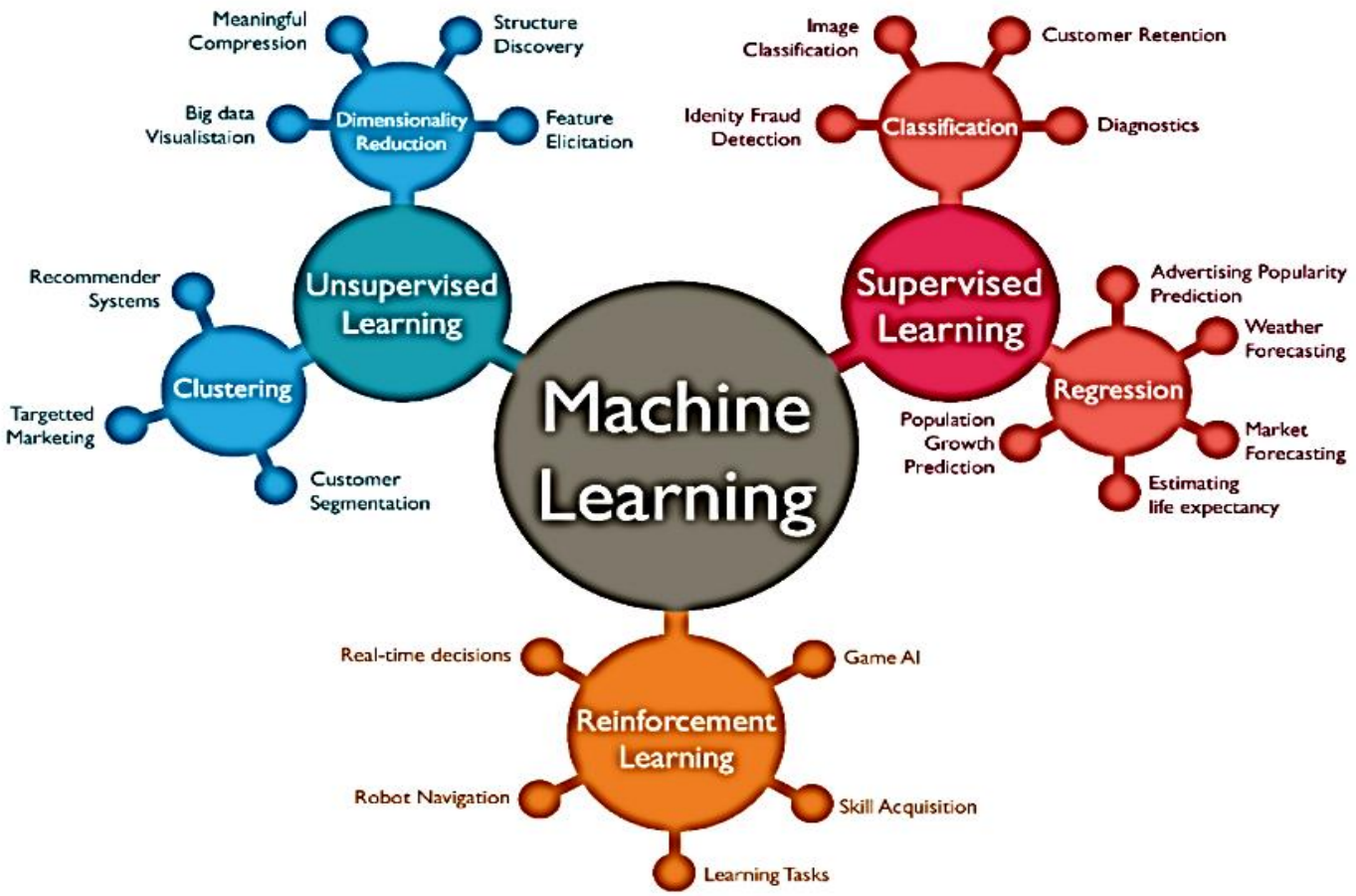
Introduction to Machine Learning

Control Tech LAB

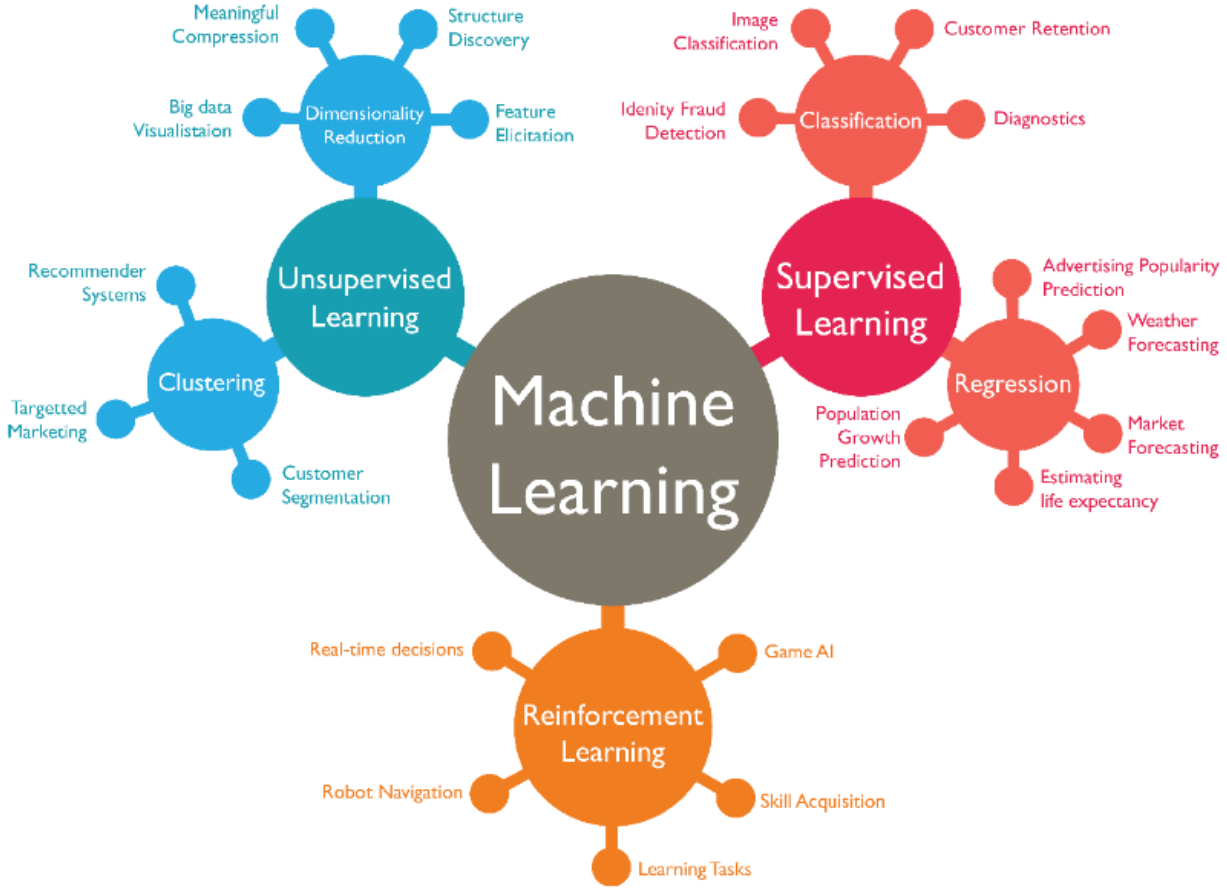


Control Tech LAB

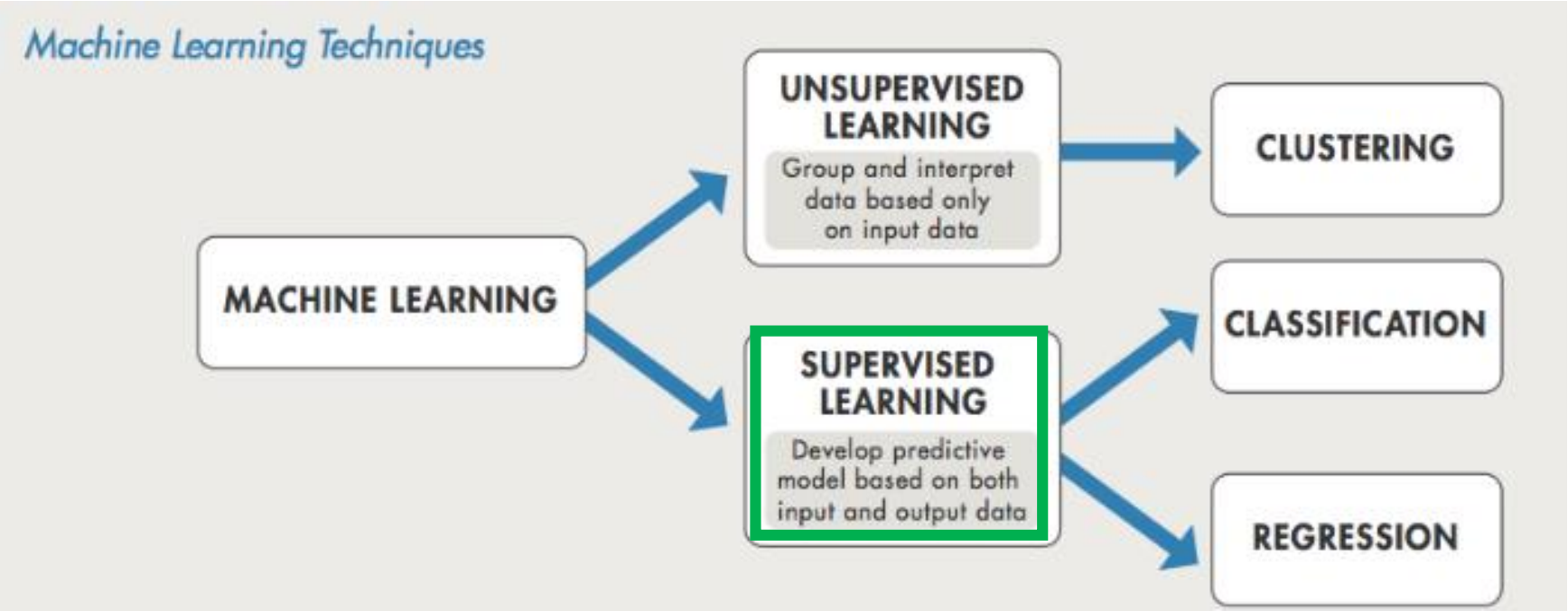
Working with:

1. Introduction to Machine Learning Presentation
2. Classification with Machine Learning and the TC Lab (A live Script that includes a video)

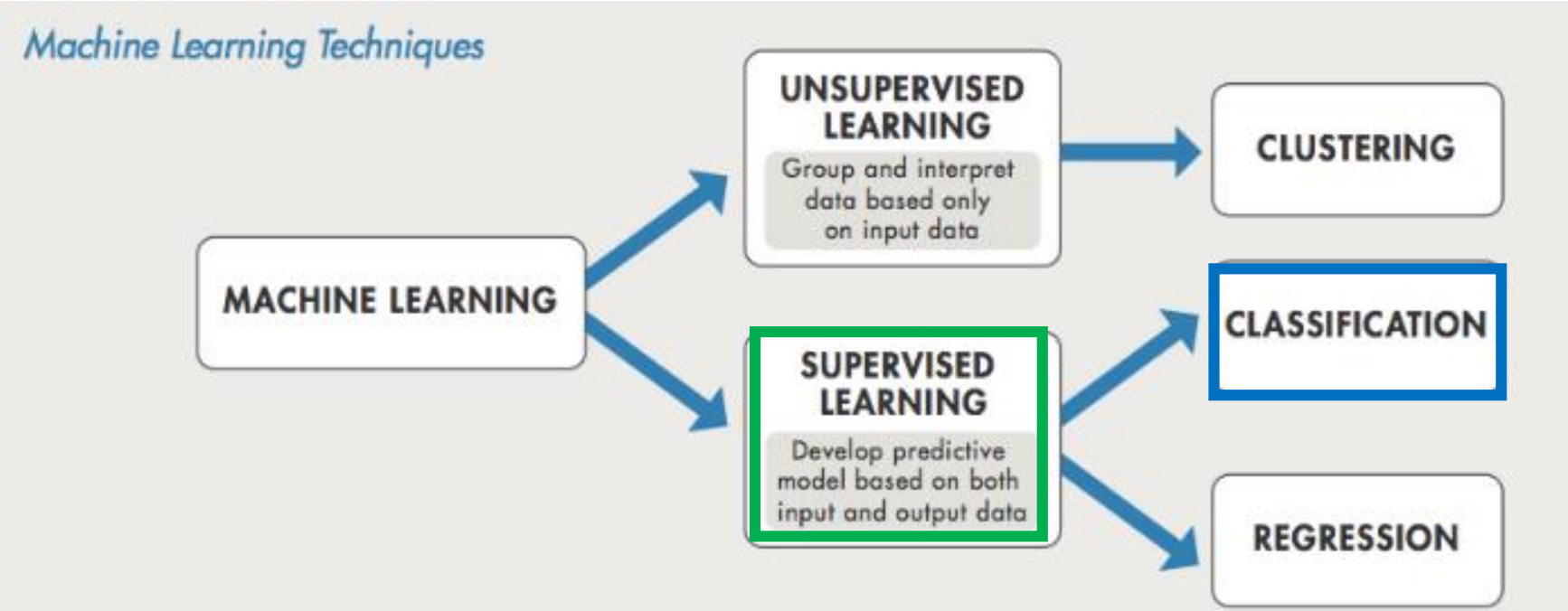
Machine Learning teaches computers to do what comes naturally to humans: learn from experience. Machine learning algorithms use computational methods to “learn” information directly from data without relying on a predetermined equation as a model. The algorithms adaptively improve their performance as the number of samples available for learning increases.



Machine learning uses two types of techniques: supervised learning, which **trains a model** on known input and output data to predict future outputs, and unsupervised learning, which finds hidden patterns or intrinsic structures in input data.



The aim of **supervised** machine learning is to **build a model that makes predictions based on evidence**. A supervised learning algorithm takes a known set of input data and known responses to the data (output) and trains a model to generate reasonable predictions for the response to new data.



***Classification* techniques** predict categorical responses, for example, whether an email is genuine or spam, or whether, a tumor is cancerous or benign. Classification models classify input data into categories. Typical applications include medical imaging, image and speech recognition, and credit scoring.

Machine Learning Onramp

6 modules | 2 hours | Languages

Learn the basics of practical machine learning methods for classification problems.

1. Overview of Machine Learning

Familiarize with machine learning concepts and the course.

2. Classification Workflow

Build a simple model to perform a classification task.

3. Importing and Preprocessing Data

Import data from multiple files.

4. Engineering Features

Calculate features from raw signals.

