive Maintenance of Turbofan Engine

Sensor data from 100 engines of the same model

Scenario 2: Have failure data

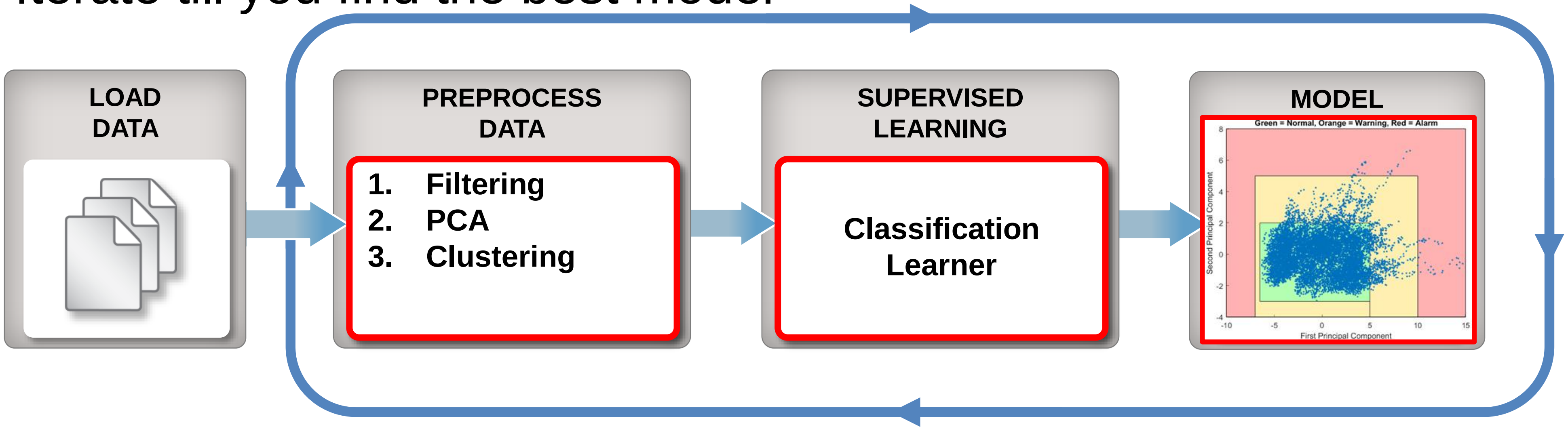
- Performing scheduled maintenance
- Failures still occurring (maybe by design)
- Search records for when failures occurred and gather data preceding the failure events
- Can we predict how long until failures will occur?



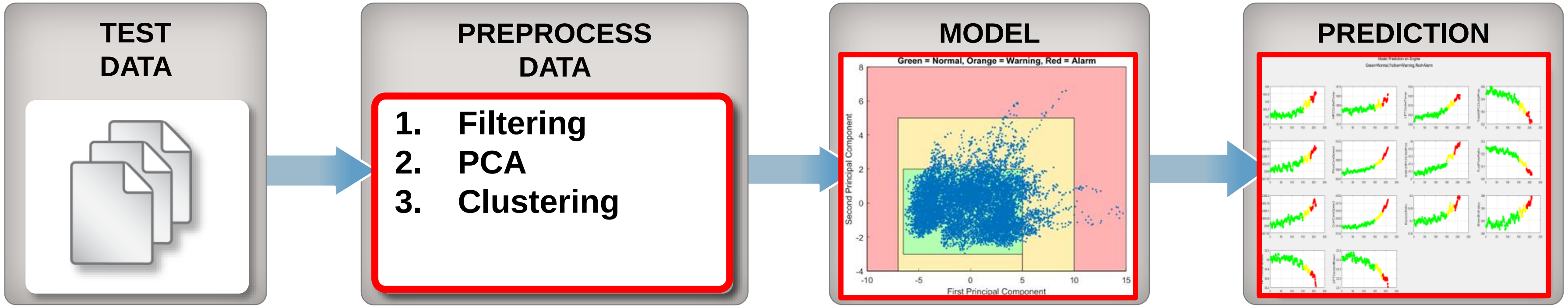
Data provided by NASA PCoE
http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/

