# OENG1116 – Modelling and Simulation of Engineering Systems

## Introduction to Modelling, Simulation and Machine Learning

***Course Lecturer:***

Dr Hamid Khayyam
Office: 251-02-34
Phone: 03 9925 4630
Email: hamid.khayyam@rmit.edu.au

# Carbon Nexus Industry (example)

**Dr Hamid Khayyam** (PhD, SMIEEE)
Research Fellow  and Team Leader  at Carbon Nexus (2013-2016)
Senior Lecturer at RMIT University (2017-now)

# Carbon Nexus Industry (questions)

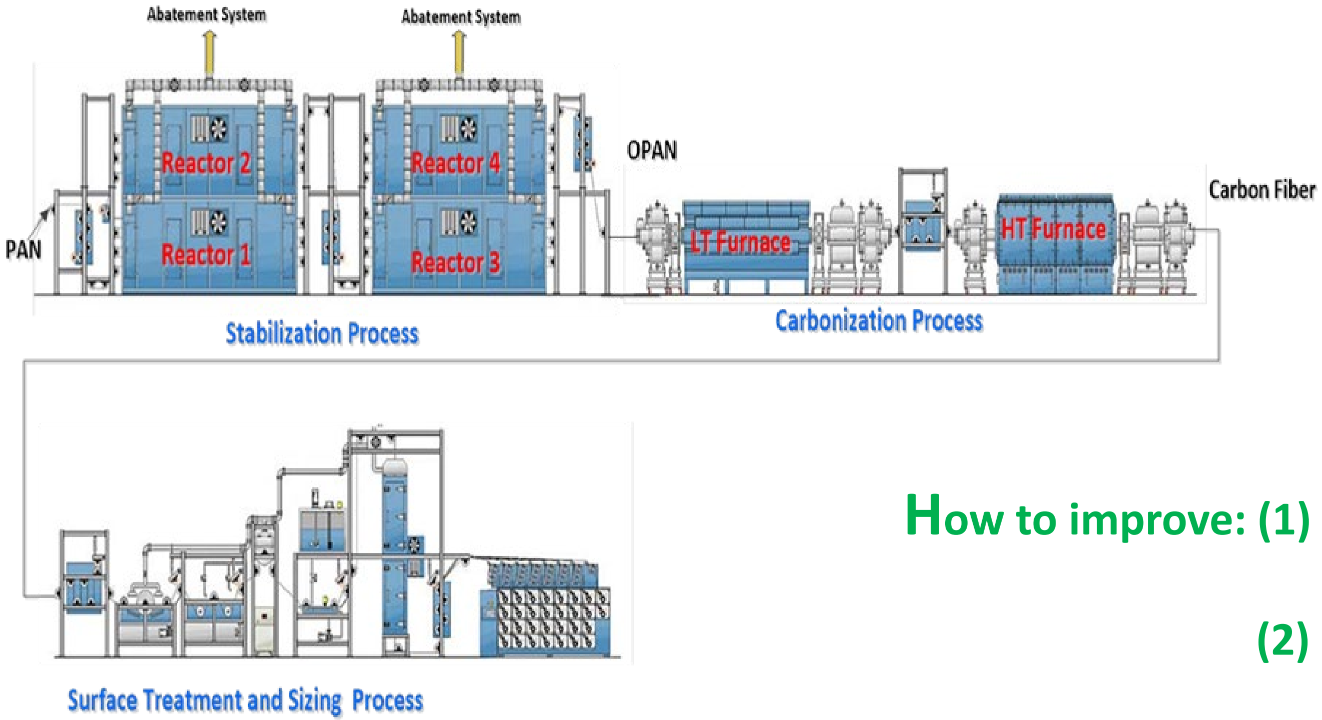



**Figure: Industrial carbon fiber manufacturing process**

How to improve: (1) Process line productivity

(2) Energy efficiency

(3) Mechanical properties

(4) Chemical properties

(5) Physical properties

……….

# Modelling and Simulation ….

– An overview of simulation modelling and analysis.

– Many critical questions are:

What is modelling?

What is simulation?

What is simulation modelling and analysis?

What types of problems are suitable for simulation?

What are the benefits and pitfalls in modelling and simulation?

# Modelling is ….

- Modelling is the process of producing a model and a model is a representation of the construction and working of some system of interest.

- A model is similar to but simpler than the system it represents. One purpose of a model is to enable the analyst to predict the effect of changes to the system.

- On the one hand, a model should be a close approximation to the real system.

- A good model is very close to realism and simplicity.

- An important issue in modelling is model validity. Model validation techniques include simulating the model under known input conditions and comparing model output with system output.

# Mathematical Models are Relationships and Variables

**Linear vs. Nonlinear:** The definition of linearity and nonlinearity is dependent on context, and linear models may have nonlinear expressions in them.

**Static** (time is not taken into account) **vs. Dynamic** (time-varying interactions among variables are taken into account).

**Explicit vs. Implicit:** If all of the input parameters of the overall model are known, and the output parameters can be calculated by a finite series of computations, the model is said to be explicit.

**Discrete vs. Continuous:** A discrete model treats objects as discrete, such as the particles in a molecular model or the states in a statistical model;

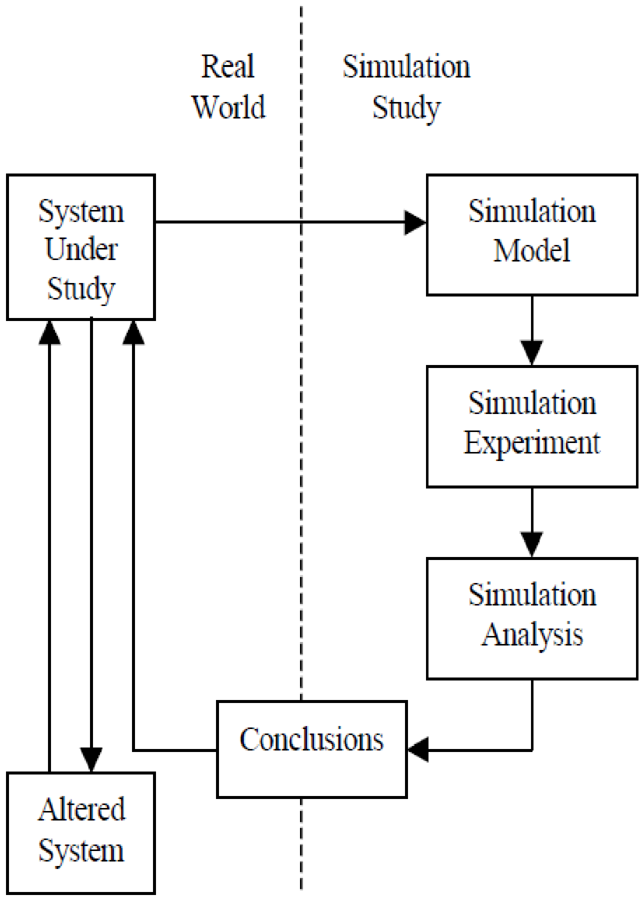
**Deterministic** (input and output variables are fixed values) **vs. Stochastic/Heuristic** (at least one of the input or output variables is probabilistic);

# Simulation is ….

- A simulation of a system is the operation of a model of the system.
- The operation of the model can be studied, and hence, properties concerning the behaviour of the actual system or its subsystem can be inferred.
- Simulation is a tool to evaluate the performance of a system, existing or proposed, under different configurations of interest and over long periods of real time.
- Simulation is used before an existing system is altered or a new system built, to reduce the chances of failure to meet specifications,

# Simulation Steps:

1. Identify the problem.
2. Formulate the problem.
3. Collect and process real system data.
4. Formulate and develop a model.
5. Validate the model.
6. Document model for future use.
7. Select appropriate experimental design.
8. Establish experimental conditions for runs.
9. Perform simulation runs.
10. Interpret and present results.
11. Recommend further course of action.



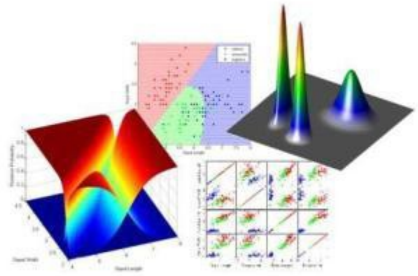

Simulation Study Schematic

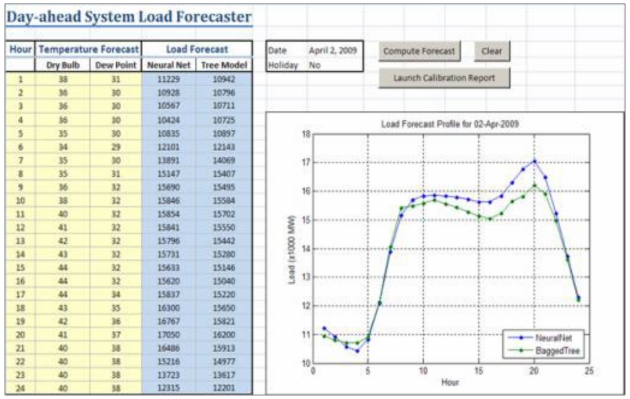
# Is Machine Learning Useful for Modelling?

- Machine Learning  tries to make predictions or decisions from data.
- Machine Learning can be used to automatically propose new models from data.
- Machine Learning is a method of data analysis that automates analytical model building.
- It is a branch of Artificial Intelligence (AI) based on the idea that systems can learn from data.
- It identify patterns and make decisions with minimal human intervention.