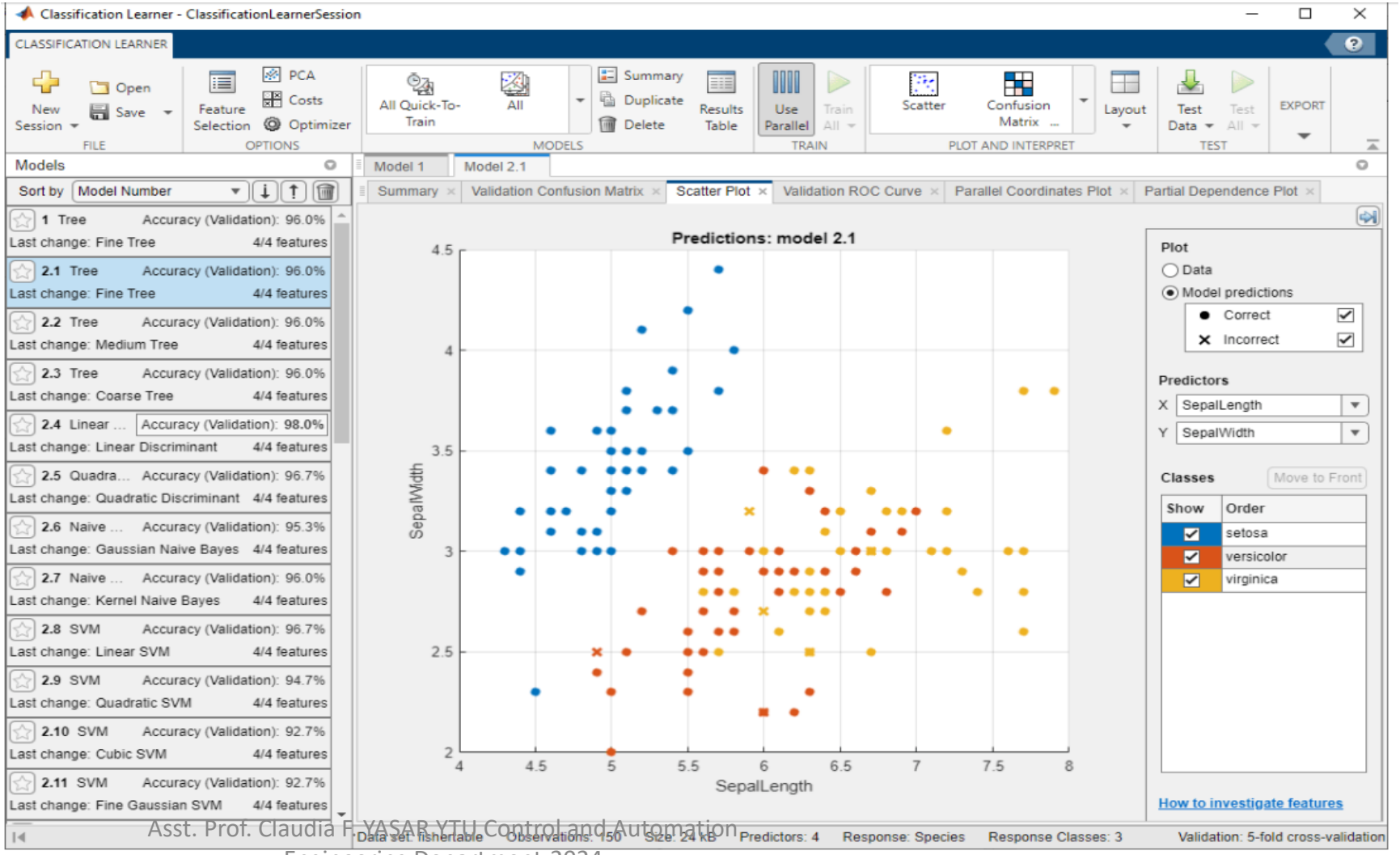
5. Classification Models

Train and use machine learning models to make predictions.

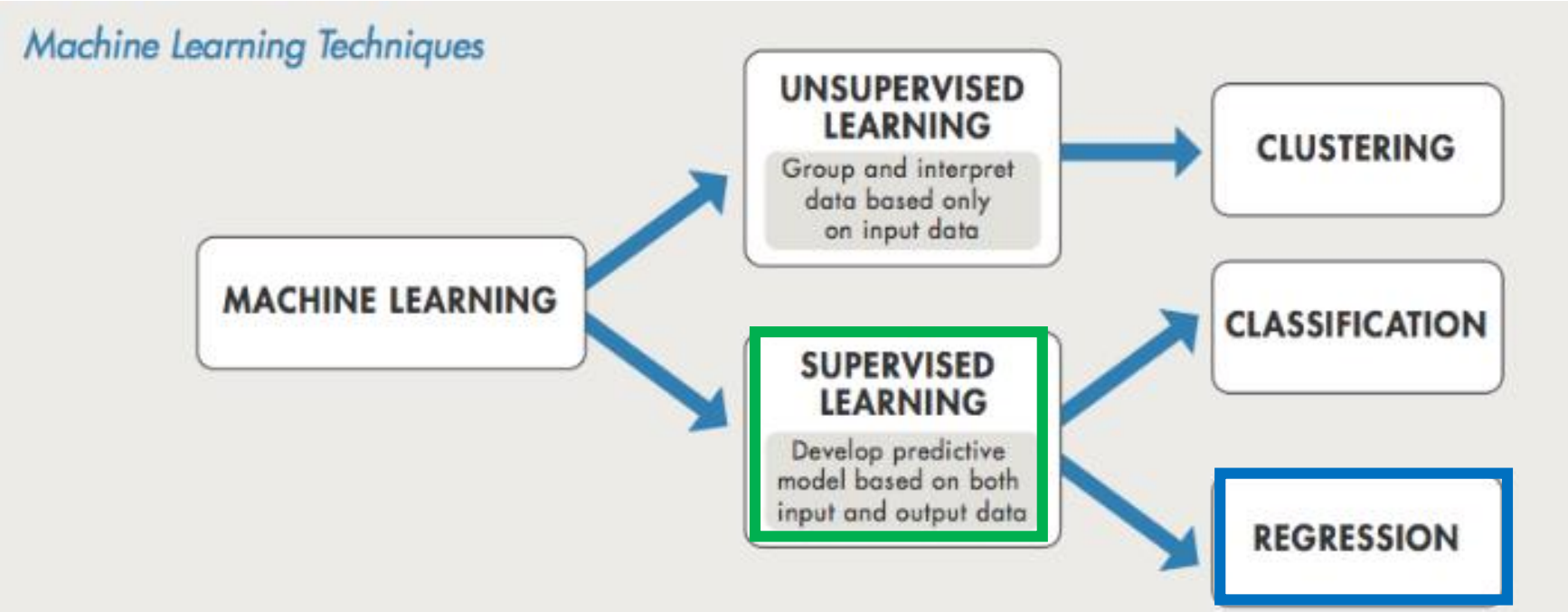
https://matlabacademy.mathworks.com/details/machine-learning-onramp/machinelearning

Train Classification Models in Classification Learner App

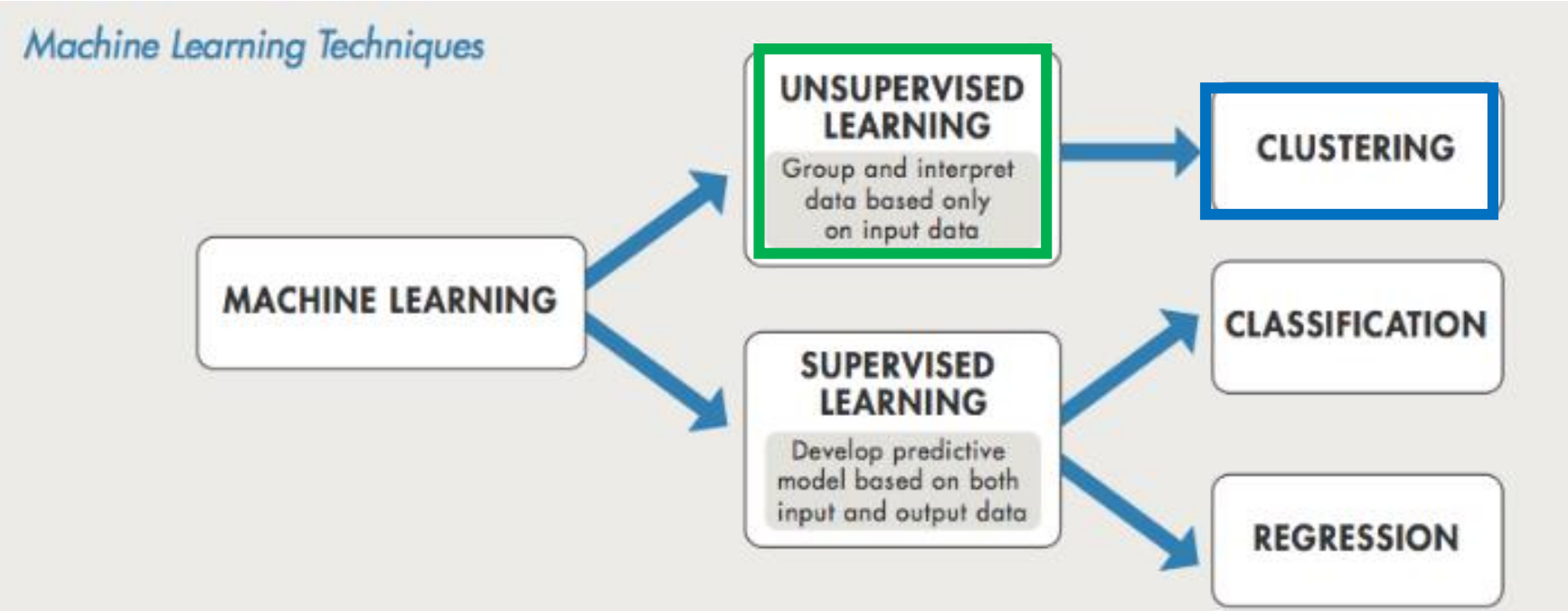
Use the Classification Learner app to train models to classify data using supervised machine learning. The app lets you explore supervised machine learning interactively using various classifiers.



