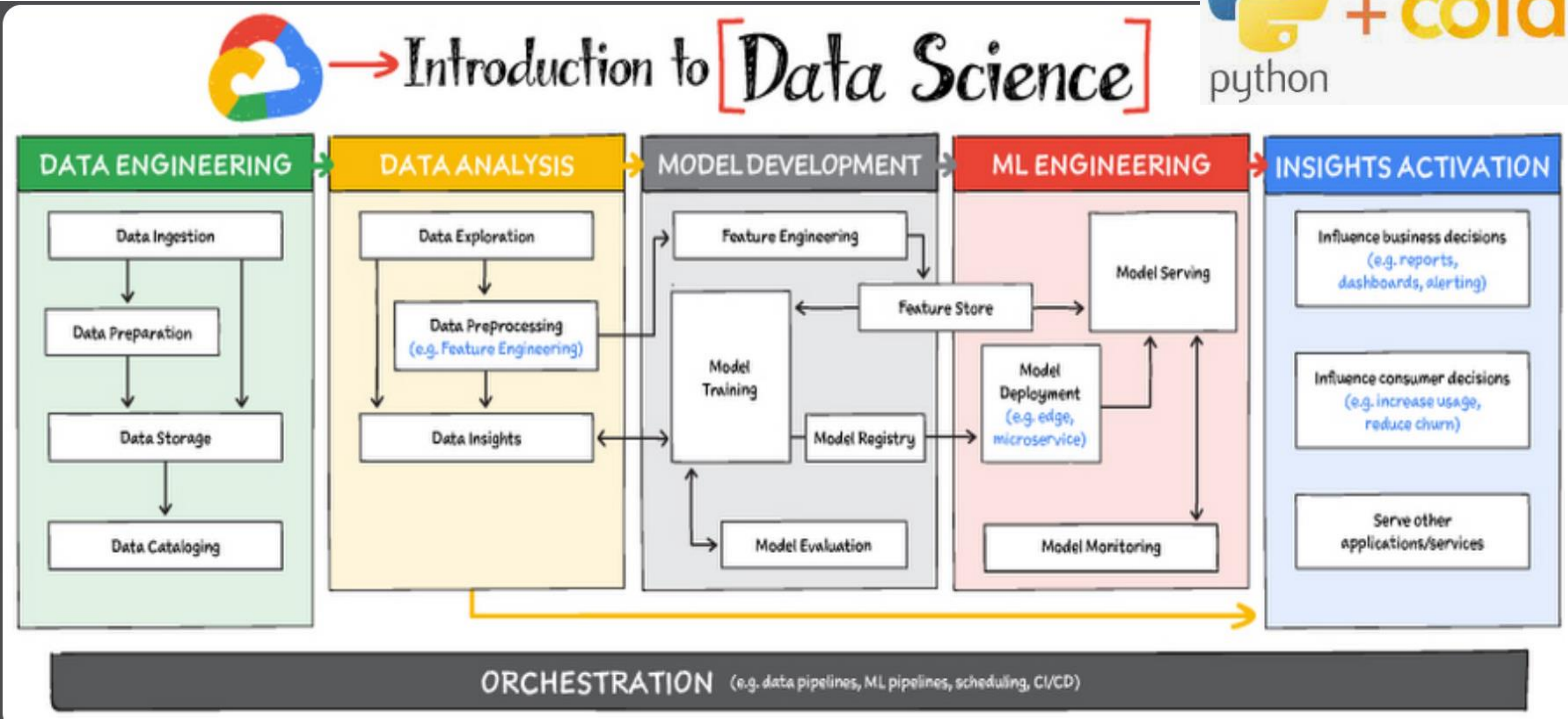
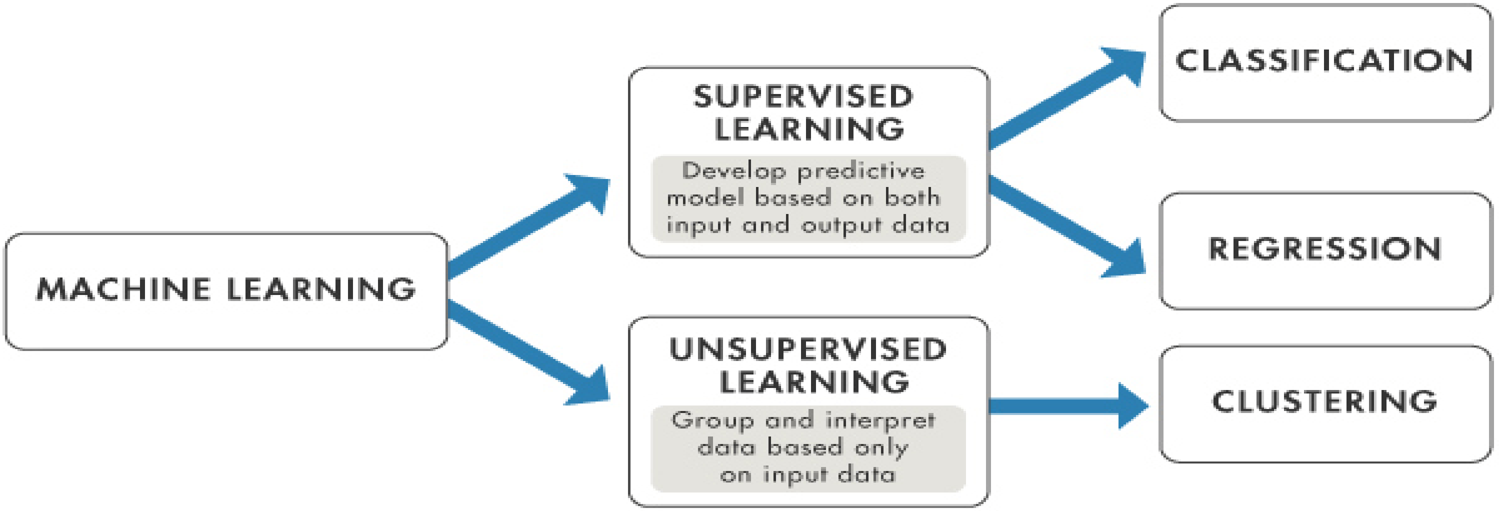


Machine learning

Intelligent Control Systems KOM5101	Preparation + Homework	Matlab Drive
1 Introduction to intelligent control systems (knowledge-based vs data-driven systems)	Select the project from https://github.com/mathworks/MathWorks-Excellence-in-Innovation#mathworks-excellence-in-innovation-projects	https://drive.matlab.com/sharing/c1f9073b-a0b0-4966-95b0-c107691878da
2 Computational thinking tools	Work with the Virtual Hardware and Labs for Control. Solve the following Labs Lab4_PositionAnalysis.mlx Lab3_PositionControl.mlx Lab2_VehicleModel.mlx Lab1_CruiseControl.mlx	https://drive.matlab.com/sharing/77e65af2-6ffd-4709-a0bd-c36e0fbe50df
3 Dynamical systems modelling	1. Study and Obtain the state space model of the crane system. 2. Study and Obtain the state space model of the Lateral Vehicle Dynamics: bicycle model with two degrees of freedom, lateral position and yaw angle.	https://drive.matlab.com/sharing/31e0ba39-b3f8-402c-a428-6a5b9d620081
4 Model Predictive Control MPC	Study and work with the MPC models explained. Use the MPC Toolbox of Matlab and the apmonitor server. Learn how to work with the drivingScenarioDesigner. Program the MPC algorithms using Simulink and Live scripts. Modify Models and MPC parameter and settings.	https://drive.matlab.com/sharing/398fa9fa-4650-4316-ab2b-0d228b24f48c
5 Machine Learning	 Machine Learning Onramp 6 modules | 2 hours | Languages Learn the basics of practical machine learning methods for classification problems.	

