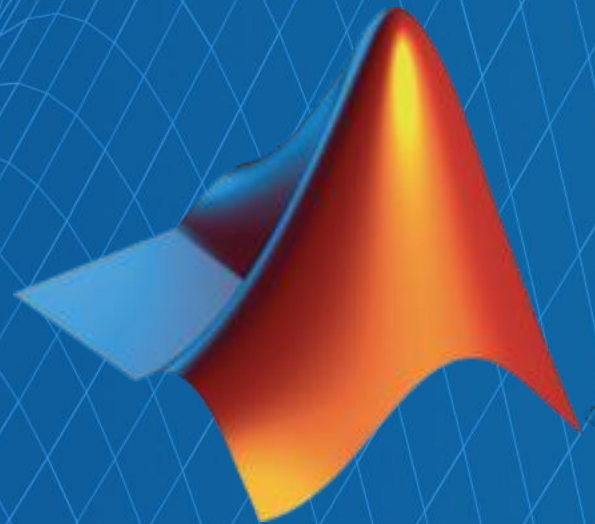


Machine Learning with MATLAB

A hands-on MATLAB workshop

Workshop Setup



*Should take
5 minutes !*

Setup overview

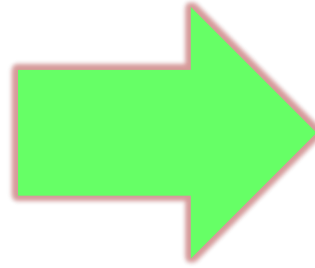
- Background:
 - Today we'll be using MATLAB Online
 - A version of MATLAB that runs in your web Browser
 - But please use the special MATLAB Online link that I will share with you shortly.

- STEPS
 1. Make sure you have a MathWorks Account
 2. Use the special MATLAB Online link
 3. Copy the Workshop files

*Should take
5-10 minutes*

A Word on Browser support

**Highly recommended to
use Google chrome**



Google Chrome

e.g. some Apps are only supported in Chrome

Stop Stare bookmark



https://www.mathworks.com/licensecenter/classroom/MO_3467150/

Set-Up Instructions

Step 1: Login with your MathWorks Account used to register for the event

Login to your MathWorks account at <https://drive.matlab.com/login>

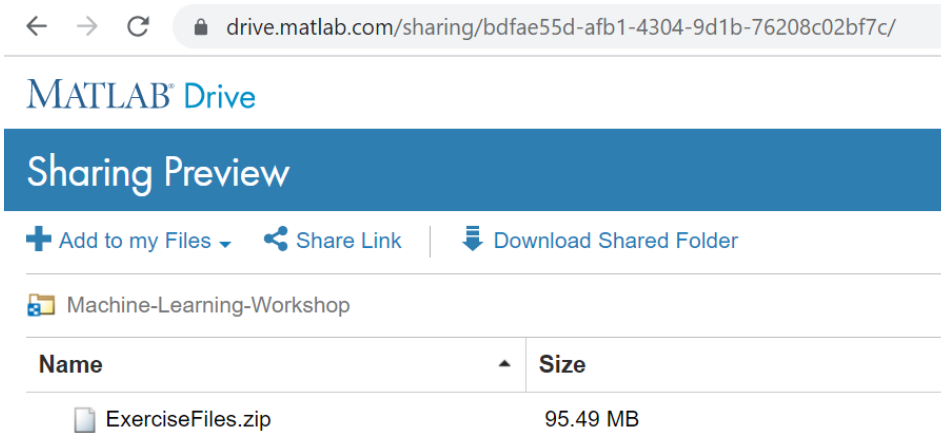
- Use the email address that you submitted to register for the event.

If creating a new account, visit <https://www.mathworks.com/mwaccount/register>

Step 2: Copy Workshop Files

- Shared files are available at the following address:
 - <https://drive.matlab.com/sharing/80ac7fef-7fb0-4f6c-b1e7-f3b60a111a1c>
- Note:** If you are unable to access the above link, wait 30 minutes and try again.

Step 2a



drive.matlab.com/sharing/bdfae55d-afb1-4304-9d1b-76208c02bf7c/

MATLAB Drive

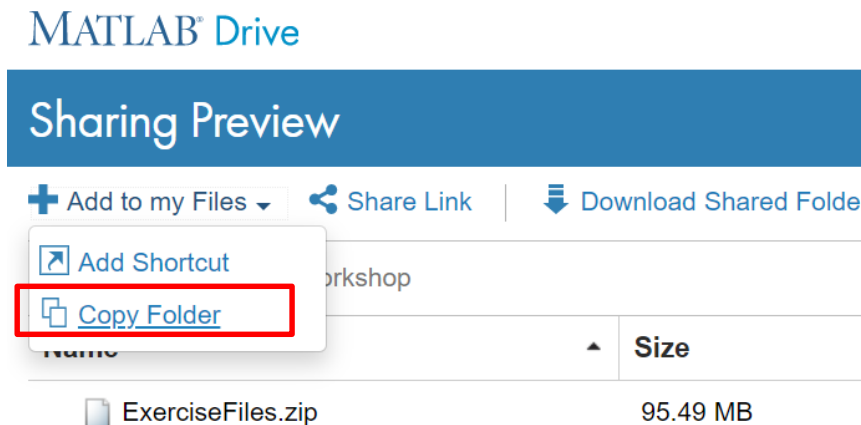
Sharing Preview

+ Add to my Files | Share Link | Download Shared Folder

Machine-Learning-Workshop

Name	Size
ExerciseFiles.zip	95.49 MB

Step 2b



MATLAB Drive

Sharing Preview

+ Add to my Files | Share Link | Download Shared Folder

Machine-Learning-Workshop

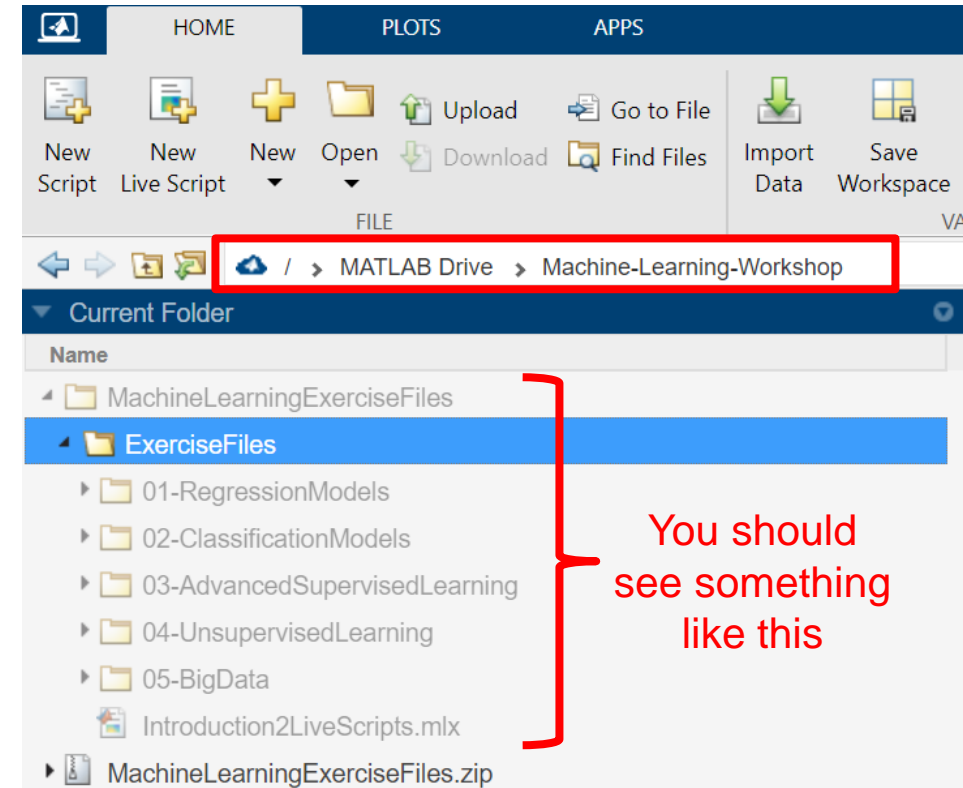
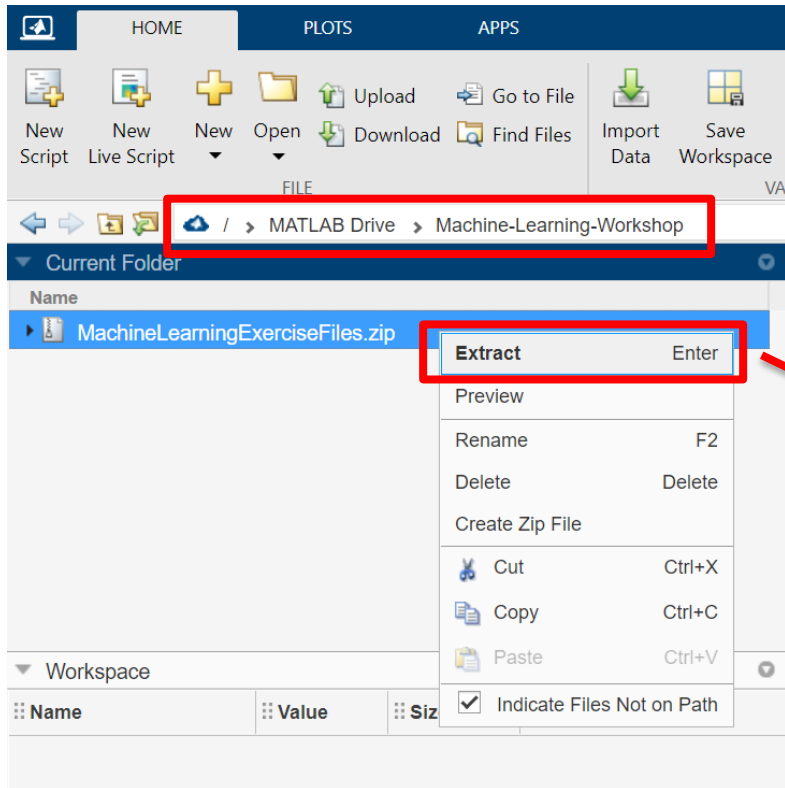
Name	Size
ExerciseFiles.zip	95.49 MB

A red box highlights the 'Copy Folder' option in the context menu.

Set-Up Instructions

Step 3: Log into the Workshop MATLAB Online and Confirm Web Browser

- Visit the following URL and login to access MATLAB Online
 - https://www.mathworks.com/licensecenter/classroom/MO_3467150/
- **Note:** If you are unable to login or access the above link, wait a few minutes and try again. If having issues with your browser, Chrome has been tested and usually works well.



- Your current folder browser should have the folder you copied over.

Statistics and Machine Learning Toolbox

157 shipping examples to explore

The screenshot shows the MathWorks Help Center page for "Statistics and Machine Learning Toolbox – Examples". The page is titled "R2021a" and features a navigation menu on the left with categories like "Descriptive Statistics and Visualization", "Probability Distributions", and "Hypothesis Tests". The main content area is divided into several sections, each with a representative plot and a brief description:

- Visualizing Multivariate Data:** Visualize multivariate data using various statistical plots. Many statistical analyses involve only two variables: a predictor variable and a response variable.
- Access Data in Dataset Array Variables:** Work with dataset array variables and their data.
- Select Subsets of Observations:** Select an observation or subset of observations from a dataset array.
- Sort Observations in Dataset Arrays:** Sort observations (rows) in a dataset array using the command line.
- Regression Using Dataset Arrays:** Perform linear and stepwise regression analyses using dataset arrays.
- Exploratory Analysis of Data:** Explore the distribution of data using descriptive statistics.

230 functions/classes

The screenshot shows the MathWorks Help Center page for "Statistics and Machine Learning Toolbox – Functions". The page is titled "R2021a" and features a navigation menu on the left with categories like "Descriptive Statistics and Visualization", "Managing Data", and "Data Types". The main content area is divided into several sections, each with a list of functions and their descriptions:

- Descriptive Statistics and Visualization:**
 - Managing Data:**
 - Data Import and Export:**

<code>caseread</code>	Read case names from file
<code>casewrite</code>	Write case names to file
<code>tblread</code>	Read tabular data from file
<code>tblwrite</code>	Write tabular data to file
<code>tdfread</code>	Read tab-delimited file
<code>xptread</code>	Create table from data stored in SAS XPORT format file
 - Data Types:**

<code>nominal</code>	(Not Recommended) Arrays for nominal data
<code>ordinal</code>	(Not Recommended) Arrays for ordinal data
<code>dummyvar</code>	Create dummy variables
<code>gplotmatrix</code>	Matrix of scatter plots by group
<code>grp2idx</code>	Create index vector from grouping variable
<code>gscatter</code>	Scatter plot by group
 - Dataset Arrays:**

<code>mat2dataset</code>	(Not Recommended) Convert matrix to dataset array
<code>cell2dataset</code>	(Not Recommended) Convert cell array to dataset array
<code>struct2dataset</code>	(Not Recommended) Convert structure array to dataset array
<code>table2dataset</code>	(Not Recommended) Convert table to dataset array
<code>dataset2cell</code>	(Not Recommended) Convert dataset array to cell array

Fun fact:

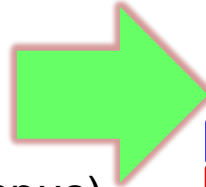
The Campus License is almost at every university in AU/NZ !

AU: 37 out of 40

NZ: 6 out of 8

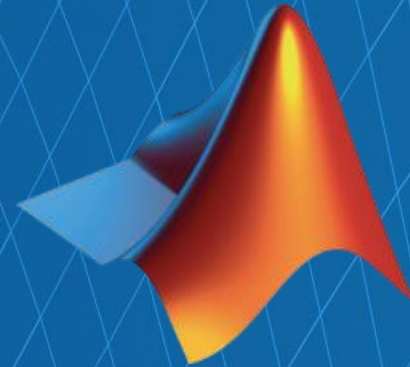
So ?

- Every student
- **Every Product**
- Every Computer(campus)
- Every Computer(personal)



MATLAB	Global Optimization Toolbox	Polyspace Code Prover	Simulink Coder
Simulink	GPU Coder	Powertrain Blockset	Simulink Compiler
5G Toolbox	HDL Coder	Predictive Maintenance Toolbox	Simulink Control Design
Aerospace Blockset	HDL Verifier	Radar Toolbox	Simulink Coverage
Aerospace Toolbox	Image Acquisition Toolbox	Reinforcement Learning Toolbox	Simulink Design Optimization
Antenna Toolbox	Image Processing Toolbox	RF Blockset	Simulink Design Verifier
Audio Toolbox	Instrument Control Toolbox	RF Toolbox	Simulink Desktop Real-Time
Automated Driving Toolbox	Lidar Toolbox	Risk Management Toolbox	Simulink PLC Coder
AUTOSAR Blockset	LTE Toolbox	RoadRunner	Simulink Real-Time
Bioinformatics Toolbox	Mapping Toolbox	RoadRunner Asset Library	Simulink Report Generator
Communications Toolbox	MATLAB Coder	Robotics System Toolbox	Simulink Requirements
Computer Vision Toolbox	MATLAB Compiler	Robust Control Toolbox	Simulink Test
Control System Toolbox	MATLAB Compiler SDK	ROS Toolbox	SoC Blockset
Curve Fitting Toolbox	MATLAB Parallel Server	Satellite Communications Toolbox	Spreadsheet Link
Data Acquisition Toolbox	MATLAB Production Server	Sensor Fusion and Tracking Toolbox	Stateflow
Database Toolbox	MATLAB Report Generator	SerDes Toolbox	Statistics and Machine Learning Toolbox
Datafeed Toolbox	MATLAB Web App Server	Signal Processing Toolbox	Symbolic Math Toolbox
DDS Blockset	Mixed-Signal Blockset	SimBiology	System Composer
Deep Learning HDL Toolbox	Model Predictive Control Toolbox	SimEvents	System Identification Toolbox
Deep Learning Toolbox	Model-Based Calibration Toolbox	Simscape	Text Analytics Toolbox
DSP System Toolbox	Motor Control Blockset	Simscape Driveline	UAV Toolbox
Econometrics Toolbox	Navigation Toolbox	Simscape Electrical	Vehicle Dynamics Blockset
Embedded Coder	OPC Toolbox	Simscape Fluids	Vehicle Network Toolbox
Filter Design HDL Coder	Optimization Toolbox	Simscape Multibody	Vision HDL Toolbox
Financial Instruments Toolbox	Parallel Computing Toolbox	Simulink 3D Animation	Wavelet Toolbox
Financial Toolbox	Partial Differential Equation Toolbox	Simulink Check	Wireless HDL Toolbox
Fixed-Point Designer	Phased Array System Toolbox	Simulink Code Inspector	WLAN Toolbox
Fuzzy Logic Toolbox	Polyspace Bug Finder		

Hands-on Machine Learning Workshop



What's Machine Learning About?