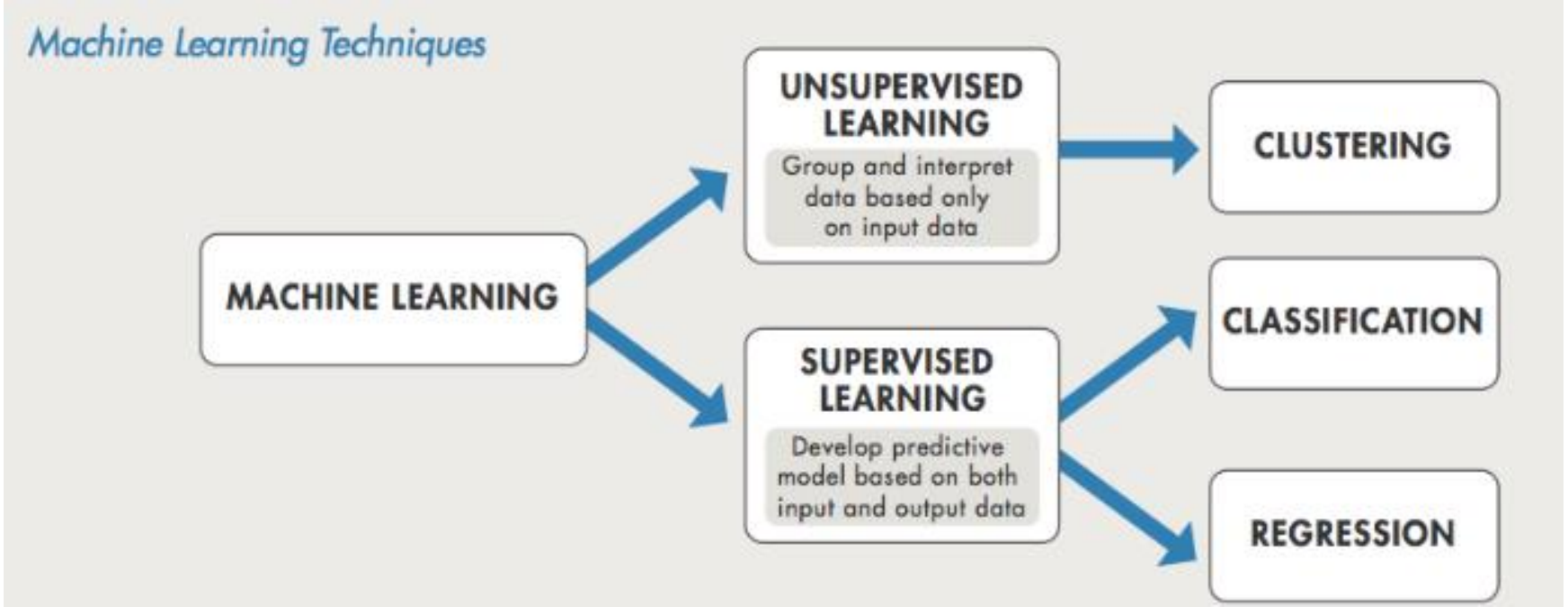
https://www.mathworks.com/help/stats/train-classification-models-in-classification-learner-app.html



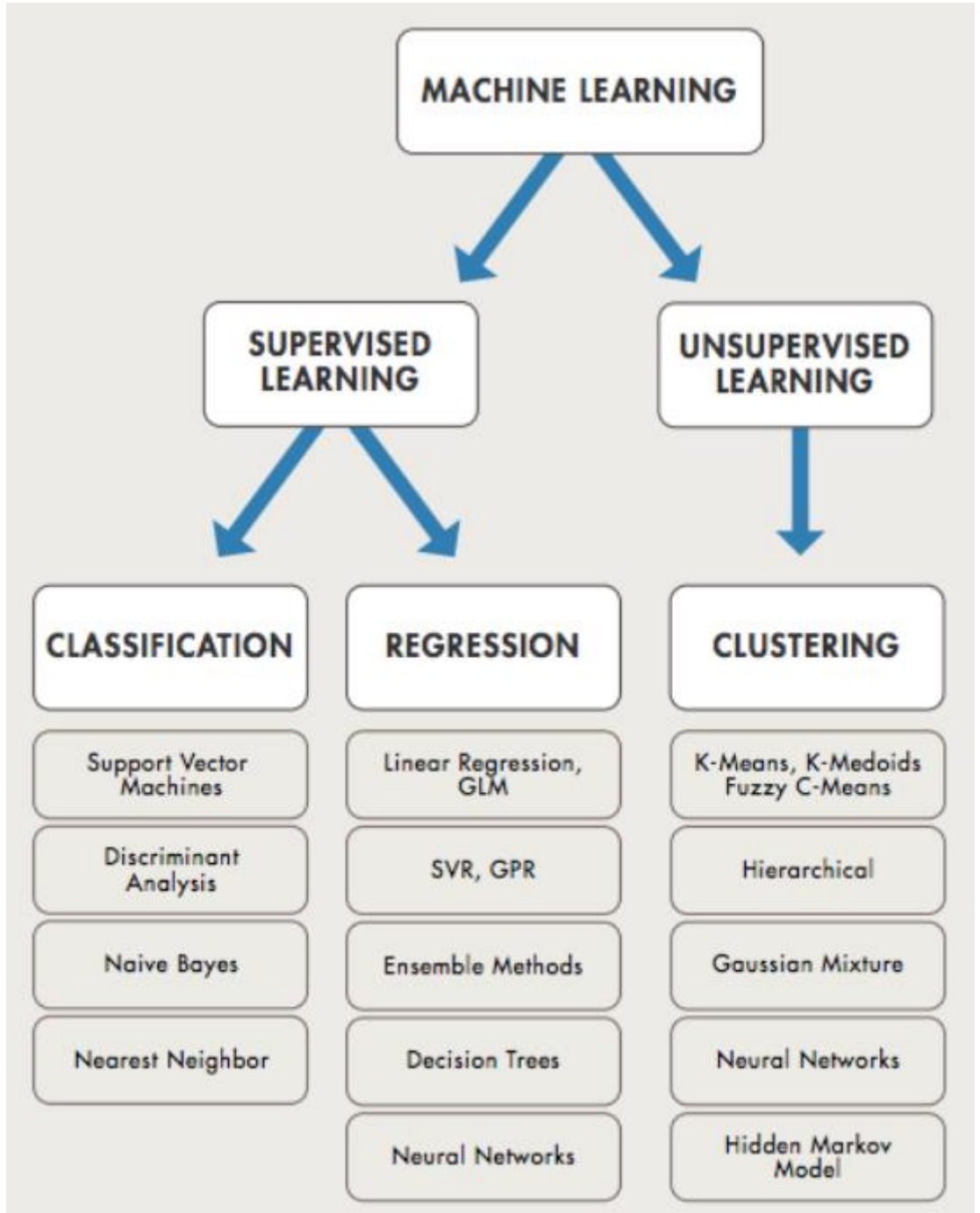
Regression techniques **predict continuous responses**, for example, changes in temperature or fluctuations in power demand. Typical applications include electricity load forecasting and algorithmic trading.