# Machine Learning
## When and where it is used?

- When to use it
  - Predict a future outcome based on
    - Historical data (many variables)
    - Specific patterns
  - Define a System that is
    - Based on inputs and outputs from the system
    - complex to define using governing equations
      *(e.g., black-box modeling)*
- Examples (engineering)
  - Pattern recognition *(speech, images)*
  - Control system (*robots, autonomous vehicle)*
  - Energy forecasting (*load, price*)
  - Biomedical (*tumour detection, drug discovery*)

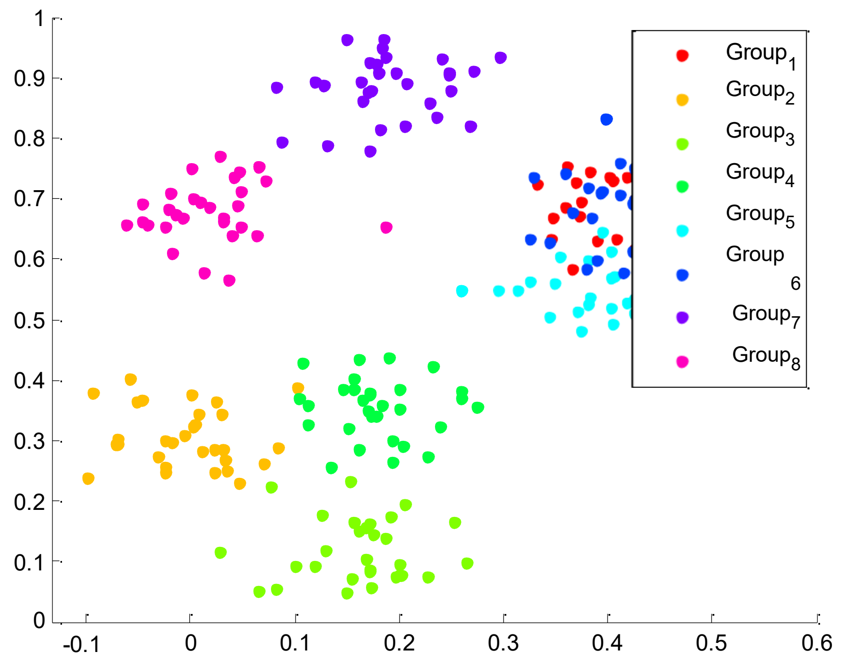





| | AAA | AA | A | BBB | BB | B | CCC | D |
|---|---|---|---|---|---|---|---|---|
| AAA | 93.68% | 5.55% | 0.59% | 0.18% | 0.00% | 0.00% | 0.00% | 0.00% |
| AA | 2.44% | 92.60% | 4.03% | 0.73% | 0.15% | 0.00% | 0.00% | 0.06% |
| A | 0.14% | 4.18% | 91.02% | 3.90% | 0.60% | 0.08% | 0.00% | 0.08% |
| BBB | 0.03% | 0.23% | 7.49% | 87.86% | 3.78% | 0.39% | 0.06% | 0.16% |
| BB | 0.03% | 0.12% | 0.73% | 8.27% | 86.74% | 3.28% | 0.18% | 0.64% |
| B | 0.00% | 0.00% | 0.11% | 0.82% | 9.64% | 85.37% | 2.41% | 1.64% |
| CCC | 0.00% | 0.00% | 0.00% | 0.37% | 1.84% | 6.24% | 81.88% | 9.67% |
| D | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100.00% |

# Basic Concepts in Machine Learning

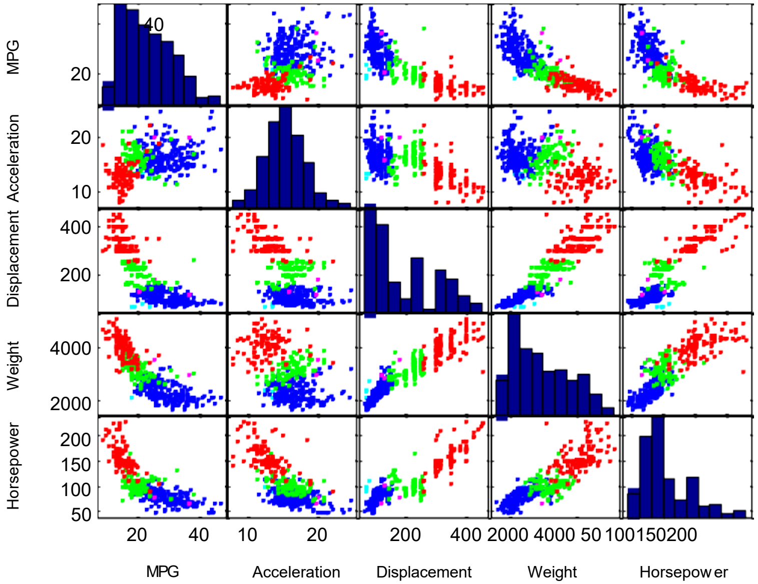
- Start with an initial set of data
- "**Learn**" from this data
  - **"Train"** your algorithm with this data
- Use the resulting **model** to **predict** outcomes for **new** data sets



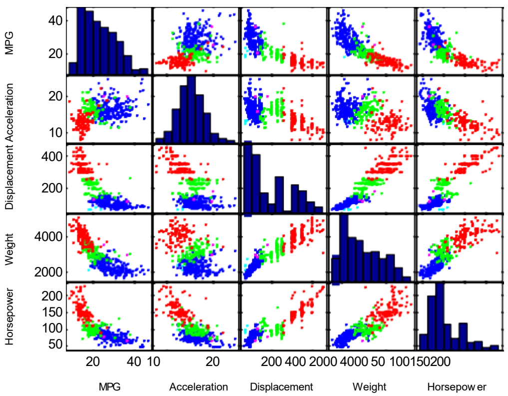
# Machine Learning Process

# Exploratory Data Analysis

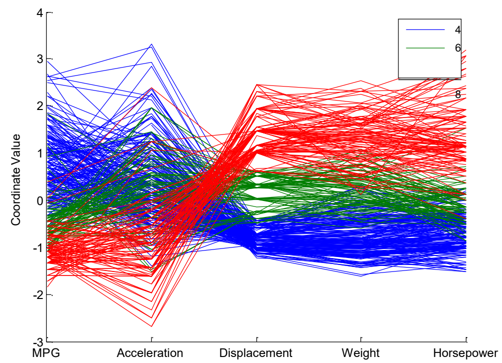
- Gain insight from visual examination
  - Identify trends and interactions
  - Detect patterns
  - Remove outliers
  - Shrink data
  - Select and pare predictors
  - Feature transformation



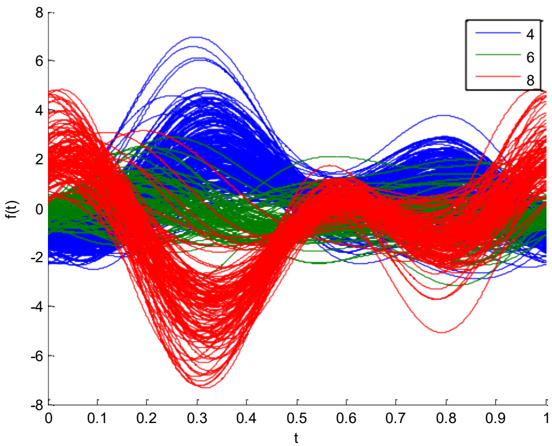
# Data Exploration
## Interactions Between Variables