Machine Learning Workflo

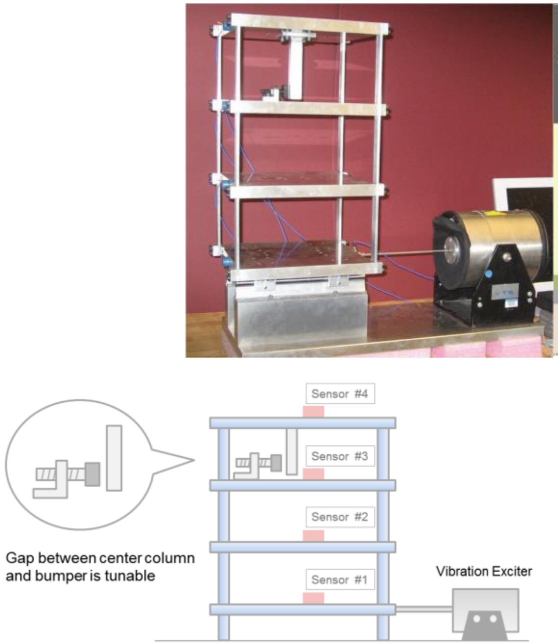
Train: Iterate till you find the best model



Predict: Integrate trained models into applications

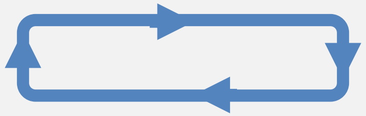


SHMTools
Los Alamos National Laboratory



http://www.lanl.gov/projects/national-security-education-center/engineering/software/shm-data-sets-and-software.php

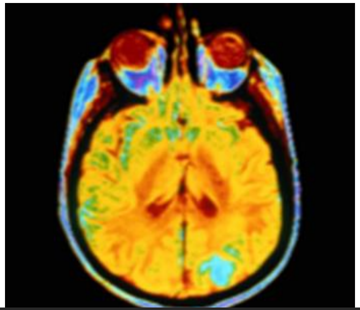
MATLAB Strengths for Machine Learning

Challenge	Solution
Data diversity	**Extensive data support** Import and work with signal, images, financial, Textual, geospatial, and several others formats
Lack of domain tools	**High-quality libraries** Industry-standard algorithms for Finance, Statistics, Signal, Image processing & more
Time consuming	**Interactive, app-driven workflows** Focus on machine learning, not programing
Avoid pitfalls Over Fitting, Speed-Accuracy-Complexity	**Integrated best practices** Model validation tools built into app Rich documentation with step by step guidance
	**Flexible architecture for customized workflows** Complete machine learning platform

Deep Learning is Ubiquitous

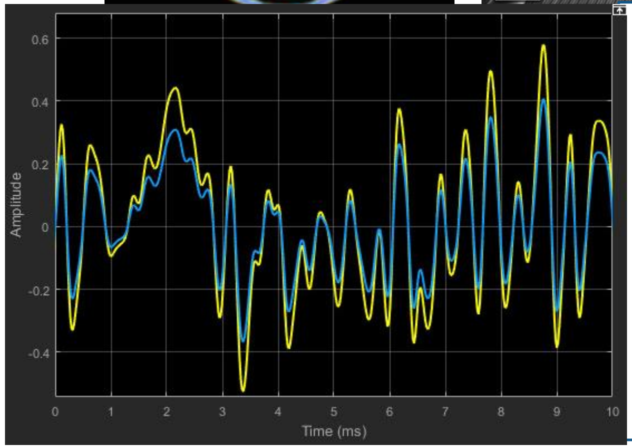
Computer Vision

- Pedestrian and traffic sign detection
- Landmark identification
- Scene recognition
- Medical diagnosis and drug discovery




Text and Signal Processing

- Speech Recognition
- Speech & Text Translation



Robotics & Controls

and many more…

What is Deep Learning ?