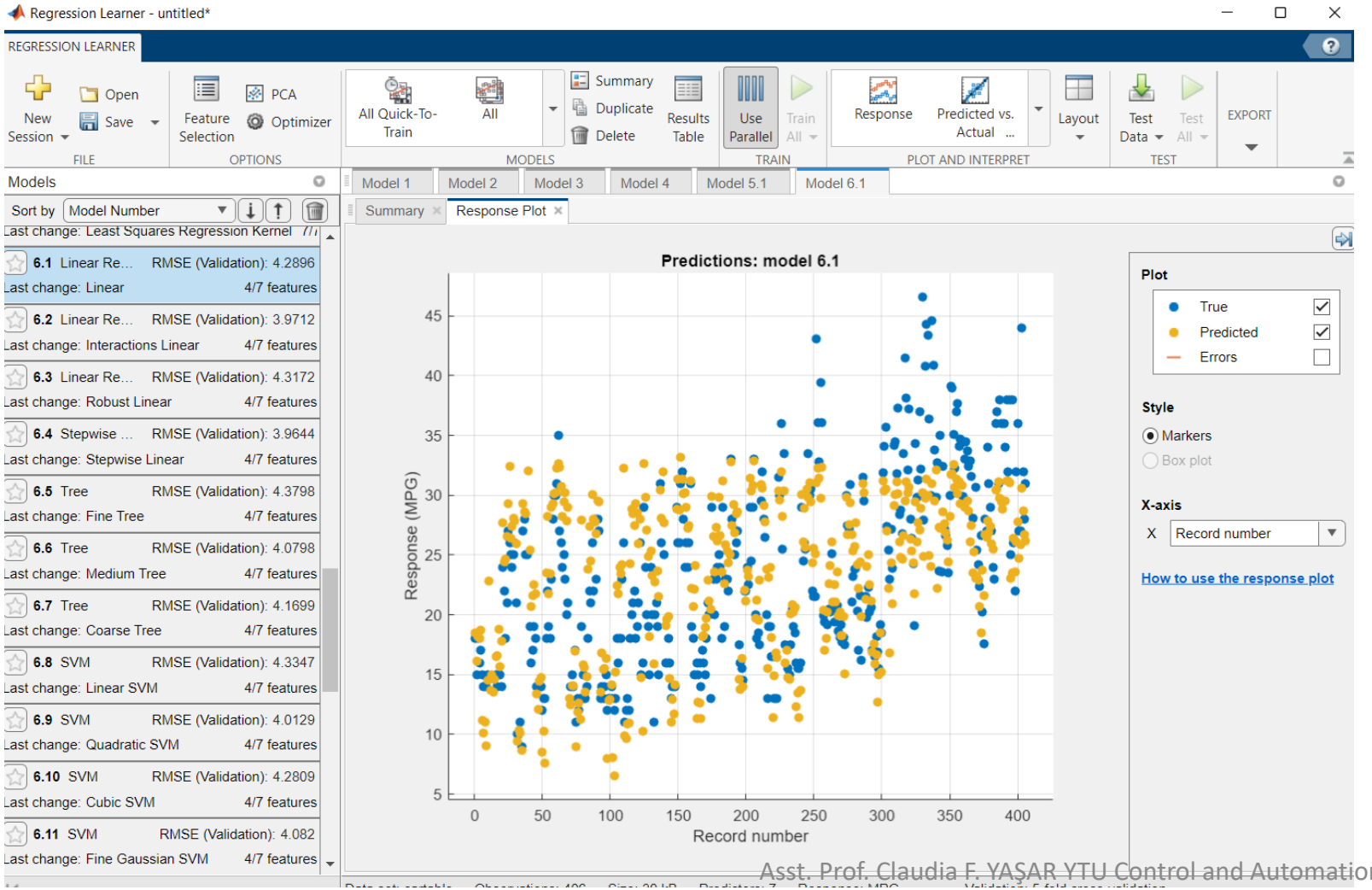
Unsupervised learning **finds hidden patterns or intrinsic structures in data**. It is used to **draw inferences from datasets.** ***Clustering*** **is** the most common unsupervised learning technique. It is used for exploratory data analysis to find hidden patterns or groupings in data. Clustering applications include gene sequence analysis, market research, and object recognition.

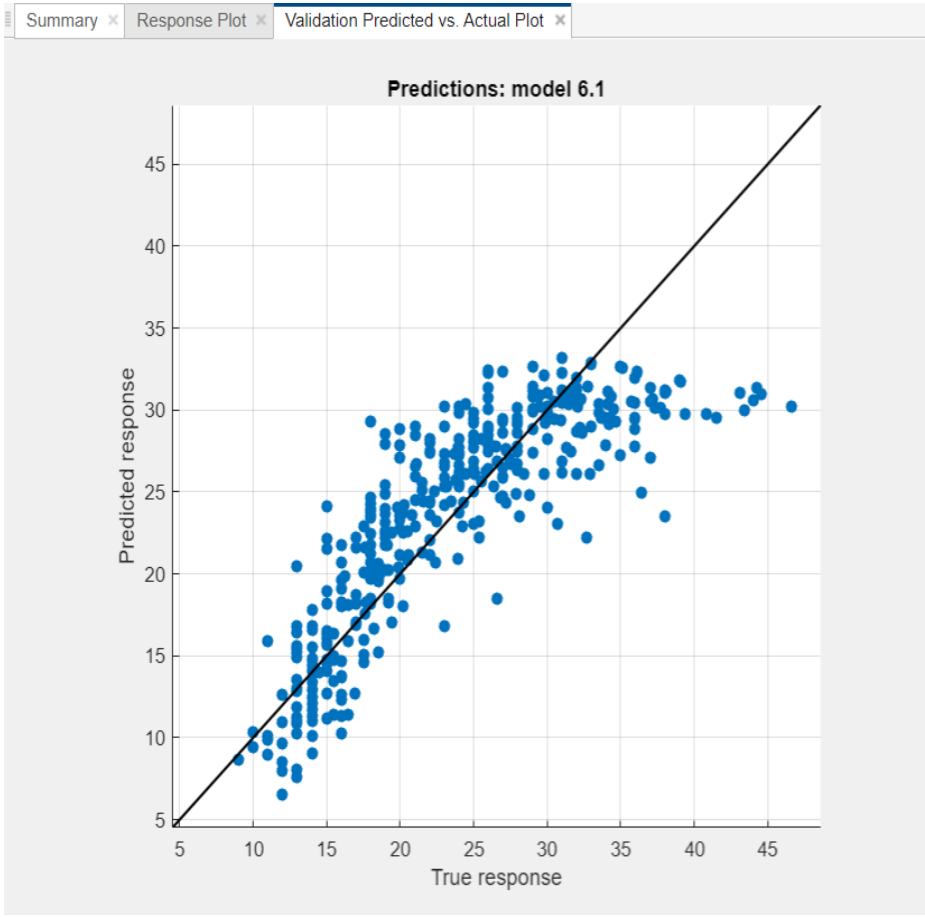


Regression to predict continuous responses	Use the Regression Learner app to automatically train a selection of models and help you choose the best. You can generate MATLAB code to work with scripts and other function options. For more options, you can use the command-line interface.	Statistics and Machine Learning Toolbox	Train Regression Models in Regression Learner App Regression Functions
Clustering	Use cluster analysis functions.	Statistics and Machine Learning Toolbox	Cluster Analysis


MACHINE LEARNING
SUPERVISED LEARNING
UNSUPERVISED LEARNING
CLASSIFICATION
REGRESSION
CLUSTERING
Support Vector Machines
Discriminant Analysis
Naive Bayes
Nearest Neighbor
Linear Regression, GLM
SVR, GPR
Ensemble Methods
Decision Trees
Neural Networks
K-Means, K-Medoids Fuzzy C-Means
Hierarchical
Gaussian Mixture
Neural Networks
Hidden Markov Model

Train Regression Models in Regression Learner App including linear regression models, regression trees, Gaussian process regression models, support vector machines, kernel approximation, ensembles of regression trees, and neural network regression models. You can explore your data, select features, specify validation schemes, and evaluate results.





There is no best method or one size fits all !

Selecting the Right Algorithm

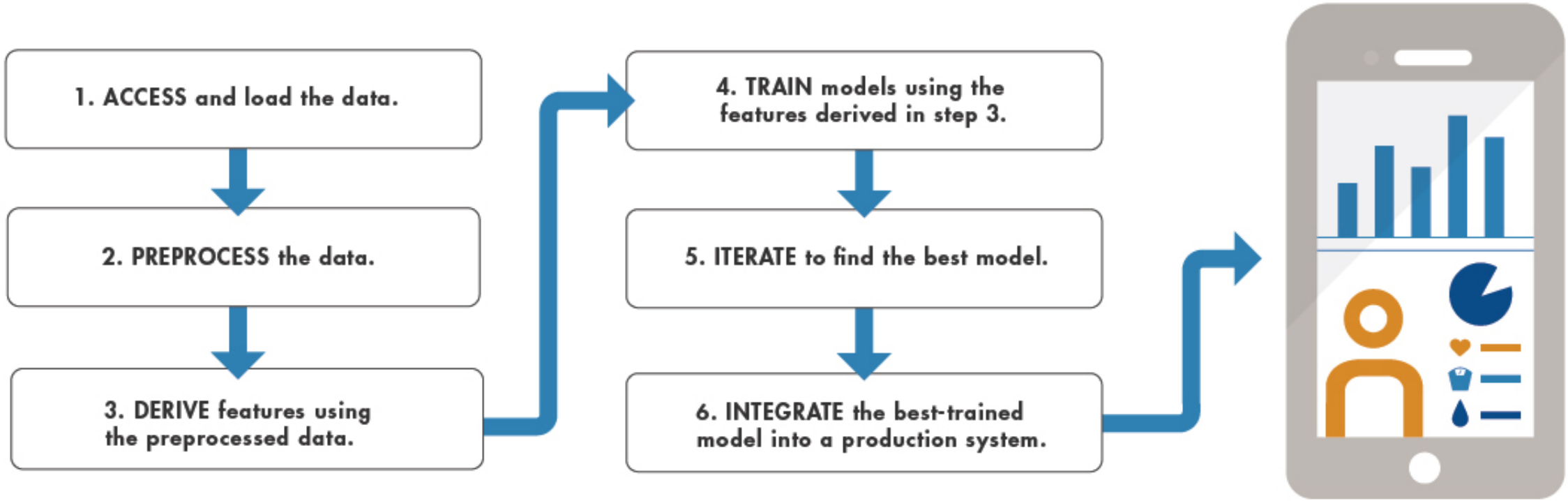
There are dozens of supervised and unsupervised machine learning algorithms, each taking a different learning approach.

