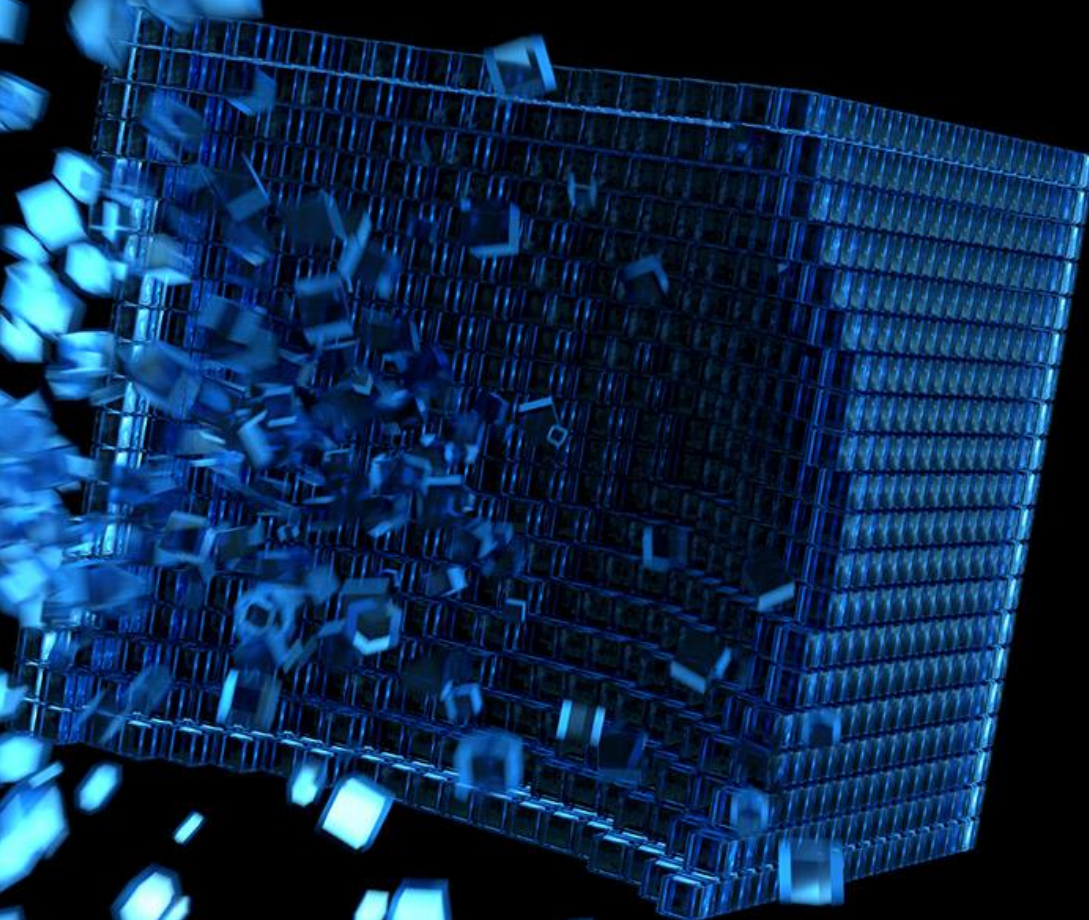


Anomaly – Everything but normal?!?