Deep learning performs **end-end learning** by learning **features, representations and tasks** directly from **images, text and sound**

Traditional Machine Learning

Manual Feature Extraction

Classification

Machine Learning

Car ✓
Truck ✗
•
•
Bicycle ✗

Deep Learning approach

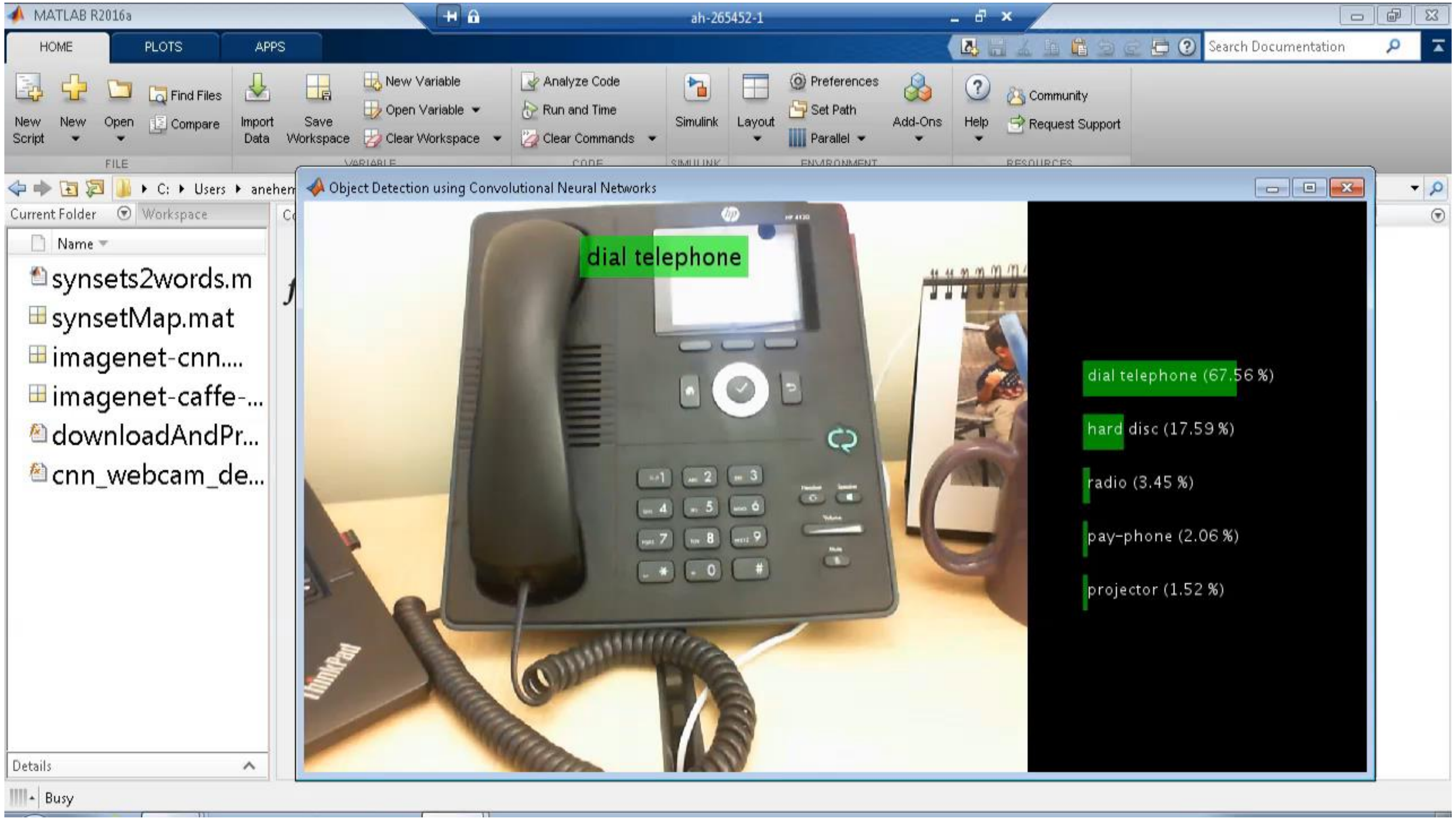
Convolutional Neural Network (CNN)

End-to-end learning

Feature learning + Classification

Car ✓
Truck ✗
•
•
Bicycle ✗

Demo : Live Object Recognition with Webcam



Why is Deep Learning so Popular ?