Source: <https://xkcd.com/1838/>

↻ Internet of Shit Retweeted



Computer Facts
@computerfact

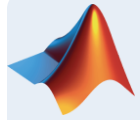


concerned parent: if all your friends jumped off a bridge would you follow them?

machine learning algorithm: yes.

2:20 PM · Mar 15, 2018

Agenda



Machine learning introduction

- Supervised machine learning models
 - Predicting fuel economy (Regression)
 - Human activity learning (Classification)
 - Feature engineering, AutoML, Interpretability
- Unsupervised learning (optional)
- Working with big data (optional)
- Deploying Machine Learning Algorithms
- Wrap-up / Resources Available to ANU

Machine Learning is Everywhere

Automobile	Industrial Automation	CES & Aero Defense	Energy & Finance
 <p data-bbox="300 711 486 746">Tire Wear</p> 	 <p data-bbox="728 711 1072 801"><u>Overlay metrology improvement</u></p> 	 <p data-bbox="1294 711 1638 801"><u>Telecom customer churn prediction</u></p> 	 <p data-bbox="1875 711 2142 801"><u>Forecasting & Risk Analysis</u></p> 
 <p data-bbox="249 1243 435 1333"><u>Detect Oversteer</u></p> 	 <p data-bbox="728 1239 1034 1329"><u>Building energy use optimization</u></p> 	 <p data-bbox="1319 1243 1689 1333">Engine Health (Pred Maintenance)</p> 	 <p data-bbox="1905 1243 2091 1333"><u>Portfolio Allocation</u></p> 

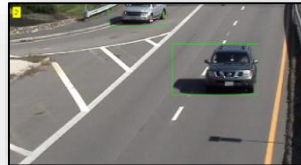
What is Machine Learning?

Ability to learn from data without being explicitly programmed

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



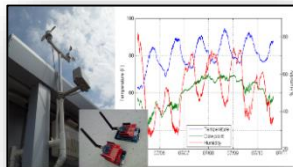
Energy Load Forecasting



Stock Market Prediction

update as more data becomes available

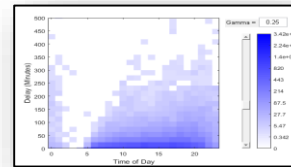
Solution needs to scale



IoT Analytics



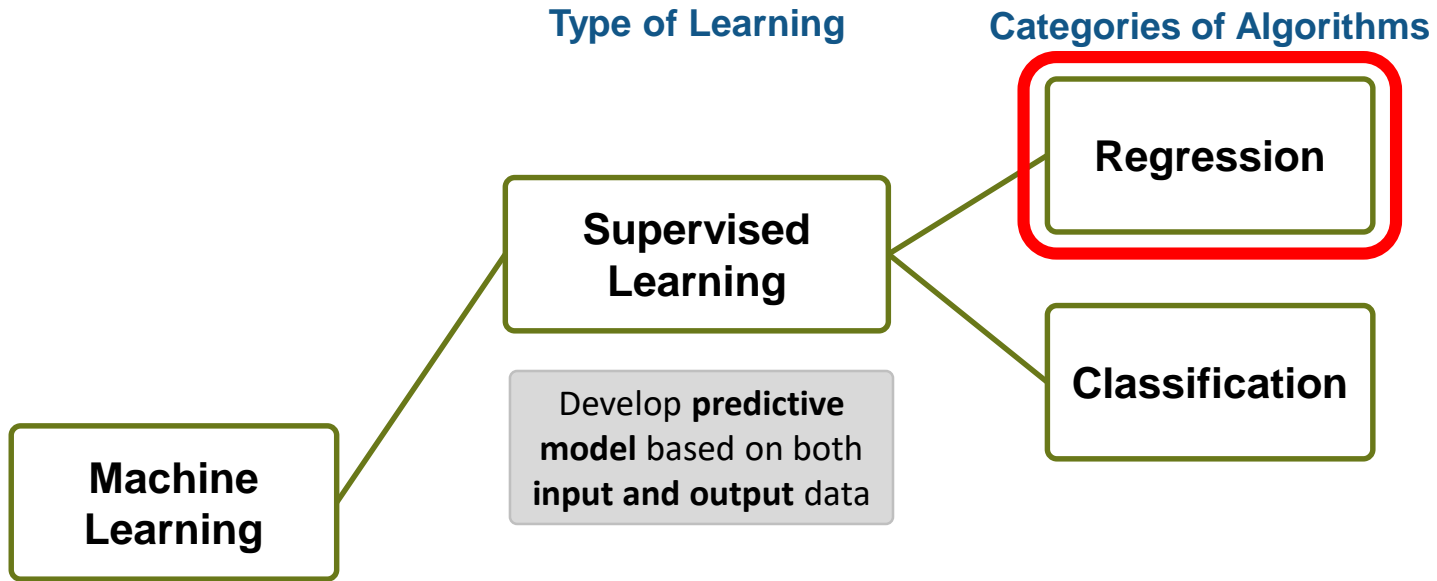
Taxi Availability



Airline Flight Delays

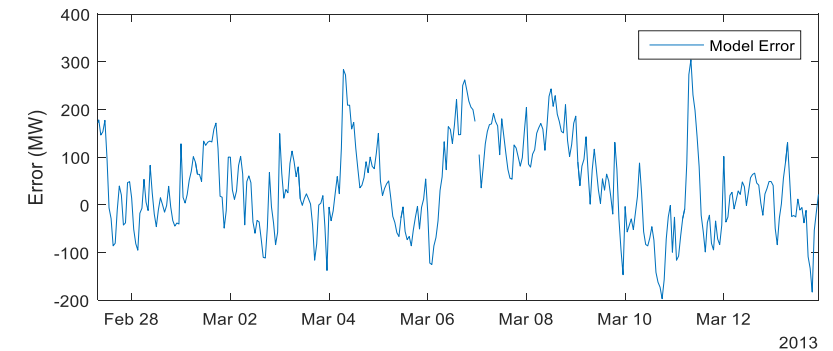
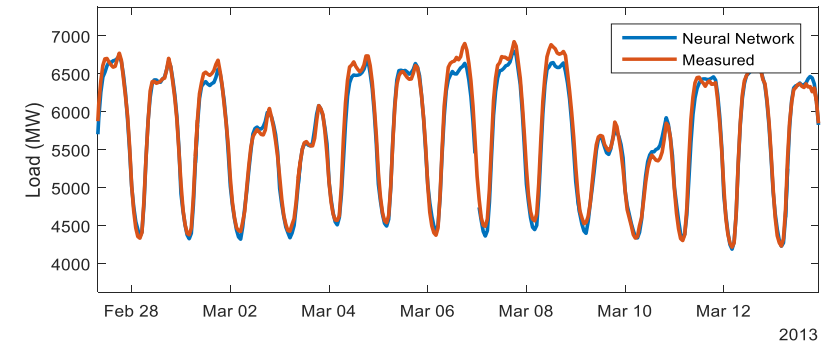
learn efficiently from very large data sets

Types of Machine Learning

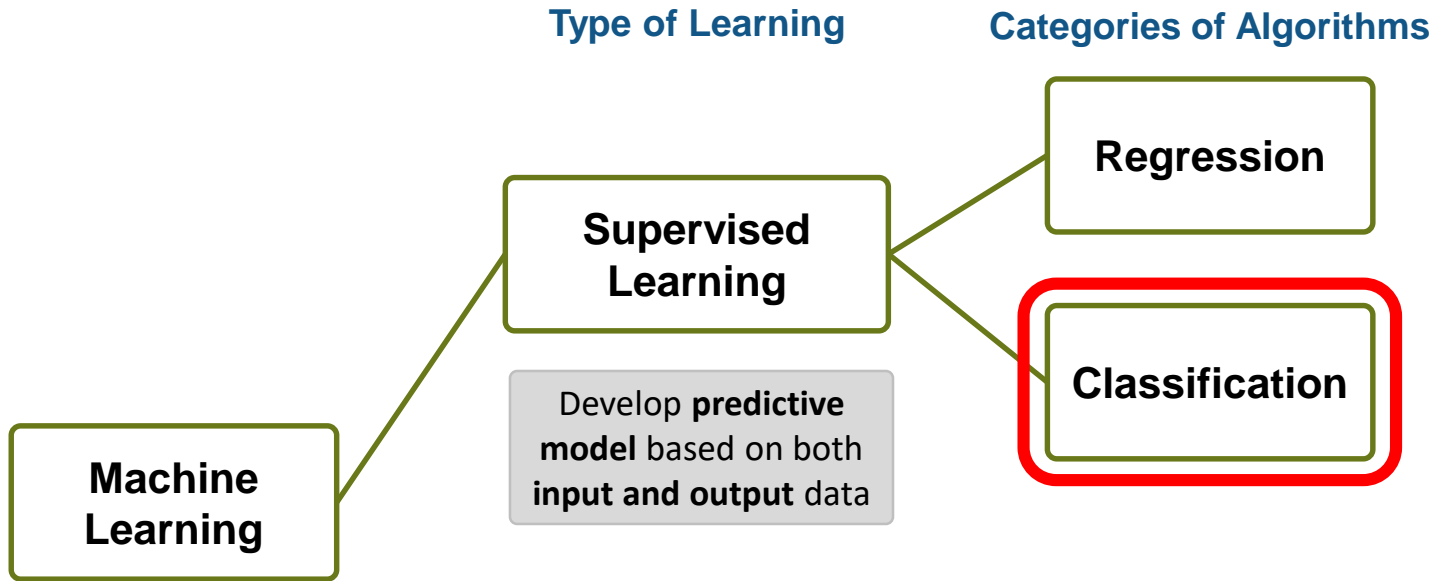


Objective:

Easy and accurate computation of day-ahead system load forecast





Types of Machine Learning



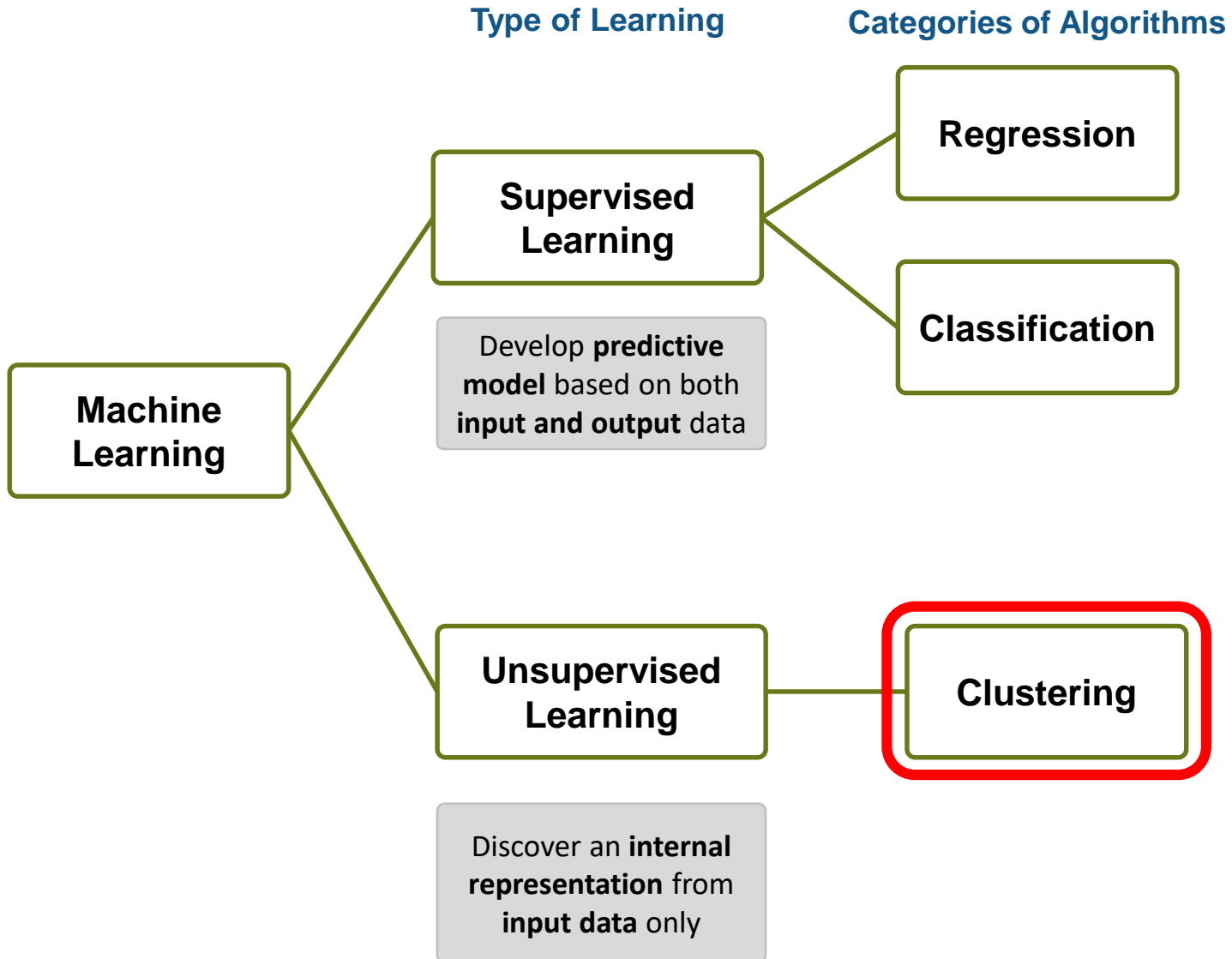
Objective:

Train a classifier to classify human activity from sensor data

Data:

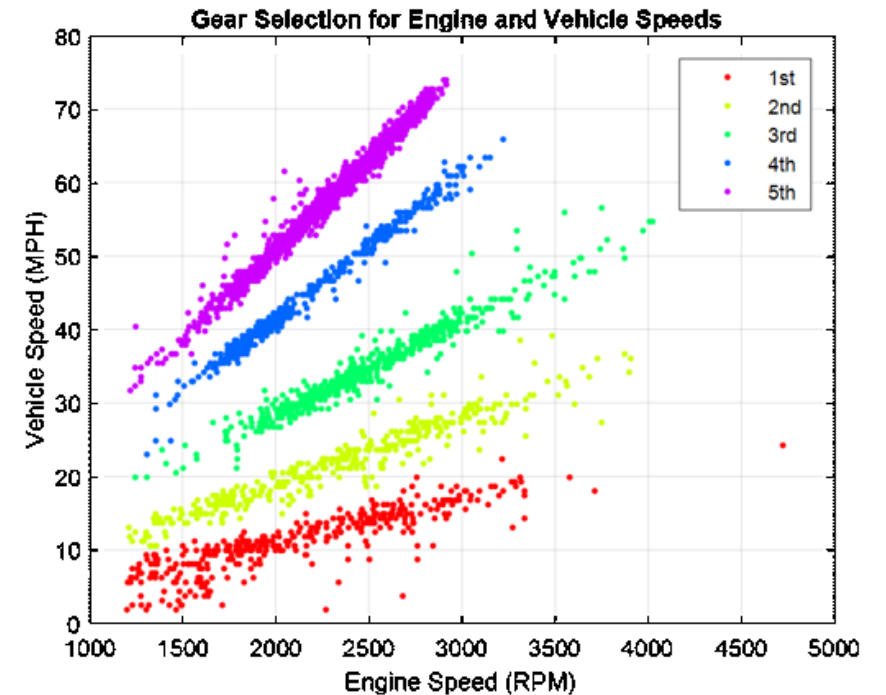
Inputs	3-axial Accelerometer 3-axial Gyroscope	
Outputs		

Types of Machine Learning



Objective:

Given data for engine speed and vehicle speed, identify clusters



MATLAB Supports the Entire Machine Learning Workflow

Access and explore data

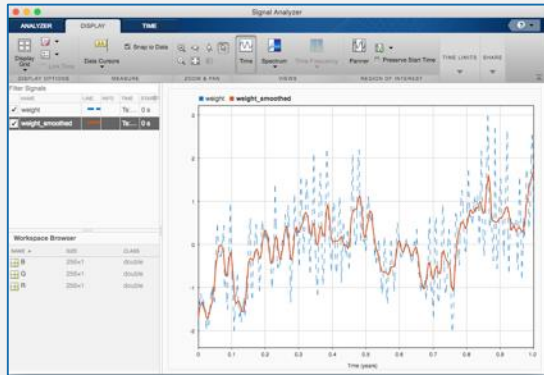
Preprocessing

Feature Engineering

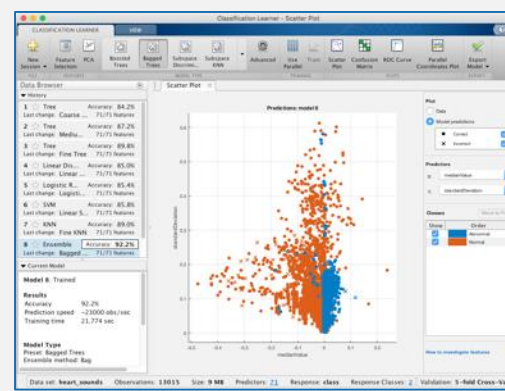
Model Training

Model Tuning

Integrate Analytics



Datatypes and tools for missing data, outliers, time-alignment, etc.



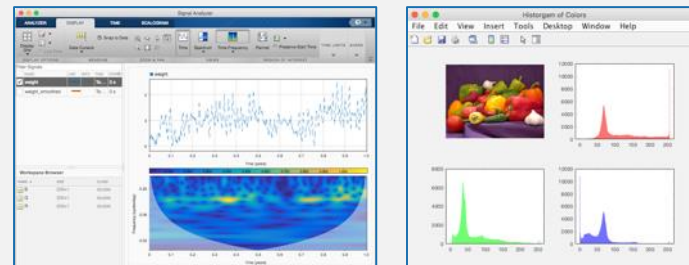
Machine Learning apps



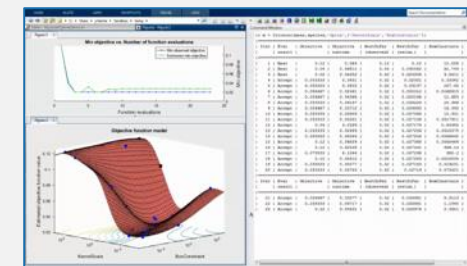
C/C++ Code Generation and Enterprise IT Integration

Region	OutageTime	Loss	Customers	RestorationTime	Cause
SouthWest	2002-02-01 12:18	418.9772218	1820159.482	2002-02-07 18:50	winter storm
SouthEast	2003-01-23 00:49	510.1399497	212035.3001		winter storm
SouthEast	2003-02-07 21:15	289.4035493	142938.6282	2003-02-17 08:14	winter storm
West	2004-04-08 05:44	414.8053524	340371.0338	2004-04-08 06:10	equipment fault
MidWest	2002-03-16 06:18	186.4367788	212754.0	0	ice storm
West	2003-06-18 02:49	0	0.0	2003-06-18 10:54	attack
West	2004-06-20 14:39	231.2847226		2004-06-20 19:16	equipment fault

Text files, spreadsheets, databases, binary files, data feeds, web, cloud storage

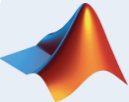


Domain-specific techniques for Signals, Images, Video, Audio, and Text



Automated Parameter Tuning

Agenda

- Machine learning introduction
- Supervised machine learning models
 -  – Predicting fuel economy (Regression)
 - Human activity learning (Classification)
 - Feature engineering, AutoML, Interpretability
- Unsupervised learning (optional)
- Working with big data (optional)
- Deploying Machine Learning Algorithms
- Wrap-up / Resources Available to ANU

Regression Example: Predicting Fuel Economy

Goal: Train a model to predict the fuel economy of a vehicle (MPG)

- Build initial models without any coding
- Don't need to be a regression expert

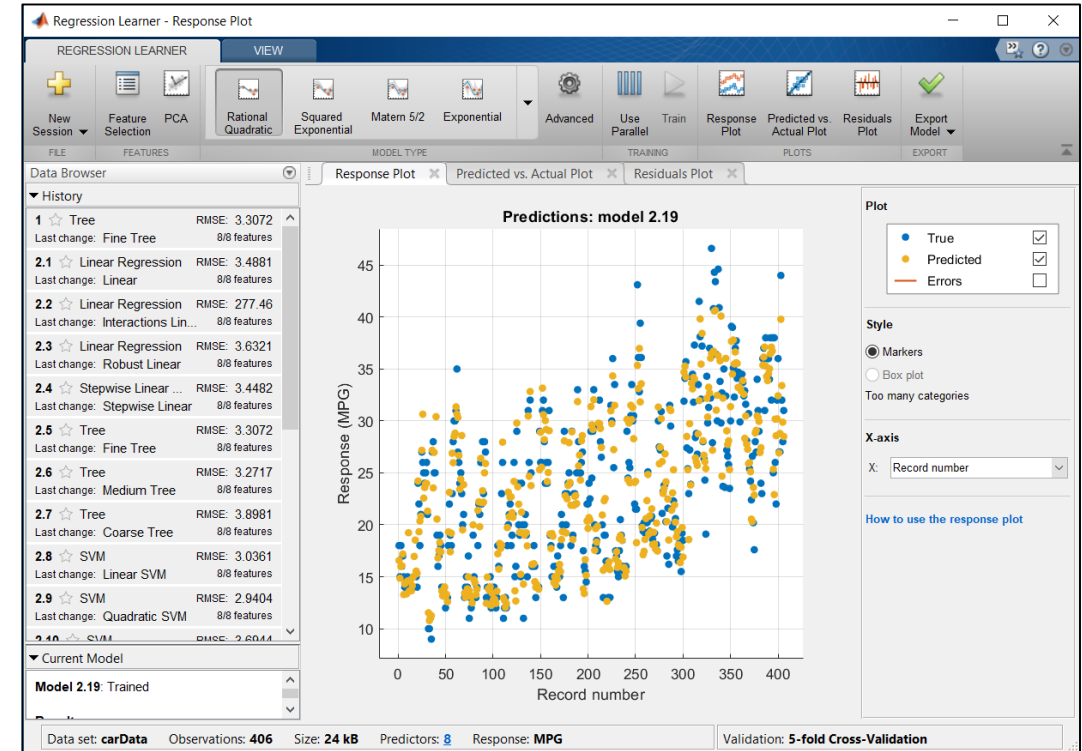
Data:

Predictors	Vehicle Horsepower, Weight, Cylinders, Year, etc.
Response	Miles Per Gallon (MPG)

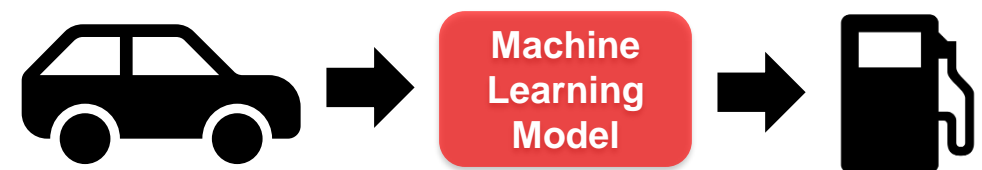
Approach:

- Load data in MATLAB
- Train and compare regression algorithms
- Test model on new vehicle data

[MATLAB Doc - Classification](#)



[MATLAB Doc - Classification Learner App](#)



Regression Example: Predicting Fuel Economy

Goal: Train a model to predict the fuel economy of a vehicle (MPG)

- Build initial models with
- Don't need to be a regression

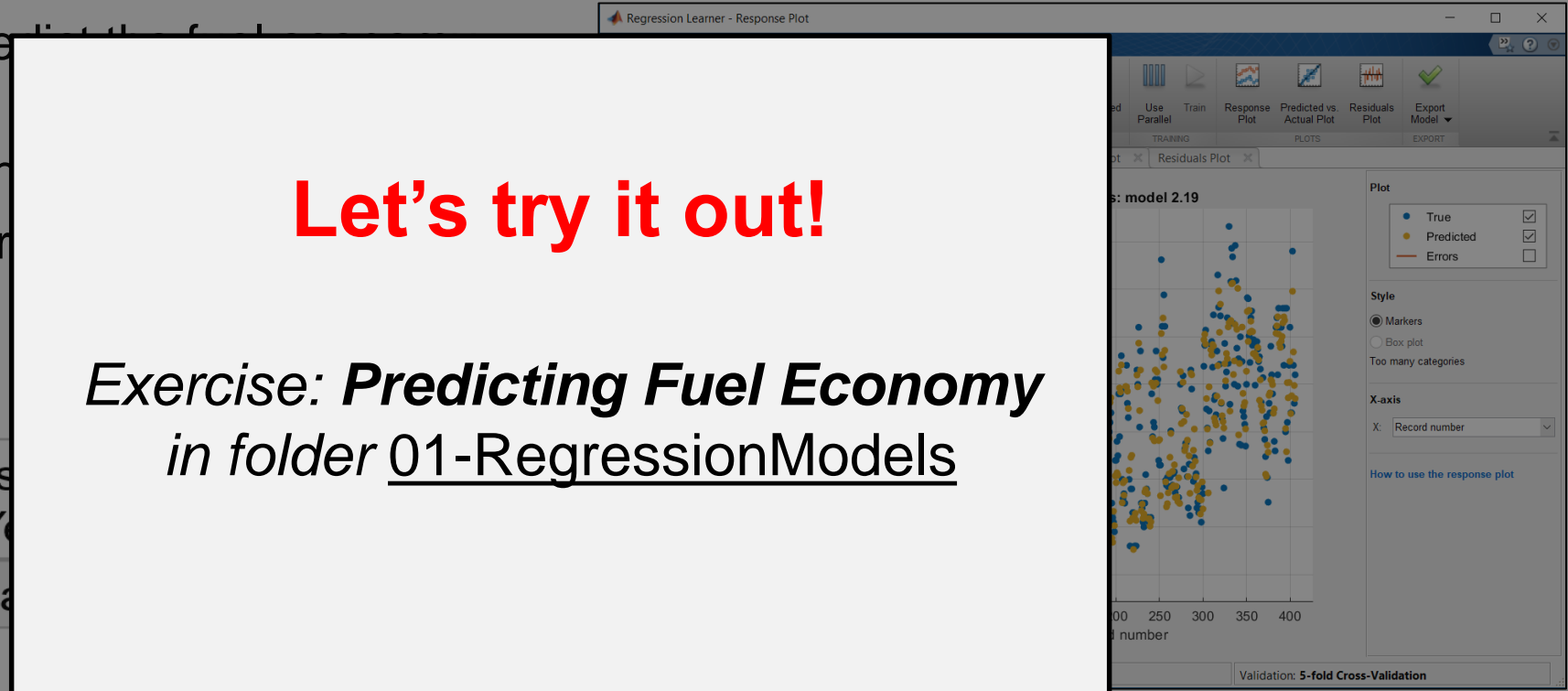
Data:

Predictors	Vehicle Horsepower, Number of Cylinders, Year
Response	Miles Per Gallon

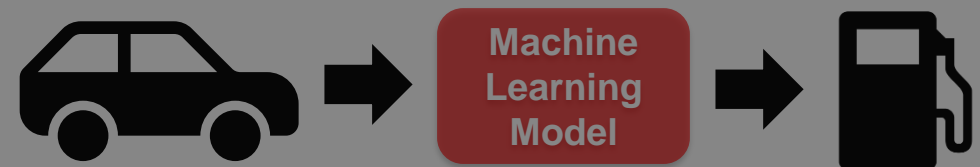
Approach:

- Load data in MATLAB
- Train and compare regression algorithms
- Test model on new vehicle data

[MATLAB Doc - Classification](#)



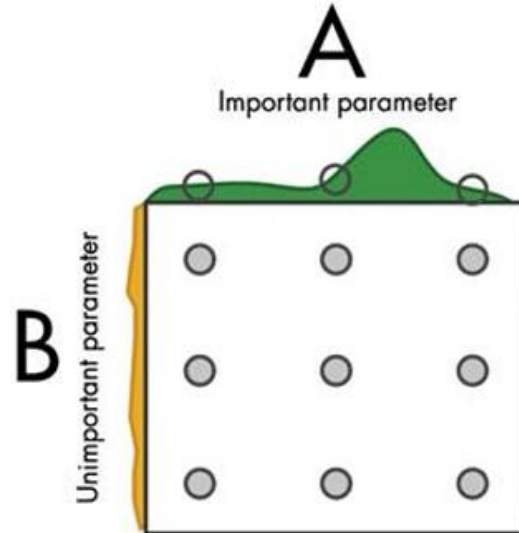
[MATLAB Doc - Classification Learner App](#)



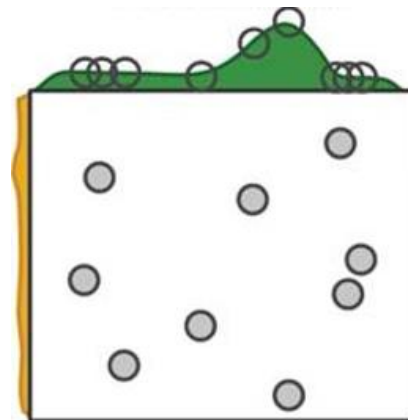
Hyperparameter Tuning

Why? – Model “knobs” (hyperparameters) need to be set properly for optimal performance

Standard:
Grid Search



Better:
Random Search



Best: Bayesian Optimization

- Bayesian model indicates impact of change
- Model picks “good” point to try next
- Much more efficient!
- Scale to multi-cores (using PCT) for larger datasets