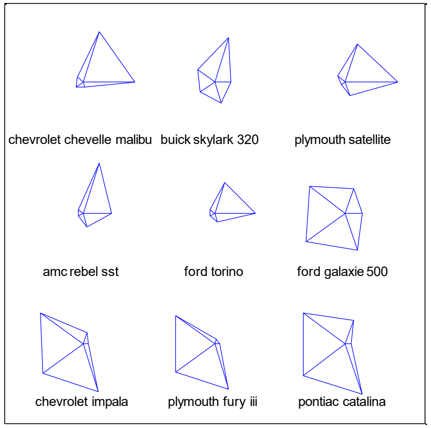

**Plot Matrix by Group**



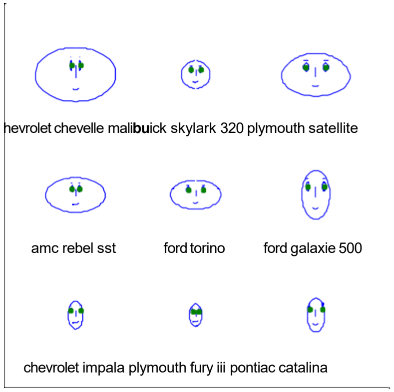

**Parallel Coordinates Plot**




**Andrews' Plot**



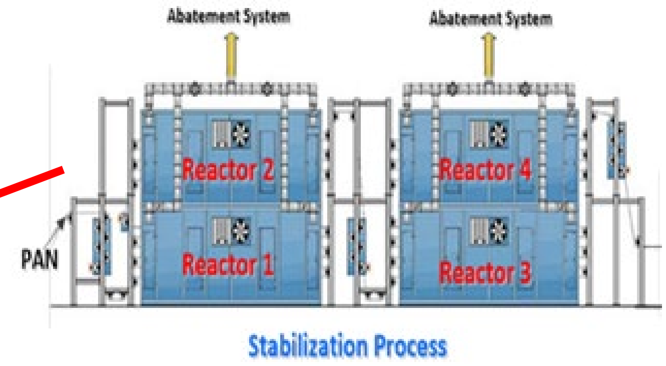

**Glyph Plot**

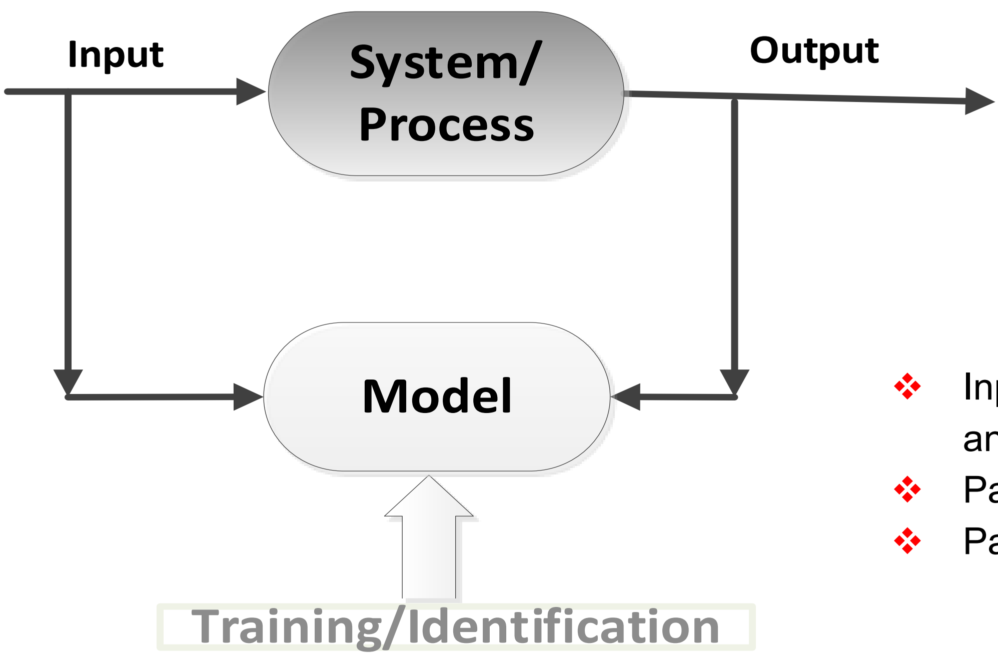



**Chernoff Faces**

# Data Modelling and Machine Learning

- **Given inputs and outputs and find actual data**

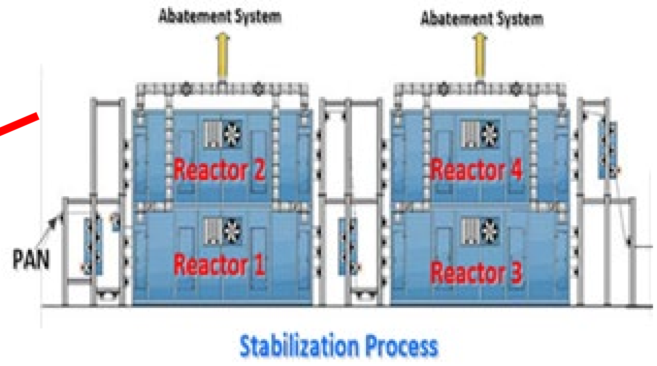


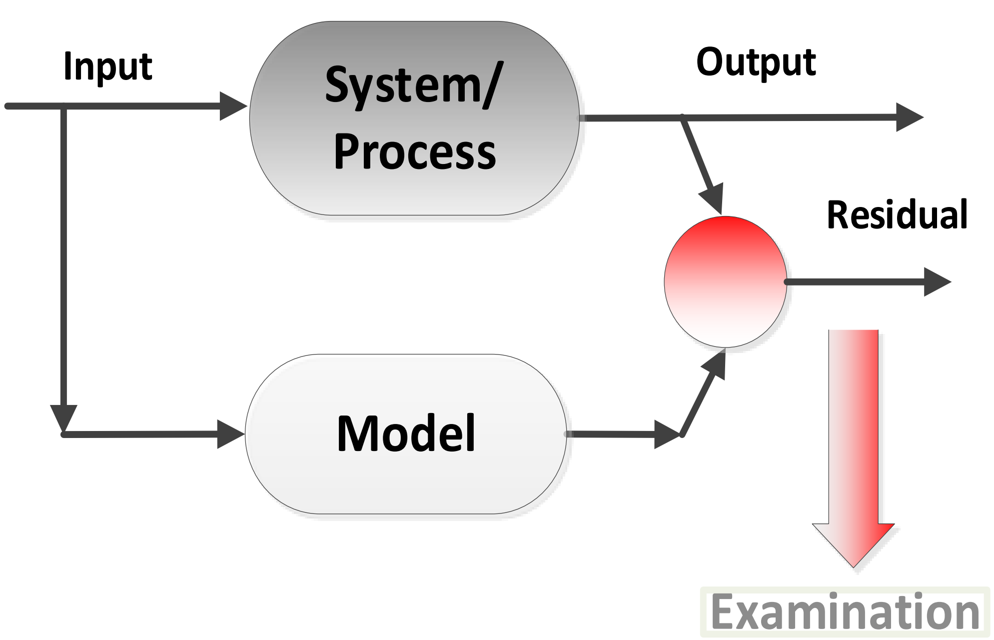




- Input/output data from instrument measurements and expert perceptions
- Parametric model structure
- Parameter estimation

# Data Modelling and Machine Learning

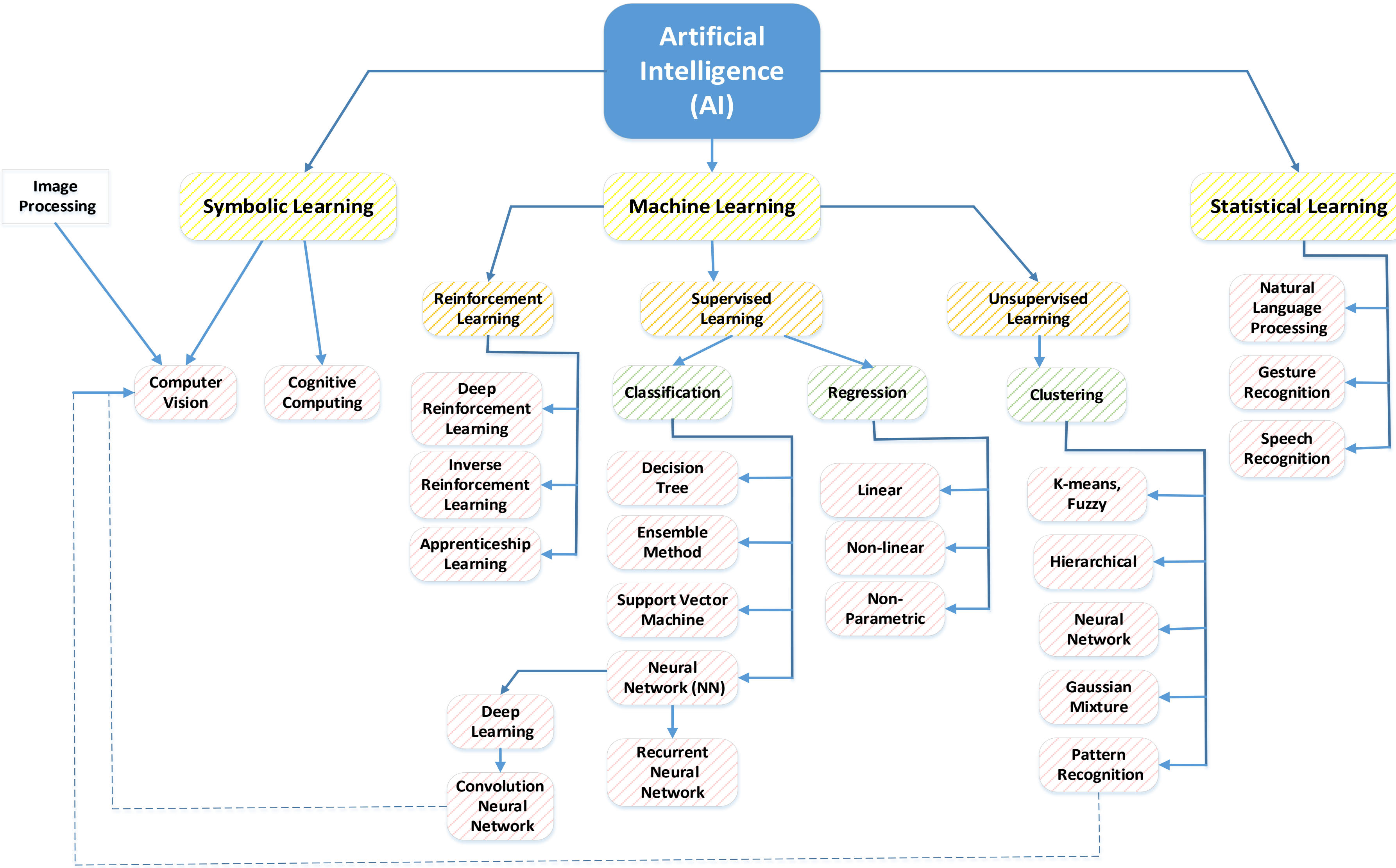
## Validation



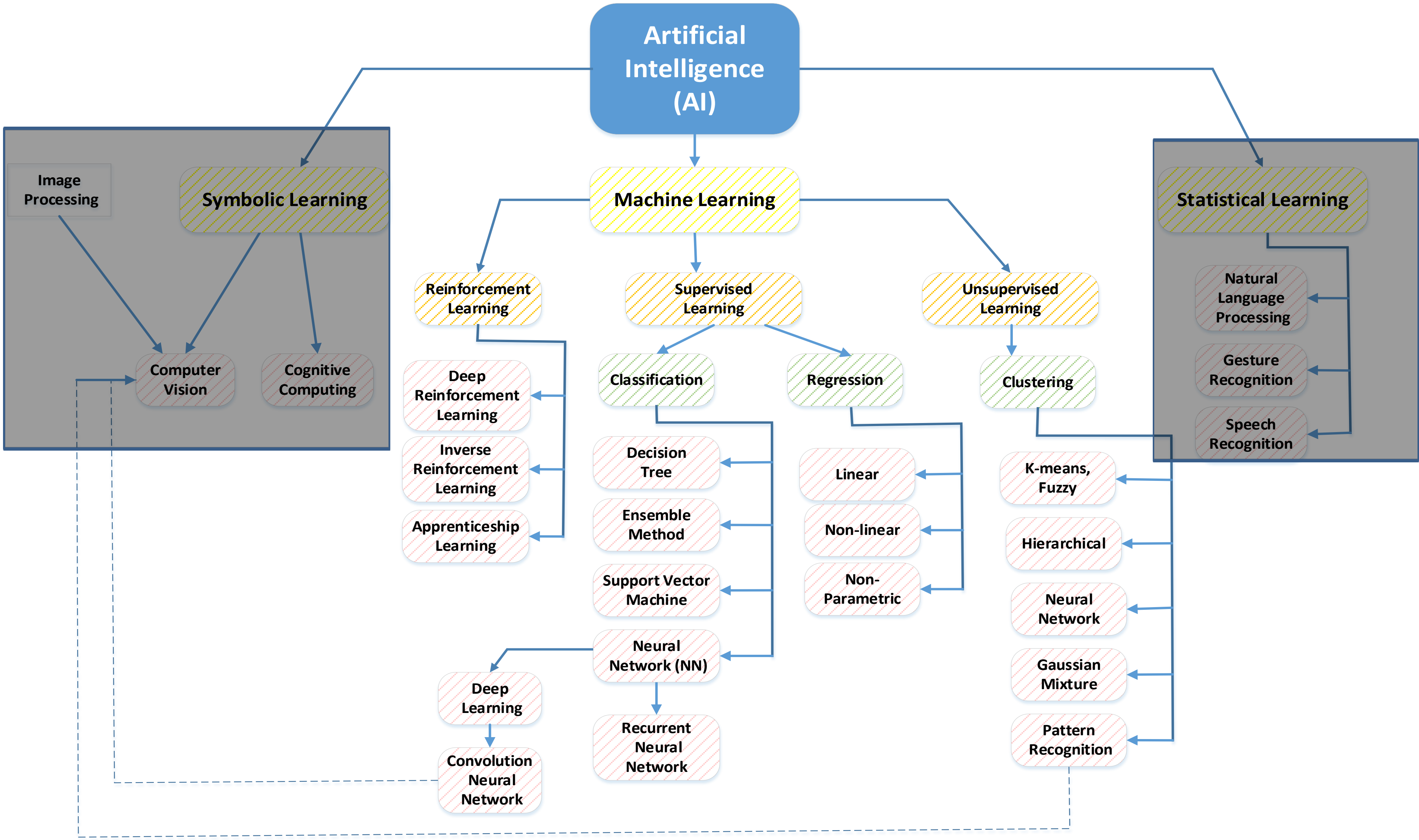

- Examine residuals
- Correlation tests
- A valid model's residuals should be reduced to uncorrelated sequence with zero mean and finite variance

# Machine Learning is a branch of Artificial Intelligence



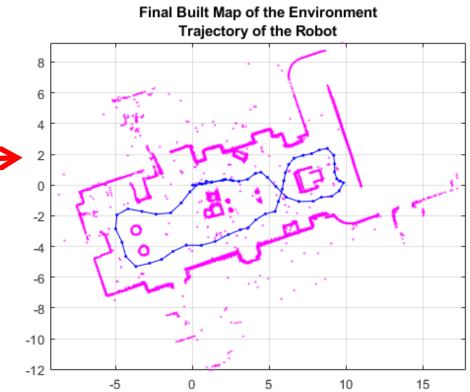

H. Khayyam, et al., Artificial Intelligence and Internet of Things for Autonomous Vehicles, Nonlinear Approaches .., 2019, Springer