Andrea Spichtinger, Syskron GmbH



Motivation



„Predictive Maintenance“



Avoid unplanned downtimes



Data-driven exchange of wear parts instead of time-based

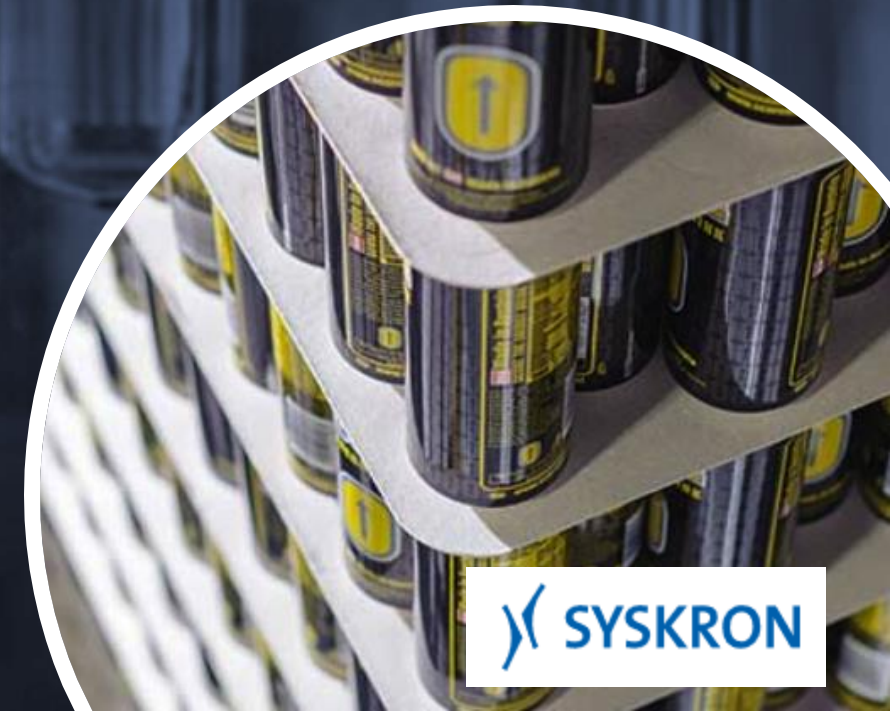


Every machine is unique

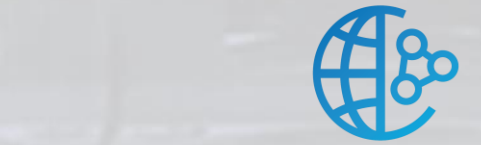


Very limited labels

→ Combination of Transfer Learning and Anomaly Detection



Definition Anomaly detection



= Outlier Detection, Novelty Detection, Deviation Detection, Exception Mining, Change Detection, ...

‘An anomaly is an observation which deviates so much from other observations as to arouse suspicions that it was generated by a different mechanism.’

(Hawking, 1980)

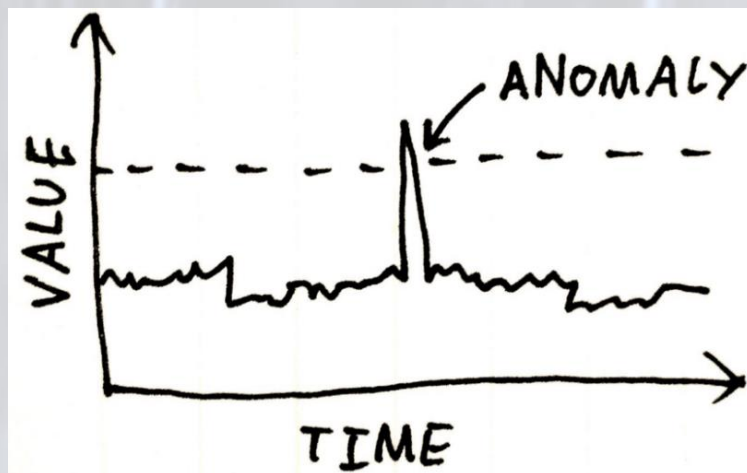
‘Anomalies [...] may or may not be harmful.’

(Esling, 2012)