Classification Learner - Confusion Matrix

CLASSIFICATION LEARNER VIEW

GET STARTED

DECISION TREES

DISCRIMINANT ANALYSIS

Now available inside the (Classification/Regression) Learner app as “Optimizable” model

Hyperparameter Tuning Workflow inside Learner Apps

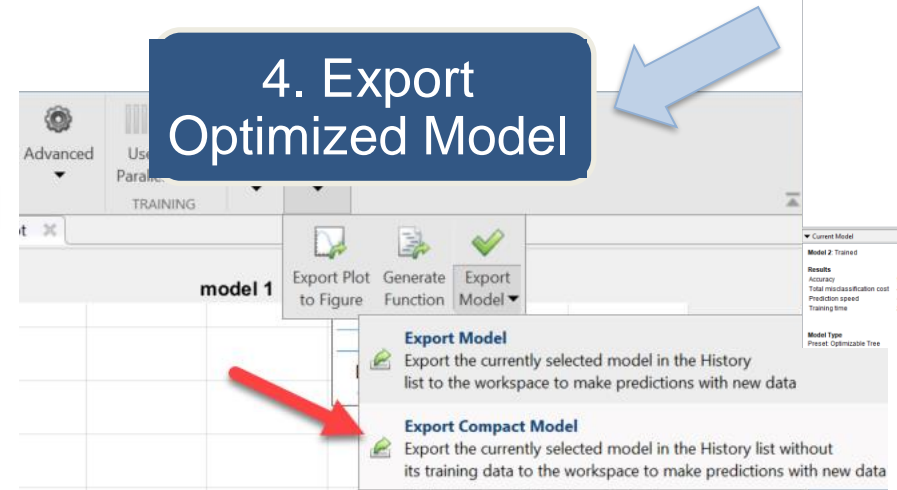
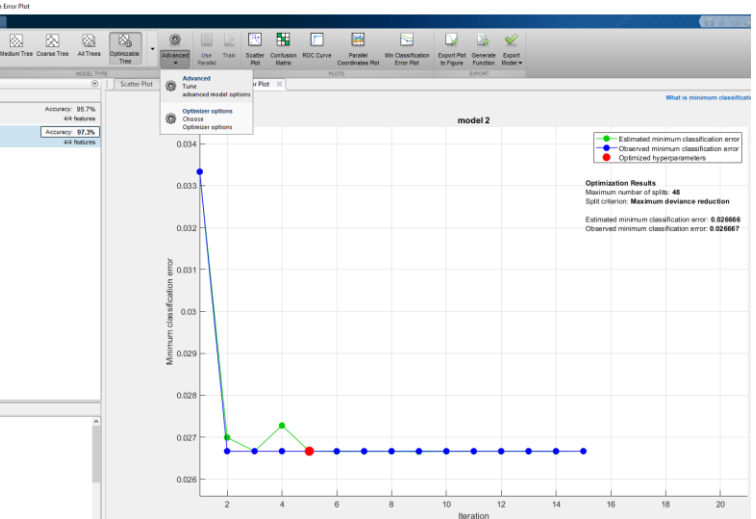
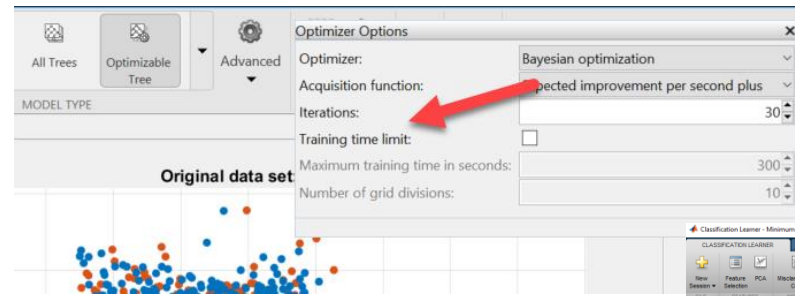
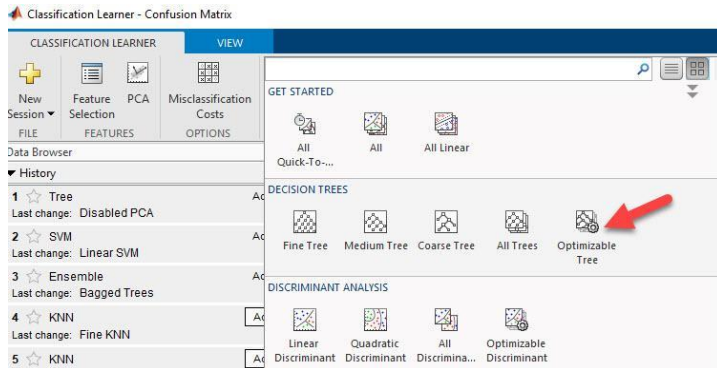
1. Choose "Optimizable" model from gallery

2. Adjust Optimizer Options (control runtime!)

3. "Train": Bayesian Optimization iterates

4. Export Optimized Model

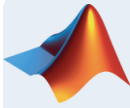
5. Iterate OR Prepare for Integration





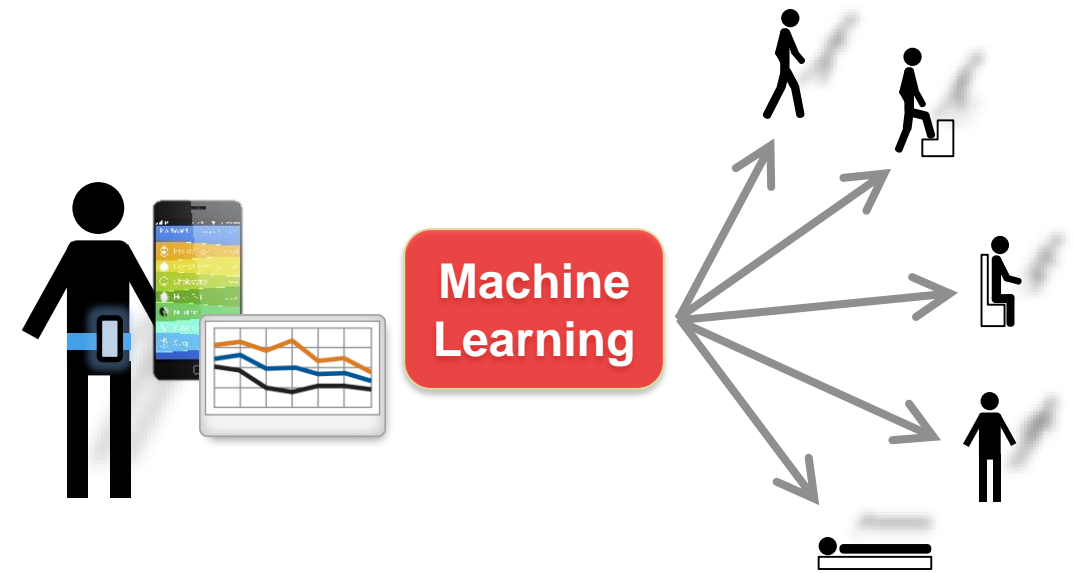
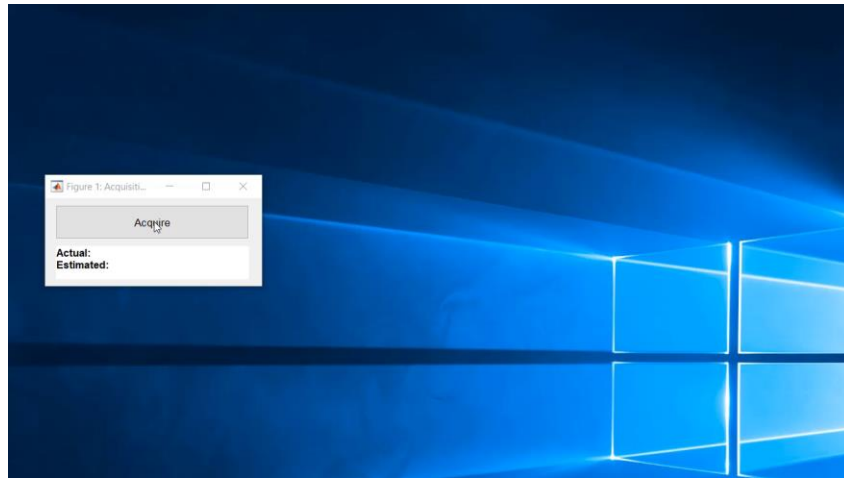
**“essentially, all models are wrong,
but some are useful”
– George Box**

Agenda

- Machine learning introduction
- Supervised machine learning models
 - Predicting fuel economy (Regression)
 -  – Human activity learning (Classification)
 - Feature engineering, AutoML, Interpretability
- Unsupervised learning (optional)
- Working with big data (optional)
- Deploying Machine Learning Algorithms
- Wrap-up / Resources Available to ANU

Human Activity Learning using Smartphones

Example task: Create a model to classify human activity from sensor data



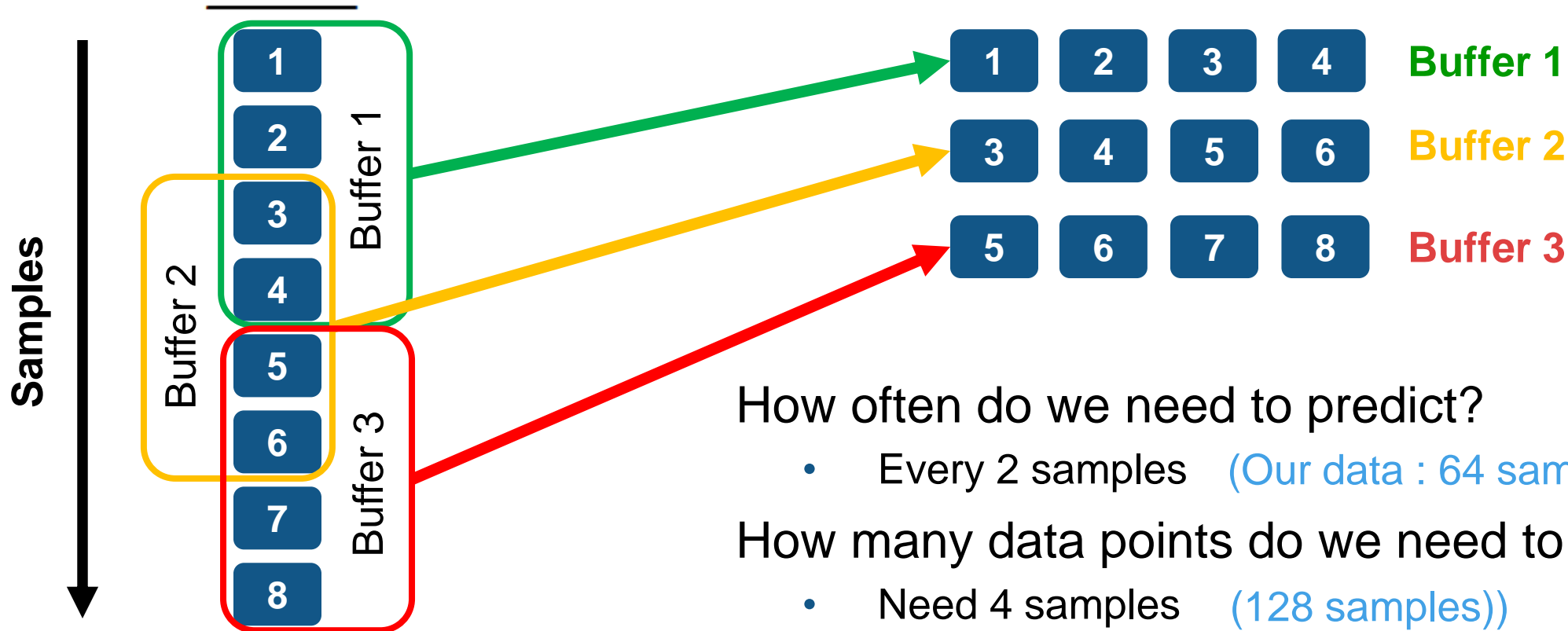
Dataset courtesy of:

Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. *Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine*. International Workshop of Ambient Assisted Living (IWAAL 2012). Vitoria-Gasteiz, Spain. Dec 2012 <http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

Signal Buffering

x

Why? – Calculate features on “chunks” during which signal doesn’t change (much), increase S/N (in feature)!



How often do we need to predict?

- Every 2 samples (Our data : 64 samples)

How many data points do we need to predict?

- Need 4 samples (128 samples)
- Create overlapping buffers of 4 points (64 samples)

Compute features (e.g. mean) on each buffer

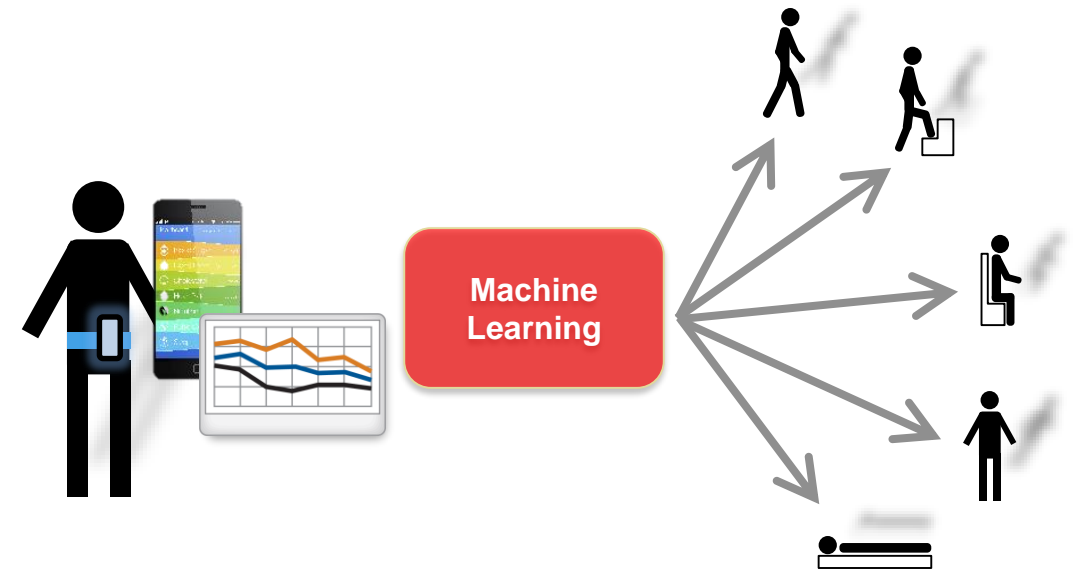
Exercise 2: Human Activity Learning using Smartphones

Goal: create initial models

- Buffering helps a lot (and computing means)
- Hyperparameter tuning generally helps (a bit)

Approach:

- Load buffered data
- Extract statistical features
- Compare various machine learning models (interactively)
- Optimize model using hyperparameter tuning



Total Subjects	30
trainData	25
validateData	3
testData	2

} Combined to held-out validation set

Dataset courtesy of:

Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. *Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine*. International Workshop of Ambient Assisted Living (IWAAL 2012). Vitoria-Gasteiz, Spain. Dec 2012 <http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

Exercise 2: Human Activity Learning using Smartphones

Goal: create initial model

- Buffering helps a lot
- Hyperparameter tuning

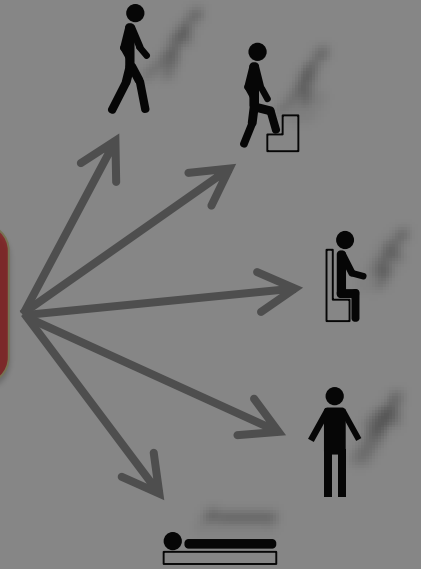
Approach:

- Load buffered data
- Extract statistical features
- Compare various models (interactively)
- Optimize model using hyperparameter tuning

Let's try it out!

