

The background of the slide features a complex network diagram. It consists of numerous small, dark grey circular nodes connected by thin, light grey lines. These lines form a dense web of interconnected paths, with some areas appearing more clustered than others. The overall effect is one of a dynamic, interconnected system, possibly representing a data network or a structural model. The network is most prominent on the left side of the slide, where it is lighter, and fades into the dark teal background on the right.

# Condition Monitoring of Structures, Machines and Processes

## Tutorial 07

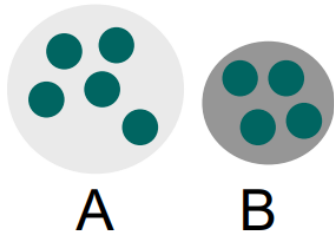
# 7

## MATLAB: CLASSIFICATION

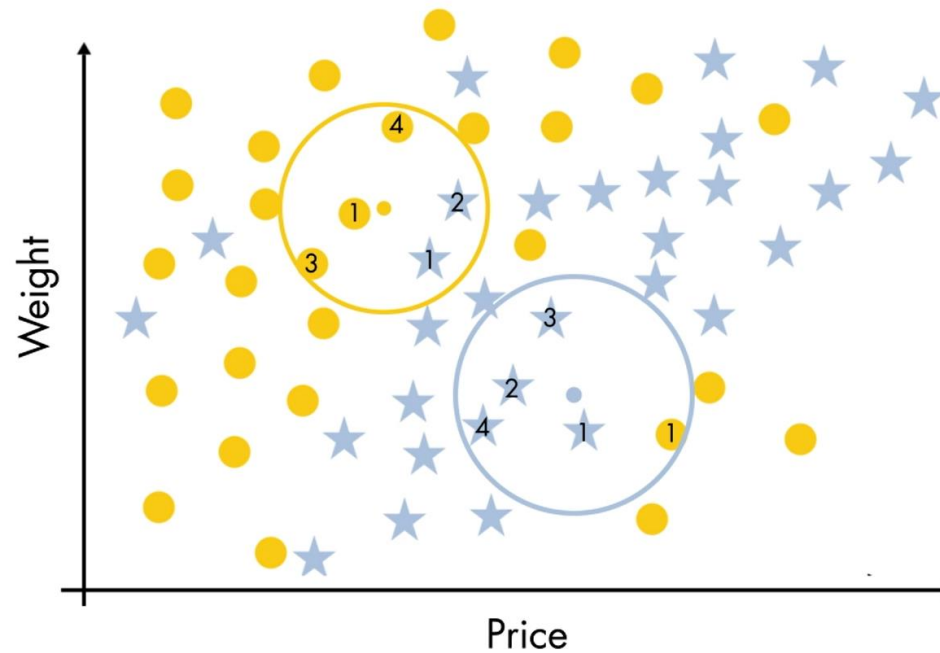
# Classification

## Classification

to which category  
does the entry  
belong?



Example: k-nearest neighbours,  $k = 4$



# Exercise 15: Classification

---

- Organizing data:
  - Download the 6 data, representing the extracted features (rms, peak frequency, partial power, etc..) for ultrasonic signal of a bicycle hub operating in gear 1, 2, 3 in a loaded and unloaded state.
  - After the import, split the data into a training dataset (500 observations) and a testing dataset (100 observations). Label and concatenate the dataset for the unloaded state into a 1500x24 Matrix and 300x24 Matrix. (23 Features and 1 Column for the labeling)
  - Remark: To avoid bias in the ML-Model it is important to work with evenly distributed data across the different states

# Exercise 15: Classification

---

- Analyzing Features:
  - To obtain a first impression of specific features, the commands `gscatter(Features1,Feature2,Color(1-3)/Label)` and `boxplot()` (<https://www.mathworks.com/help/stats/boxplot.html>) could be useful, visualizing the statistics and the corresponding label in one plot.
    - Apply `gscatter` and `boxplot()` to different features.
  - To identify the importance of features, Matlab provides useful functions like `fscchi2()` or `fscmrnr()`.(<https://www.mathworks.com/help/stats/feature-selection.html>)
    - Using these functions, try to identify the most important features and visualize them as in the previous exercise.
    - In particular, Gear 1 and 2 seem to be hard to distinguish. Identify the feature, which helps best to distinguish these gears.

## Exercise 15: Classification

---

- Building a model
  - Use the `fitcknn()` command to build a k-nearest neighbor model based on the trainingdata. The default value of k is 1. You can change it using `fitcknn(X,Y,'NumNeighbors',k)`
  - Use the `predict()` function to evaluate the model. That is, calculate the proportion of correct predictions based on the testdata by dividing the number of correct predictions by the total number of predictions