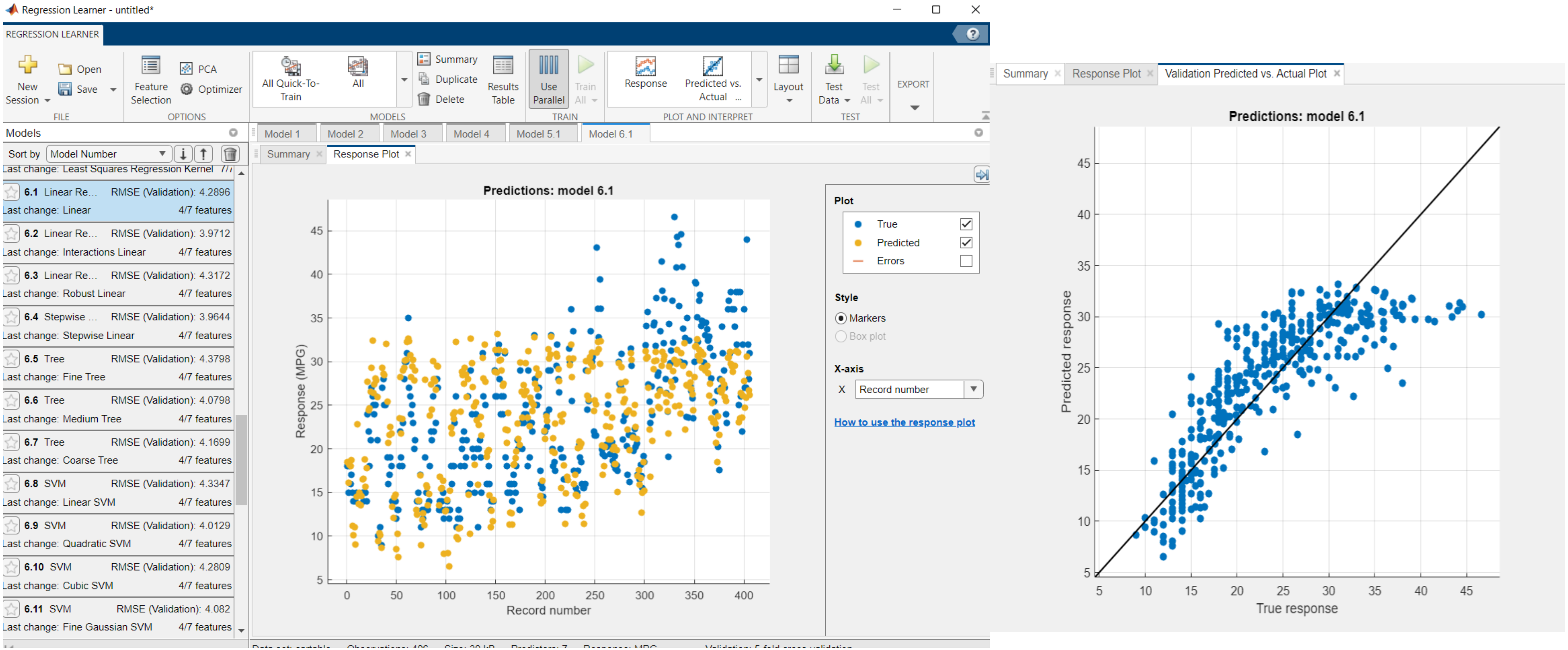
	Intelligent Control Systems KOM5101	Preparation + Homework	Matlab Drive
6	Data-Driven Modelling	Use the FOPDT example and do your own model estimation. Work with FOPDT live scripts :OPDT_Lab/L06_Assignment_graphical **Use the 2nd_order_linear model and obtain the regression parameters**	https://drive.matlab.com/sharing/71cc50d0-e79e-47e3-b91e-9ba1aa1cf78b
7	Data-Driven Modelling	Use the Hybrid Moving Horizon Estimation 2nd order MIMO System of the Tclab. Use the Classification Learner App and the Regression Learner App.	https://drive.matlab.com/sharing/be234c54-3734-431c-be84-ebb27e1047e6



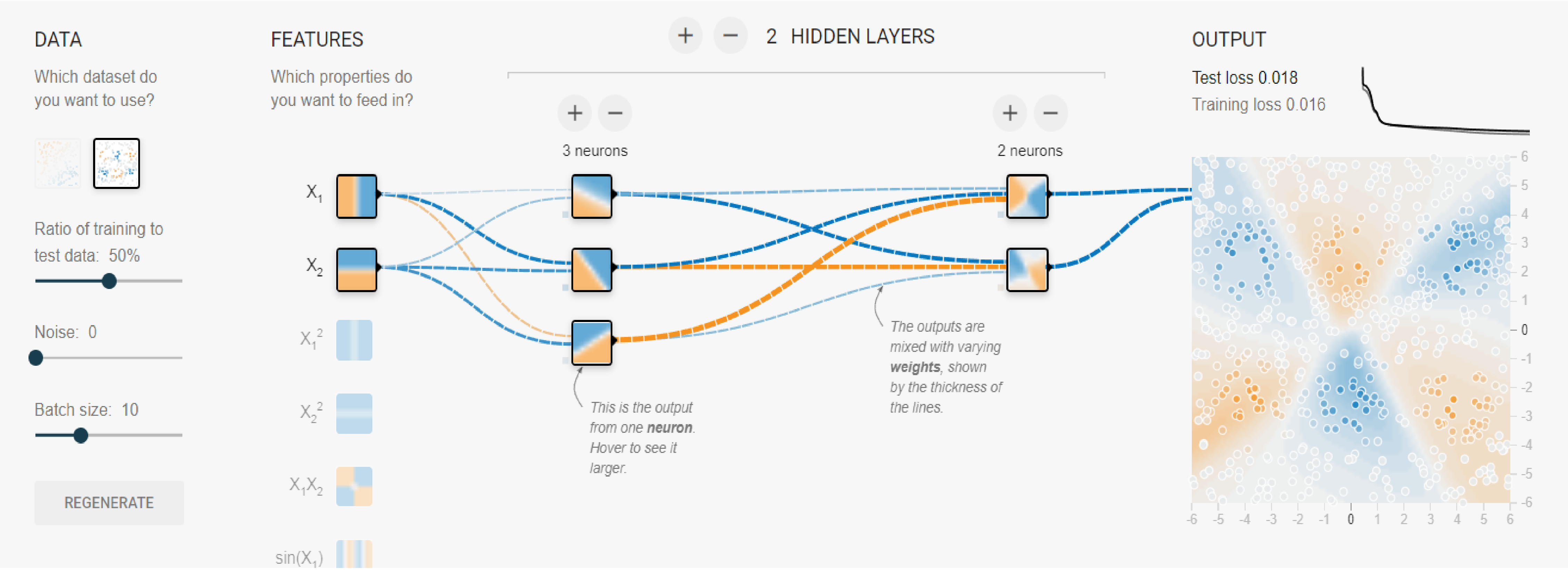
Regression to predict continuous responses	Use the Regression Learner app to automatically train a selection of models and help you choose the best. You can generate MATLAB code to work with scripts and other function options. For more options, you can use the command-line interface.	Statistics and Machine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

Train Regression Models in Regression Learner App

You can use Regression Learner to train regression models including linear regression models, regression trees, Gaussian process regression models, support vector machines, kernel approximation, ensembles of regression trees, and neural network regression models. In addition to training models, you can explore your data, select features, specify validation schemes, and evaluate results.



playground.tensorflow



https://playground.tensorflow.org/#activation=tanh&batchSize=10&dataset=circle&regDataset=reg-plane&learningRate=0.03&regularizationRate=0&noise=0&networkShape=4,2&seed=0.73206&showTestData=false&discretize=false&percTrainData=50&x=true&y=true&xTimesY=false&xSquared=false&ySquared=false&cosX=false&sinX=false&cosY=false&sinY=false&collectStats=false&problem=classification&initZero=false&hideText=false

Selecting the Right Algorithm

There are dozens of supervised and unsupervised machine learning algorithms, and each takes a different approach to learning. There is no best method or one size fits all. 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 overfit data by modeling 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.