Exercise:
humanActivityClassification.mlx
in folder 02-ClassificationModels

Machine Learning



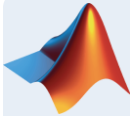
	30
	25
validateData	3
testData	2

} Combined to held-out validation set

Dataset courtesy of:

Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. *Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine*. International Workshop of Ambient Assisted Living (IWAAL 2012). Vitoria-Gasteiz, Spain. Dec 2012 <http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

Agenda

- Machine learning introduction
- Supervised machine learning models
 - Predicting fuel economy (Regression)
 - Human activity learning (Classification)
 -  – Feature engineering
 - AutoML
 - Interpretability
- Unsupervised learning (optional)
- Working with big data (optional)
- Deploying Machine Learning Algorithms

Feature Engineering

Using domain knowledge to create features for machine learning algorithms

“... is the art part of data science”

Sergey Yurgenson
(Kaggle Master)

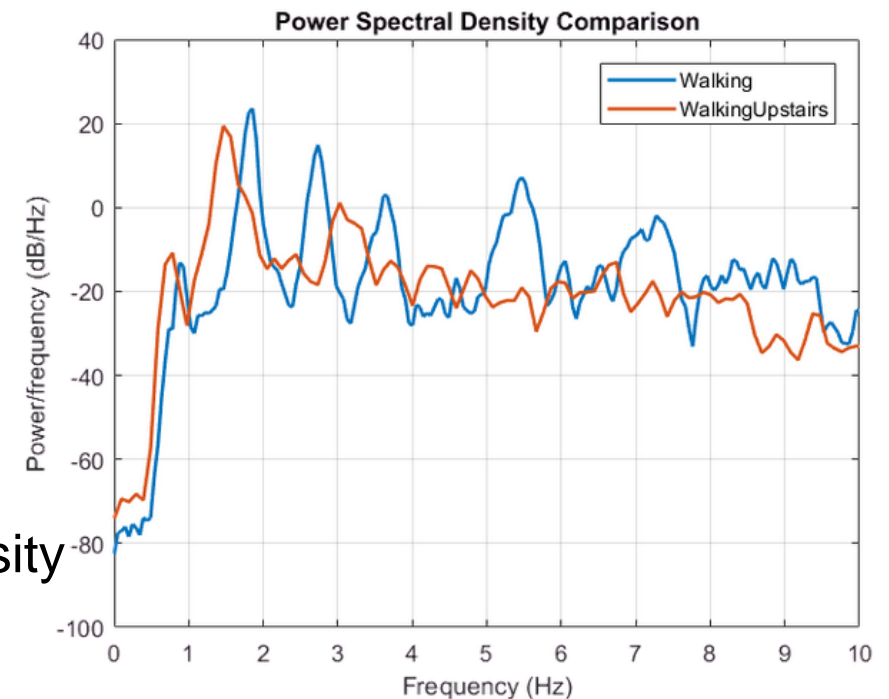


Feature transformation: reduce dimensionality

Feature selection: subset of relevant features

Possible feature engineering ideas:

- Additional statistics – PCA, NCA etc.
- Signal Processing Techniques – power spectral density, wavelets etc.
- Image Processing Techniques – bag of words, pixel intensity etc.
- Get creative!



[How to use Diagnostic Feature Designer](#) [12 min video]

Diagnostic Feature Designer App

Predictive Maintenance Toolbox

Why? – Explore and discover techniques for feature engineering without writing MATLAB code

Approach:

- Use signal processing to extract time and frequency domain features
- Select the most relevant features for training algorithms

Key features:

- Extract, visualize, and rank features from sensor data
- Use both statistical and dynamic modeling methods
- Work with out-of-memory data

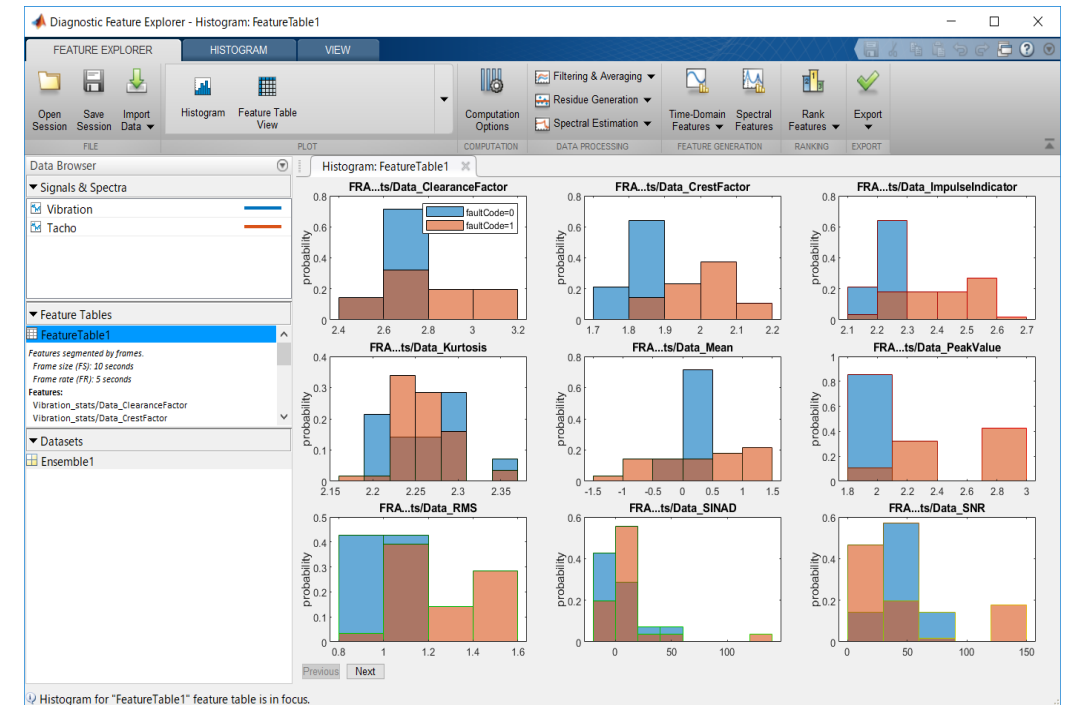


Exercise 3 – Feature Engineering for human activity

Goal: Explore different techniques for feature engineering

Approach:

- Use signal processing techniques to extract time domain and signal features
- Use feature selection technique to reduce the set of features to the most relevant
- Browse examples in MATLAB documentation for different applications



Exercise 3 – Feature Engineering for human activity

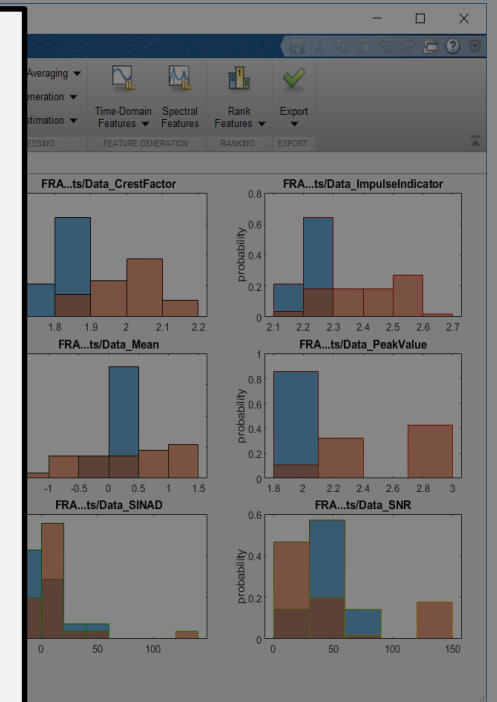
Goal: Explore different techniques for feature engineering

Approach:

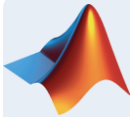
- Use signal processing to extract time domain features
- Use feature selection to reduce the set of features to those most relevant
- Browse examples in the documentation for different applications

Let's try it out!

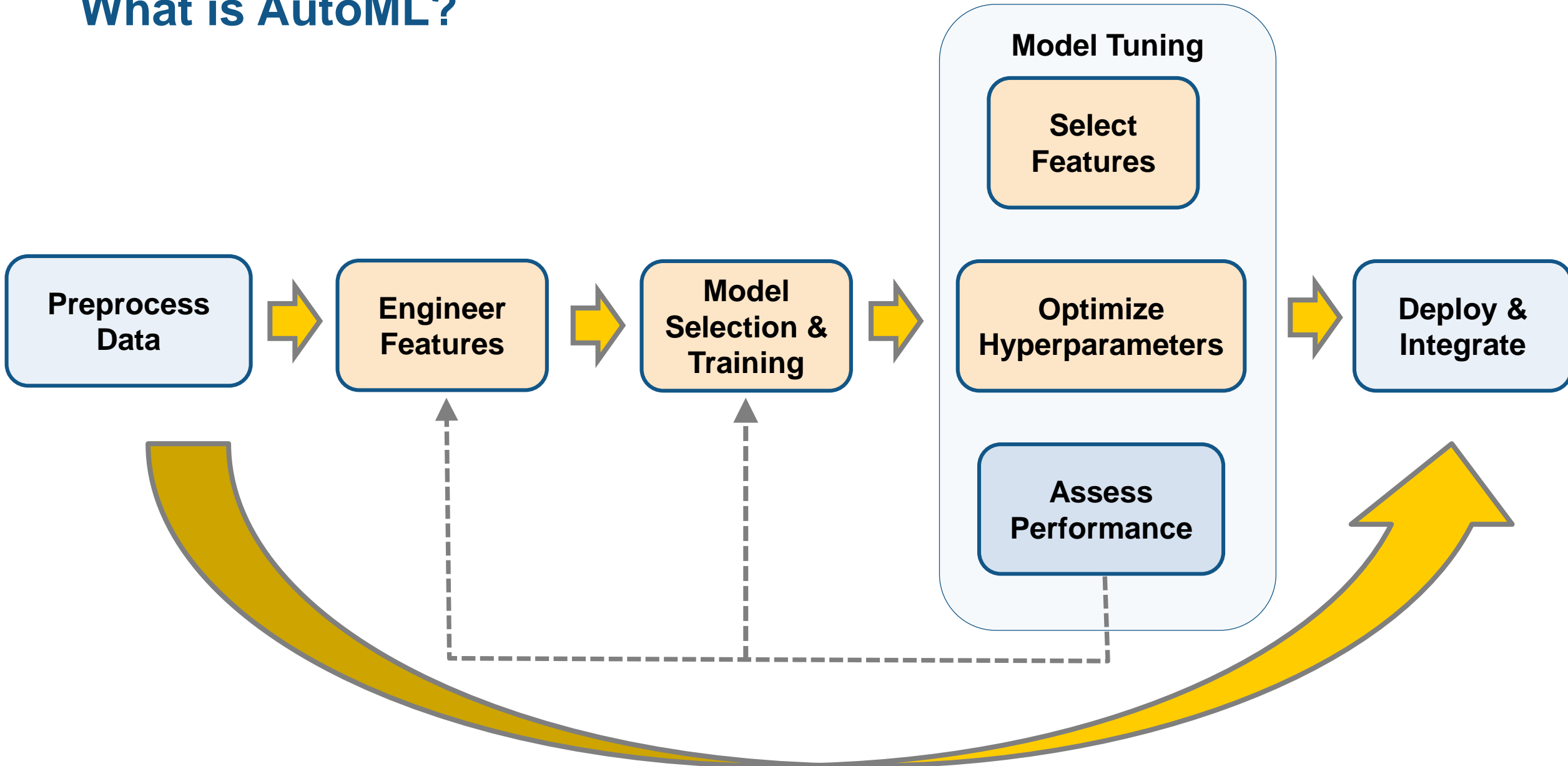
*Exercise: **featureEngineering.mlx**
in folder 03-FeatureEngineering*



Agenda

- Machine learning introduction
- Supervised machine learning models
 - Predicting fuel economy (Regression)
 - Human activity learning (Classification)
 - Feature engineering
 -  – AutoML
 - Interpretability
- Unsupervised learning (optional)
- Working with big data (optional)
- Deploying Machine Learning Algorithms

What is AutoML?



Feature Generation with Wavelet Scattering

Why? – Obtain good features “automagically”, without domain knowledge

What are Wavelets?

- Instead of decomposing signal into complete sinus waves, decompose into “wavelets”
- [Tech Talks explaining Wavelets](#) [4 videos]
- [Example](#)
- This conceptually looks like this:

**Slide Wavelet
across Signal**



Wavelet Scattering Framework [\[Bruna and Mallat 2013\]](#)

- Automatic Feature Extraction
- Great starting point if you don't have a lot of data
- Reduces data dimensionality and provides compact features

Works with both Signal and Image data [\[Texture example, Digit Classification\]](#)

Wavelet Scattering Nuts and Bolts



Pseudo-Code:

```
sf = waveletScattering(SignalLength) ;  
Loop over signal  
    waveletFeature = featureMatrix(sf,signal)  
    Append waveletFeature to feature table  
    Add labels  
end
```

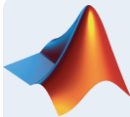
Additional Resources:

[Wavelet scattering Tech talk](#) [4 min video]

[Wavelet scattering for ECG](#) [doc example]

[Blog about Wavelet scattering](#) on towardsdatascience.com

Agenda

- Machine learning introduction
- Supervised machine learning models
 - Predicting fuel economy (Regression)
 - Human activity learning (Classification)
 - Feature engineering
 - AutoML
-  – Interpretability
- Unsupervised learning (optional)
- Working with big data (optional)
- Deploying Machine Learning Algorithms

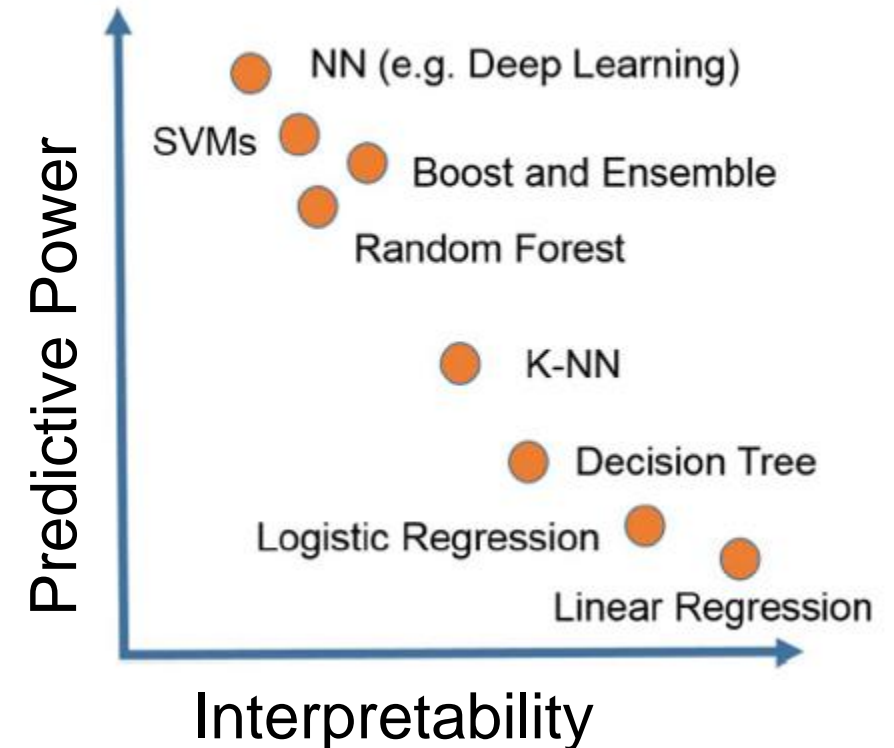
Interpretability and Explainability

Both terms describe the process of making “blackbox” models understandable

- “Interpretability”: primarily “classic” machine learning, causality of specific model decisions
- “**Explainable AI**” often refers to AI=Deep Learning, sometimes explaining how model works