Finding the right algorithm is partly based on trial and error - even highly experienced data scientists cannot tell whether an algorithm will work without trying it out.

Highly flexible models tend to over fit data by modelling minor variations that could be noise. Simple models are easier to interpret but might have lower accuracy. Therefore, choosing the right algorithm requires trading off one benefit against another, including model speed, accuracy, and complexity.

Trial and error is at the core of machine learning. If one approach or algorithm does not work, you try another.

The following systematic machine learning workflow can help you tackle machine learning challenges



Classification with Machine Learning and TC Lab

REAL-TIME EXPERIMENTS - TC Lab -

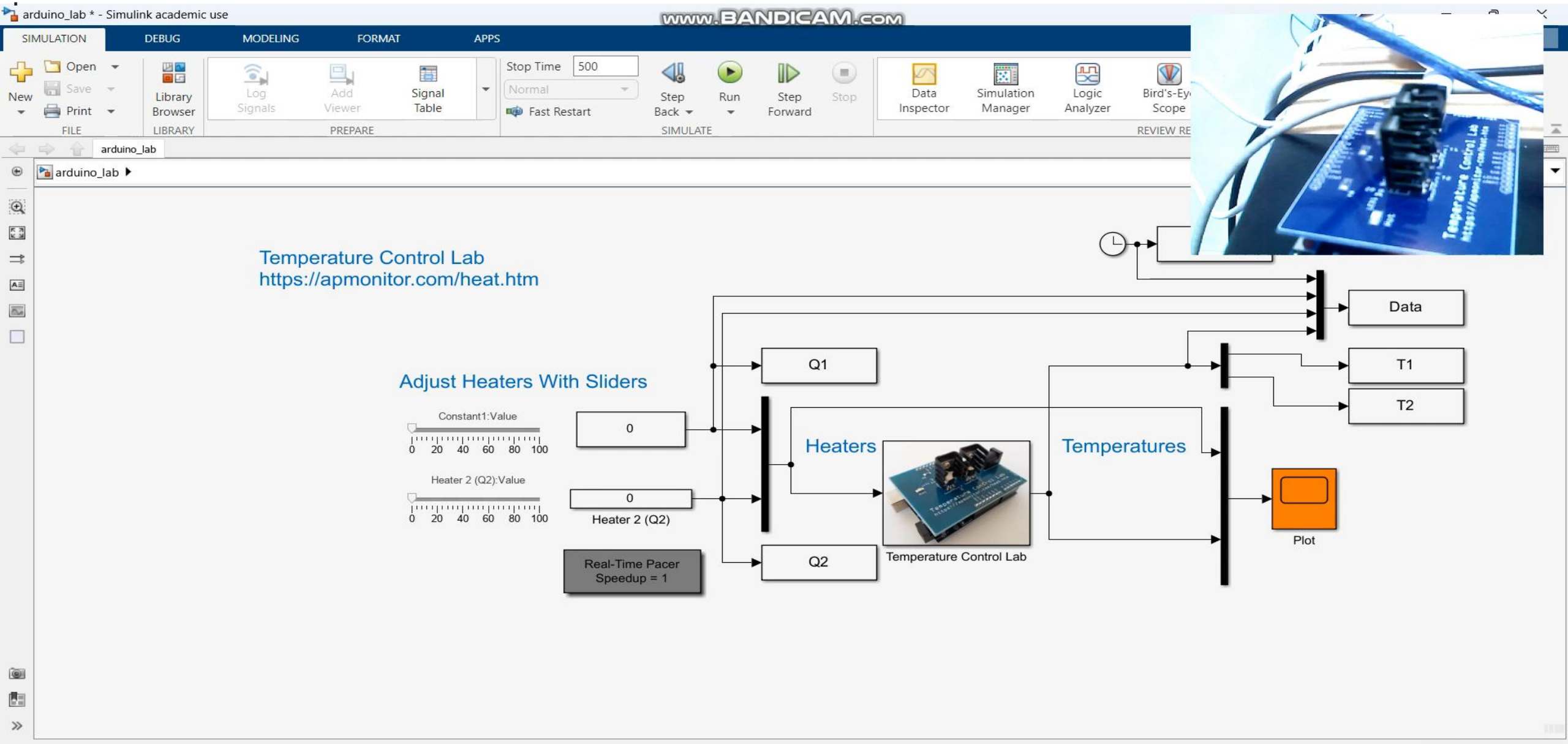
Discover the power of supervised machine learning with the Classification Learner app, augmented with insights from the TC Lab. Effortlessly train classification models from data exploration to model refinement, while also leveraging diverse classifiers for an intuitive learning experience. Unveil the latent potential within your datasets with the seamless integration of the TC Lab.

- Generate a Small Data Set
- Processing the collected data
- Loading Large Data
- Processing the loaded data
- Use the Classification Learner

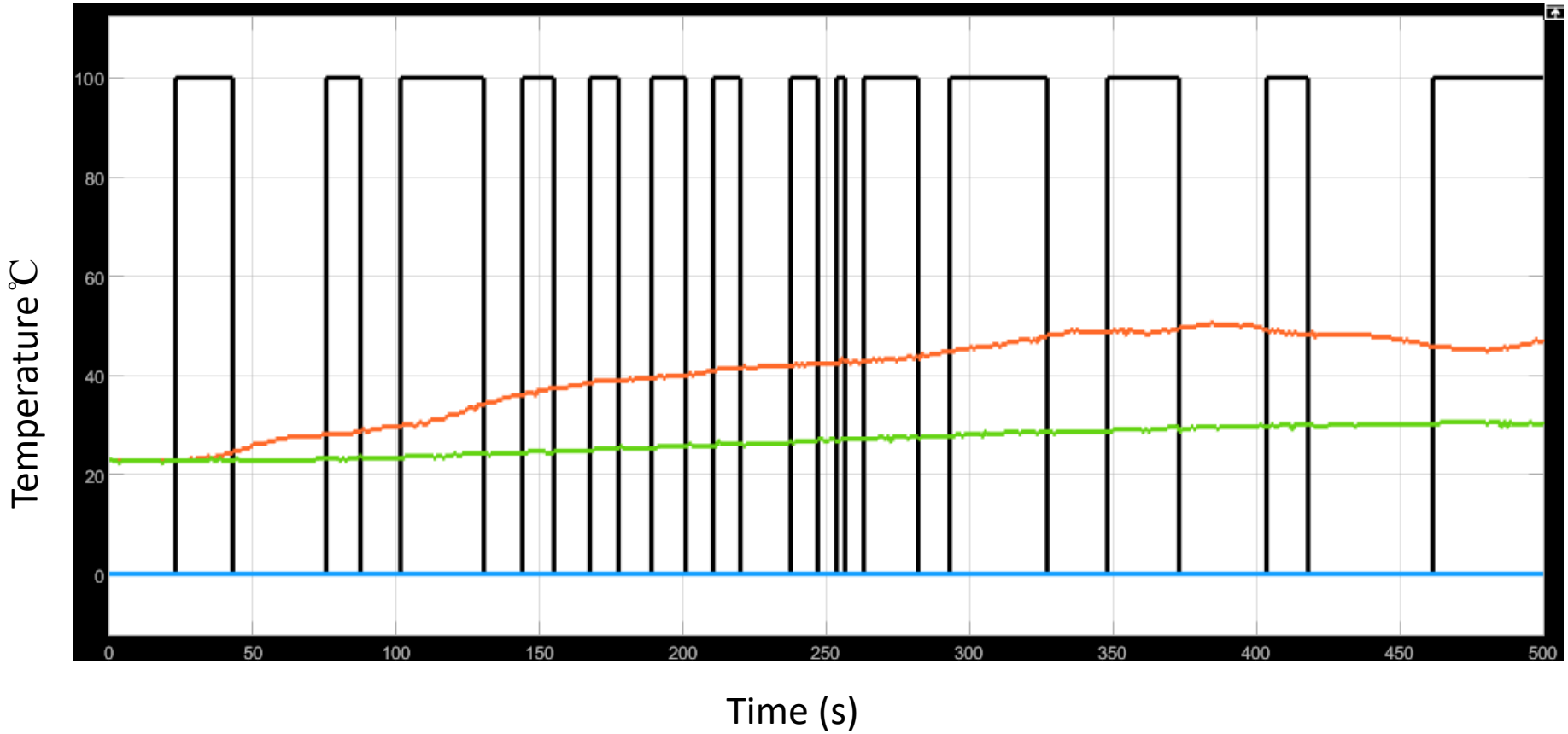
Use the Live Script: **Classification with Machine Learning and the TC Lab**

