Machine Learning for Predictive Maintenance: A Multiple Classifier Approach

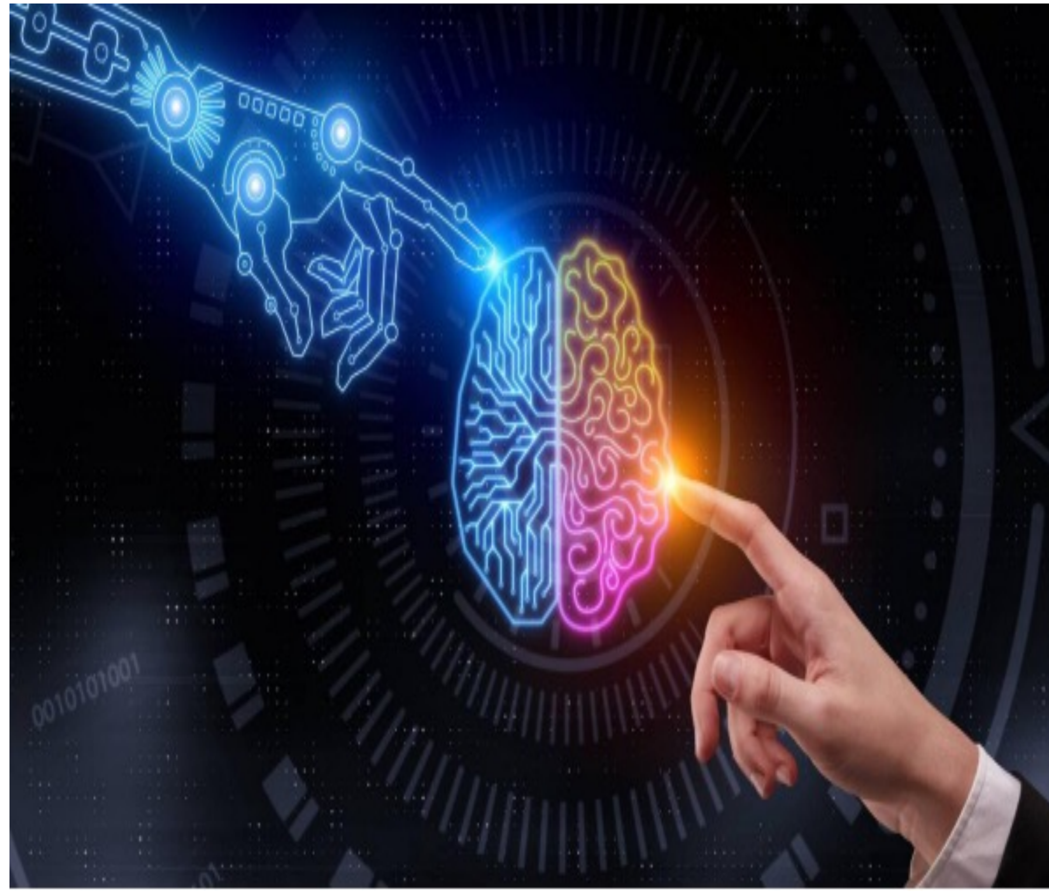
Ambuj Kumar Mondal

28-March-2022

PADERBORN UNIVERSITY

Agenda

- PdM
  - Intelligence of Prediction
  - Maintenance types
  - Importance of PdM
- Machine Learning in PdM
  - Supervised Learning
  - Unsupervised Learning
  - Semi-supervised Learning
- Multiple Classifier Approach
  - Tradeoff Optimization
  - MC PdM algorithm
  - Classification algorithms
- Use case and Experiment
  - Semiconductor manufacturing
  - Challanges & future tasks
- Conclusion

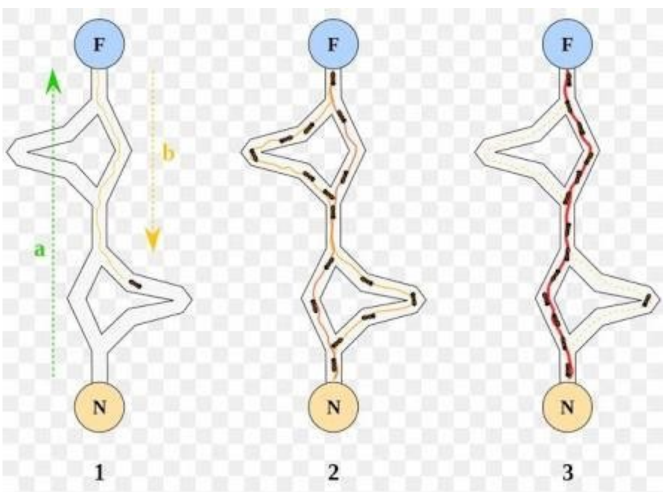


[pic0]