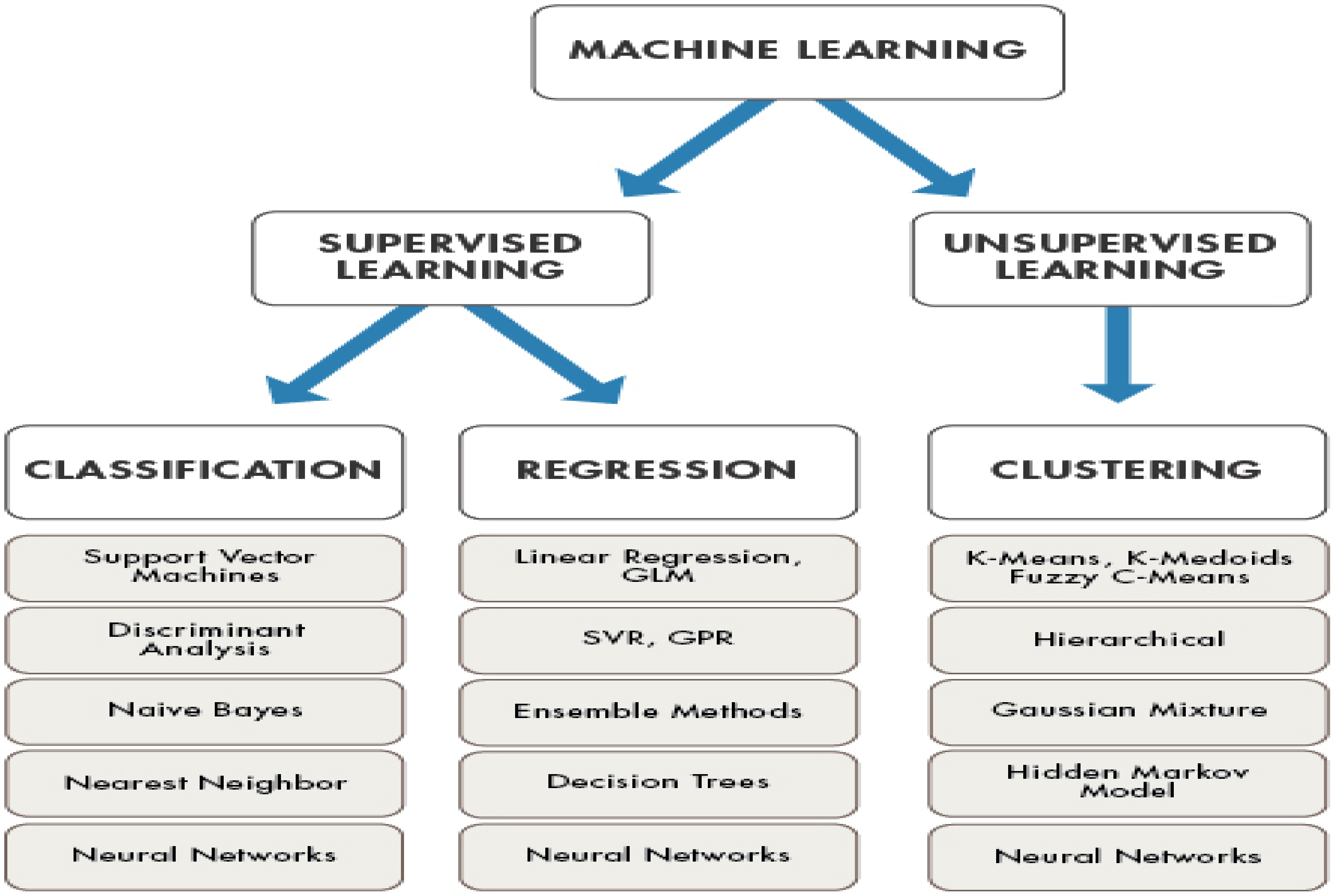
Model Statistics

Model 2: Tree
Status: Trained

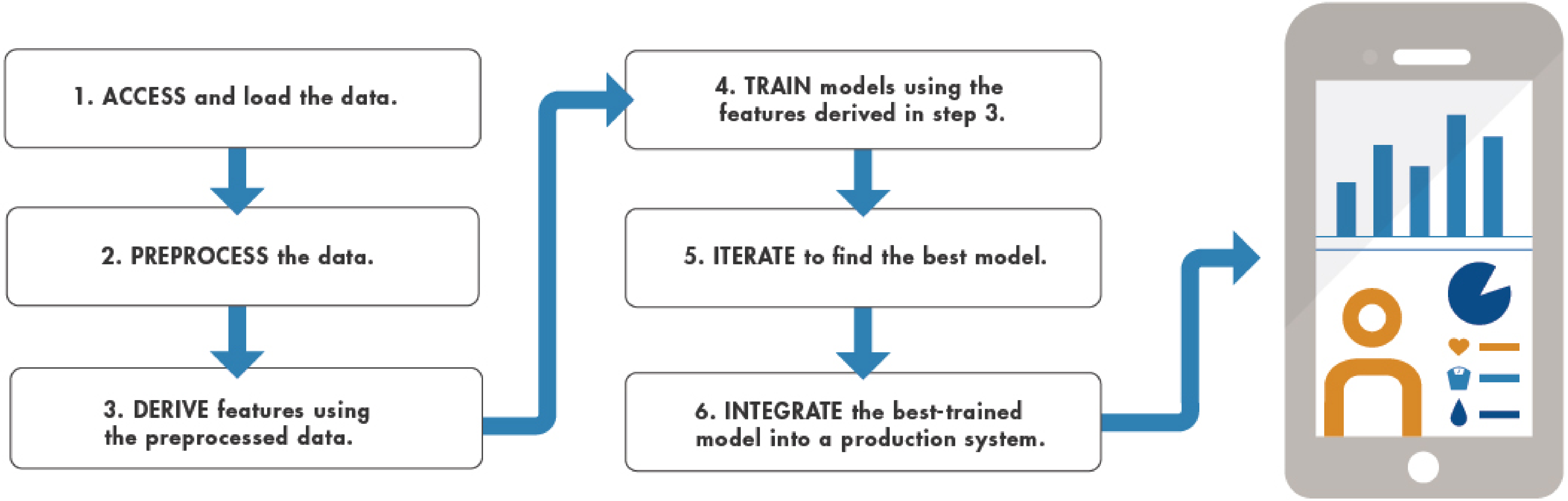
Training Results

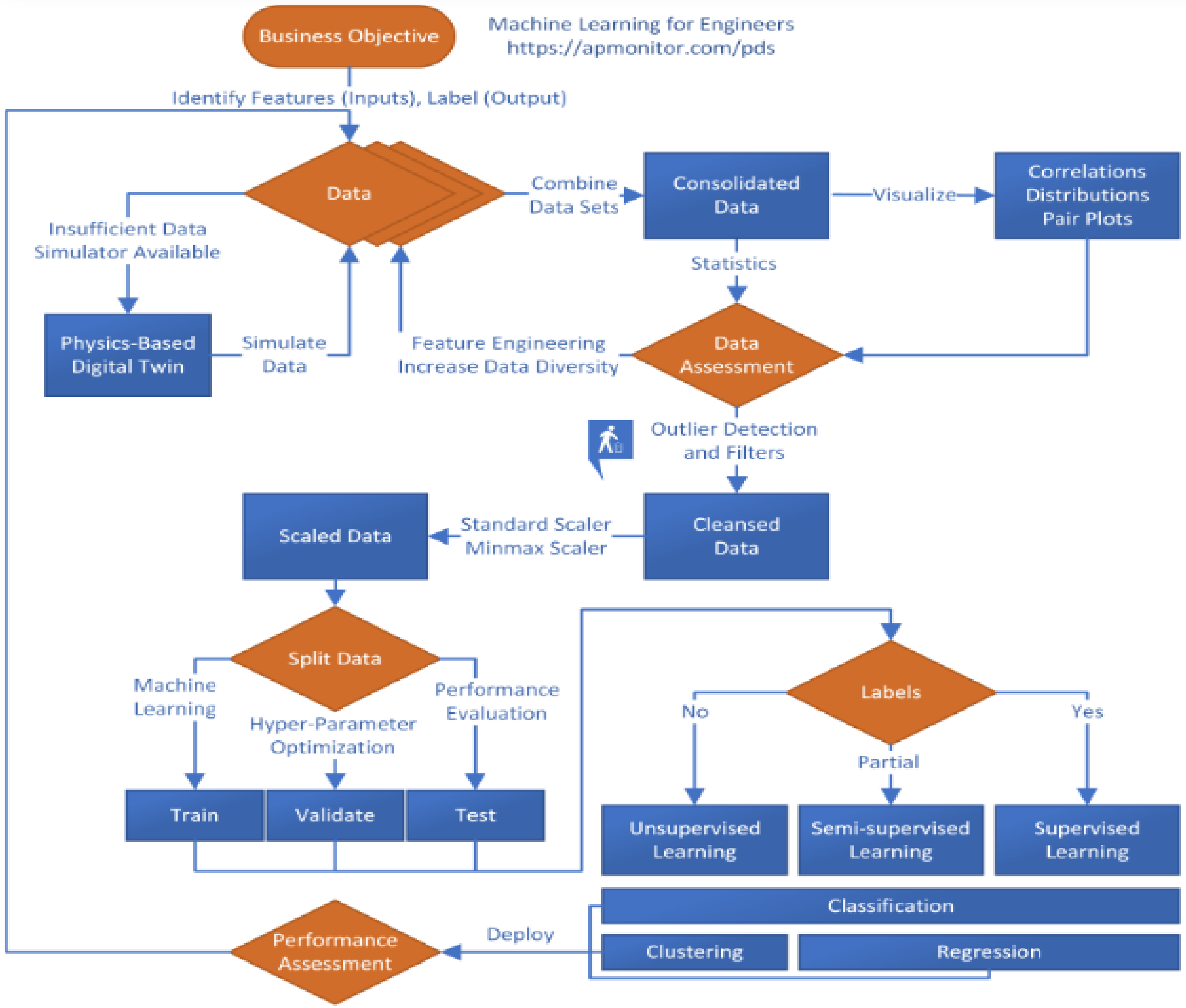
RMSE (Validation)	3.2821
R-Squared (Validation)	0.82
MSE (Validation)	10.772
MAE (Validation)	2.3731
Prediction speed	~5000 obs/sec
Training time	3.5947 sec

Statistic	Description	Tip	
RMSE	Root mean squared error. The RMSE is always positive and its units match the units of your response.	Look for smaller values of the RMSE.	$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}$
R-Squared	Coefficient of determination. R-squared is always smaller than 1 and usually larger than 0. It compares the trained model with the model where the response is constant and equals the mean of the training response. If your model is worse than this constant model, then R-Squared is negative.	Look for an R-Squared close to 1. $R^2 = 1 - \frac{SS_{RES}}{SS_{TOT}} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$	
MSE	Mean squared error. The MSE is the square of the RMSE.	Look for smaller values of the MSE.	$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \tilde{y}_i)^2$