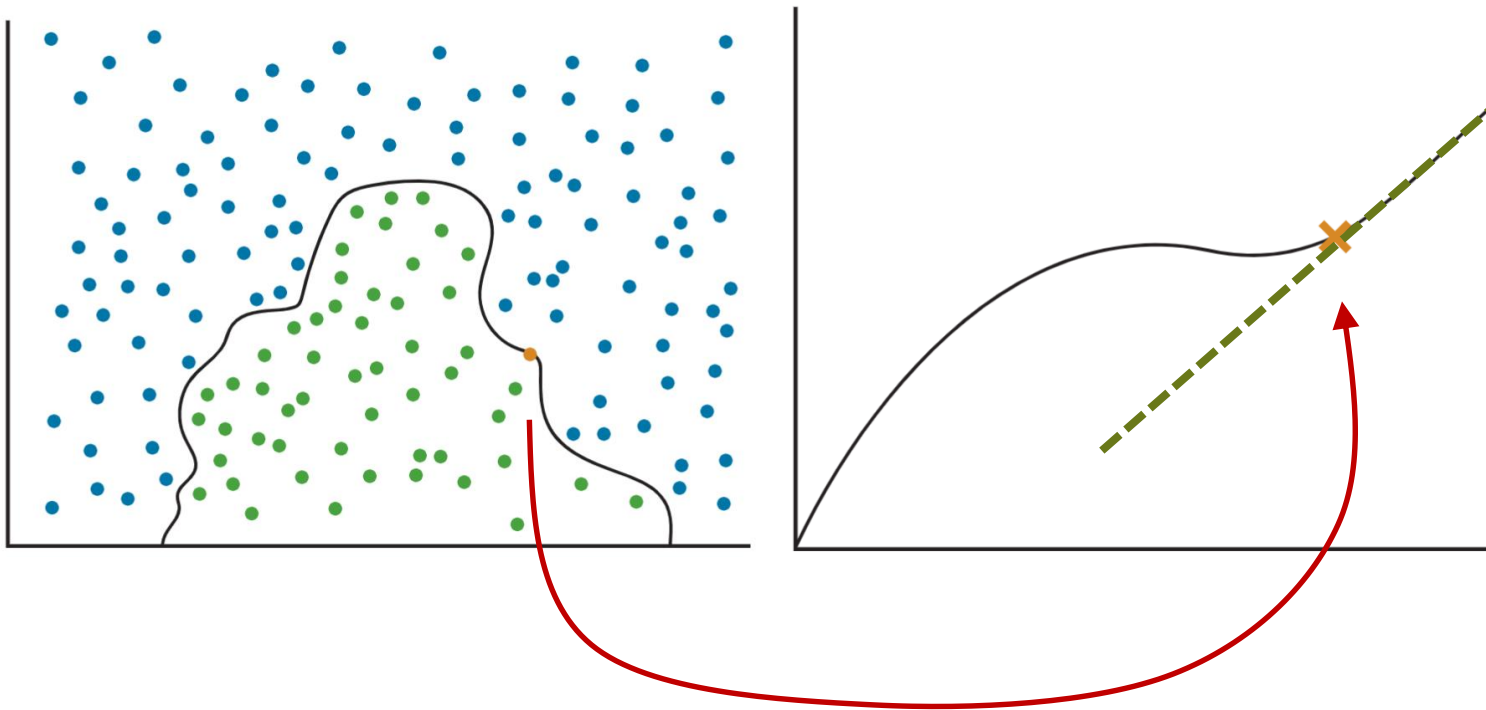
Why Interpretability?

1. Overcome “blackbox” model
 - Not acceptable by company guidelines
 - Build trust for users unfamiliar with machine learning
 - Pick model that looks at “right” evidence
2. Regulatory requirements (Finance, Europe’s GDPR):
3. Debug models

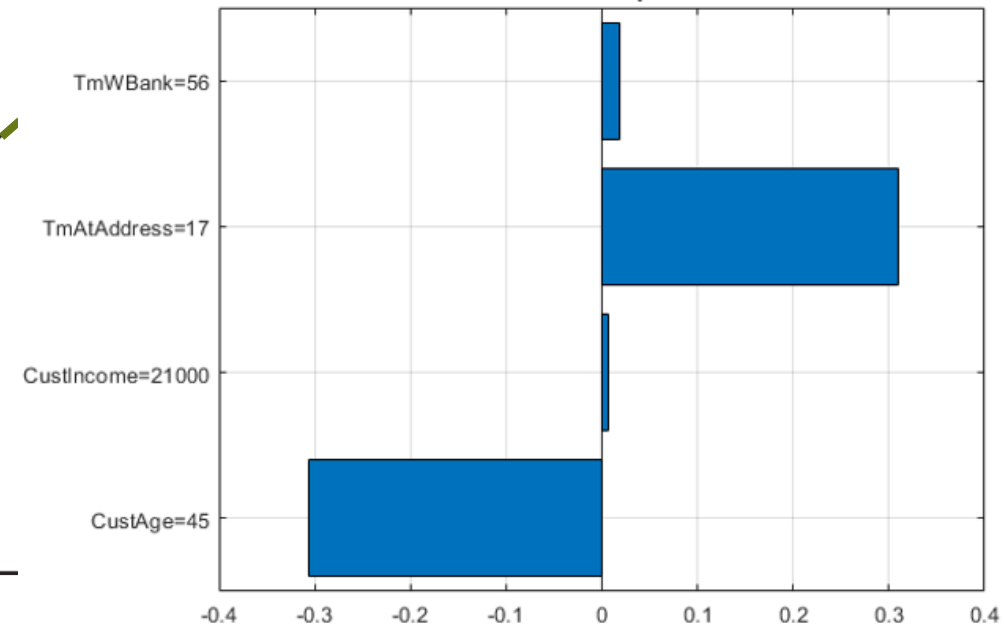


LIME = Local Interpretable Model-Agnostic Explanations

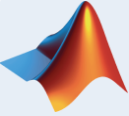
① Approximate complex model near Point of Interest with simple model



② “Explain” using weights of simple model



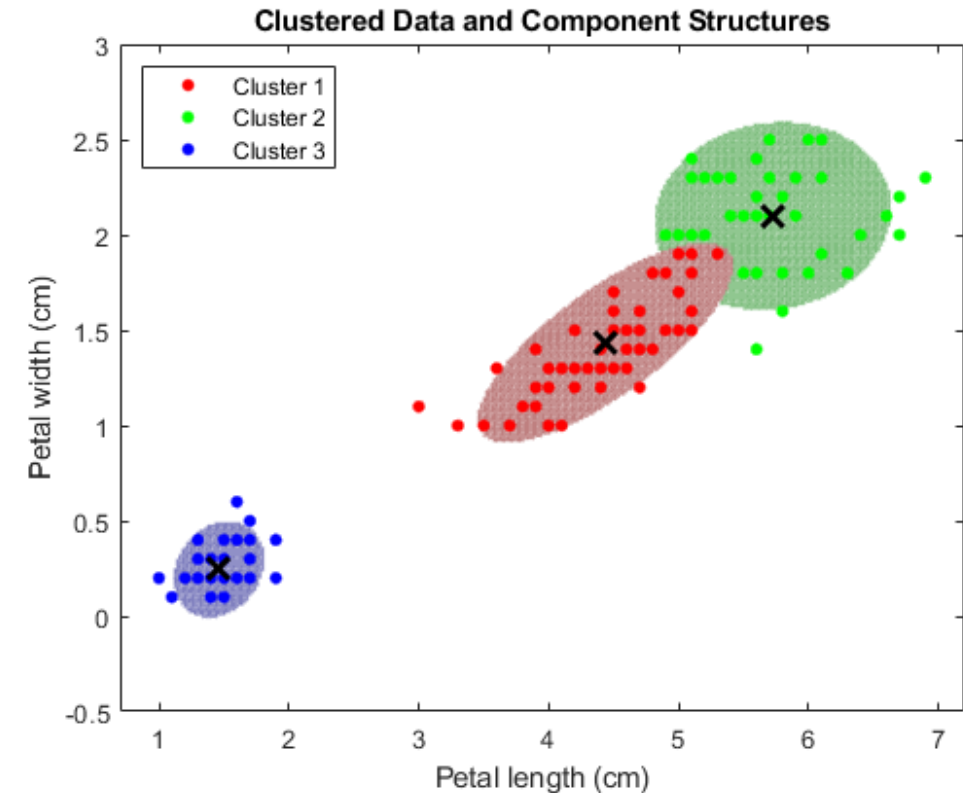
Agenda

- Machine learning introduction
- Supervised machine learning models
 - Predicting fuel economy (Regression)
 - Human activity learning (Classification)
- Feature extraction and feature selection
-  Unsupervised learning (optional)
 - Working with big data (optional)
 - Deploying Machine Learning Algorithms
 - Wrap-up / Resources Available to ANU

Clustering

Why? Discover patterns, identify possible features, check for outliers

- Segment data into groups based on similarity
 - [MATLAB doc - Cluster Analysis](#)
- Can be achieved by various algorithms
 - *k-means, k-medoids, Hierarchical, Gaussian Mixture Models, Nearest Neighbors, Hidden Markov Models*
 - [MATLAB doc - Choose Cluster Analysis Method](#)
- Is it an iterative process

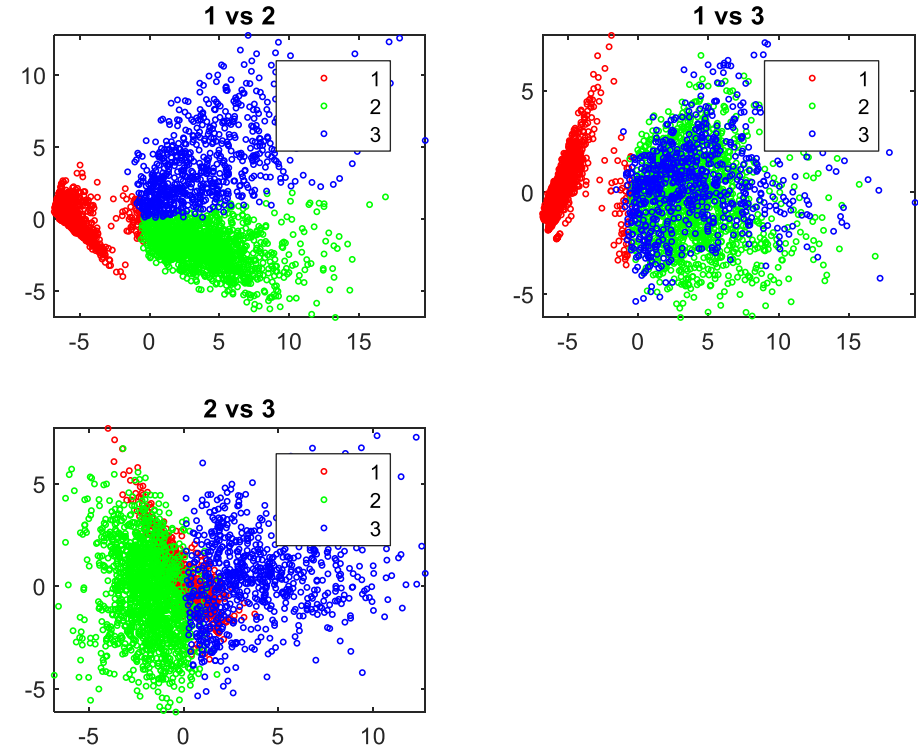


Exercise 4: Clustering Human Activity

Goal: find additional options to improve human activity classification

Approach:

- Reduce dimensionality and visualize structure of data using PCA
- Evaluate different clustering techniques to identify groups of behaviors



Exercise 4: Clustering Human Activity

Goal: find additional options to improve human activity classification

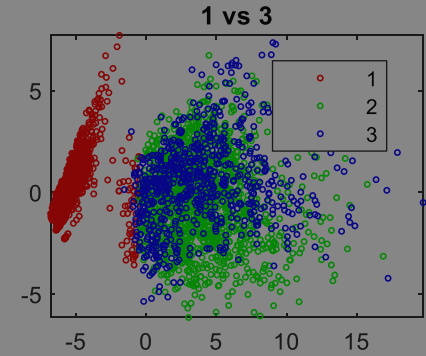
Approach:

- Reduce dimensionality and simplify structure of data
- Evaluate different clustering algorithms to identify groups of similar data points

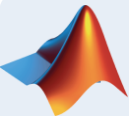
Let's try it out!