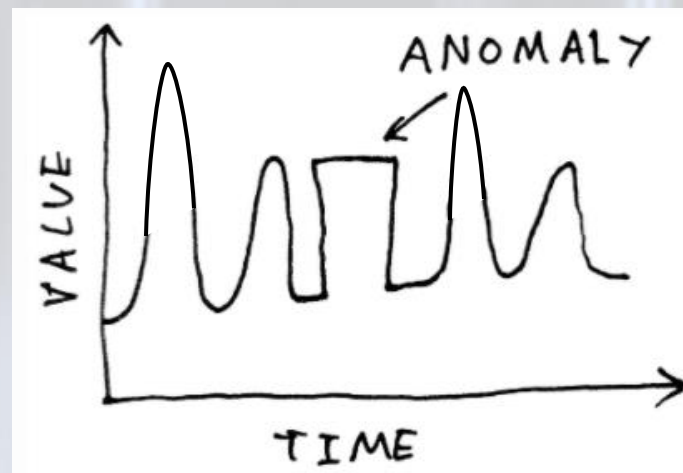
Types of Anomalies



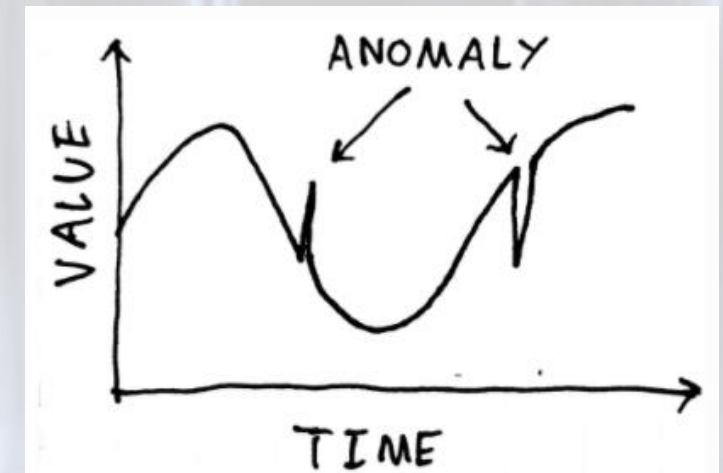
Point anomaly



Collective anomaly



Contextual anomaly

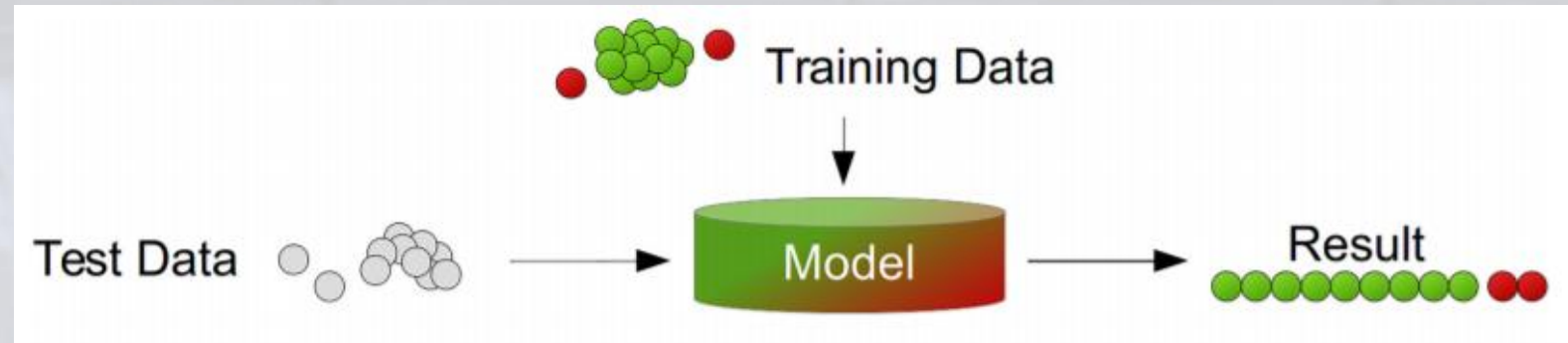


!! Important !!
Context and Domain knowledge
essential

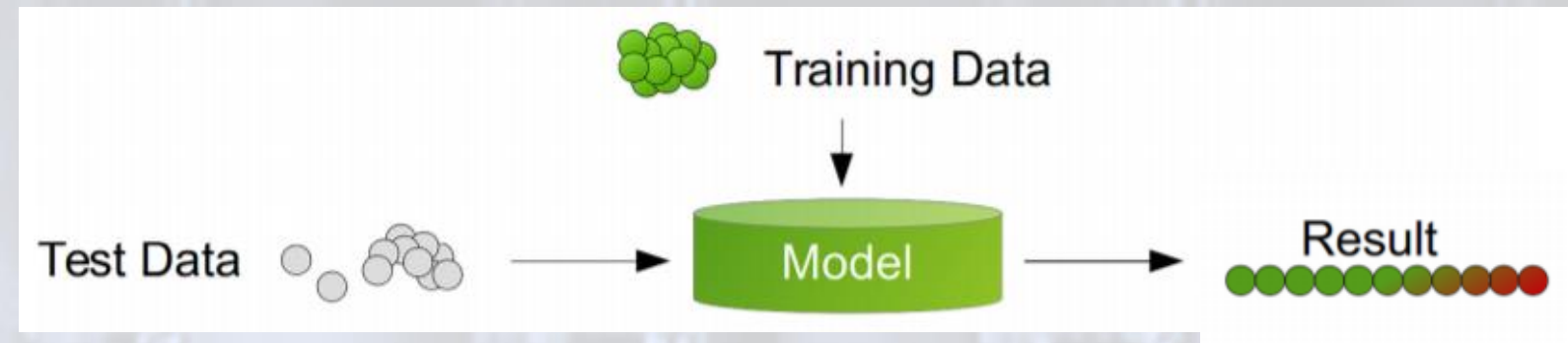
Setups of Anomaly Detection



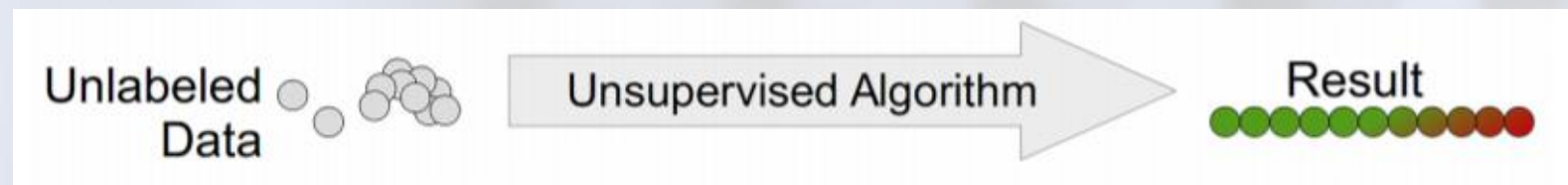
Supervised
Anomaly Detection



Semi-Supervised
Anomaly Detection



Unsupervised
Anomaly Detection



Anomaly approaches for point anomalies



Before: Feature engineering

supervised

semi-supervised

unsupervised

Anomaly algorithm

After: Transform into anomaly probability

Anomaly approaches for point anomalies



Before: Feature engineering

supervised

semi-supervised

unsupervised

Statistical

Classification

Clustering

Nearest
Neighbor

Spectral

Information
theoretic