MAE	Mean absolute error. The MAE is always positive and similar to the RMSE, but less sensitive to outliers.	Look for smaller values of the MAE.	$MAE = \frac{1}{n} \sum_{i=1}^{n} \lvert Y_i - \hat{Y}_i \rvert$


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

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





https://apmonitor.com/pds/index.php/Main/CourseSchedule

Stationarity: We do not have independence but consistency.

Data distribution depends on a difference (window) in time not location in time

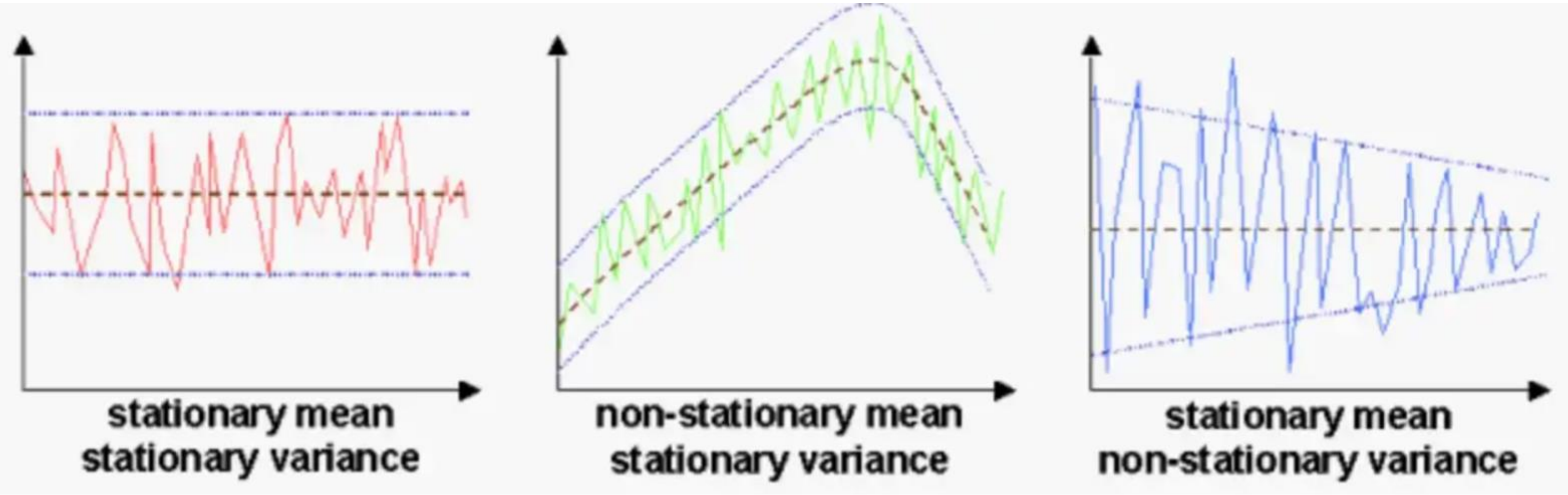
Stationarity is an important concept in time series analysis:

1. Stationarity means that **the statistical properties of a time series** do not change over time.
2. Stationarity is important because many useful analytical tools and statistical tests and models rely on it.

The ability to determine if a time series is stationary is important. This usually means being able to ascertain, with high probability, that a series is generated by a stationary process.