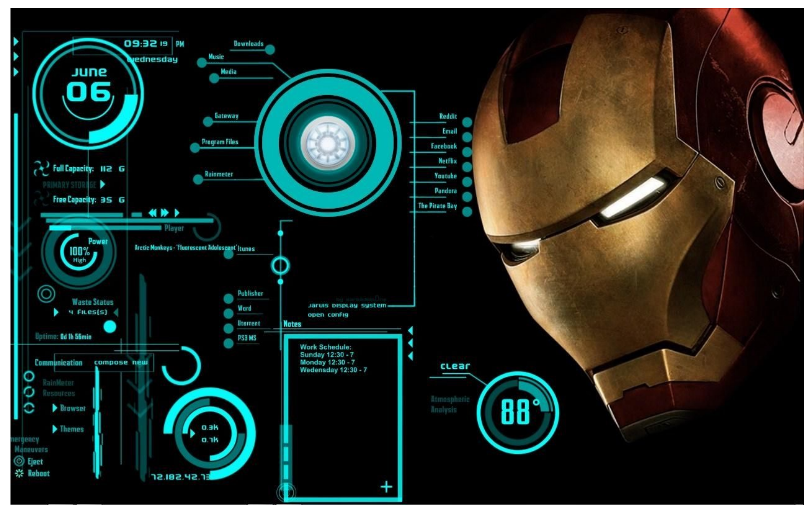
Intelligence of Prediction

- “The ability to perceive or infer **information**, and to retain it as **knowledge** to be applied towards **adaptive** behaviours within an environment or context” [wikipedia]

- **Artificial Intelligence :** is the design of artificial agents that perceive their environment and make decisions to maximise the chances of achieving a goal.[2]



[pic1]



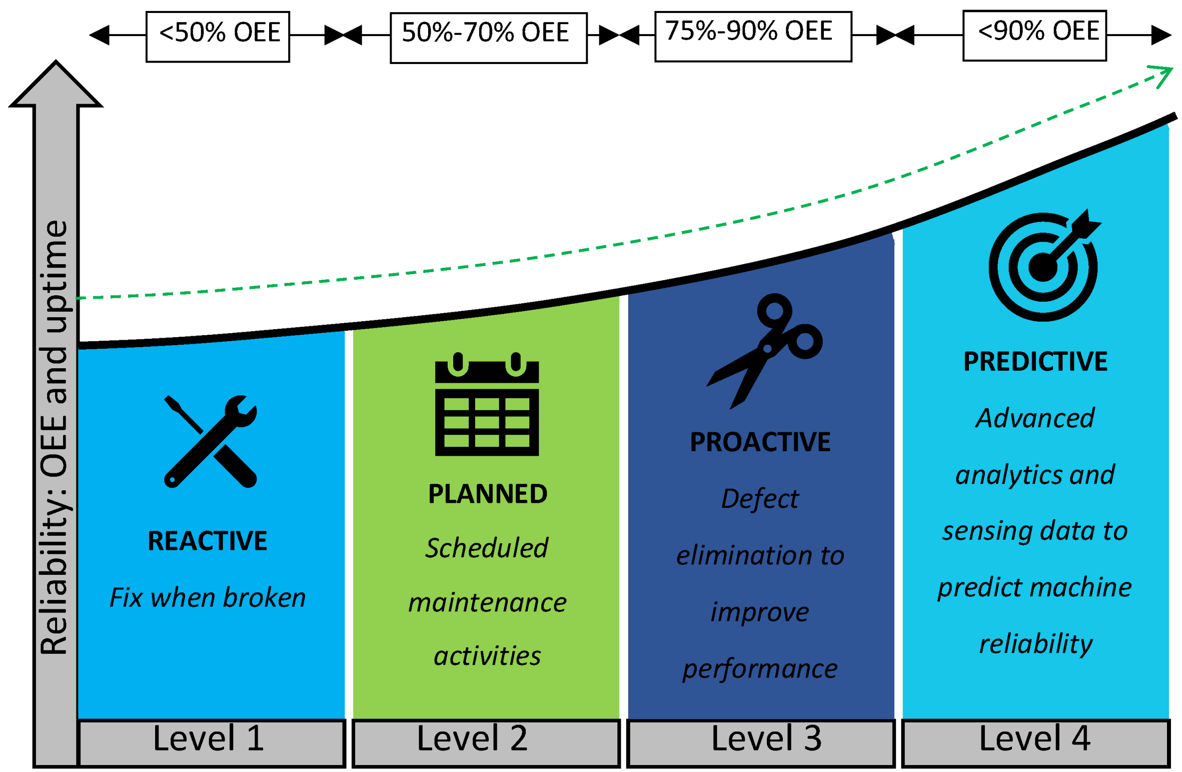
[pic2]