**After: Transform into anomaly
probability**

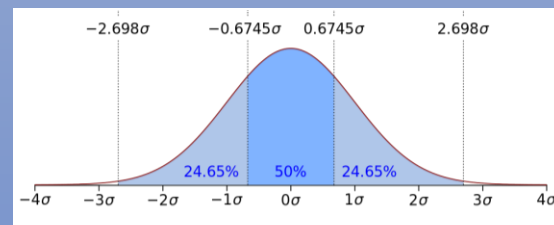
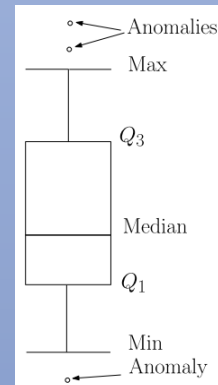
Statistical



*“Assumption: Normal data instances occur in high probability regions of a **stochastic model**, while anomalies occur in the low probability regions of the stochastic model”*

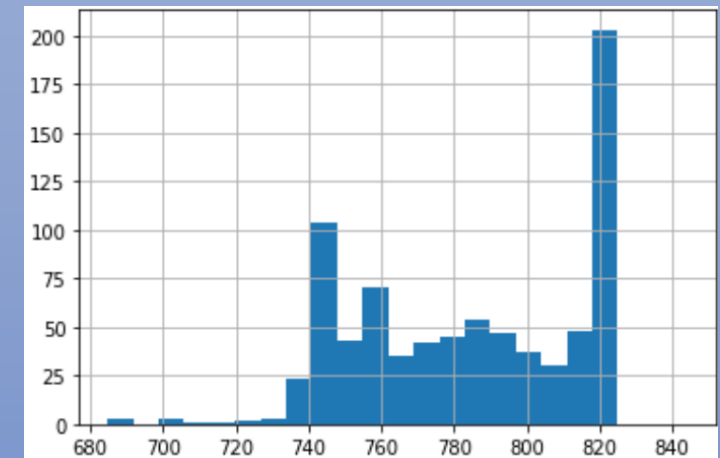
Parametric Techniques, e.g. Gaussian

- Standard Deviation
- Box plot rule
- student's t-test
- Gaussian Mixture Model
- ...



Non-Parametric Techniques

- Histograms
- Kernel functions





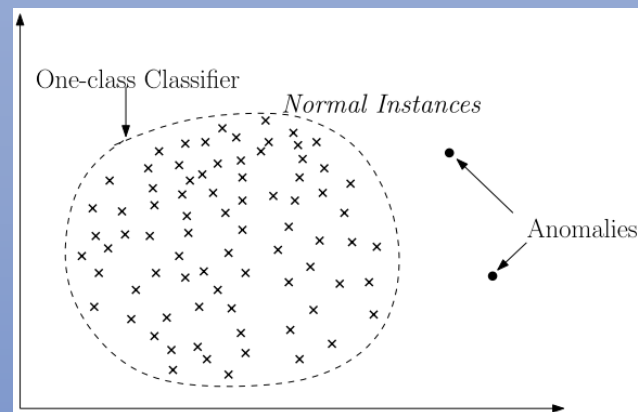
Classification

*“Assumption: A **classifier** that can distinguish between normal and anomalous classes can be learnt in the given feature space.”*

One-class Anomaly Detection

Learn discriminative boundary

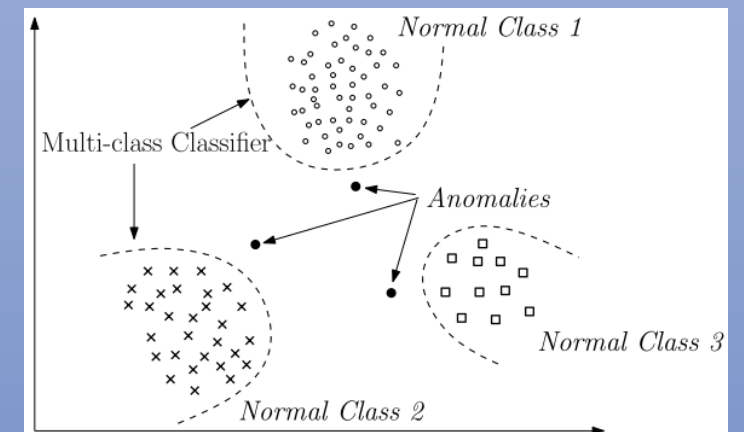
- One-class SVM
- ...