# Machine Learning is a branch of Artificial Intelligence



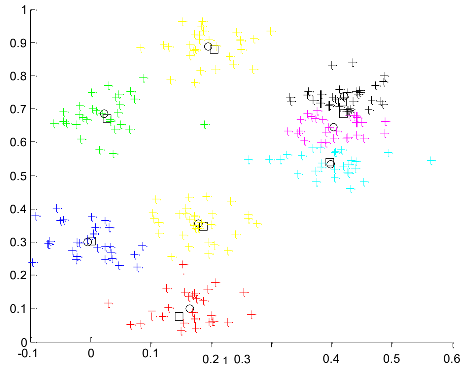

{2} H. Khayyam, et al., Artificial Intelligence and Internet of Things for Autonomous Vehicles, Nonlinear Approaches .., 2019, Springer

# Machine Learning Overview



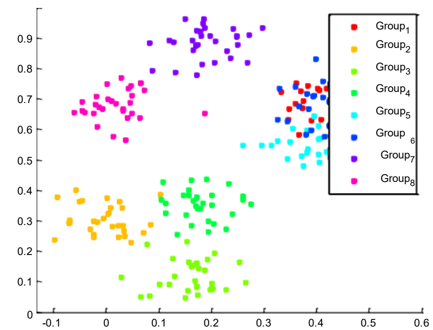

## Reinforcement Learning

- Rewards (react to an environment)



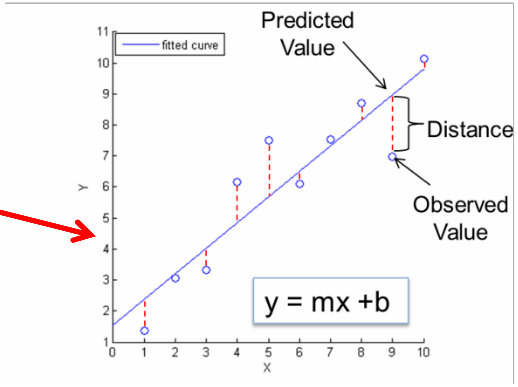
## Unsupervised Learning

- Clustering (no labels)



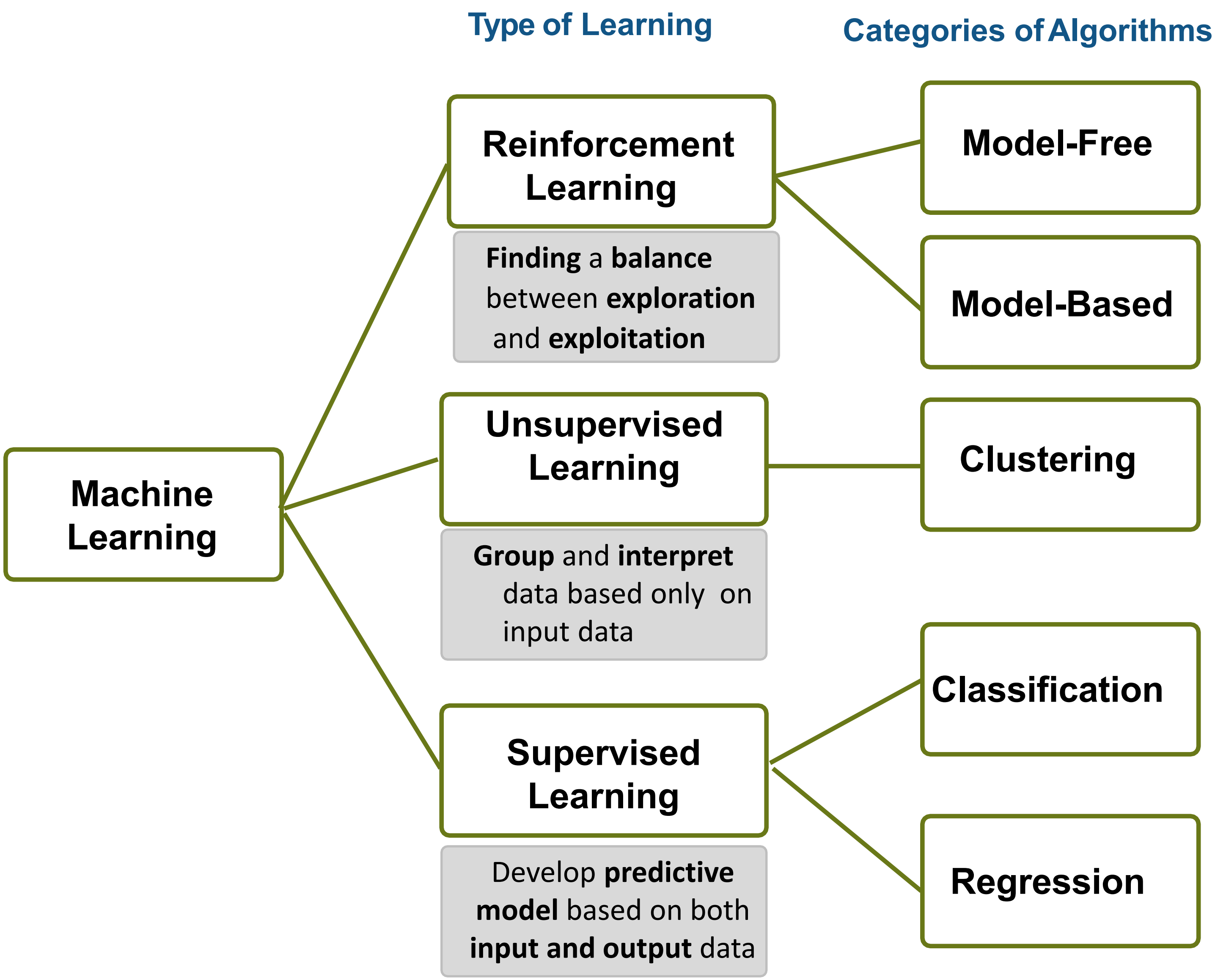

## Supervised Learning

- Classification (predefined labels)
- Regression (numerical)

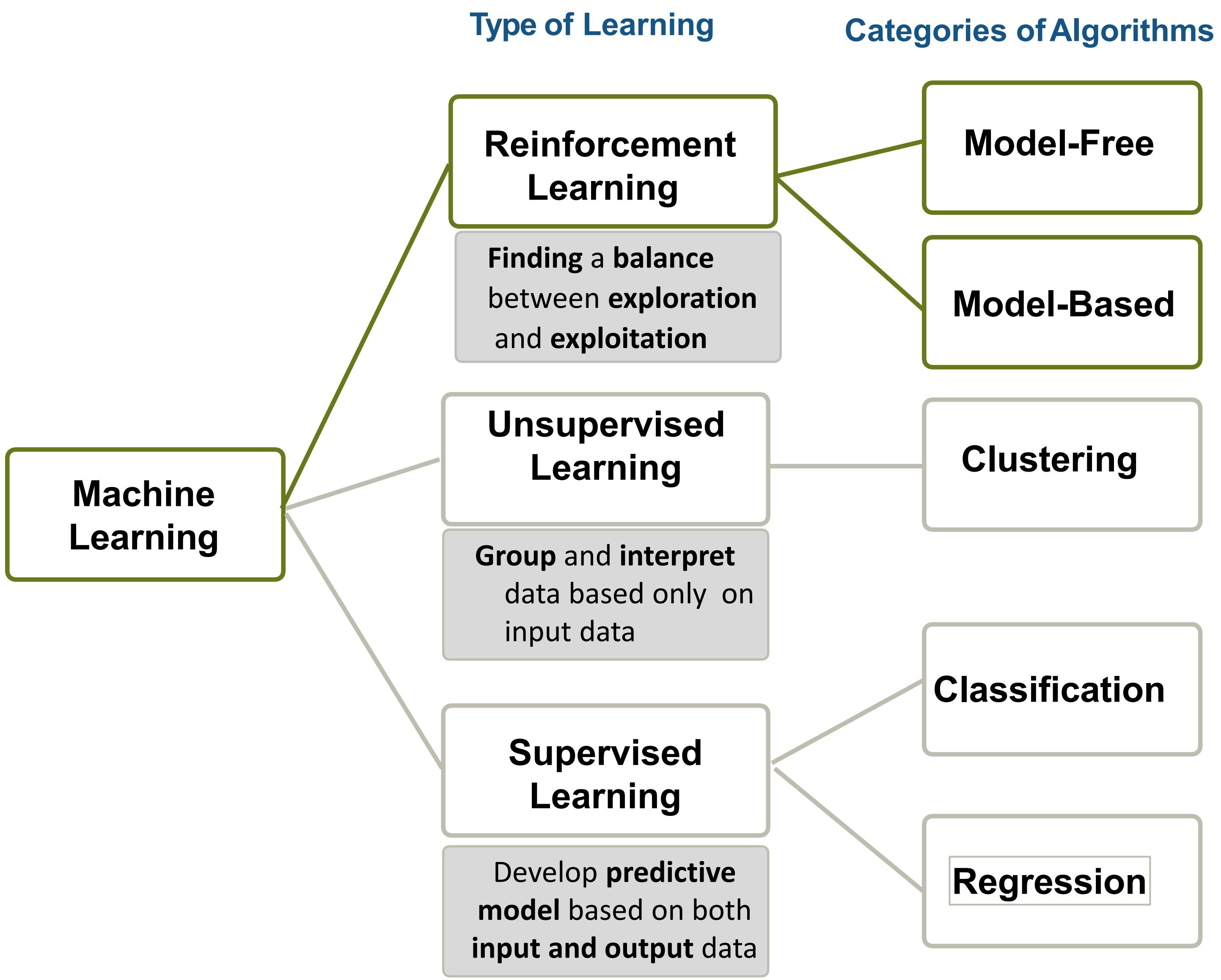



**[3] Build Machine Learning Models with a MATLAB Trial**

# Machine Learning Overview



# Machine Learning Overview



# Reinforcement Learning Algorithms


**Figure: a**

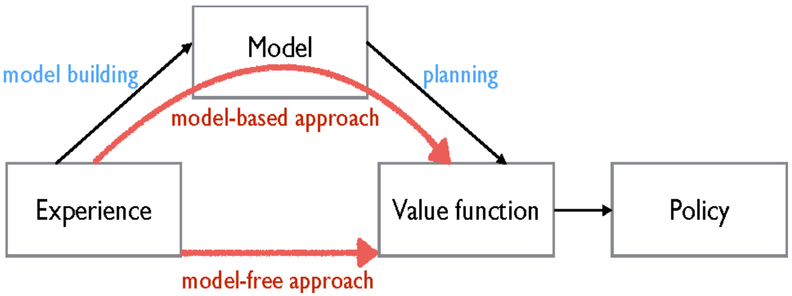



**Figure: b**

- **Model-Free: directly estimates the value function**
- **Model-Based: builds a model of the environment and then computes a value function from the model**