Maintenance Activities

- **R2F- Run to Failure**
  Descriptive
  Diagnostic

- **PvM- Preventive Maintenance**
  Scheduled maintenance
  Planned Activities
  Avoid breakdowns

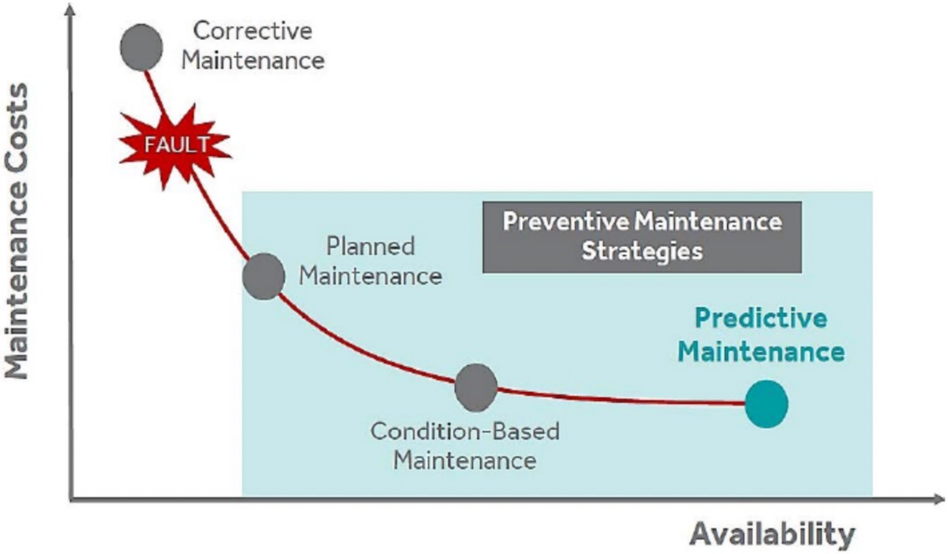
- **PdM- Predictive Maintenance**
  Failure Prevention
  Efficient Operations
  Reduced maintenances frequency



[pic3] : Different ways of Maintenance

Why Predictive Maintenance (PdM)