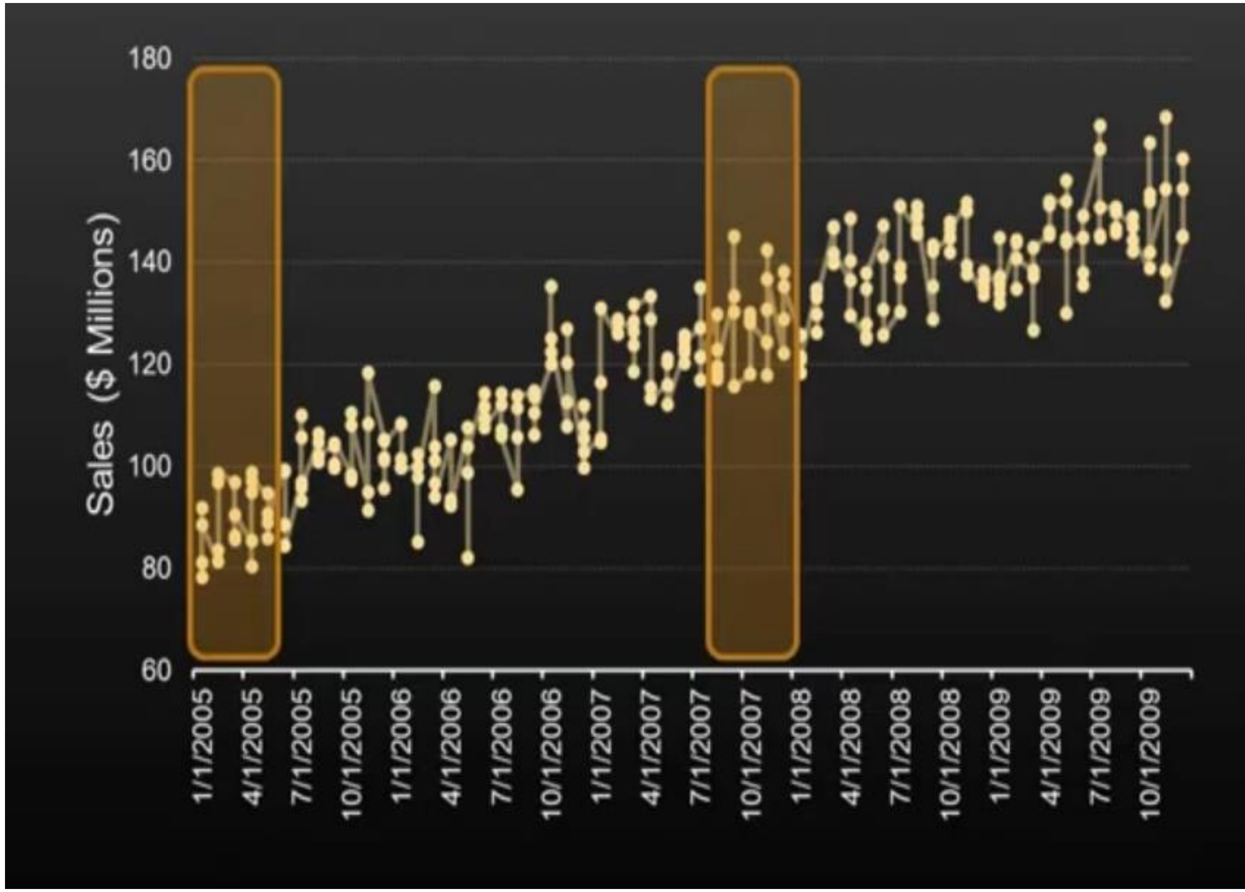
Test stationarity!

Mean, variance, autocorrelation depends **only** on difference in time, **not** location in time.