*Exercise: **clusteringHumanActivity.mlx**
in folder 04-UnsupervisedLearning*

1 vs 2



Agenda

- Machine learning introduction
- Supervised machine learning models
 - Predicting fuel economy (Regression)
 - Human activity learning (Classification)
- Feature extraction and feature selection
- Unsupervised learning (optional)
-  Working with big data (optional)
 - Deploying Machine Learning Algorithms
 - Wrap-up / Resources Available to ANU

Big Data in MATLAB: Tall Arrays

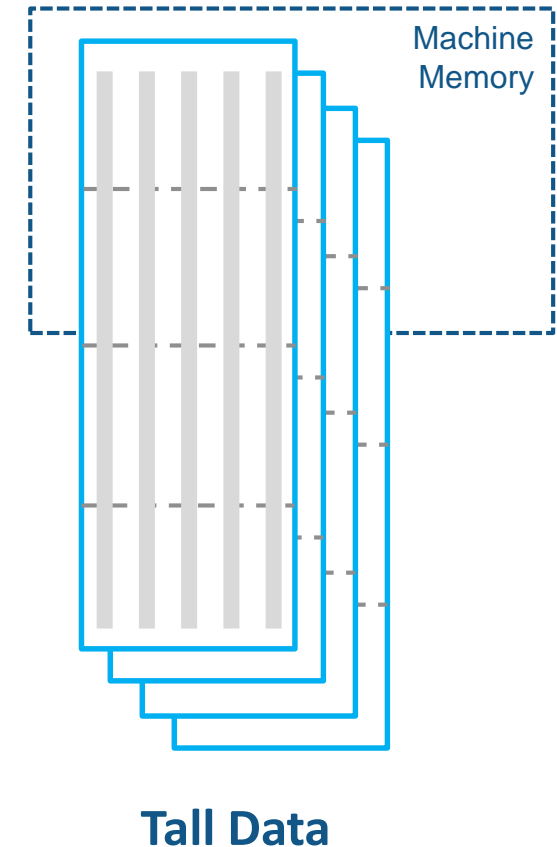
Extends the “array” data type to out-of-memory

- Use like a regular (in-memory) array in supported functions
- (With some setup) Scales processing to clusters with Spark

Applicable when:

- Data is **columnar** – with **many** rows
- Overall data size is **too big to fit into memory**
- Operations are mathematical/statistical in nature

Hundreds of functions supported in MATLAB and Statistics and Machine Learning Toolbox



Big Data Without Big Changes

One file

Access Data

```
measured = readtable('PumpData.csv');
measured = table2timetable(measured);
```

Preprocess Data

Select data of interest

```
measured = measured(timerange(seconds(1),seconds(2)),:)
```

Work with missing data

```
measured = fillmissing(measured,'linear');
```

Calculate statistics

```
m = mean(measured.Speed);
s = std(measured.Speed);
```

One hundred files (Big Data)

Access Data

```
measured = datastore('PumpData*.csv');
measured = tall(measured);
measured = table2timetable(measured);
```

Preprocess Data

Select data of interest

```
measured = measured(timerange(seconds(1),seconds(2)),:)
```

Work with missing data

```
measured = fillmissing(measured,'linear');
```

Calculate statistics

```
m = mean(measured.Speed);
s = std(measured.Speed);
```

```
[m,s] = gather(m,s);
```

Exercise 5: Predicting Tips for Cab Drivers

Goal: Create a model on a (simulated) large dataset

Approach:

- Access data spread across many files
- Preprocess and Explore data
- Train and validate a machine learning model



Exercise 5: Predicting Tips for Cab Drivers

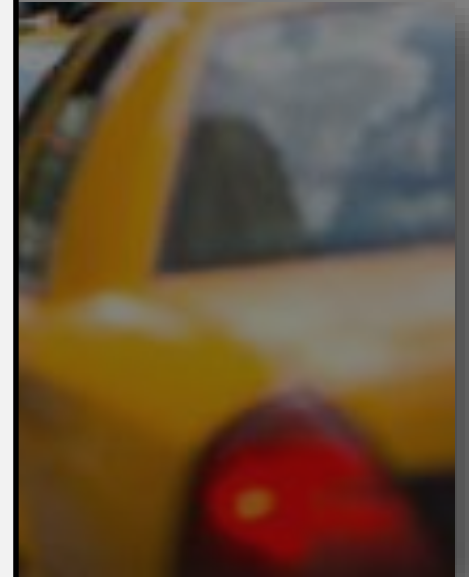
Goal: Create a model on a (simulated)
large dataset

Approach:

- Access data spread
- Preprocess and Exp
- Train and validate a model

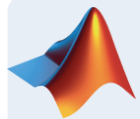
Let's try it out!

*Exercise: **predictDriverTip.mlx**
in folder 05-BigData*



Agenda

- Machine learning introduction
- Supervised machine learning models
 - Predicting fuel economy (Regression)
 - Human activity learning (Classification)
- Feature extraction and feature selection
- Unsupervised learning (optional)
- Working with big data (optional)

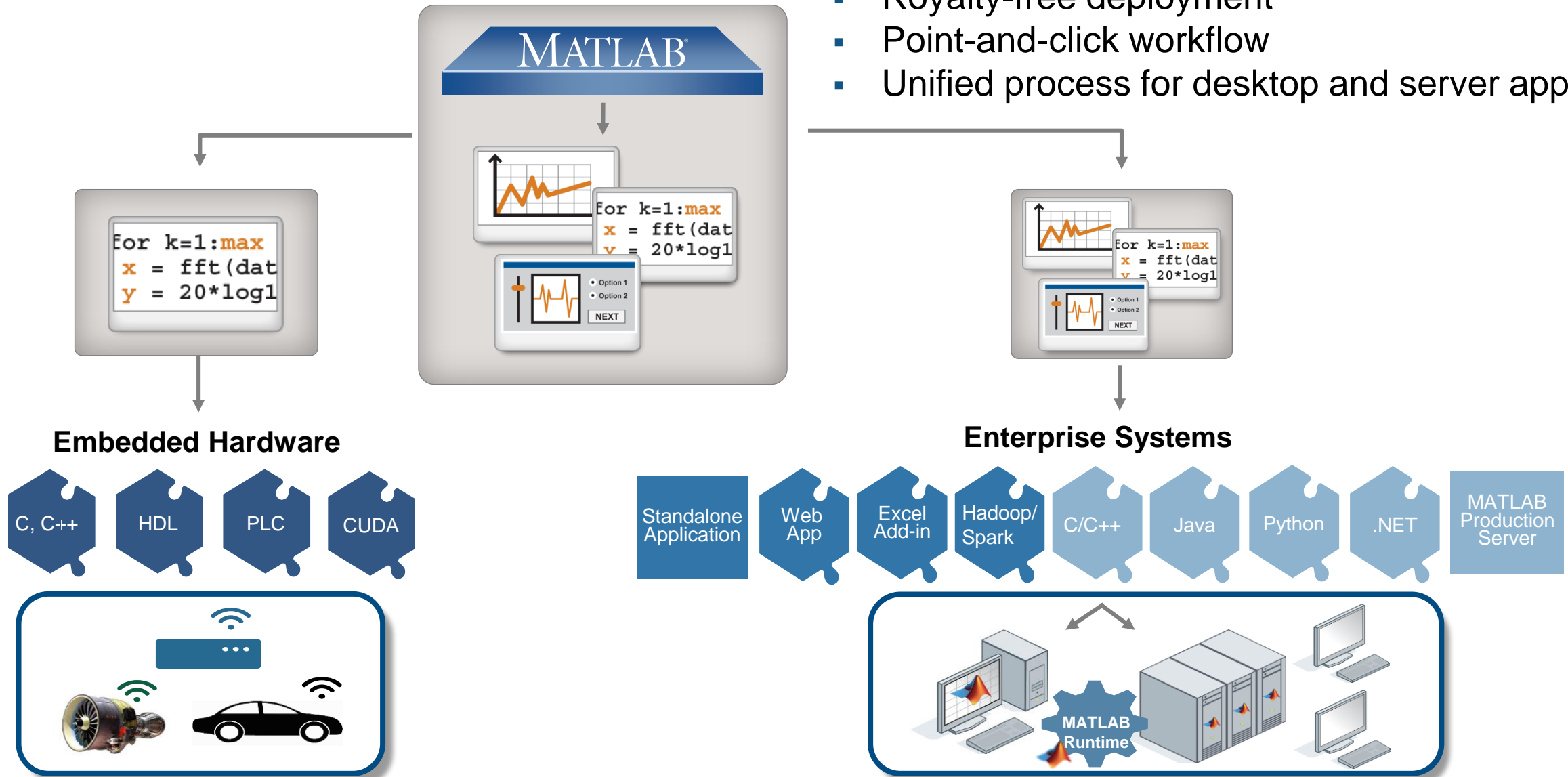


Deploying Machine Learning Algorithms

- Wrap-up / Resources Available to ANU

Deploying MATLAB Algorithms

- Royalty-free deployment
- Point-and-click workflow
- Unified process for desktop and server apps



Deploying MATLAB Algorithms

- Royalty-free deployment

desktop and server apps

Let's observe!

Try this on your own:

[MATLAB Doc - Code Generation Workflows Using MATLAB Coder App](#)

MATLAB Command:

```
>>openExample('stats/CodeGenerationWorkflowsUsing  
MATLABCoderAppExample')
```

```
for k=1:max  
x = fft(dat  
y = 20*log1
```

Embedded Hardware



Beyond traditional Machine Learning: Deep Learning

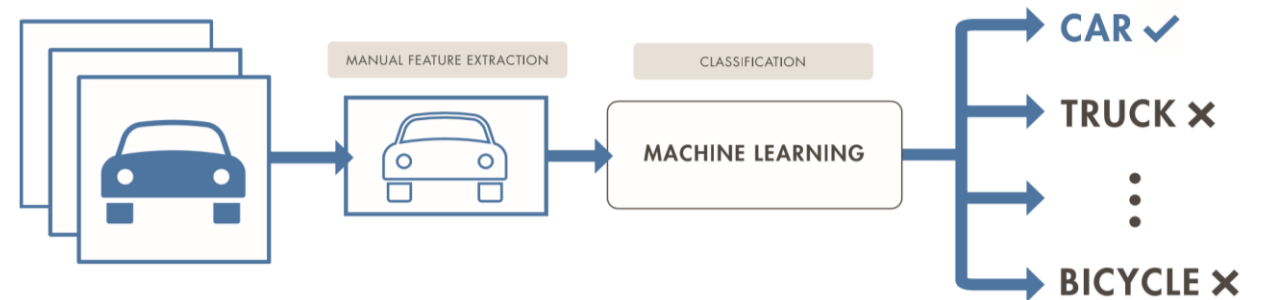
Machine Learning

Deep Learning

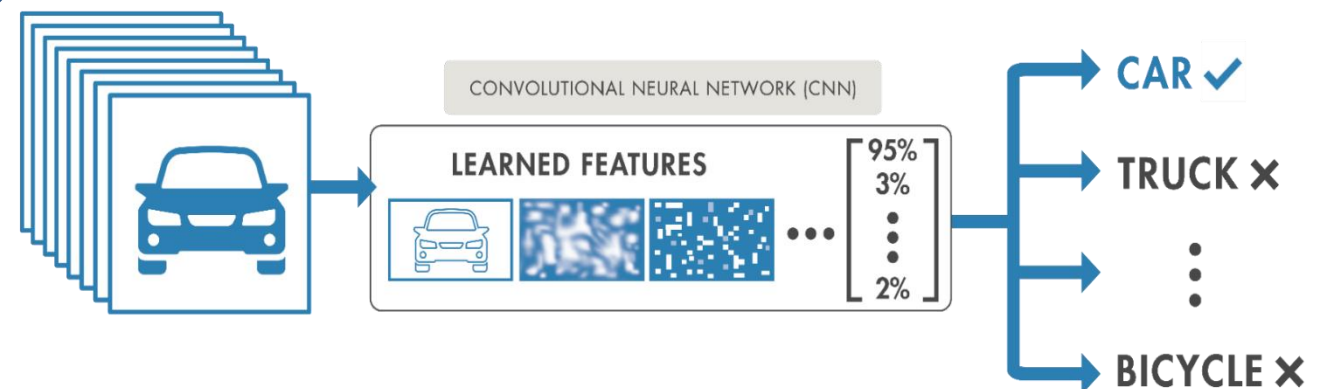
Neural Networks
with many Hidden
Layers

- Learns directly from data
- More Data = better model
- Computationally Intensive
- **Not interpretable**

Machine Learning



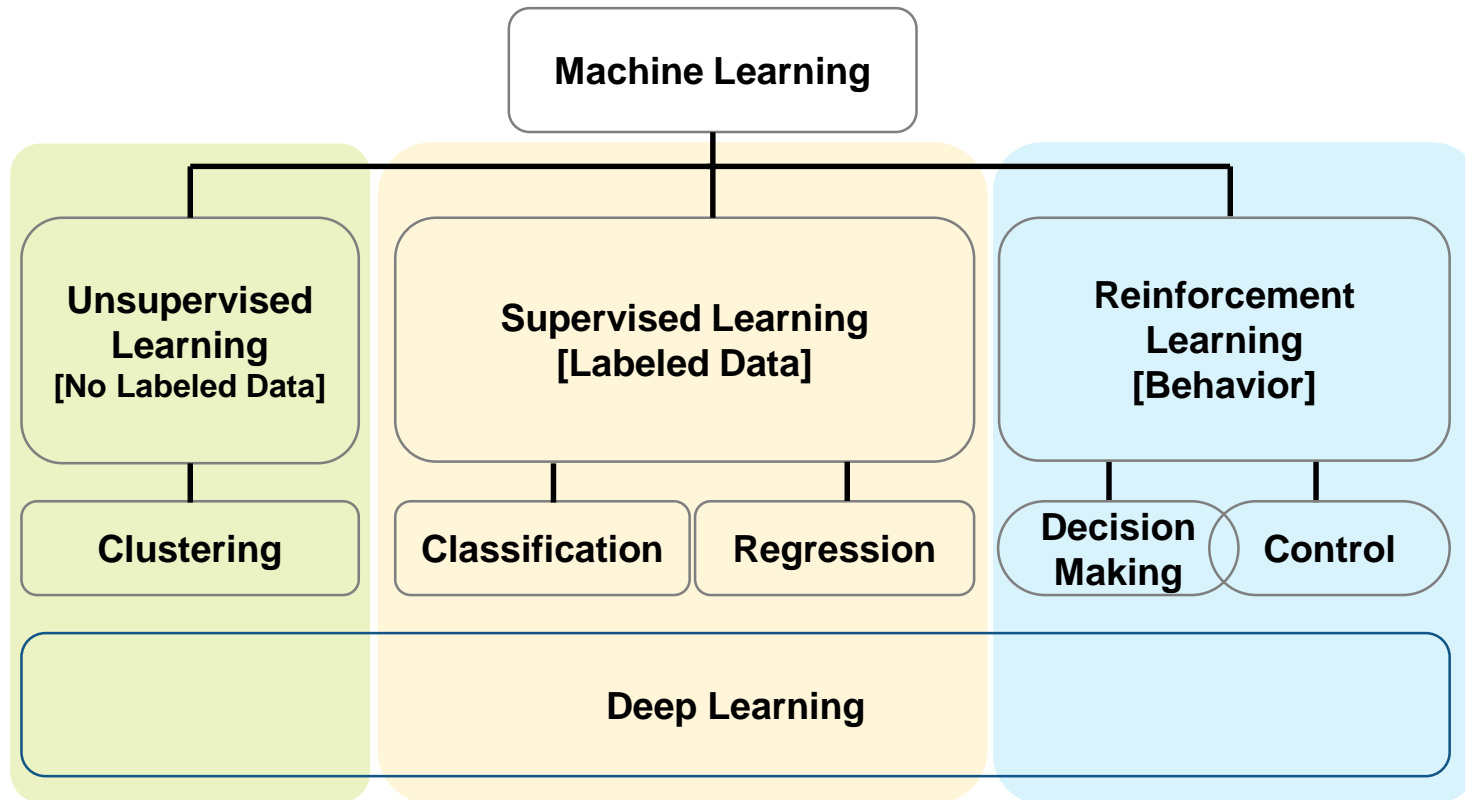
Deep Learning



[Deep Learning and Traditional Machine Learning: Choosing the Right Approach](#)

[MATLAB Doc - Example of Deep Learning for Image Recognition](#)

Beyond Machine Learning: Reinforcement Learning



Reinforcement learning:

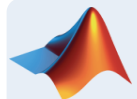
Learning through trial & error
[*interaction data*]

Complex problems typically
need deep learning

It's about learning a **behavior**
or accomplishing a **task**

Agenda

- Machine learning introduction
- Supervised machine learning models
 - Predicting fuel economy (Regression)
 - Human activity learning (Classification)
- Feature extraction and feature selection
- Unsupervised learning (optional)
- Working with big data (optional)
- Deploying Machine Learning Algorithms



Wrap-up / Resources Available to ANU

What's our AutoML?