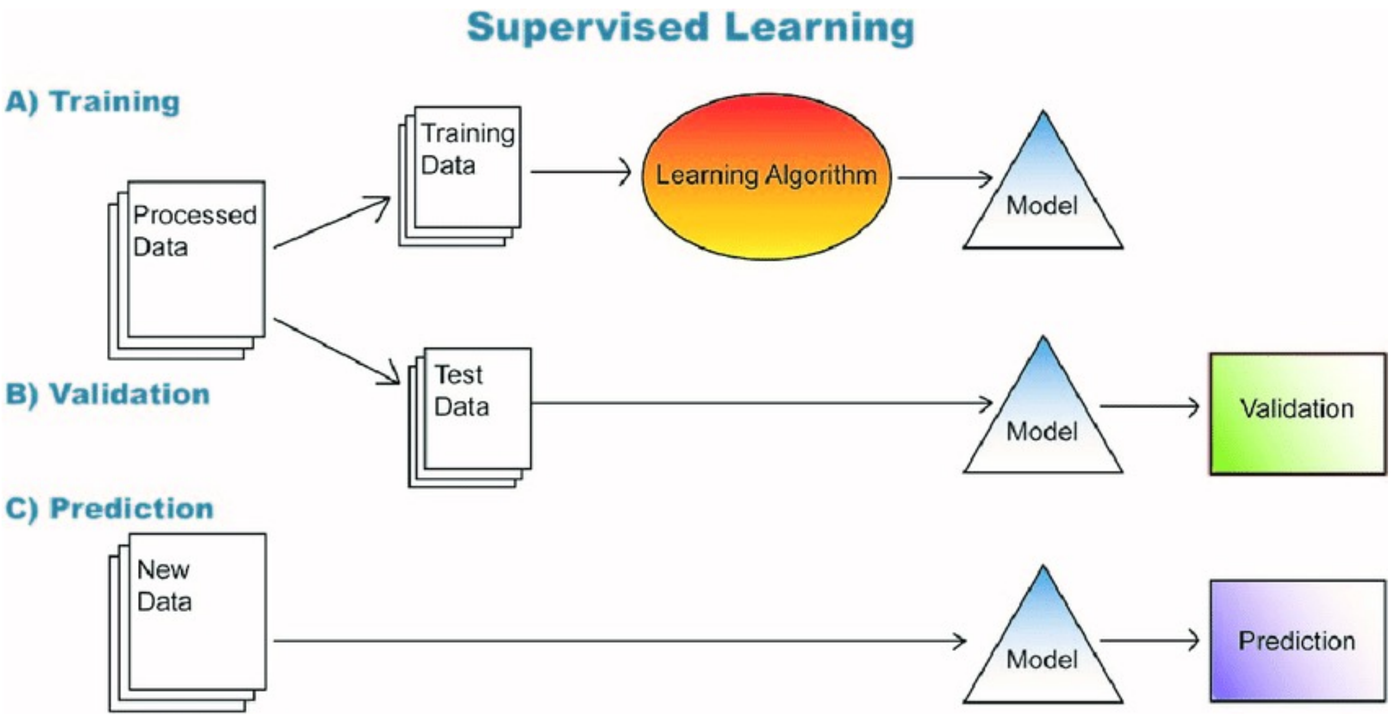
[pic4] : Benefits of Predictive Maintenance



[pic5] : Cost vs availability, different maintenance approaches [MPR18]

Machine Learning in PdM

- **Supervised Learning**
  - Needs the availability of *labelled* dataset S $\mathcal{S} = \{x_i, y_i\}_{i=1}^{n}$
  - Classification and Regression



[pic6] Supervised Learning steps

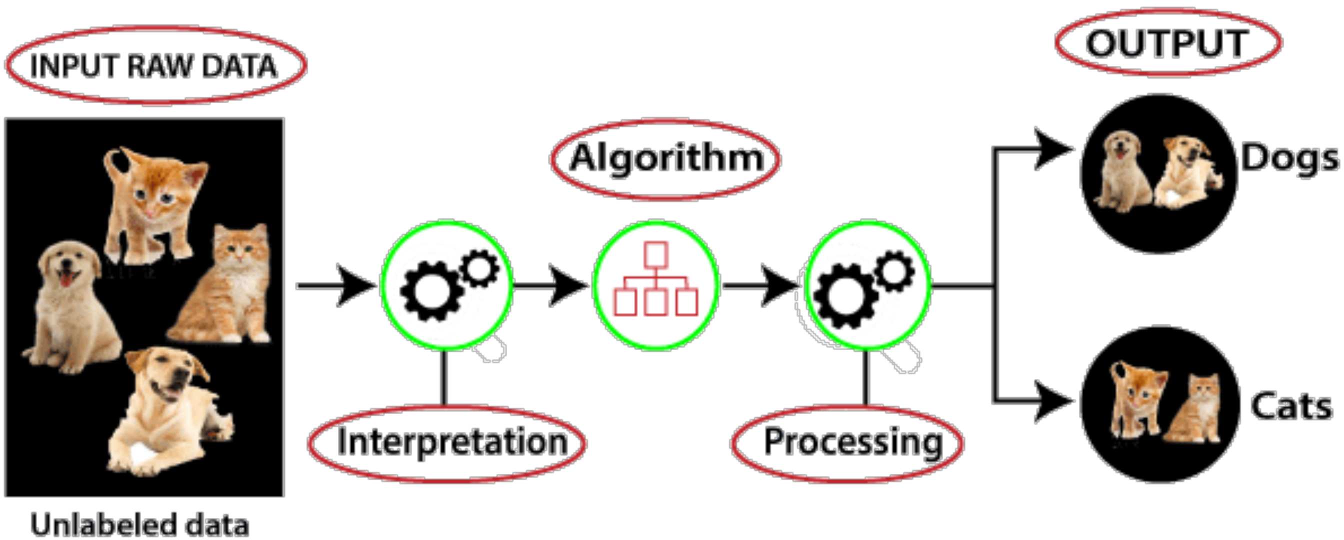
Supervised Machine Learning in PdM

- **Classification:**
  - Algorithm for accurately assigning input data into a specific class.
  - SVMs – Support Vector Machines, K-NN – K Nearest neighbours