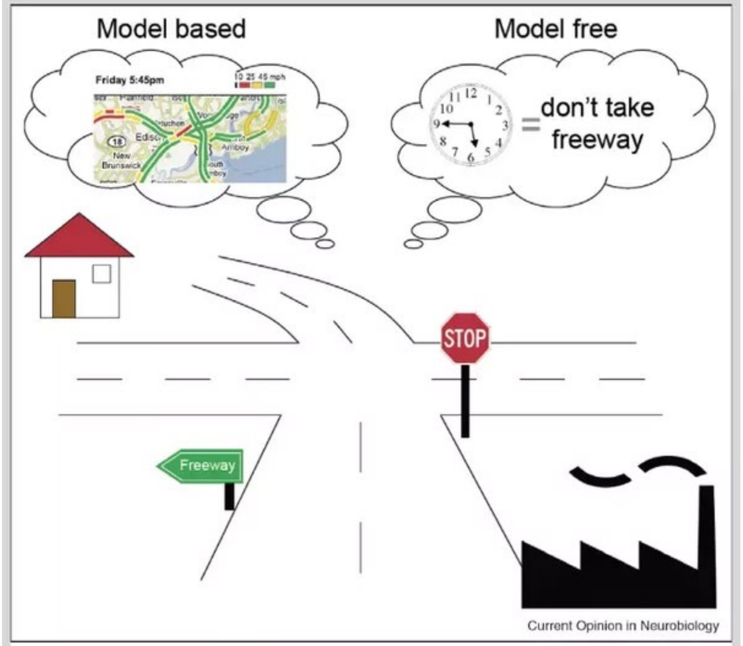
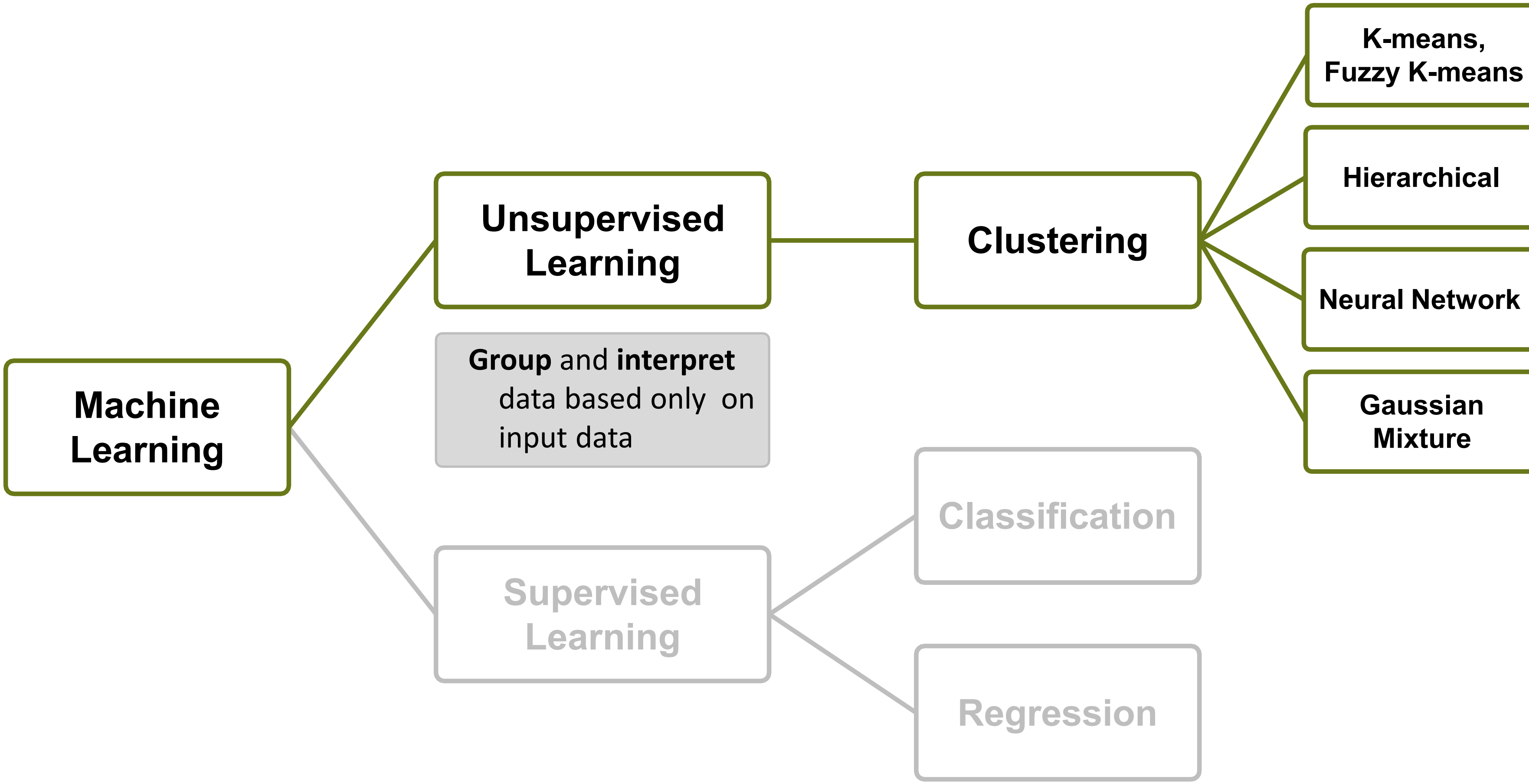



Figure 1: Two ways to choose which route to take when traveling home from work on friday evening.

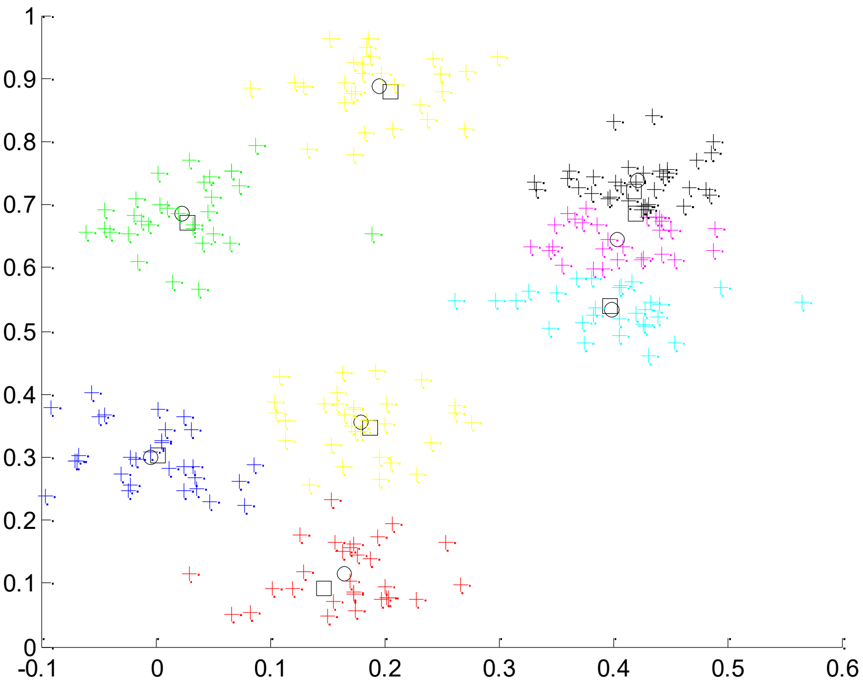
**Figure: c**

# Unsupervised Learning
## Clustering



# Clustering
## Overview

- What is clustering?
  - Segment data into groups, based on data similarity



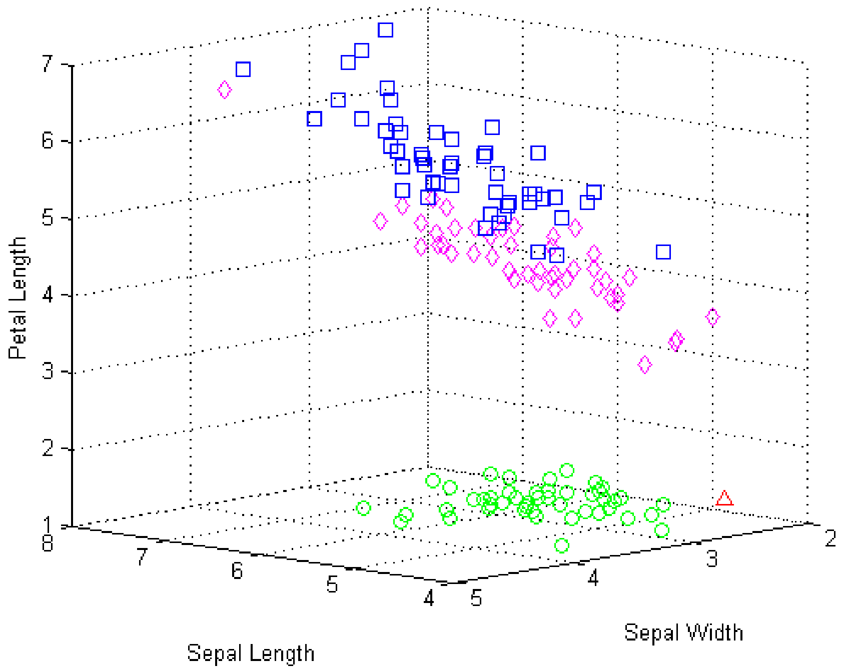

- Why use clustering?
  - Identify outliers
  - Resulting groups may be the matter of interest

- How is clustering done?
  - Can be achieved by various algorithms
  - It is an iterative process *(involving trial and error)*

# Clustering: K – Means Clustering

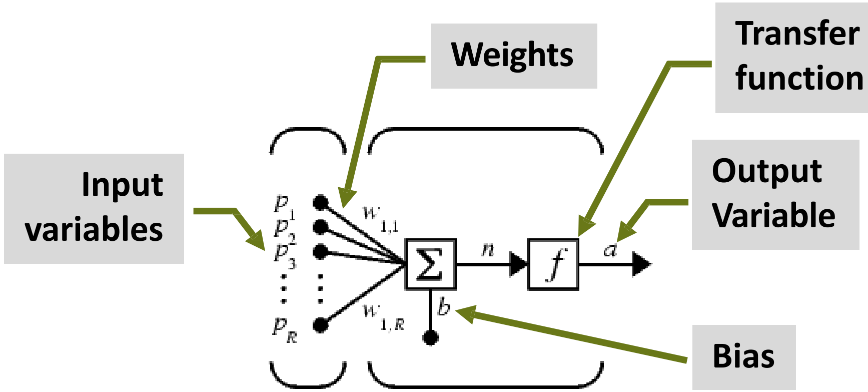
- K-means is a partitioning method
- Partitions data into ***k*** exclusive clusters
- Each cluster has a centroid (or center)