- **Results:** Achieved substantially better results on ImageNet large scale recognition challenge
  - 95% + accuracy on ImageNet 1000 class challenge

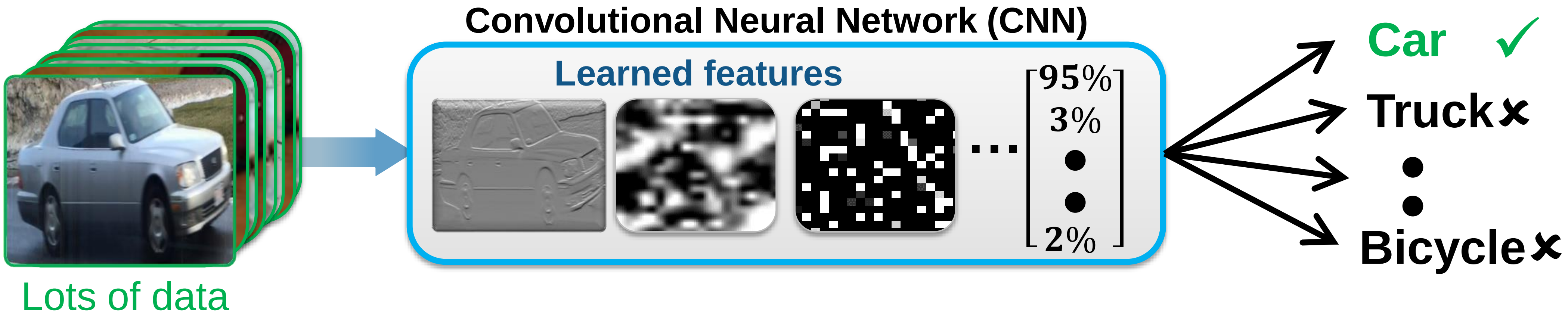
Year	Error Rate
Pre-2012 (traditional computer vision and machine learning techniques)	> 25%
2012 (Deep Learning)	~ 15%
2015 (Deep Learning)	<5 %

- **Computing Power:** GPU’s and advances to processor technologies have enabled us to train networks on massive sets of data.

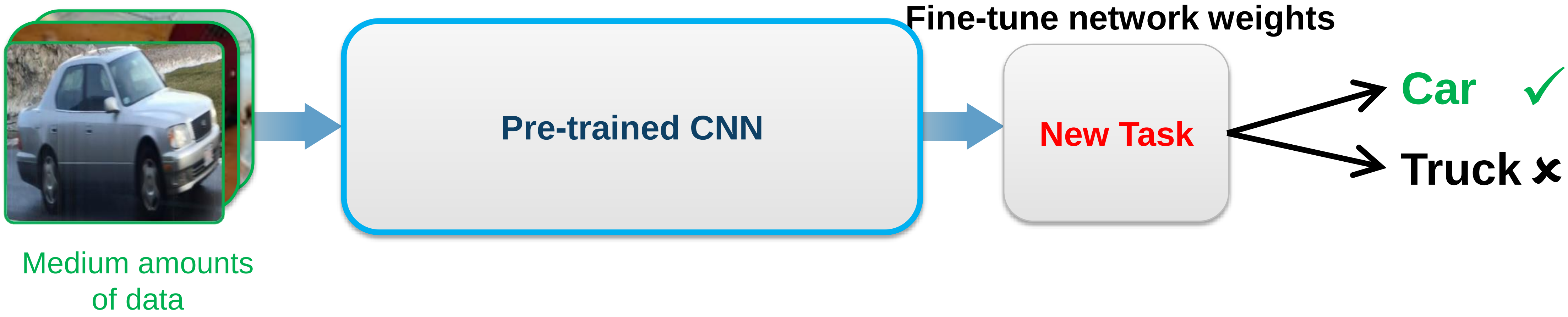
- **Data:** Availability of storage and access to large sets of labeled data
  - E.g. ImageNet , PASCAL VoC , Kaggle