- **Regression:**
  - Algorithm for understanding the relationship be- tween different dependent and independent entities
  - Linear Regression, Logistic Regression, Polynomial regression
  - ***Remaining Useful Lifetime (RuL):*** duration that a machine can run further before it faces a breakdown, provided no maintenance activity was performed. [BR10]

UnSupervised Learning in PdM

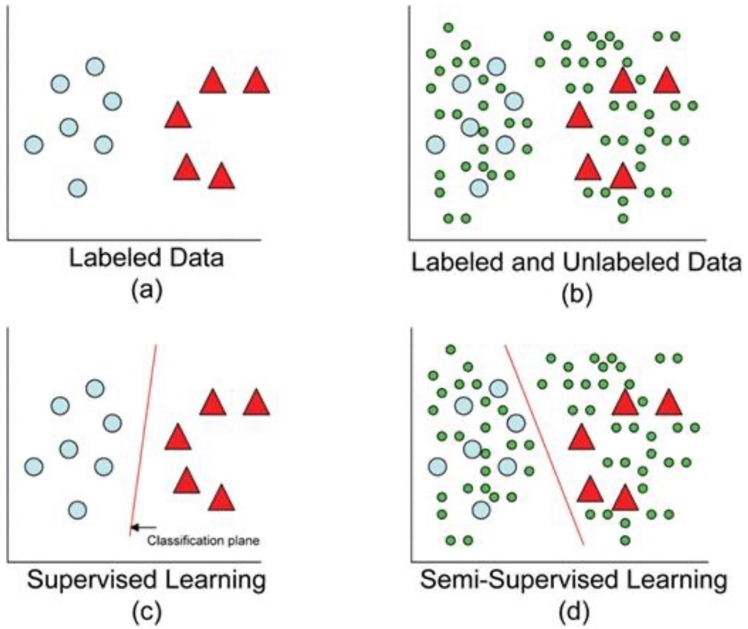
- **UnSupervised Learning:**
  - Data is given as input and the algorithm finds hidden structures in the data
  - useful in *clustering* problems and probability density estimation
  - complex compared to supervised models.
  - Need more CPU cores and processing powers.



[pic7] Unsupervised Learning steps

Semi Supervised Learning

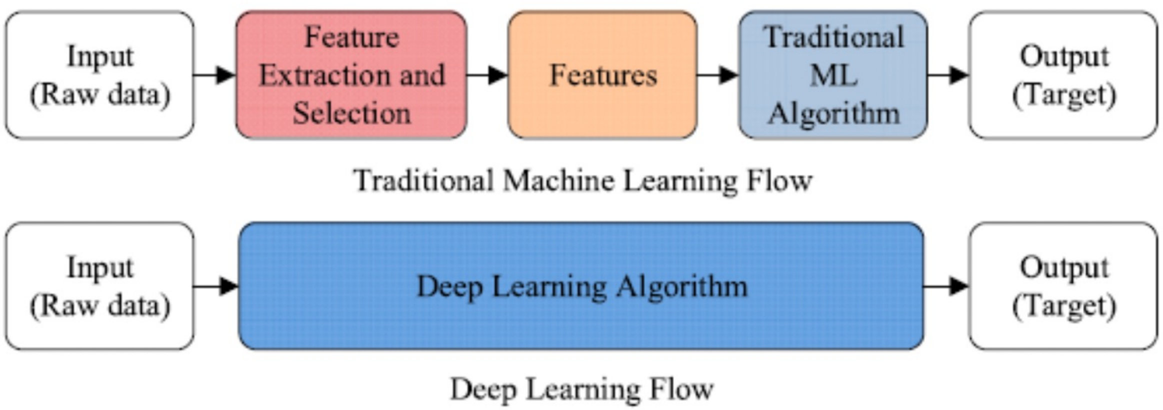
- **Practical Scenario**
  - Data for a particular state is known and most data for the complete system is not known or labelled
  - predict a target value for a specific input data set and can detect *outliers.*
  - needs proper knowledge about normal production behaviour.
  - useful in *Anomaly Detection* and One-Class classification problems.