Multi-class Anomaly Detection

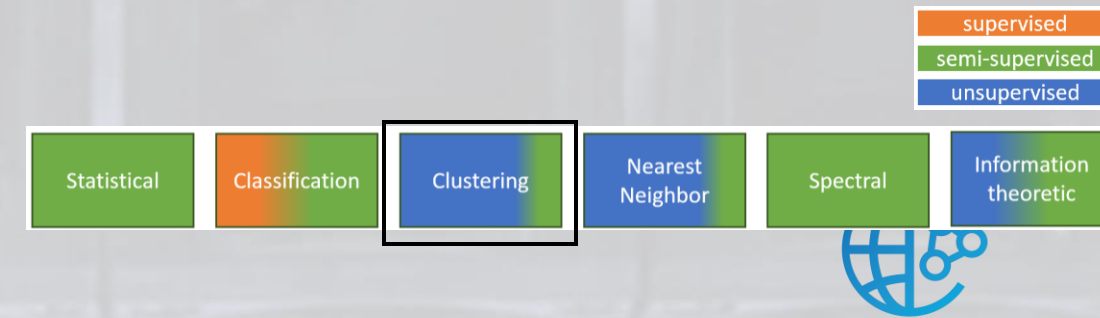
Train classifier for every class
→ If none of the classifiers confident → anomaly

- Neuronal networks
- Decision Tree
- ...



Assigns measurement to class, not probability

Clustering

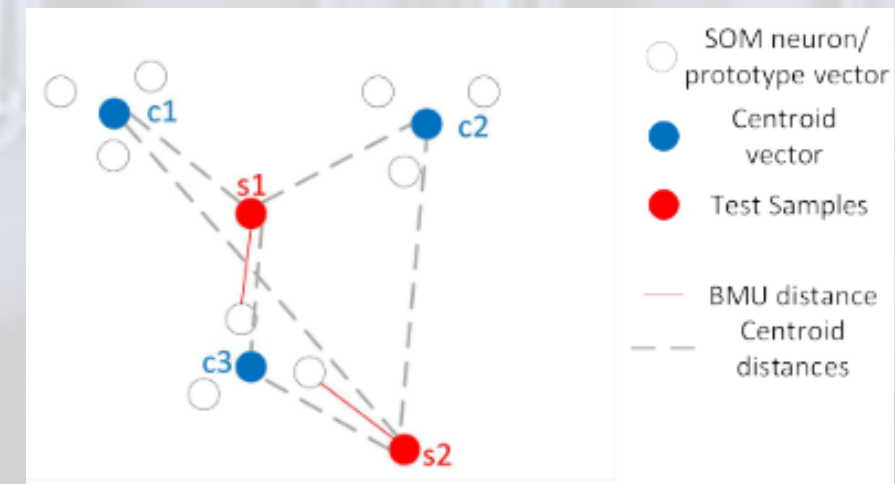
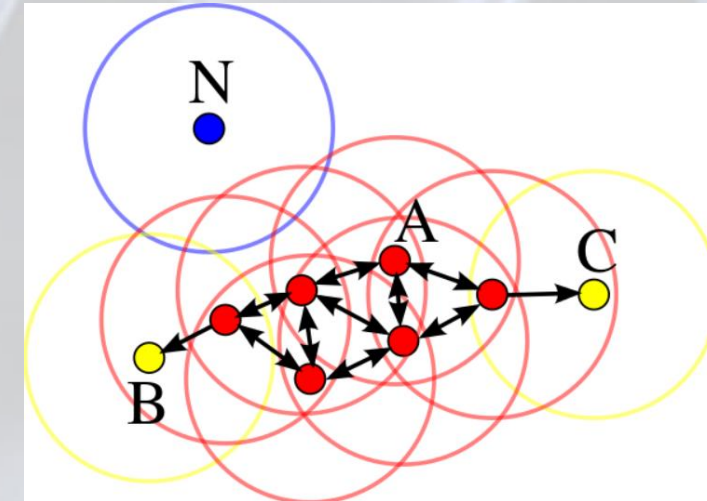


*“Assumption 1:
Normal data instances belong
to a **cluster**, while anomalies do
not belong to any cluster.”*

- DBSCAN
- OPTICS
- ...

*“Assumption 2: Normal data
instances lie close to their
closest **cluster centroid**, while
anomalies are far away from
their closest cluster centroid.”*

- Self-Organizing
Maps (SOM)
- ...



Nearest Neighbor



“Assumption: Normal data instances occur in dense neighborhoods, while anomalies occur far from their closest neighbors.”