Two Approaches for Deep Learning

1. Train a Deep Neural Network from Scratch

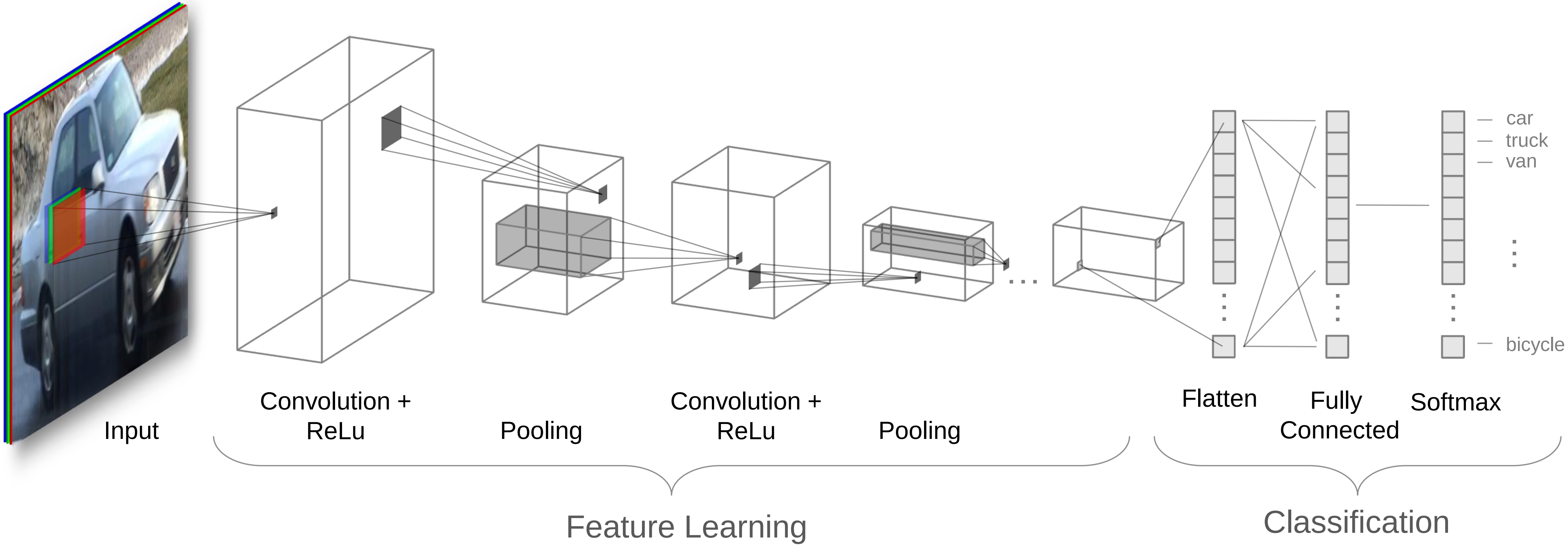


2. Fine-tune a pre-trained model (transfer learning)



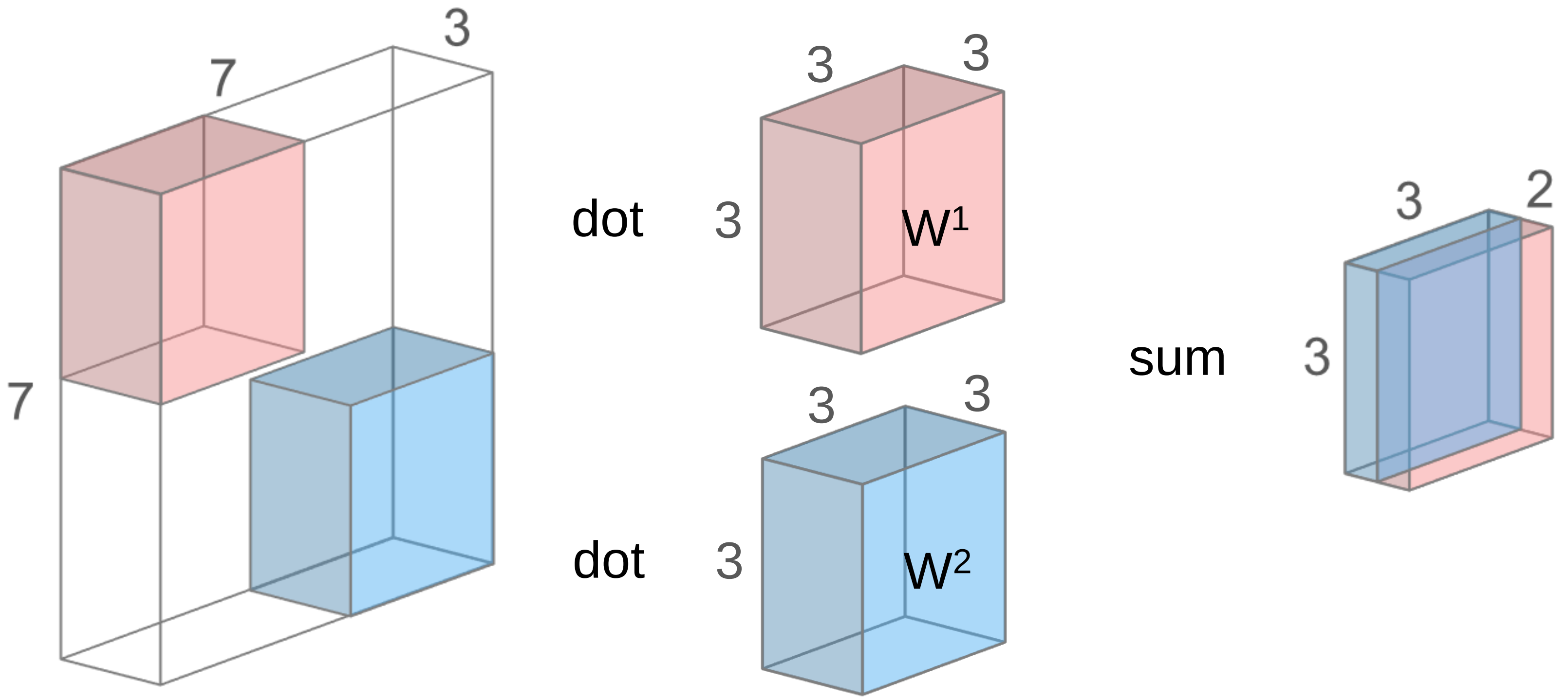
Convolutional Neural Networks

- Train “deep” neural networks on structured data (e.g. images, signals, text)