Automate main steps to minimize expertise needed and increase productivity

Access and explore data

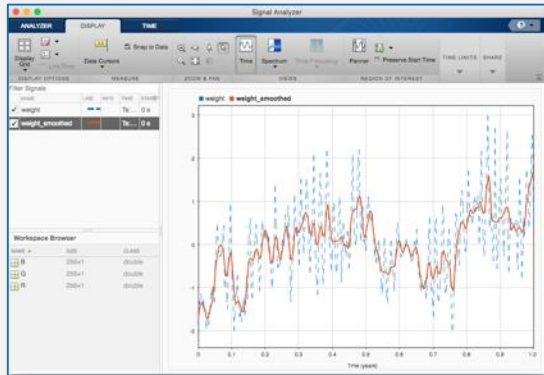
Preprocessing

Feature Engineering

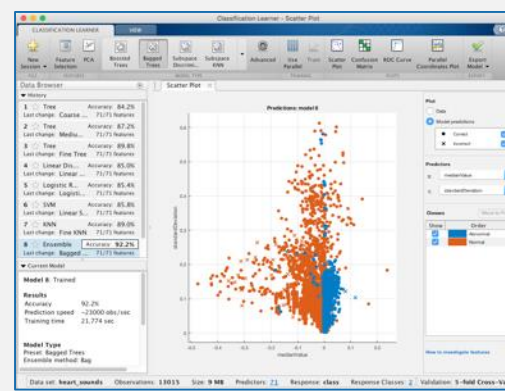
Model Training

Model Tuning

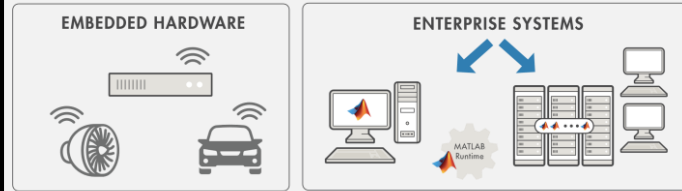
Integrate Analytics



Datatypes and tools for missing data, outliers, time-alignment, etc.



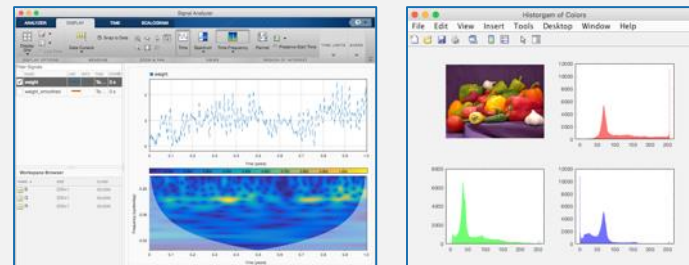
Machine Learning apps



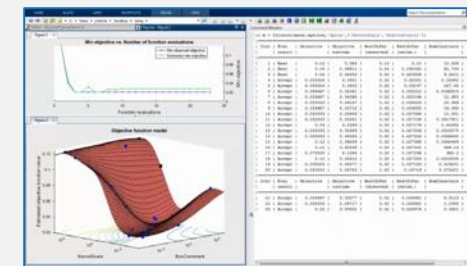
C/C++ Code Generation and Enterprise IT Integration

Region	OutageTime	Loss	Customers	RestorationTime	Cause
SouthWest	2002-02-01 12:18	418.9772218	1820159.482	2002-02-07 16:50	winter storm
SouthEast	2003-01-23 00:49	510.1399497	212035.3001		winter storm
SouthEast	2003-02-07 21:15	289.4035493	142938.6282	2003-02-17 08:14	winter storm
West	2004-04-08 05:41	414.8053524	340371.0338	2004-04-08 06:10	equipment fault
MidWest	2002-03-16 06:18	186.4367788	212754.0	Converted TdType: Number, Value: 0.0	ign storm
West	2003-06-18 02:49	0	0.0	2003-06-18 10:54	attack
West	2004-06-20 14:39	231.2947226		2004-06-20 19:16	equipment fault

Text files, spreadsheets, databases, binary files, data feeds, web, cloud storage



Domain-specific techniques for Signals, Images, Video, Audio, and Text



Automated Parameter Tuning

Using MATLAB with Other Languages

Calling Libraries Written in Another Language From MATLAB



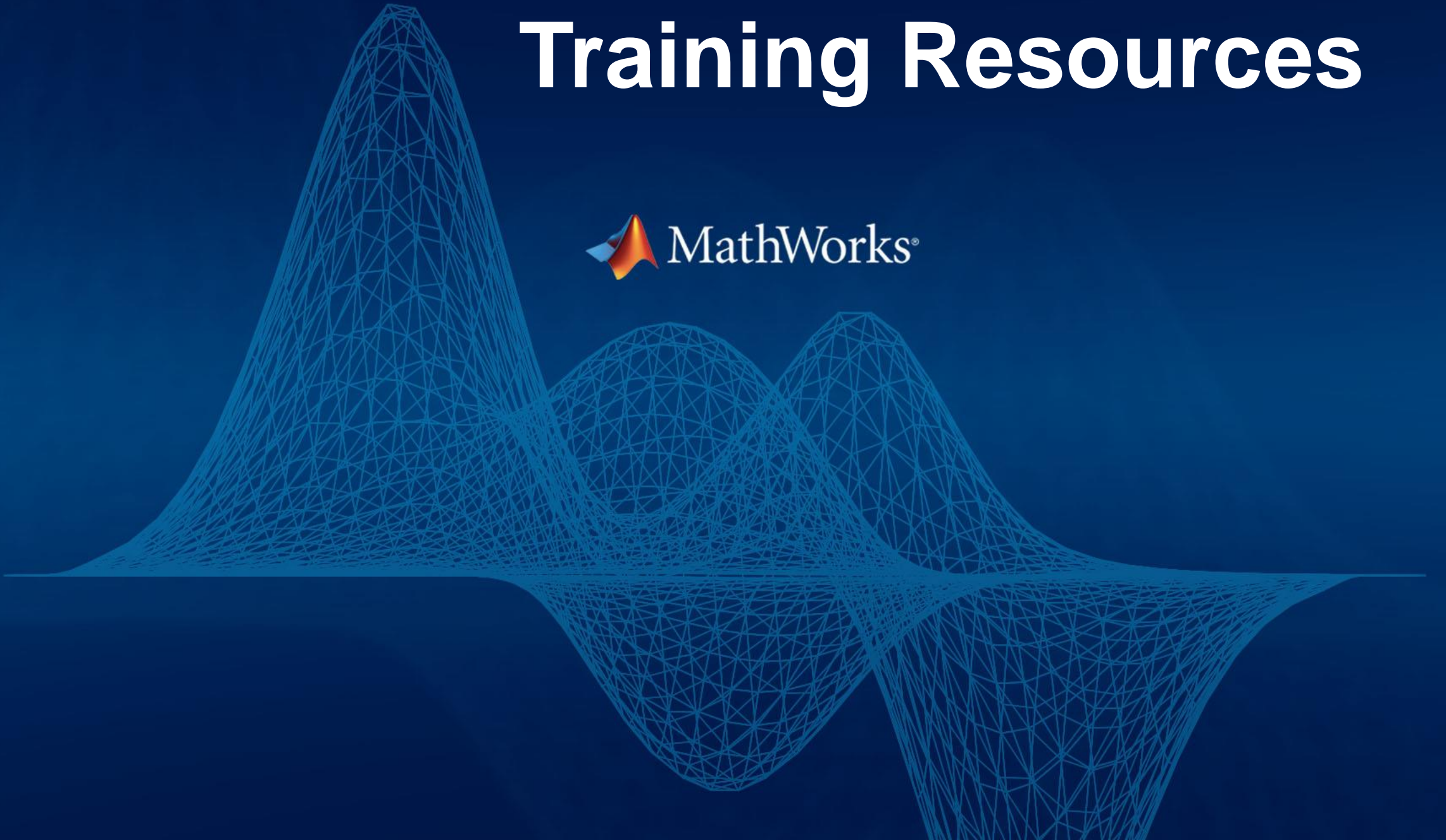
- Java
- Python
- C
- C++
- Fortran
- COM components and ActiveX[®] controls
- RESTful, HTTP, and WSDL web services

Calling MATLAB from Another Language



- Java
- Python
- C/C++
- Fortran
- COM Automation server

Training Resources



Self-Paced Online Courses

<https://matlabacademy.mathworks.com/>

Get Started

MATLAB Onramp
2 hours

Deep Learning Onramp
2 hours

Simulink Onramp
3 hours

Stateflow Onramp
Learn the basics of creating, editing, and simulating state machines in Stateflow.
2 hours

Machine Learning Onramp
Learn the basics of practical machine learning methods for classification problems.
2 hours

Free
5 courses

The screenshot shows the MATLAB Academy interface for the 'Deep Learning Onramp' course. The task pane on the left contains instructions for 'Task 2' and a 'Submit' button. The code editor shows the MATLAB code: `img1 = imread('file01.jpg');`. The workspace window displays a matrix `img1` and a preview of the image loaded into the variable.



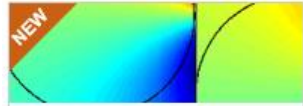
Why not do this today !

Measurable progress report and completion certificate
Interactive lessons with immediate feedback
24/7 availability

Self-Paced Online Courses

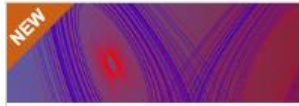
Computational Mathematics

*Available only to users at universities that offer campus-wide online training access.



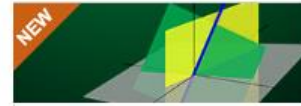
Solving Nonlinear Equations with MATLAB

1.5 hours



Solving Ordinary Differential Equations with MATLAB

2 hours



Introduction to Linear Algebra with MATLAB

1.5 hours



Introduction to Statistical Methods with MATLAB

2 hours



Introduction to Symbolic Math with MATLAB

2 hours

5 courses targeting MATLAB skills needed in the classroom

**Free
Of charge**

11 courses

Core MATLAB Functionality



MATLAB Fundamentals

20 hours



MATLAB Programming Techniques

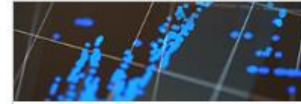
14 hours



MATLAB for Financial Applications

20 hours

Data Analytics



MATLAB for Data Processing and Visualization

7 hours



Machine Learning with MATLAB

14 hours



Deep Learning with MATLAB

14 hours

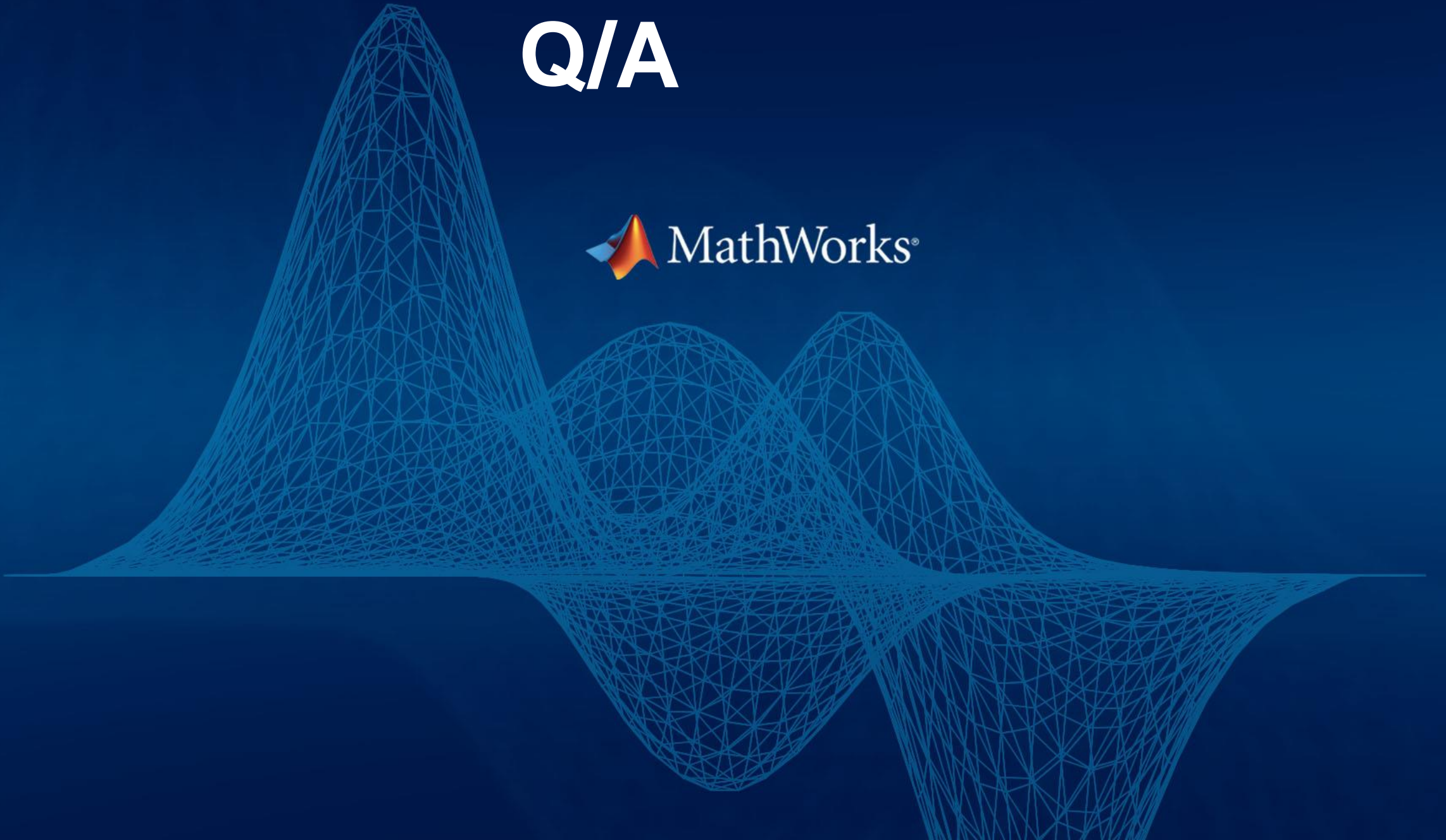
6 in-depth courses for enhancing MATLAB skills

ANU has this as part of their MATLAB Campus License
So you have FULL access to these courses



<https://matlabacademy.mathworks.com/>

Q/A



Q/A: Would you like to know more ?

- Questions

<https://www.mathworks.com/solutions/machine-learning.html>

Machine Learning

Overview | Latest Features | Resources | Customer Success

MATLAB for Machine Learning

Train models, tune parameters, and deploy to production or the edge

[Download a free trial](#)

Using MATLAB®, engineers and other domain experts have deployed thousands of machine learning applications. MATLAB makes the hard parts of machine learning easy with:

- Point-and-click apps for training and comparing models
- Advanced signal processing and feature extraction techniques
- Automatic hyperparameter tuning and feature selection to optimize model performance
- The ability to use the same code to scale processing to big data and clusters
- Automated generation of C/C++ code for embedded and high-performance applications
- Popular classification, regression, and clustering algorithms for supervised and unsupervised learning
- Faster execution than open source on most statistical and machine learning computations

Deep Learning or Machine Learning?

Deep Learning vs. Machine Learning: Choosing the Best Approach (Ebook)

MATLAB for Deep Learning

Overview | Topics for Deep Learning | Latest Features | Resources | Models

MATLAB for Deep Learning

Data preparation, design, simulation, and deployment for deep neural networks

[Download a free trial](#)

With just a few lines of MATLAB® code, you can apply deep learning techniques to your work whether you're designing algorithms, preparing and labeling data, or generating code and deploying to embedded systems.

With MATLAB, you can:

- Create, modify, and analyze deep learning architectures using apps and visualization tools.
- Preprocess data and automate ground-truth labeling of image, video, and audio data using apps.
- Accelerate algorithms on NVIDIA® GPUs, cloud, and datacenter resources without specialized programming.
- Collaborate with peers using frameworks like TensorFlow, PyTorch, and MxNet.

Top 5 Reasons to Use MATLAB for Deep Learning