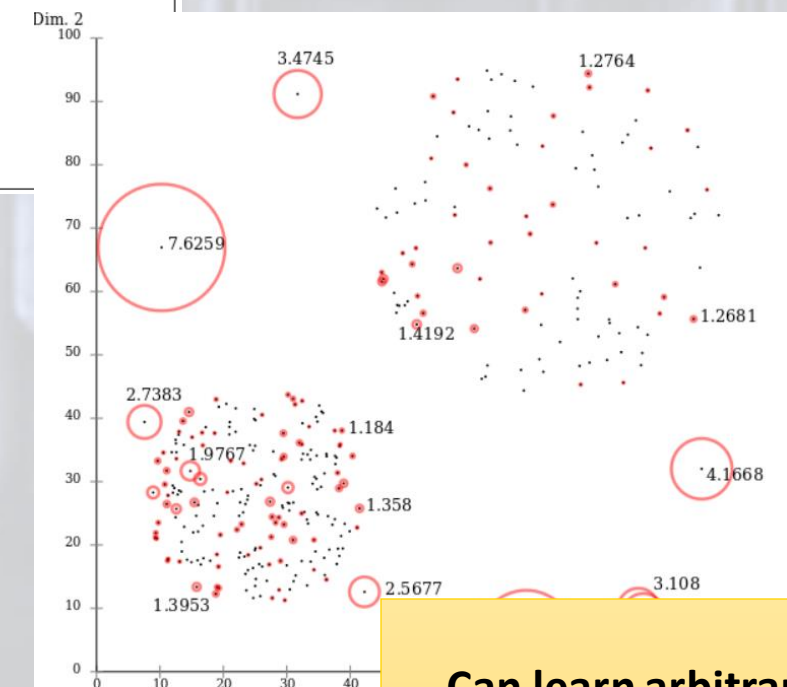
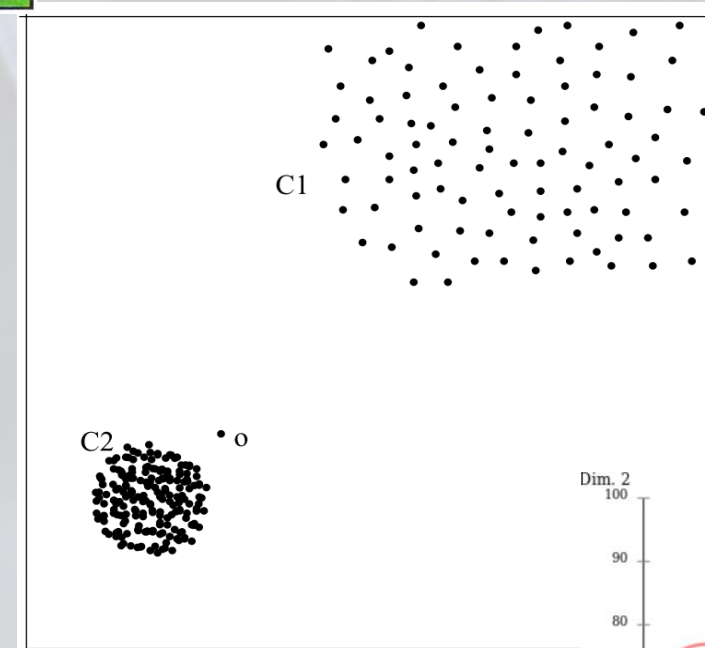


Global

- Distance to the kth nearest neighbor
- ...

Local

- Local Outlier Factor
- ...



Can learn arbitrary shapes!!

https://en.wikipedia.org/wiki/Isolation_forest/, https://de.wikipedia.org/wiki/Local_Outlier_Factor,

Chandola 2009, YSKRON

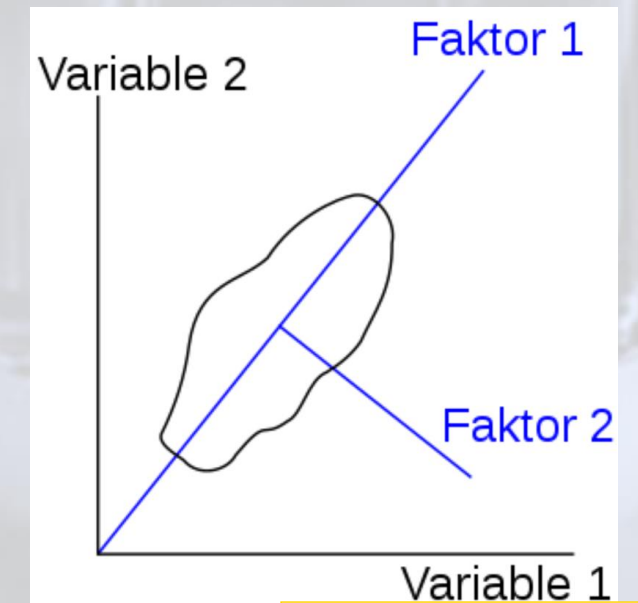
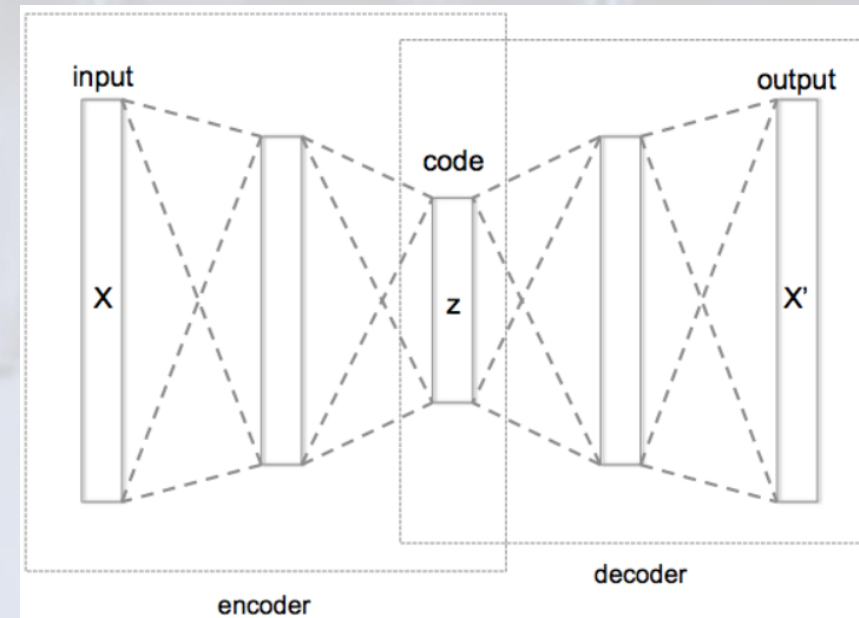
[Where is everybody? 'Where is everybody?' 'On HD164595.' | by Kopjes Kattenoppas | Medium](#)