  -Sum of distances from all objects to the center is minimized

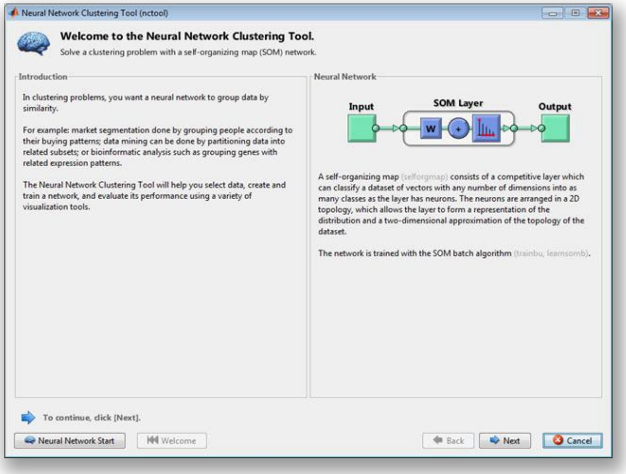


# Clustering: Neural Networks

- Networks are comprised of one or more layers
- Outputs computed by applying a nonlinear transfer function with weighted sum of inputs



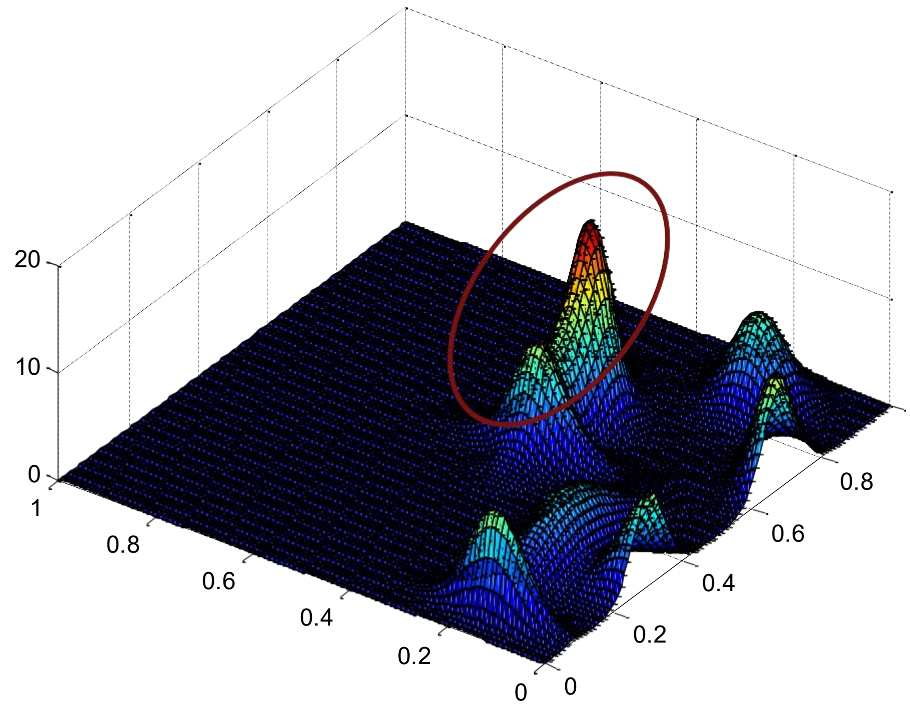

- Trained by letting the network continually adjust itself to new inputs *(determines weights)*
- Interactive apps for easily creating and training networks
- Multi-layered networks created by cascading (*provide better accuracy)*
- Example architectures for clustering:
  - Self-organizing maps
  - Competitive layers



# Clustering : Gaussian Mixture Models

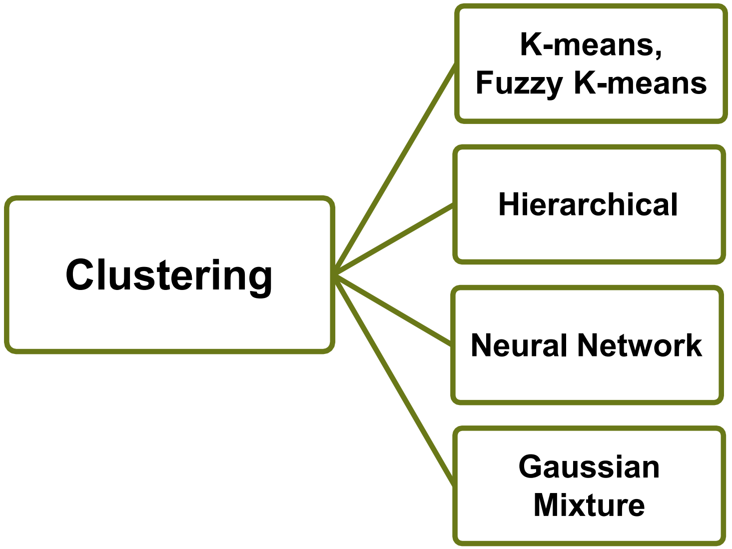
- Good when clusters have different sizes and are correlated

- Assume that data is drawn from a fixed number ***K*** of normal distributions

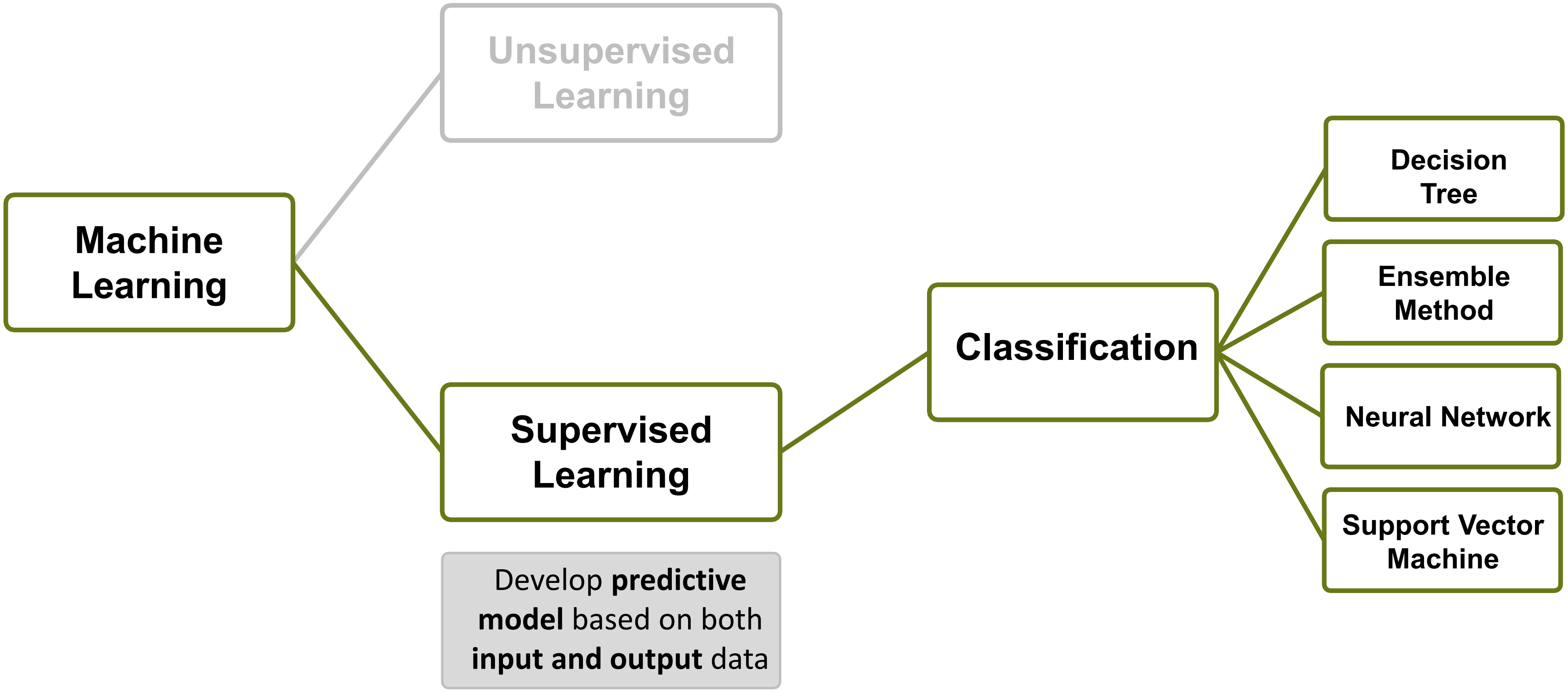


# Cluster Analysis
# Summary

- Segments data into groups, based on data similarity
- No method is perfect
  *(depends on data)*
- Process is iterative;
  explore different algorithms
- Beware of local minima
  *(global optimization can help)*

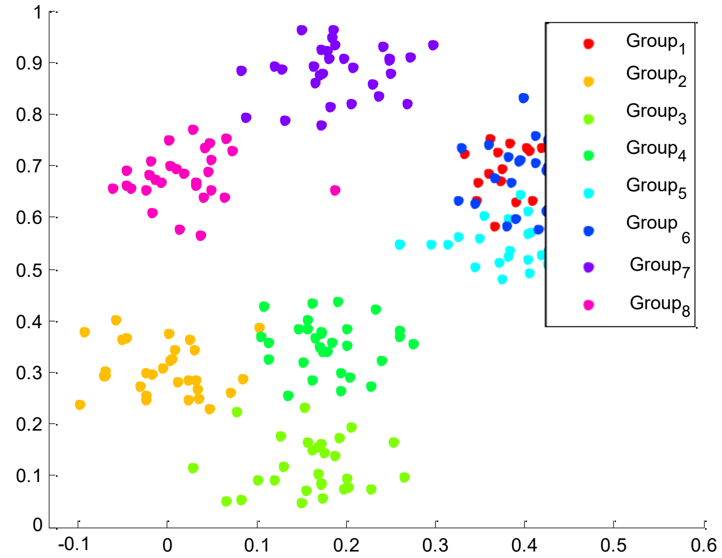


# Supervised Learning
## Classification for Predictive Modelling



# Classification
## Overview

- What is classification?
  - Predicting the best group for each point
  - “Learns” from labeled observations
  - Uses input features