Purchase the TC Lab
https://apmonitor.com/pdc/index.php/Main/PurchaseLabKit

Generate labelled data where the [TC Lab heater](#) input is either on at 100% or 0% for multiple periods (a total of *500 s*)

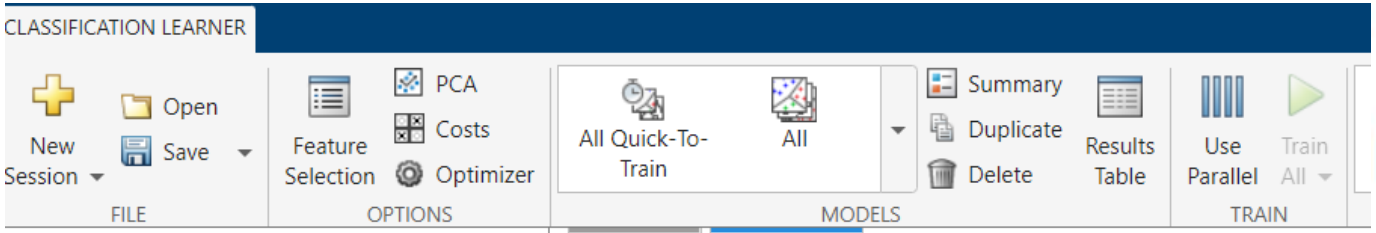


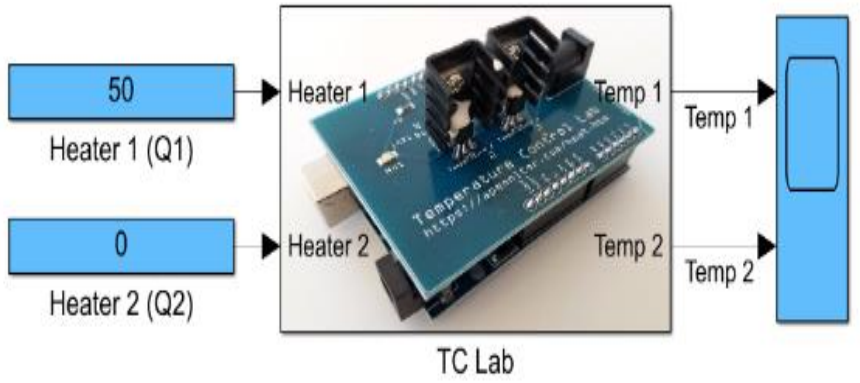
Generate labelled data where the **TC Lab heater** input is either on at 100% or 0% for multiple periods (a total of *500 s*)
.



Test

Classification Learner





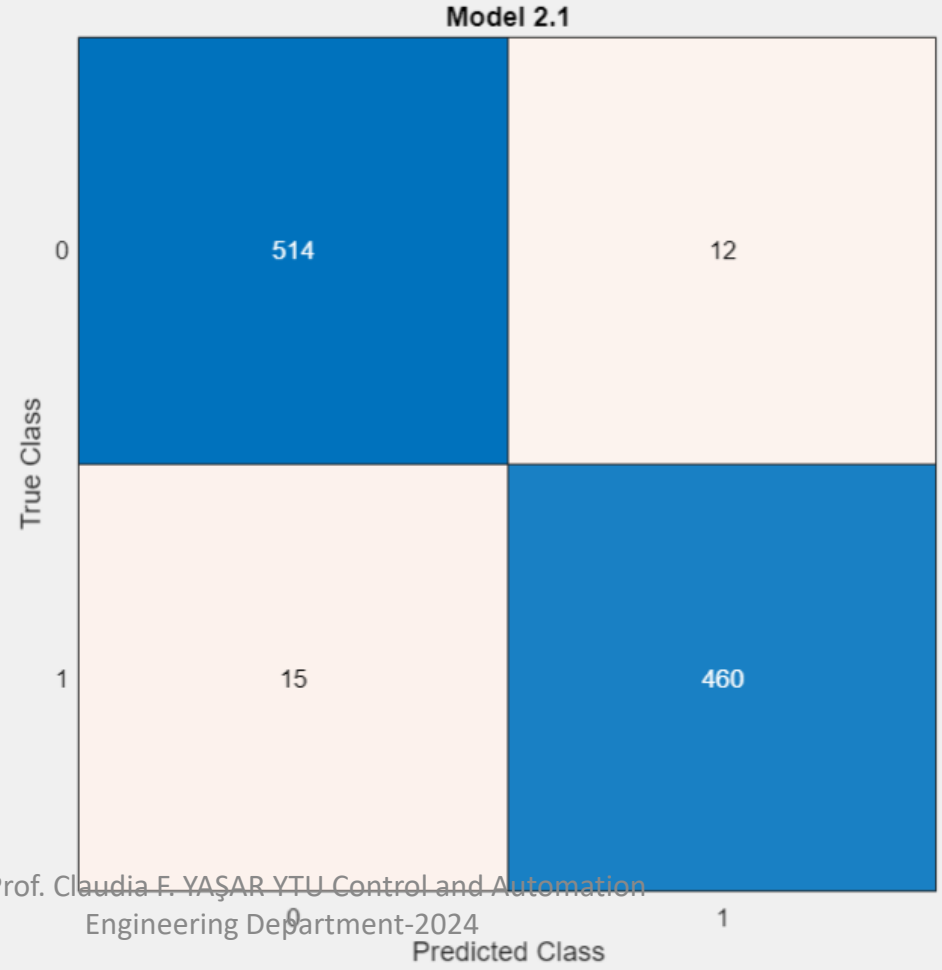
Models

Sort by Model Number

1 Tree	Accuracy (Validation): **97.3%**	
ast change: Fine Tree	6/6 features	
2.1 Tree	Accuracy (Validation): **97.3%**	
ast change: Fine Tree	6/6 features	
2.2 Tree	Accuracy (Validation): 93.8%	
ast change: Medium Tree	6/6 features	
2.3 Tree	Accuracy (Validation): 61.2%	
ast change: Coarse Tree	6/6 features	
2.4 KNN	Accuracy (Validation): 92.0%	
ast change: Fine KNN	6/6 features	
2.5 KNN	Accuracy (Validation): 79.0%	
ast change: Medium KNN	6/6 features	
2.6 KNN	Accuracy (Validation): 58.6%	
ast change: Coarse KNN	6/6 features	
2.7 KNN	Accuracy (Validation): 78.7%	
ast change: Cosine KNN	6/6 features	
2.8 KNN	Accuracy (Validation): 78.5%	
ast change: Cubic KNN	6/6 features	
2.9 KNN	Accuracy (Validation): 91.7%	
ast change: Weighted KNN	6/6 features	

Model 1 | Model 2.1

Summary | Validation Confusion Matrix



Plot

- (•) Number of observations
- ( ) True Positive Rates (TPR) / False Negative Rates (FNR)
- ( ) Positive Predictive Values (PPV) / False Discovery Rates (FDR)

What is the confusion matrix?