Spectral

*“Assumption: Data can be embedded into a **lower dimensional subspace** in which normal instances and anomalies appear significantly different.”*

**Make yourself small
and you will be great.**

- Autoencoder
- Principal Component Analysis (PCA)
- ...



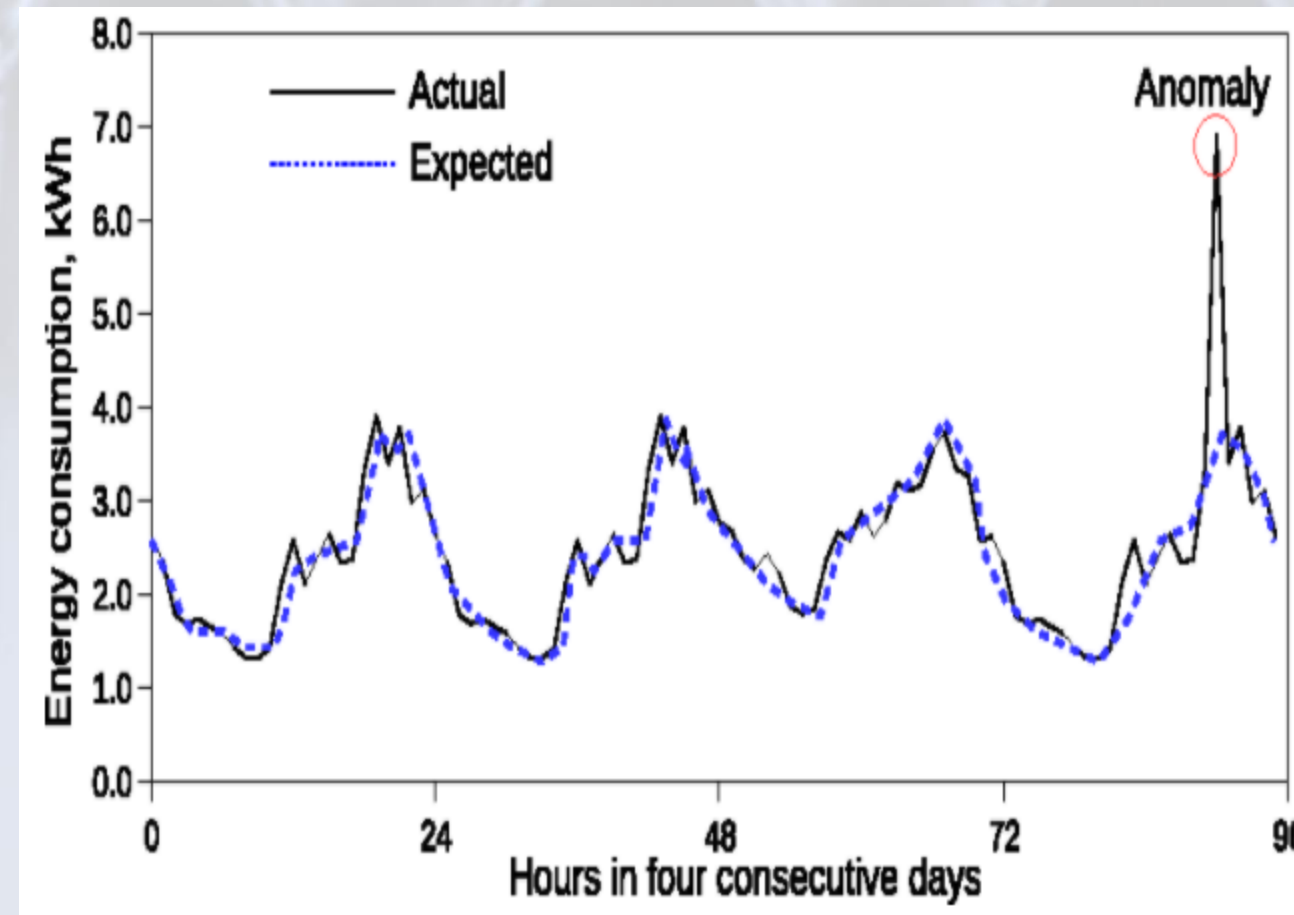
**Especially good for
higher dimensions!**