[pic8] Difference between supervised and semi supervised learning methods

Machine Learning Flow

- **Data Exploration:**
  - Sensor data, Logs, Monitoring data
- **Feature Extraction:**
  - Data Interpolation
  - mathematical value that can be used to train the machine learning algorithm
  - Manufacturing industry – thousands of possible features.
  - **Feature Selection** is important – Average, standard deviation, skewness, Kurtosis etc.



[pic9] Flow of Ml and DL [ZYW19]

Multiple Classifier (MC) in PdM

- **Classification of PdM:** [ZYW19]
  - Assume data of N maintenance cycles available
  - Define matrix X containing all the collectible information
  - Information on maintenance events is contained in the variable Y

$$n = \sum_{i=1}^{N} n_i$$

$$X = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T$$

$$Y(i) = y_i = \begin{cases} \text{F}, & \text{if iteration } i \text{ is faulty} \\ \text{NF}, & \text{if iteration } i \text{ is not faulty.} \end{cases}$$

- **Weak Form:** [ZYW19]
  - In the case of R2F, each cycle results in F - Dataset is unbalanced and skewed.
  - Only current process iteration is classified. No fault prevention
  - No operating cost optimization policy
  - Key Metrics –

    Frequency of unexpected breaks (FuB) - *percentage of failures not prevented*

    Amount of unexploited lifetime (FuL) – *number of iteration that could have run before breakdown*

MC PdM Concept

- Costs, CuB and CuL associated with for FuB anf FuL; CuB relates to unplanned interruptions in Production. $CuB >> CuL$
- an optimal tradeoff solution is sought through minimization of the total operating costs $J = FuB.CuB + FuL.CuL$
- Possible approach to prevent unexpected failures.
- Label last m iterations as F instead of considering just last failed cycle and reduce *skewness* of Dataset.
- The j^{th} classifier is associated with the labels Y^j, where considering the t^{th} process iteration of the i^{th} maintenance cycle (of length n_i)

$$y_t^{(j)} = \begin{cases} \text{NF}, & \text{if } t \leq n_i - m^{(j)} \\ \text{F}, & \text{otherwise} \end{cases}$$

Tradeoff Optimization

- N_m samples for F and n-N_m samples for the class NF. This can be repeated with k different values of horizon *m.*
- given the current costs CUB(t) and CUL(t) at time t, the MC PdM suggests a corrective action when the j^{th} classifier

$$\begin{aligned} j^* &= \arg\min_{j=1,\ldots,k} J^{(j)}(t) \\ &= \arg\min_{j=1,\ldots,k} \rho_{\mathsf{UB}}^{(j)} c_{\mathsf{UB}}(t) + \rho_{\mathsf{UL}}^{(j)} c_{\mathsf{UL}}(t) \end{aligned}$$

- **Algorithm 1:** [ZYW19]

- In the performance of each classifier is evaluated using repeated random subsampling validation, also known as *Monte Carlo cross validation (MCCV)*
- *MCR[%]* = Percentage of misclassified samples
- Tuning by cross validation with MCR

Mc PdM Algorithm

- CuB and CuL are provided by the user, and may be changed at each evaluation of the PdM module
  - Unexpected breakdown brings full maintenance cycle
  - Valuable new data are available to update the MC PdM module and related performance metrics

- It is important to retain the facility to update the MC PdM module following deployment.