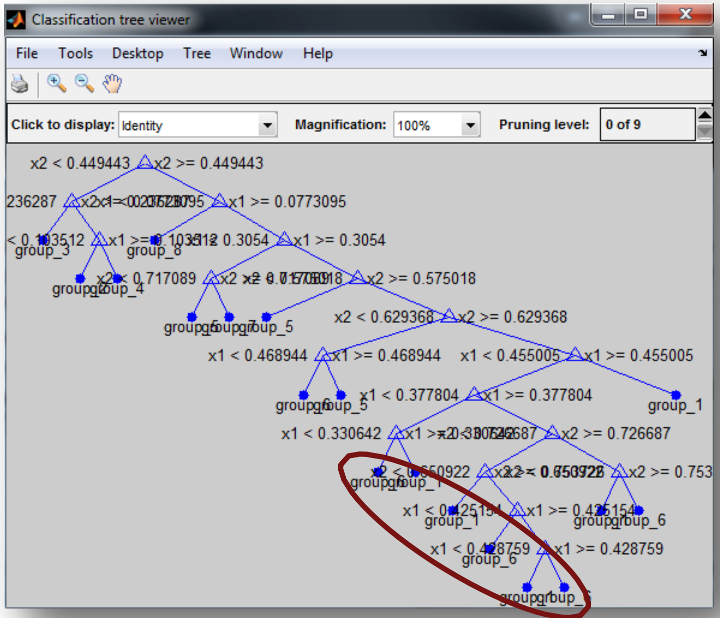



- Why use classification?
  - Accurately group data never seen before

- How is classification done?
  - Can use several algorithms to build a predictive model
  - Good training data is critical

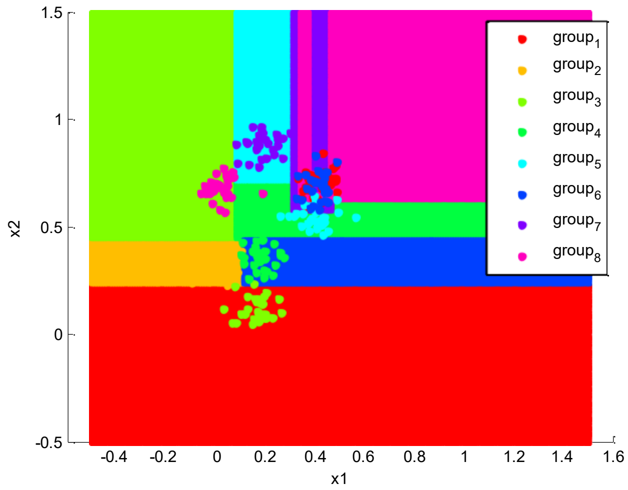
# Classification - Decision Trees

- Builds a tree from training data
  - Model is a tree where each node is a logical test on a predictor
- Traverse tree by comparing features with threshold values
- The “leaf” of the tree  specifies the group



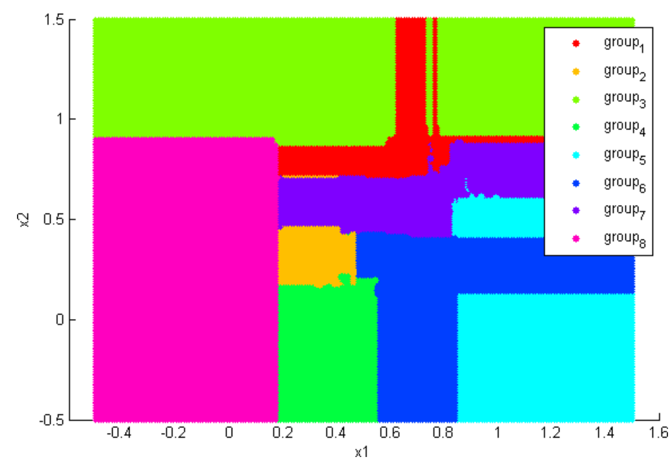
# Classification - Ensemble Learners
## Overview

- Decision trees are "weak" learners
  - Good to classify data used to train
  - Often not very good with new data
  - Note rectangular groups



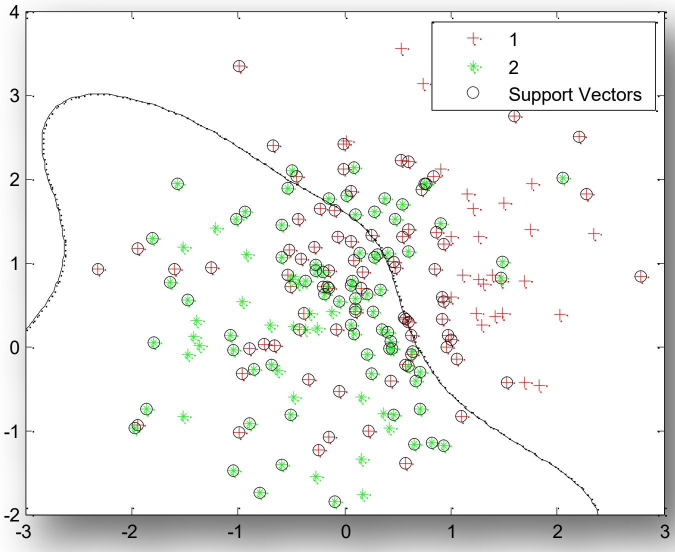

- What are ensemble learners?
  - Combine many decision trees to create a "strong" learner
  - Uses "bootstrapped aggregation"

- Why use ensemble methods?
  - Classifier has better predictive power
  - Note improvement in cluster shapes



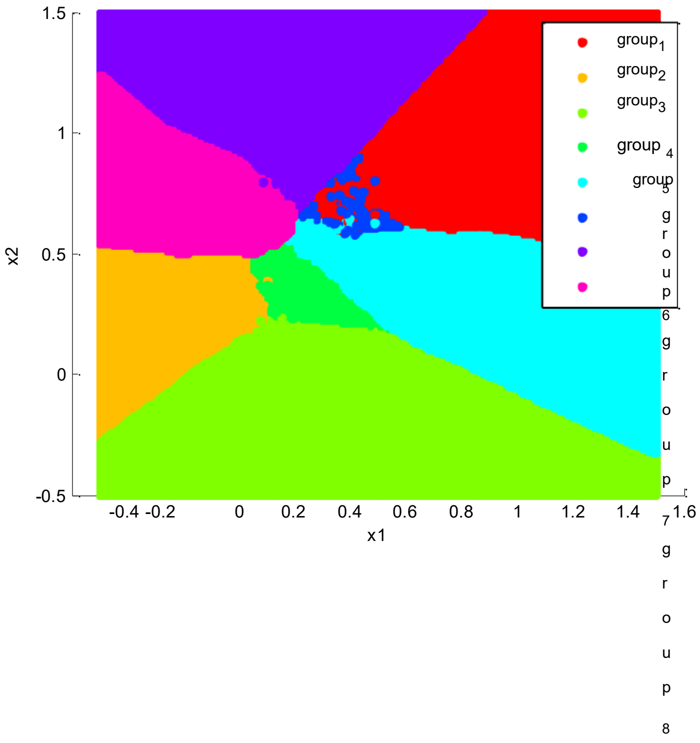
# Classification - Support Vector Machines
## Overview

- Good for modeling with complex boundaries between groups
  - Can be very accurate
  - No restrictions on the predictors
- What does it do?
  - Uses non-linear “kernel” to calculate the boundaries
  - Can be computationally intensive
- Version in Statistics Toolbox only classifies into two groups

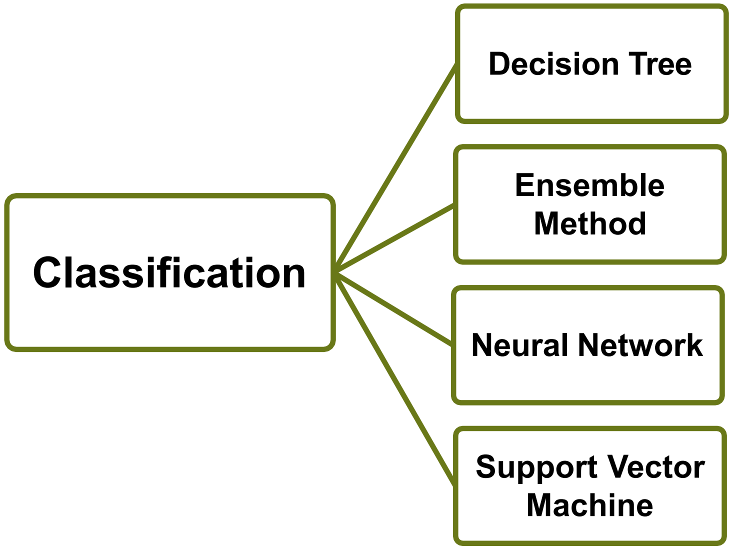


# K-Nearest Neighbor Classification

- One of the simplest classifiers
- Takes the ***K*** nearest points from the training set, and chooses the majority class of those ***K*** points
- No training phase – all the work is done during the application of the model