Information theoretic

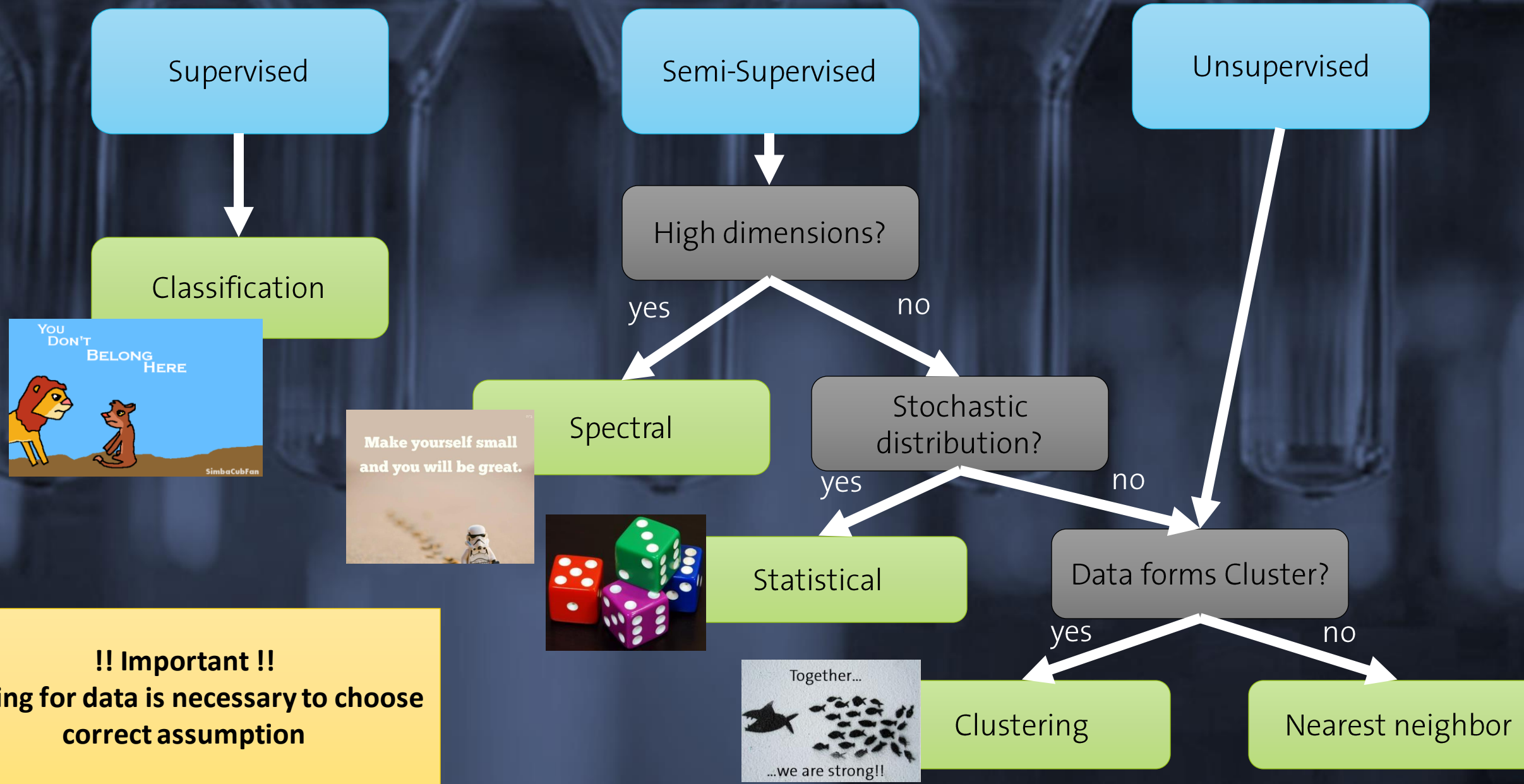
*“Assumption: Anomalies in data induce irregularities in the **information content** of the data set.”*

- Kolomogorov complexity, e.g. size of compressed file
- Entropy measures, e.g. Kullback-Leibler divergence
- ...

Contextual anomalies: Prediction approach



Personal Essence (simplified)



!! Important !!
Feeling for data is necessary to choose
correct assumption

Thanks for your
attention!