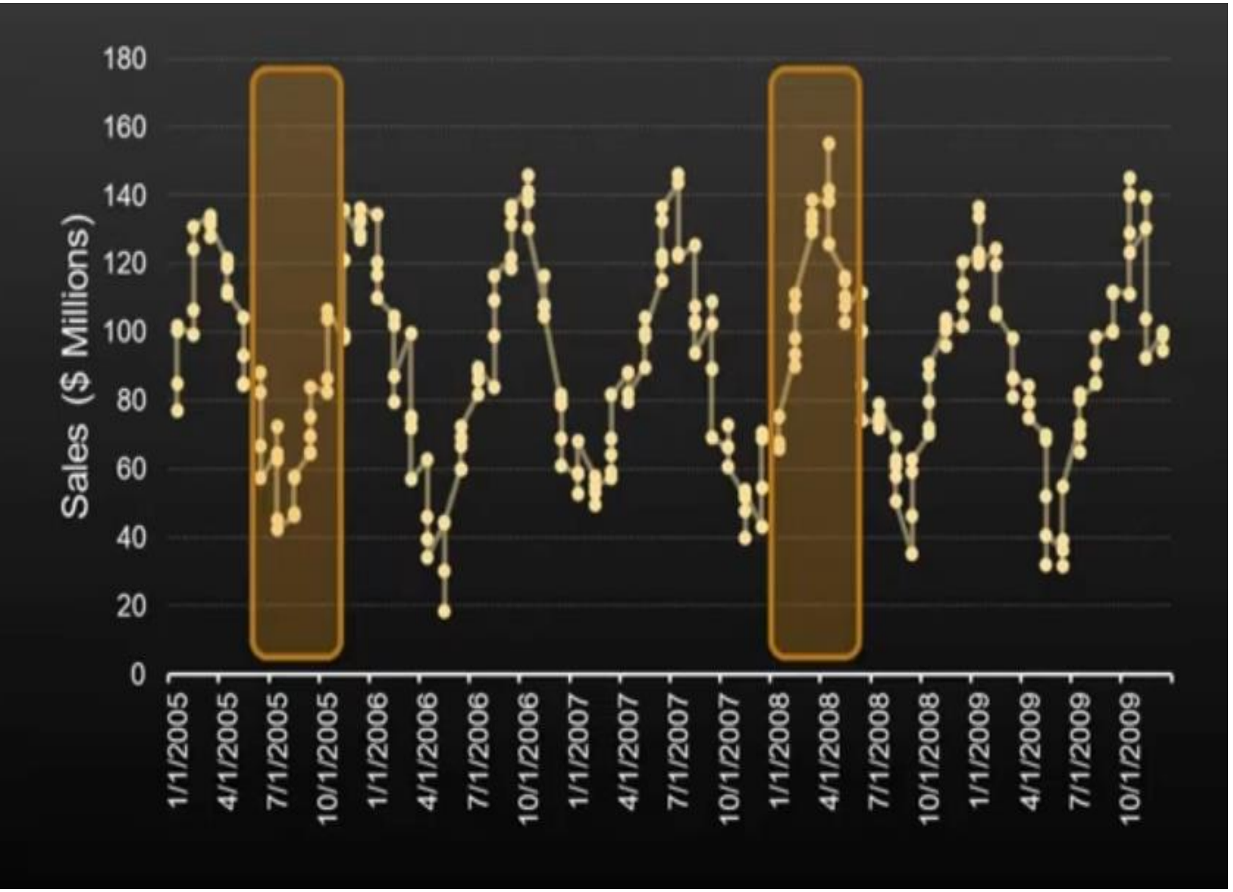


Not stationarity !
They do not have the same mean

Trending

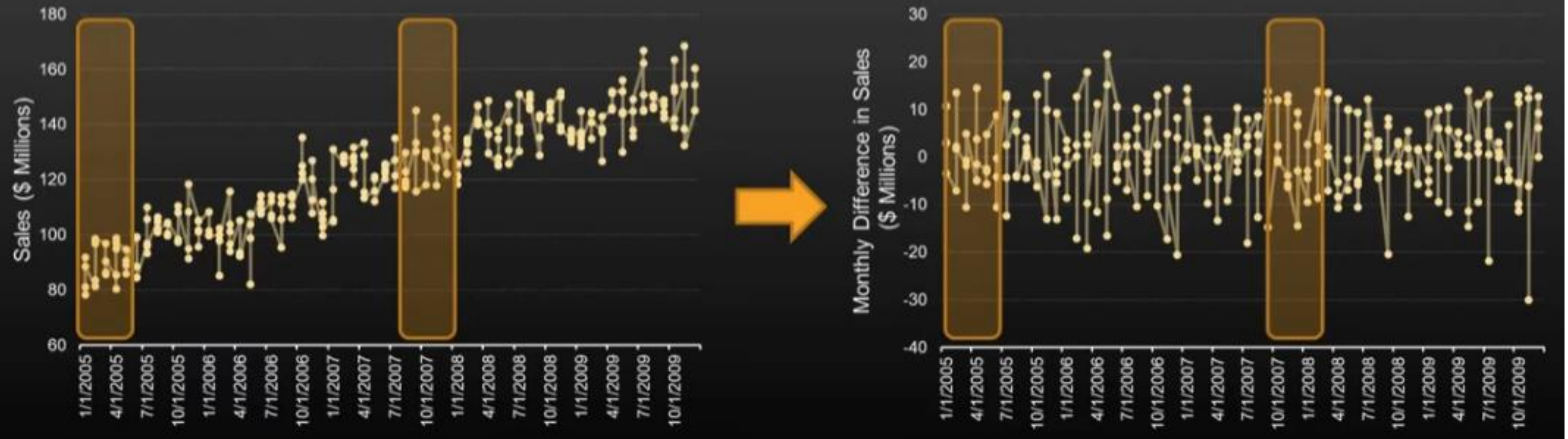


Seasonality



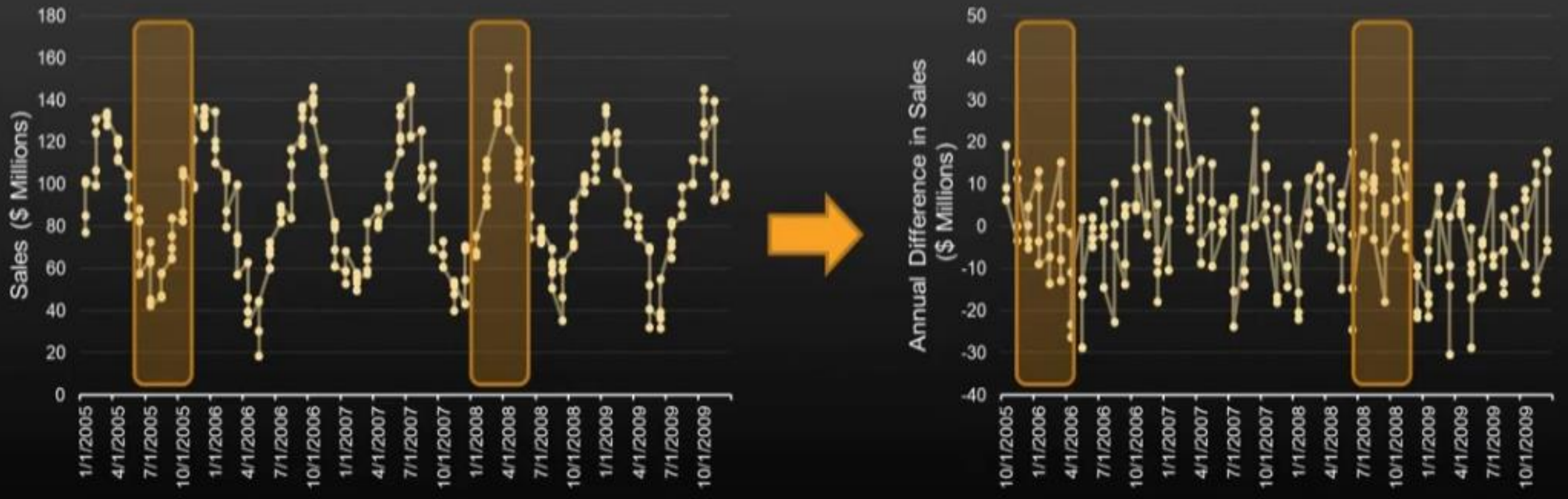
Stationarity!

- **Trend** – look at difference between current point and previous one: $Y_t - Y_{t-1}$



Stationarity!

- **Season** – look at difference between current point and the same point in the previous season: $Y_t - Y_{t-s}$



Correlation

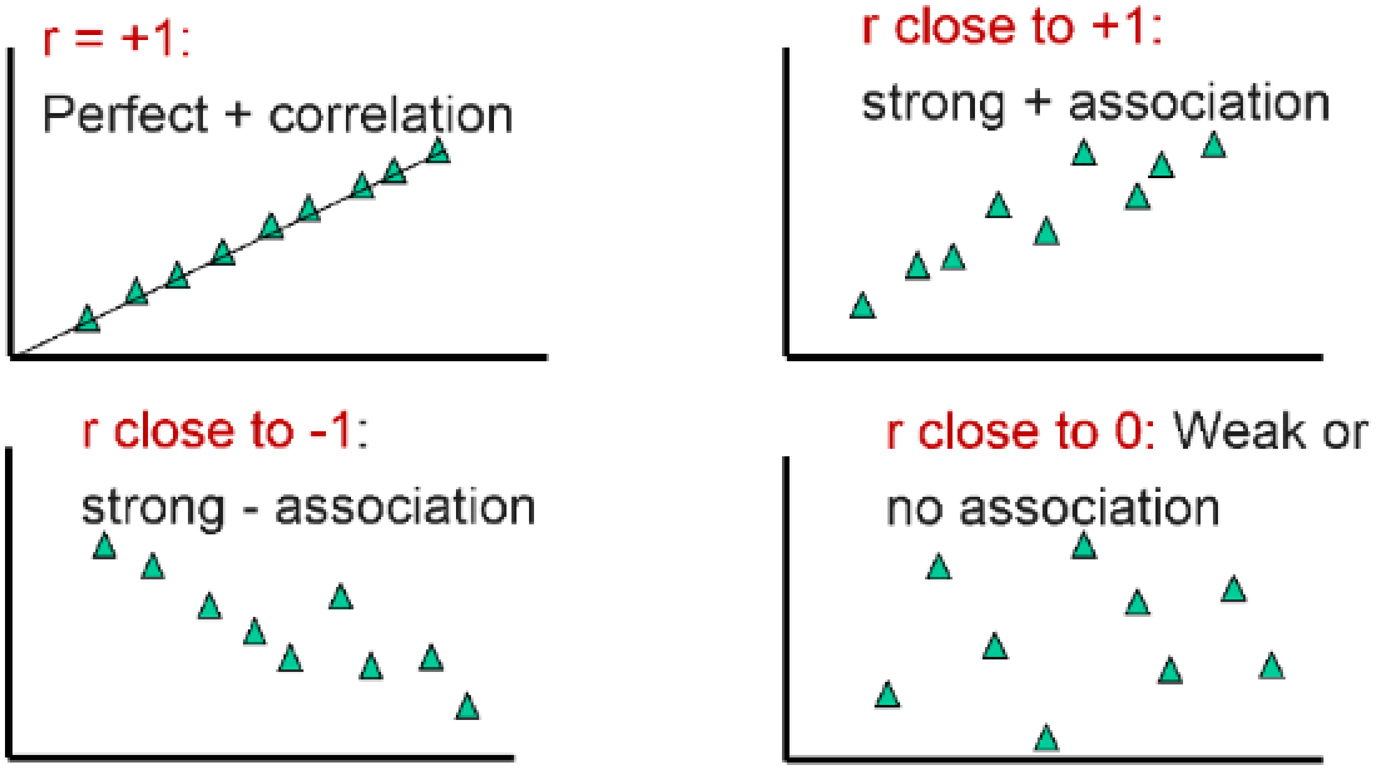
Correlation is a single statistic or data point, whereas regression is the entire equation with all of the data points that are represented with a line.