- Implements Feature Learning: Eliminates need for “hand crafted” features
- Trained using GPUs for performance



Convolution Layer

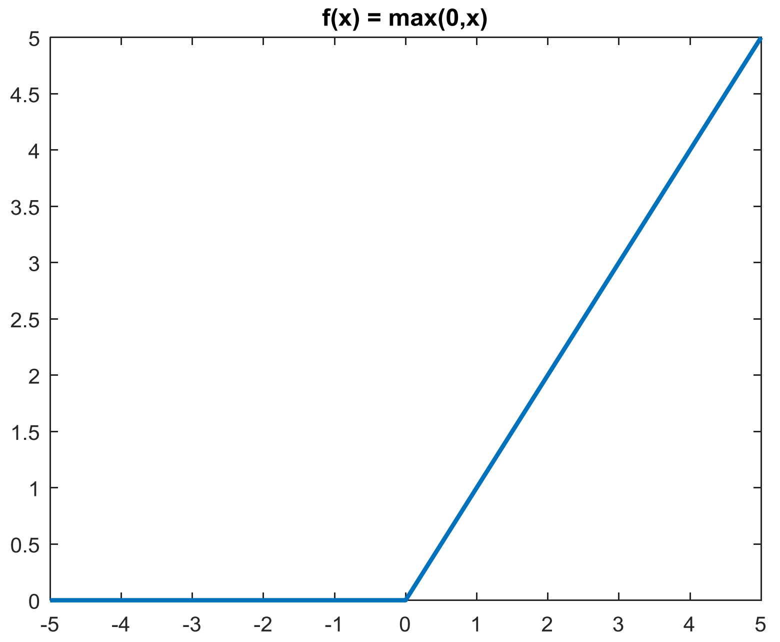
- Core building block of a CNN
- Convolve the filters sliding them across the input, computing the dot product