Algorithm 1. MC PdM Training

Data: $X, Y, k, q_1, q_2, Q_1, Q_2, \{m^{(j)}\}_{j=1}^{k}$

Result: Classifiers $\{f^{(j)}\}_{j=1}^{k}$ and PdM performances

1. Let $\rho_{\mathrm{UB}} = [\,\cdot\,]$ and $\rho_{\mathrm{UL}} = [\,\cdot\,]$ (empty vectors)

for $j = 1\,to\,k$ **do**

  2. Compute $Y^{(j)}$ as in Eq. (5) **for** $i = 1\,to\,Q_1$ **do**

    3. Randomly split the maintenance cycles between training and validation samples, keeping the ratio q_1

    4. Compute the MCR of the classifier with different hyper-parameters

  5. Chose the hyper-parameters based on the averaged MCR over the Q_1 simulations

  6. Compute $f^{(j)}$ with the selected hyper-parameters

  for $i = 1\,to\,Q_2$ **do**

    7. Randomly split the maintenance cycles between training and validation, keeping the ratio q_2

    8. Compute $f^{(j)}$ with the selected hyper-parameters

    9. Compute ρ_{UB} and ρ_{UL} of $f^{(j)}$

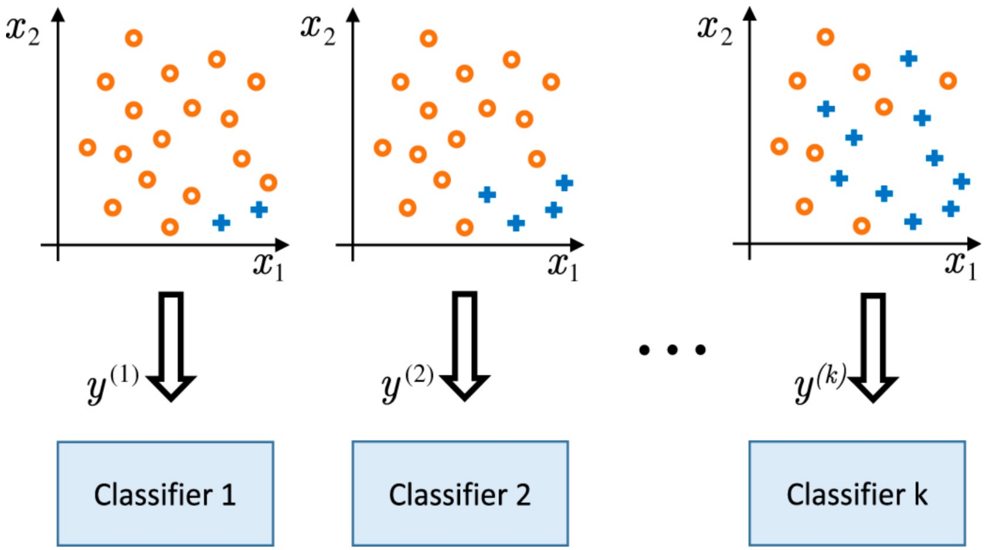
  10. Compute ρ^*_{UB} and ρ^*_{UL} as averaged over the Q_2 simulations

  11. $\rho_{\mathrm{UB}} = [\rho_{\mathrm{UB}}; \rho^*_{\mathrm{UB}}]$ and $\rho_{\mathrm{UL}} = [\rho_{\mathrm{UL}}; \rho^*_{\mathrm{UL}}]$

K Classifiers

- Performance of the MC PdM methodology increases with the number of classifiers k

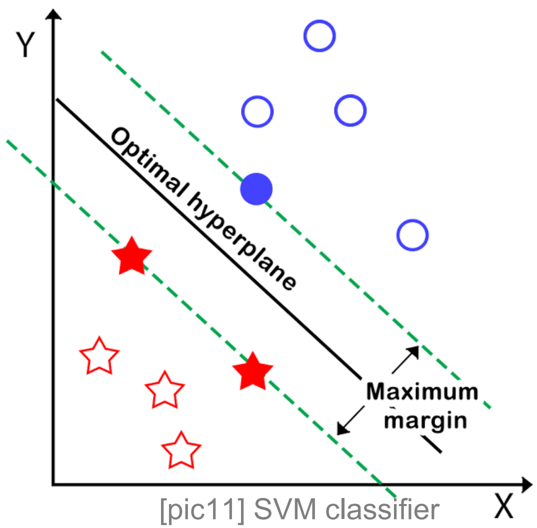
- Define the failure horizon, depends on the nature of the fault under investigation



[pic10] Two-dimensional example of k classifiers. [ZYW19]