Correlation shows the relationship between the two variables, while regression allows us to see how one affects the other.

Difference between Correlation and Regression

Basis For Comparison	Correlation	Regression
Meaning	Correlation is a statistical measure that determines the association or co-relationship between two variables.	Regression describes how to numerically relate an independent variable to the dependent variable.
Usage	To represent a linear relationship between two variables.	To fit the best line and to estimate one variable based on another.
Dependent and Independent variables	No difference	Both variables are different.
Indicates	Correlation coefficient indicates the extent to which two variables move together.	Regression indicates the impact of a change of unit on the estimated variable (y) in the known variable (x).
Objective	To find a numerical value expressing the relationship between variables.	To estimate values of random variables on the basis of the values of fixed variables.

Correlation



$$r_{xy} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma(x_i - \bar{x})^2 \, \Sigma(y_i - \bar{y})^2}}$$

r_{xy} = correlation coefficient between x and y

x_i = the values of x within a sample

y_i = the values of y within a sample

$\bar{x}$ = the average of the values of x within a sample

$\bar{y}$ = the average of the values of y within a sample

Autoregressive Models

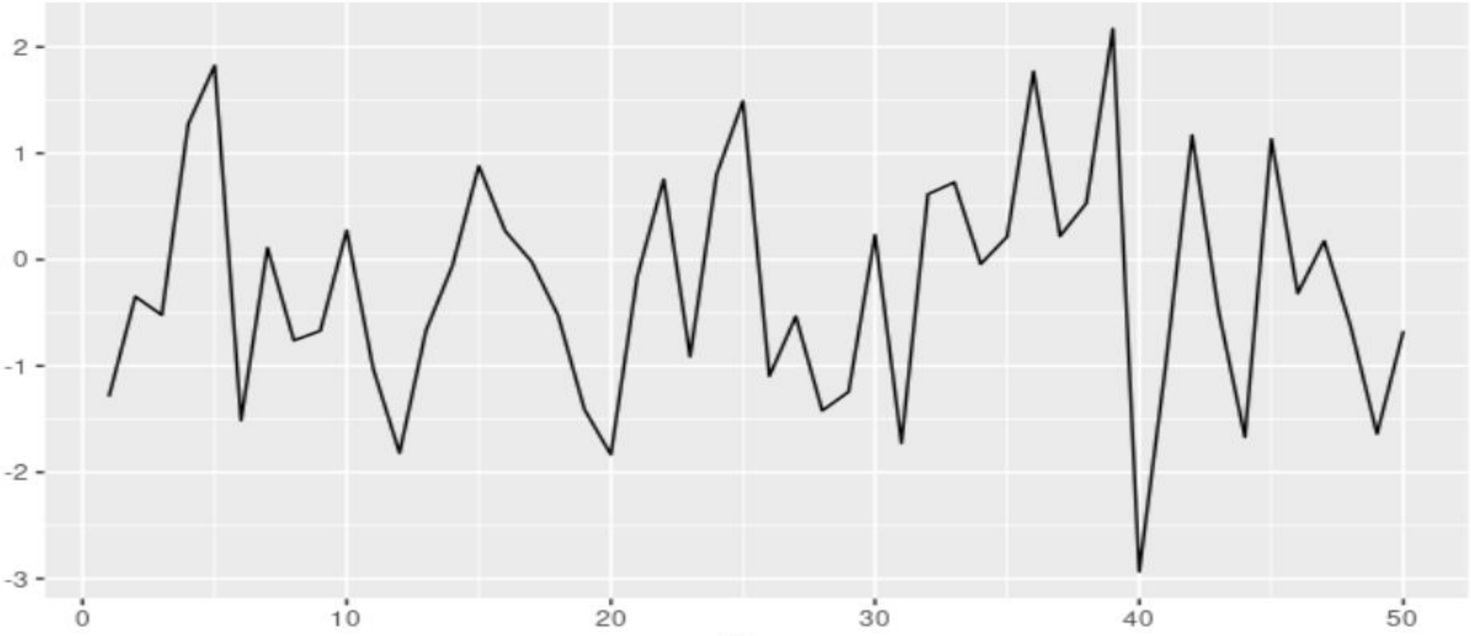
ARTime Series Model

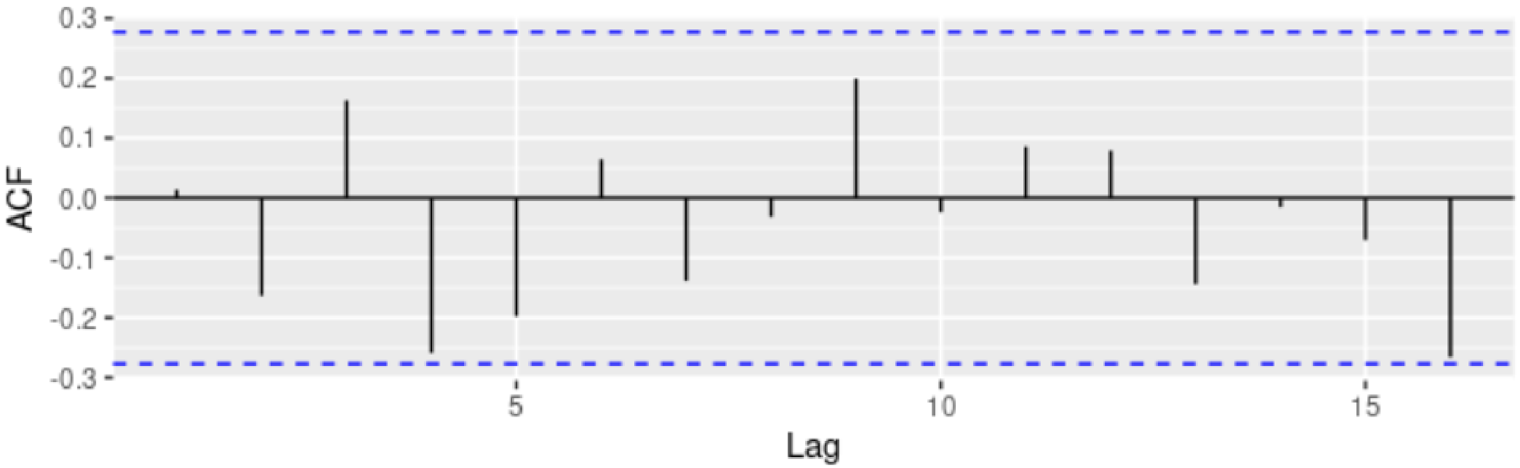
In a multiple regression model, we forecast the variable of interest using a linear combination of predictors. In an autoregression model, we forecast the variable of interest using a linear combination of *past values of the variable*. The term *auto*regression indicates that it is a regression of the variable against itself. Thus, an autoregressive model of order p can be written as

$$y_t = \omega + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t,$$

where e_t is white noise. This is like a multiple regression but with *lagged values* of y_t as predictors. We refer to this as an **AR(**p**) model**, an autoregressive model of order p.

Time series that show no autocorrelation are called white noise





For white noise series, we expect each autocorrelation to be close to zero. Of course, they will not be exactly equal to zero as there is some random variation. For a white noise series, we expect 95% of the spikes in the ACF to lie within $\pm 2/\sqrt{T}$ where T is the length of the time series. It is common to plot these bounds on a graph of the ACF (the blue dashed lines above). If one or more large spikes are outside these bounds, or if substantially more than 5% of spikes are outside these bounds, then the series is probably not white noise.

LSTM (Long Short Term Memory) networks are a special type of RNN (Recurrent Neural Network) that is structured to remember and predict based on long-term dependencies that are trained with time-series data. An LSTM repeating module has some interacting components.