- Intuition: learn filters that activate when they "see" some specific feature

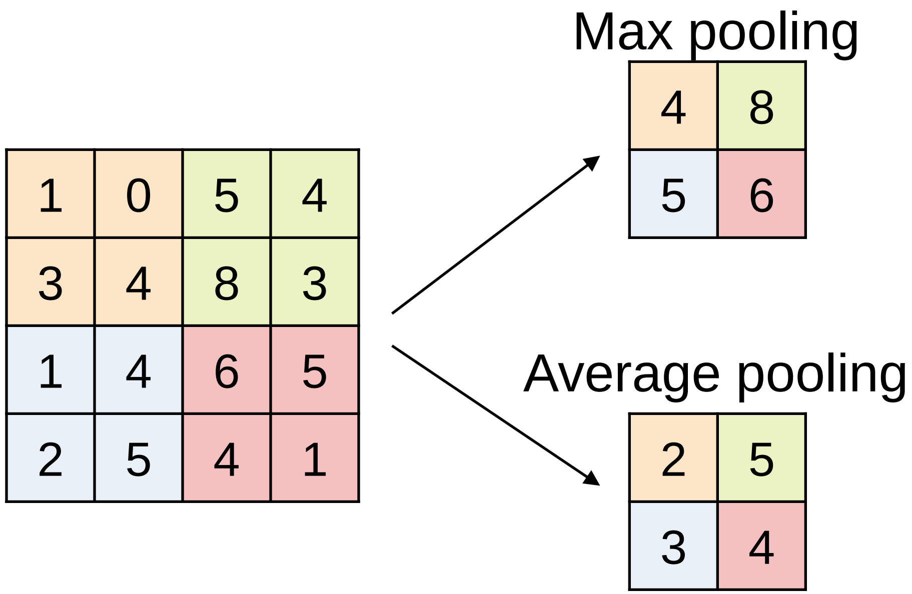
Rectified Linear Unit (ReLU) Layer

- Frequently used in combination with Convolution layers
- Do not add complexity to the network
- Most popular choice: $\boldsymbol{f(x) = max(0, x)}$, activation is thresholded at 0

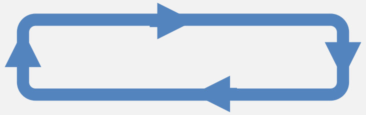


Pooling Layer

- Perform a **downsampling** operation across the spatial dimensions
- Goal: progressively decrease the size of the layers
- Max pooling and average pooling methods
- Popular choice: Max pooling with 2x2 filters, Stride = 2



Challenges using Deep Learning for Computer Vision

Steps	Challenge
Importing Data	Managing large sets of labeled images
Preprocessing	Resizing, Data augmentation
Choosing an architecture	Background in neural networks (deep learning)
Training and Classification	Computation intensive task (requires GPU)
Iterative design	

Demo

Fine-tune a pre-trained model (transfer learning)

New Data

Pre-trained CNN
(AlexNet – 1000 Classes)