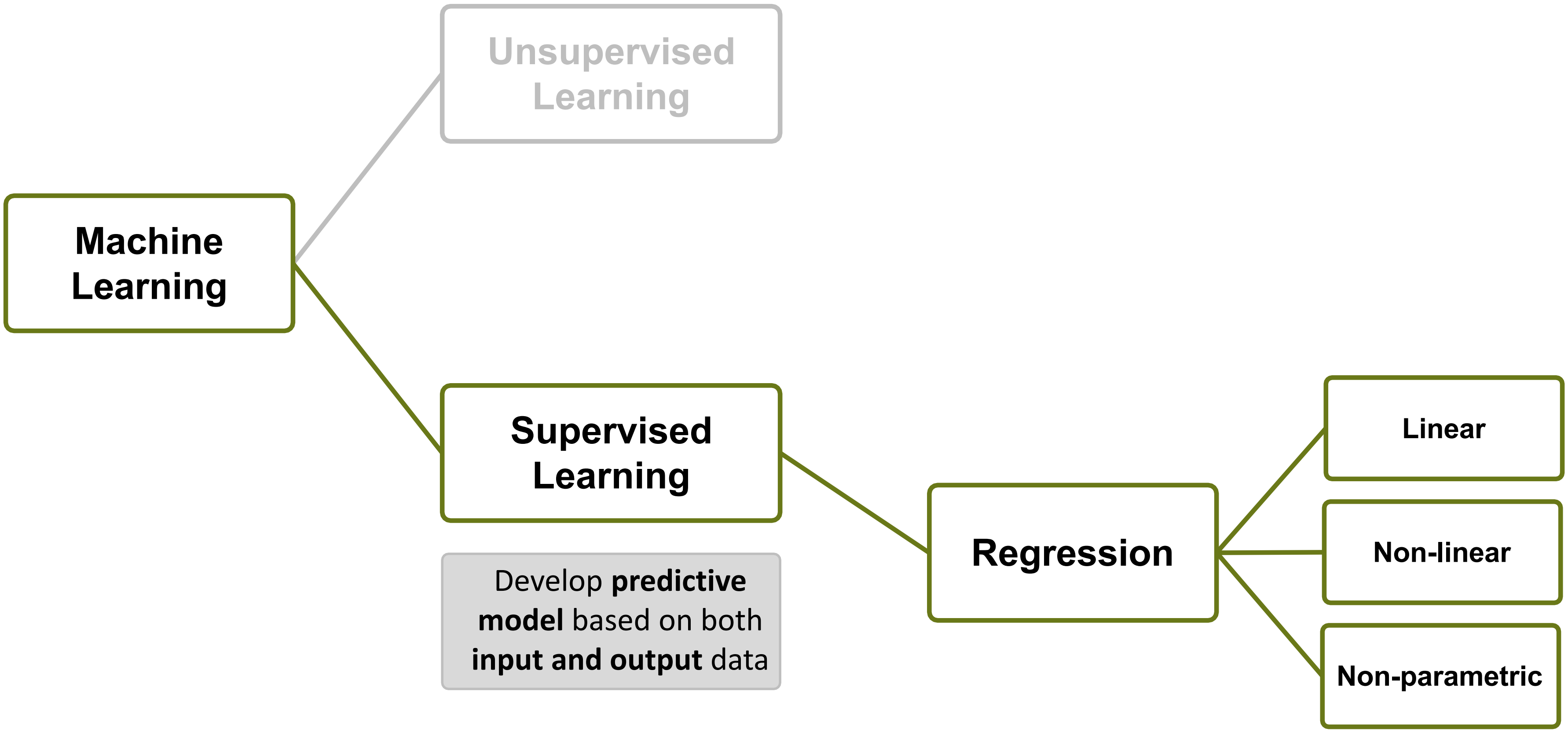


# Classification
## Summary



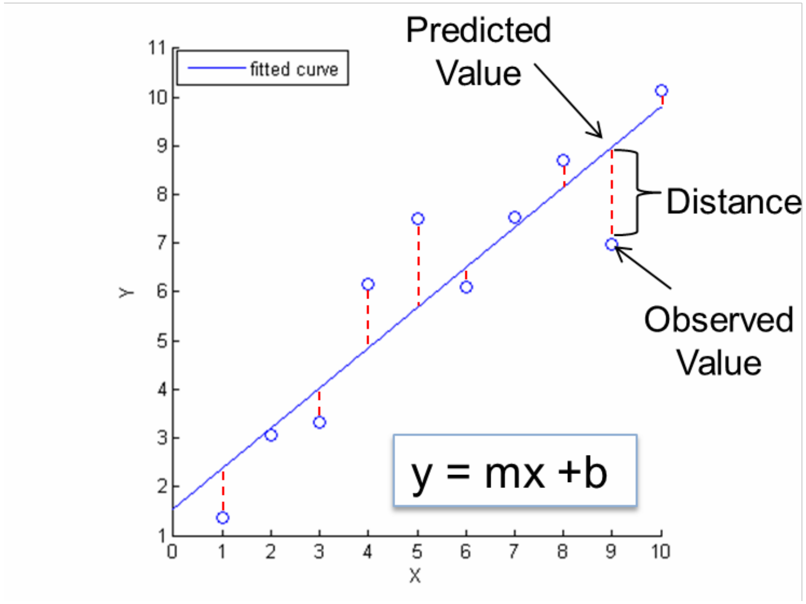

- No absolute best method
- Simple does not mean inefficient
- Watch for overfitting
  - Decision trees and neural networks may overfit the noise
  - Use ensemble learning and cross-validation
- Parallelize for speedup

# Supervised Learning
## Regression for Predictive Modelling



# Regression

- Why use regression?
  - Predict the continuous response for new observations




- Type of predictive modeling
  - Specify a model that describes *y* as a function of *x*; *y=f(x)*
  - Estimate coefficients that minimize the difference between predicted and actual

- You can apply techniques from earlier sections with regression as well

# Linear Regression

- *Y* is a *linear* function of the regression coefficients
- Common examples:

| | |
|---|---|
| **Straight line** | $Y = B_0 + B_1X_1$ |
| **Plane** | $Y = B_0 + B_1X_1 + B_2X_2$ |
| **Polynomial** | $Y = B_0 + B_1X_1 + B_2X_1^2 + B_3X_1^3$ |
| **Polynomial with cross terms** | $Y = B_0 + B_1X_1^2 + B_2(X_1 * X_2) + B_3X_2^2$ |

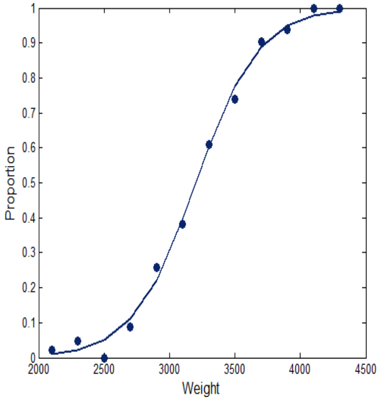
# Nonlinear Regression

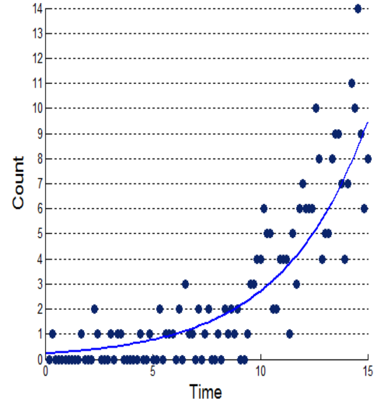
- *Y* is a *nonlinear* function of the regression coefficients
- Syntax for formulas:

| | |
|---|---|
| **Fourier Series** | `y ~ b0 + b1*cos(x*b3) + b4*sin(x*b3)` |
| **Exponential Growth** | `@(b,t)(b(1)*exp(b(2)*t)` |
| **Logistic Growth** | `@(b,t)(1/(b(1) + exp(-b(2)*x)))` |

# Generalized Linear Models

- Extends the linear model
  - Define relationship between model and response variable
  - Model error distributions other than normal
- Logistic regression
  - Response variable is binary (true / false)
  - Results are typically expressed as an odd's ratio
- Poisson regression
  - Model count data (non-negative integers)
  - Response variable comes from a Poisson distribution



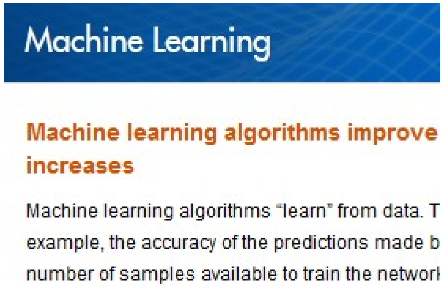



# Machine Learning with MATLAB

- Interactive environment
  - Visual tools for exploratory data analysis
  - Easy to evaluate and choose best algorithm
  - Apps available to help you get started
    *(e.g,. neural network tool, curve fitting tool)*

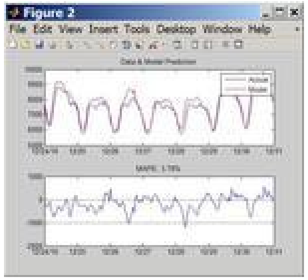
- Multiple algorithms to choose from
  - Clustering
  - Classification
  - Regression

# Learn More : Machine Learning with MATLAB

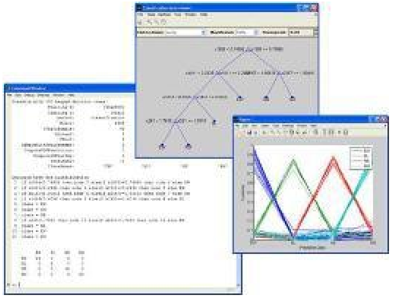



http://www.mathworks.com/discovery/machine-learning.html

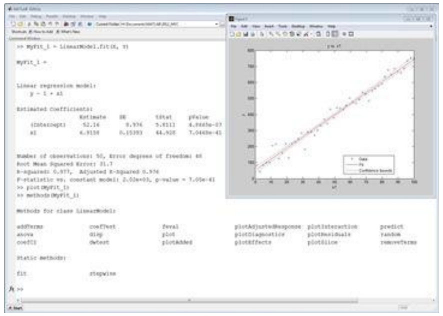
Data Driven Fitting with MATLAB



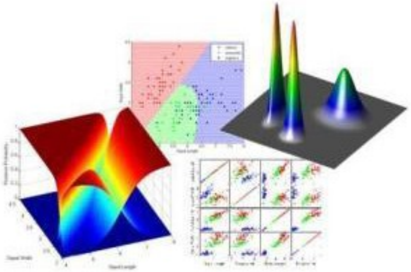

Classification with MATLAB



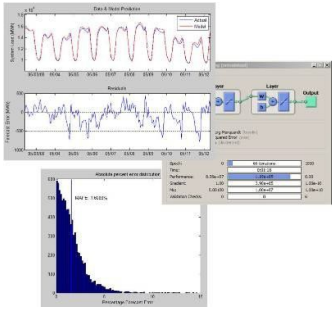
Regression with MATLAB



Multivariate Classification in the Life Sciences



Electricity Load and Price Forecasting



Credit Risk Modeling with MATLAB

# Machine Learning with MATLAB

- Model Evaluation/Performance

Lower values are better. Zero means no error.

$$MSE = \frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n} \quad (1)$$

*where Xobs is observed values and Xmodel is modelled values.*

– Root Mean Square Error (RMSE) is the average root squared difference between outputs and targets.
Lower values are better.
Zero means no error.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}} \quad (2)$$

*where Xobs is observed values and Xmodel is modelled values.*

– Regression (R) values measure the correlation between outputs and targets.
An R value of 1 means a close relationship, 0 a random relationship

$$R = \frac{n(\sum x_o x_m) - (\sum x_o)(\sum x_m)}{\sqrt{[n\sum x_o^2 - (\sum x_o)^2][n\sum x_m^2 - (\sum x_m)^2]}} \quad (3)$$

*where Xo is observed values and Xm is modelled values.*

# References

[1] H. Khayyam, G. Golkarnarenji, R. N Jazar "*Limited data modelling approaches for engineering applications*" Book Nonlinear approaches in Engineering Applications, Pages 345-379 Publisher Springer (2018)

[2] H. Khayyam, et al., "*Artificial Intelligence and Internet of Things for Autonomous Vehicles*", Book Nonlinear Approaches in Engineering Applications, Publisher Springer (2019).

[4] B. Crawford, H. Khayyam, A.S. Milani, R.N. Jazar, "Big Data Modeling Approaches for Engineering Applications" Nonlinear Approaches in Engineering Applications, 307-365 (2019)

[5] Build Machine Learning Models with a MATLAB Trial (2016)

[6] Mastering Machine Learning, A Step-by-Step Guide with MATLAB (2019)

[7] K. Asadi " Strengths, Weaknesses, and Combinations of Model-based and Model-free Reinforcement Learning (2015)

# References

## Download

*MATLAB Code for the Heart Sounds Classification  Application*

## Watch

*Machine Learning Using Heart Sound Classification Example*  (22:03)

## Read

*Big Data with MATLAB*

*What Is Deep Learning?*

*Introducing Deep Learning with MATLAB* *Parallel Computing on the Cloud with MATLAB* *Predictive Analytics*