Classification Algorithms

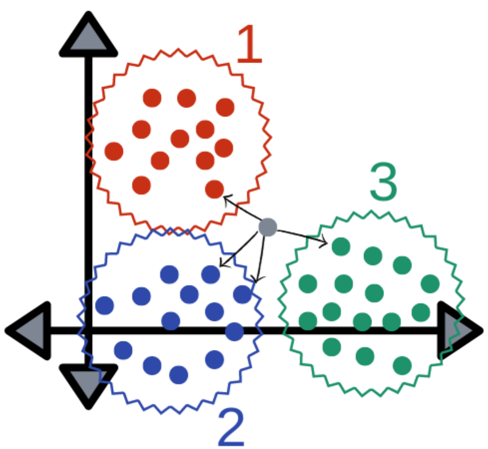
- **Support Vector Machines (SVMs)**
  - A *hyperplane* is defined such that it separates the two classes F and NF with the highest margin.
  - In real classification problems, classes are often non separable, i.e., the two categories overlap
  - With the liberty of wrong classification of some data points to the other class, SVMs can still be used for classification.
  - The computational cost for training a nonlinear SVM is generally between $\boldsymbol{O(n2)}$ **and** $\boldsymbol{O(n3)}$



[pic11] SVM classifier

Classification Algorithms

- **K Nearest Neighbours (k-NN)**
  - requires just computation of distances between samples.
  - each point of the input space is labelled according to the labels of its k closest neighbouring samples
  - choice of k is usually data-driven (often decided though cross-validation
  - larger values of k reduce the effect of noise on the classification
  - optimized algorithms for computing k-NN have a computational cost of $O(\log n)$.



[pic12] k-NN algorithm

Use Case and Experiment

- **Semi conductor Manufacturing**
  - replacing tungsten filaments used in ion implantation
  - Every time a filament is changed, the tool is down for approx- imately 3 h making this the most important maintenance issue
  - high values of pressure, voltage, and filament current can drastically reduce the lifetime of the filament

- **Data Description**
  - N = 33 maintenance cycles and total of n = 3671 batches
  - A total of 31 variables having have a time-series evolution during each run
  - extracted six features for each of the 31 time series
    maximum, minimum, average, variance, skewness, kurtosis

Experiment Results

- **Semi conductor Manufacturing**
- PdM and MC PdM outperform PvM approaches
- For FuL, MC PdM classifiers achieve similar performance to the PdM approaches
- In case of FuB, MC PdM-svm consistently achieved the best results
- MC PdM-svm guarantees better performance than MC PdM-knn but is more time consuming

Challanges and Future

- Data is crutial, Data exploration is very difficult in manufacturing industry
- MC PdM is available only for labelled data.

- Using Unsupervised learning methods such as Deep Neural Networks
- e-Maintenance, remote maintenance
- Prescriptive Maintenance – Automated decision making, self healing systems
- Sustainable and Ecological growth

Conclusion

- PdM avoids unnecessary scheduled maintenance of healthy systems unlike Preventive maintenance approaches.
- PdM targets surgical maintenance and help reducing downtimes and optimize operational cost.
- Multiple ML classifiers work in parallel to exploit the knowledge at each process iteration in order to enhance decision making.
- MC PdM module is robust to variations in the operating costs associated with unexpected breaks and unexploited equipment part lifetime.
- k classifiers enhance process understanding and allow cost aware maintenance management decisions.
- SVMs offer superior performance to k-NN classifiers.
- Choice of the fault horizon m strongly affects the performance of the corresponding classifier in MC PdM.

Thank you

References

- *Papers-*
  - **[SSP+14]** Susto et.al. Machine learning for predictive maintenance: A multiple classifier approach, pages 812– 820, 2014
  - **[ZYW19]** Zhang et.al. Data-driven methods for predictive maintenance of industrial equipment: A survey , 2019
  - **[MPR18]** Motaghare et.al. Predictive mainte- nance architecture, pages 1–4. IEEE, 2018
  - **[Mob02]** R Keith Mobley. An introduction to predictive maintenance. Elsevier, 2002.

- *Images-*
  - [pic0] https://images.app.goo.gl/6KdrSqQNHngmQ2an9
  - [pic1] https://images.app.goo.gl/PnQtCRahUS7sjKNj9
  - [pic2] https://www.pinterest.com/pin/411446115928434072/
  - [pic3] https://images.app.goo.gl/Y6DwVDsqh6odgQjJ6
  - [pic4] https://images.app.goo.gl/vFGHvDcHgu2KGsxd6
  - [pic6] https://images.app.goo.gl/ibrZFqd914G6kCS28
  - [pic7] https://images.app.goo.gl/2CAtAQSptQTaNzA5A
  - [pic8] https://images.app.goo.gl/ggYKYPoVqJxuTdUbA
  - [pic11] https://images.app.goo.gl/VbYcLdTVmMjvhxGv7
  - [pic12] https://images.app.goo.gl/LW8bnJFEHLvBax2q6