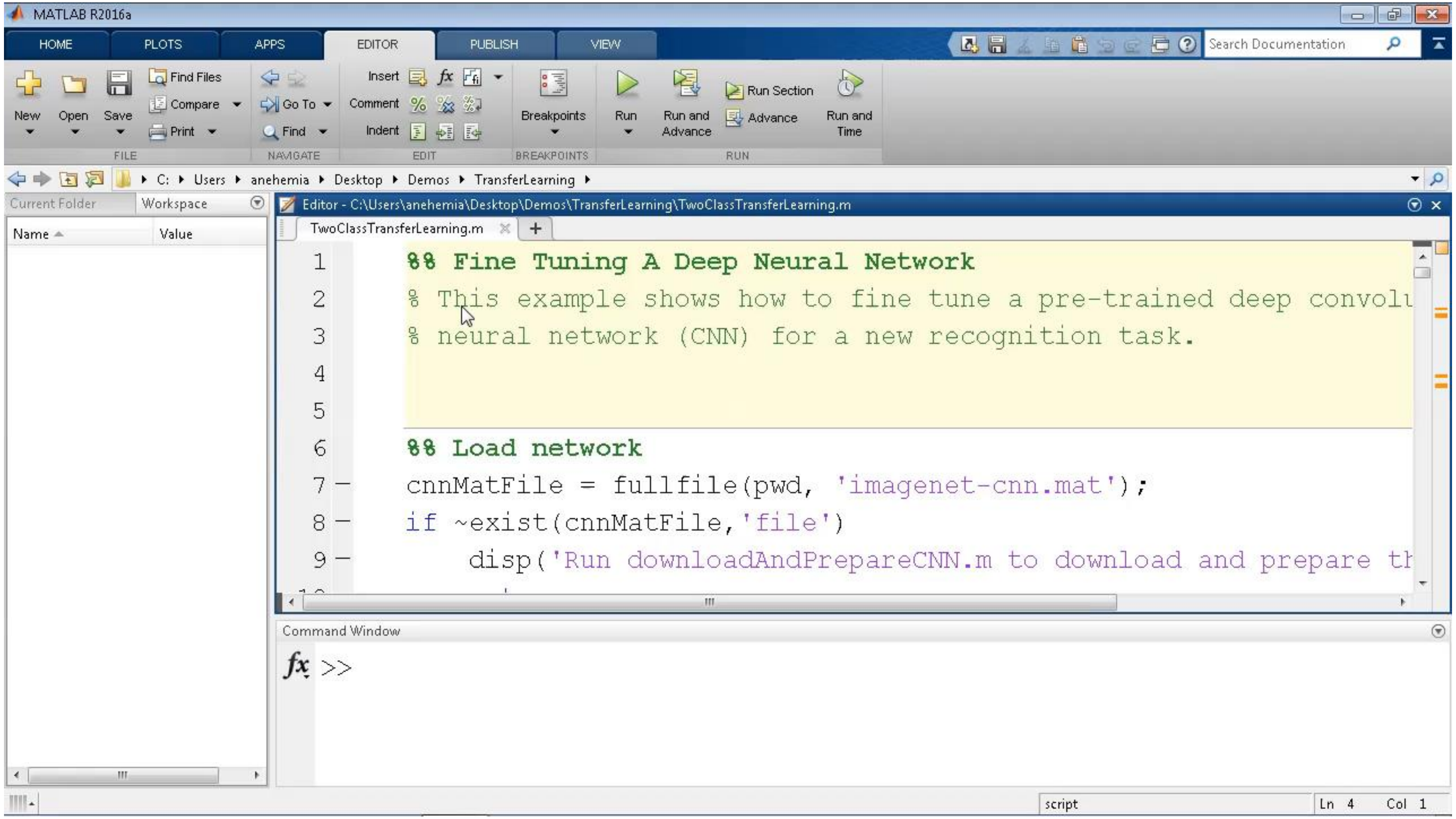
Car

SUV

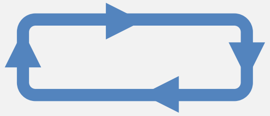
New Task – 2 Class Classification

Demo

Fine-tune a pre-trained model (transfer learning)



Addressing Challenges in Deep Learning for Computer Vision

Challenge	Solution
Managing large sets of labeled images	**`imageSet`** or **`imageDataStore`** to handle large sets of images
Resizing, Data augmentation	**`imresize`**, **`imcrop`**, **`imadjust`**, **`imageInputLayer`**, etc.
Background in neural networks (deep learning)	Intuitive interfaces, well-documented architectures and examples
Computation intensive task (requires GPU)	Training supported on GPUs No GPU expertise is required
	Automate. Offload computations to a cluster and test multiple architectures

Key Takeaways

- MATLAB enables engineers and data scientists to quickly create, test and implement predictive maintenance programs

- Predictive maintenance
  - Saves money for equipment operators
  - Increases reliability and safety of equipment
  - Creates opportunities for new services that equipment manufacturers can provide

- Consider Deep Learning when:
  - Accuracy of traditional classifiers is not sufficient
    - *ImageNet classification problem*
  - You have a pre-trained network that can be fine-tuned
  - Too many image categories (100s – 1000s or more)
    - *Face recognition*