Train

Classification Learner

n Learner - untitled*

CLASSIFICATION LEARNER

New Session | Open | Save — FILE
Feature Selection | PCA | Costs | Optimizer — OPTIONS
All Quick-To-Train | All | Summary | Duplicate | Delete | Results Table — MODELS
Use Parallel | Train All — TRAIN
Scatter | Confusion Matrix ... — PLOT AND INTERPRET
Layout
Test Data | Test All — TEST
EXPORT

Models

Sort by Model Number

1 Tree	Accuracy (Validation): **87.6%**	Last change: Fine Tree — 6/6 features
2.1 Tree	Accuracy (Validation): **87.6%**	Last change: Fine Tree — 6/6 features
2.2 Tree	Accuracy (Validation): 78.9%	Last change: Medium Tree — 6/6 features
2.3 Tree	Accuracy (Validation): 73.6%	Last change: Coarse Tree — 6/6 features
2.4 KNN	Accuracy (Validation): 81.9%	Last change: Fine KNN — 6/6 features
2.5 KNN	Accuracy (Validation): 77.8%	Last change: Medium KNN — 6/6 features
2.6 KNN	Accuracy (Validation): 74.6%	Last change: Coarse KNN — 6/6 features
2.7 KNN	Accuracy (Validation): 77.8%	Last change: Cosine KNN — 6/6 features
2.8 KNN	Accuracy (Validation): 77.7%	Last change: Cubic KNN — 6/6 features

Model 1 | Model 2.1

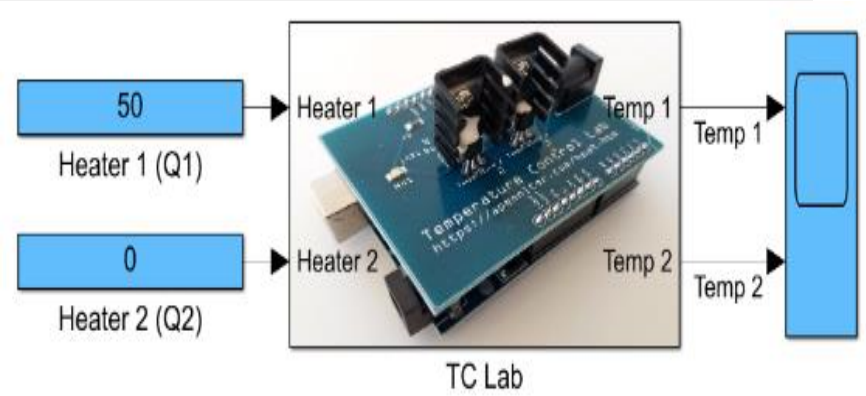
Summary | Validation Confusion Matrix

Model 2.1: Tree
Status: Trained

Training Results

Accuracy (Validation)	87.6%
Total cost (Validation)	448
Prediction speed	~67000 obs/sec
Training time	4.595 sec

- **Model Hyperparameters**
- **Feature Selection: 6/6 individual features selected**
- **PCA: Disabled**
- **Misclassification Costs: Default**
- **Optimizer: Not applicable**



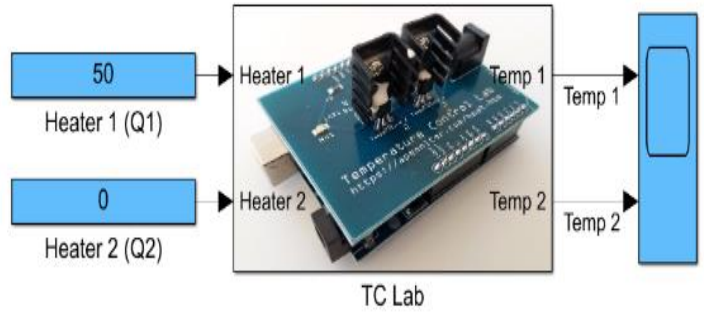
TC Lab

Classification Learner - untitled*

CLASSIFICATION LEARNER

New Session | Open | Save | Feature Selection | PCA | Costs | Optimizer | All Quick-To-Train | All | Summary | Duplicate | Delete | Results Table | Use Parallel | Train All | Scatter

FILE | OPTIONS | MODELS | TRAIN | PLOT AND



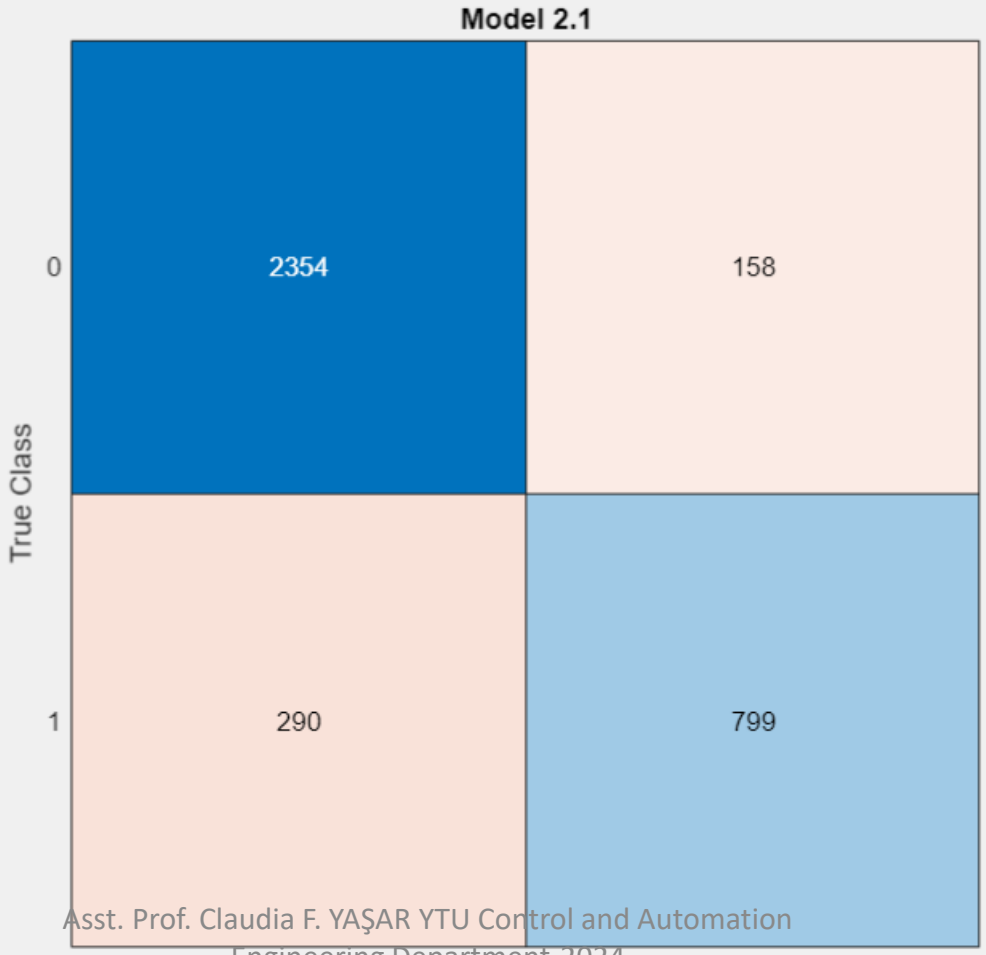
Models

Sort by Model Number

1 Tree	Accuracy (Validation): **87.6%**	
Last change: Fine Tree	6/6 features	
2.1 Tree	Accuracy (Validation): **87.6%**	
Last change: Fine Tree	6/6 features	
2.2 Tree	Accuracy (Validation): 78.9%	
Last change: Medium Tree	6/6 features	
2.3 Tree	Accuracy (Validation): 73.6%	
Last change: Coarse Tree	6/6 features	
2.4 KNN	Accuracy (Validation): 81.9%	
Last change: Fine KNN	6/6 features	
2.5 KNN	Accuracy (Validation): 77.8%	
Last change: Medium KNN	6/6 features	
2.6 KNN	Accuracy (Validation): 74.6%	
Last change: Coarse KNN	6/6 features	
2.7 KNN	Accuracy (Validation): 77.8%	
Last change: Cosine KNN	6/6 features	
2.8 KNN	Accuracy (Validation): 77.7%	
Last change: Cubic KNN	6/6 features	
2.9 KNN	Accuracy (Validation): 83.0%	
Last change: Weighted KNN	6/6 features	

Model 1 | Model 2.1

Summary | Validation Confusion Matrix



Plot

- (•) Number of observations
- ( ) True Positive Rates (TPR) / False Negative Rates (FNR)
- ( ) Positive Predictive Values (PPV) / False Discovery Rates (FDR)

What is the confusion matrix?