- Data Generation and Preparation
- LSTM Model Build
- LSTM Model Training
- LSTM Prediction Validation
- LSTM Forecast Validation

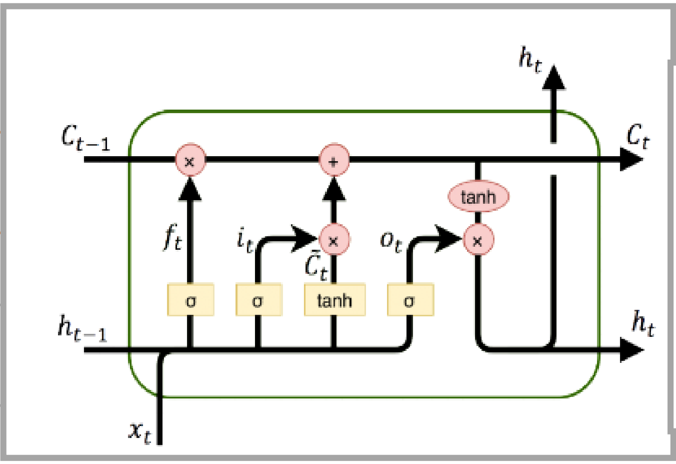
Data Preparation

Data preparation for LSTM networks involves consolidation, cleansing, separating the input window and output, scaling, and data division for training and validation.

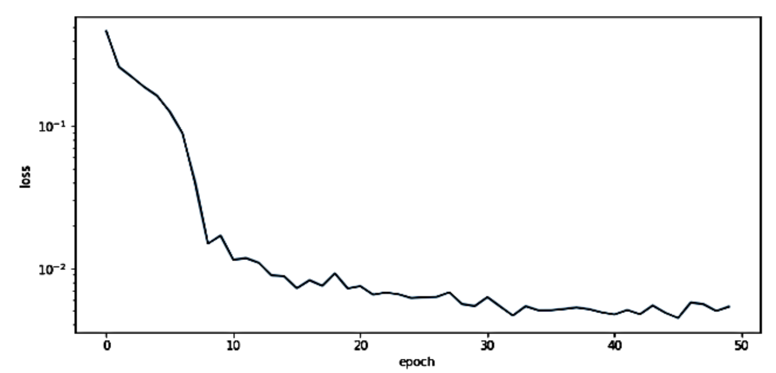
- Consolidation - consolidation is the process of combining disparate data (Excel spreadsheet, PDF report, database, cloud storage) into a single repository.
- Data Cleansing - bad data should be removed and may include outliers, missing entries, failed sensors, or other types of missing or corrupted information.
- Inputs and Outputs - data is separated into inputs (prior time-series window) and outputs (predicted next value). The inputs are fed into a series of functions to produce the output prediction. The squared difference between the predicted output and the measured output is a typical loss (objective) function for fitting.
- Scaling - scaling all data (inputs and outputs) to a range of 0-1 can improve the training process.
- Training and Validation - data is divided into training (e.g. 80%) and validation (e.g. 20%) sets so that the model fit can be evaluated independently of the training. Cross-validation is an approach to divide the training data into multiple sets that are fit separately. The parameter consistency is compared between the multiple models.

LSTM (Long Short Term Memory) networks are a special type of RNN (Recurrent Neural Network) that is structured to remember and predict based on long-term dependencies that are trained with time-series data. An LSTM repeating module has some interacting components.

- **Data Generation and Preparation**
- **LSTM Model Build**
- **LSTM Model Training**
- **LSTM Prediction Validation**
- **LSTM Forecast Validation**

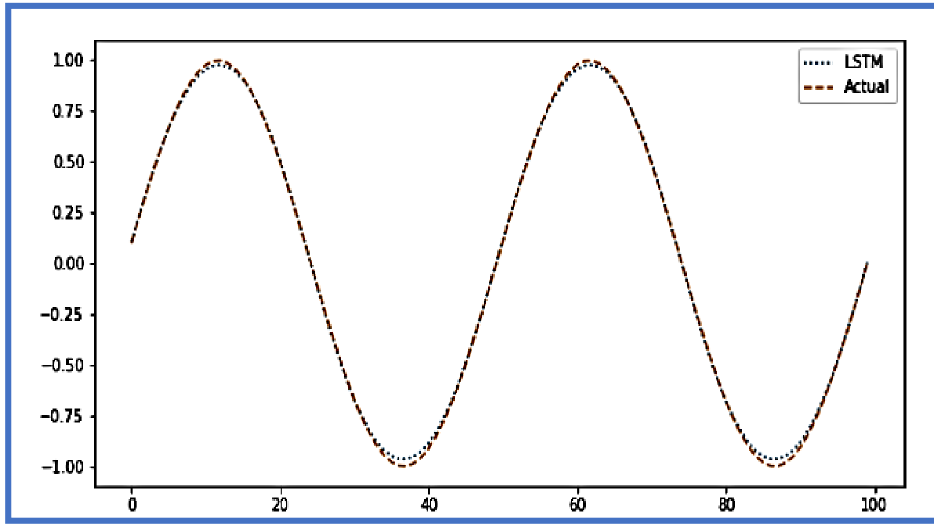


An LSTM network relates the input data window to outputs with layers. Instead of just one layer, LSTMs often have multiple layers.



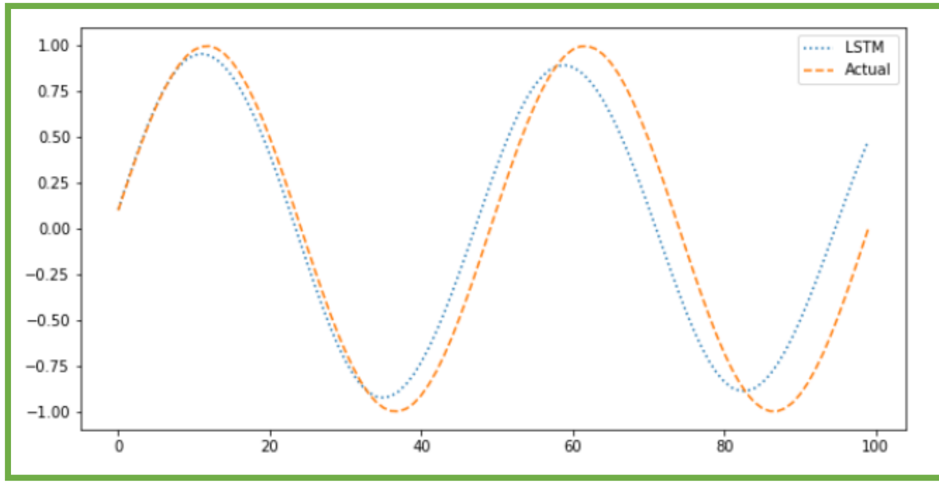
LSTM Prediction Validation

The validation test set assesses the ability of the neural network to predict based on new conditions that were not part of the training set. The validation is performed with the last 20% of the data that was separated from the beginning 80% of data.



LSTM Forecast Validation

When performing the validation it is also important to determine how the model performs with without measurements when it uses prior predictions to predict the next outcome. This is important to determine how well the model performs in a predictive application such as model predictive control where the model is projected forward over the control horizon to determine the sequence of optimal manipulated variable moves and possible future constraint violation. Generating predictions without measurement feedback is a forecast.



An Optimized Deep Learning Approach for Forecasting Temperature

https://www.canva.com/design/DAFjUD15Imc/sR9AgxdrTPi60mbf9yGPmw/edit?utm_content=DAFjUD15Imc&utm_campaign=designshare&utm_medium=link2&utm_source=sharebutton