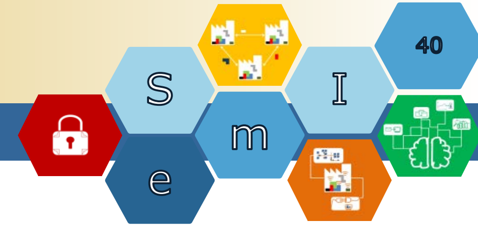
A Review of Related Work on Machine Learning in Semiconductor Manufacturing and Assembly Lines

DI Darko Stanisavljevic | VIRTUAL VEHICLE
DI Michael Spitzer | VIRTUAL VEHICLE

i-KNOW'16
18.-19.10.2016, Graz

The project SemI40 is co-funded by grants from Austria, Germany, Italy, France, Portugal and the ENIAC Joint Undertaking and is coordinated by Infineon Technologies Austria AG

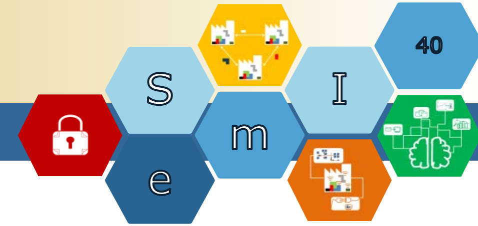
Agenda



- Introduction
- Motivation Machine Learning
- Examples of **semiconductor manufacturing process** and production on **automated assembly lines**
- Methodology and gained insights
- Conclusion

ECSEL Joint Undertaking
Electronic Components and Systems for European Leadership

Introduction



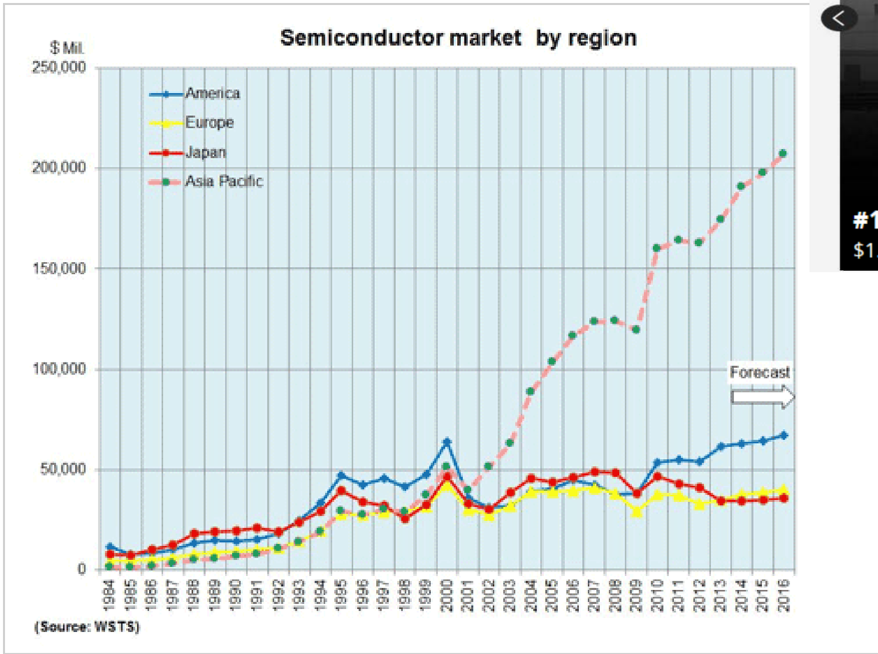
- Focus on:
  - semiconductor manufacturing and
  - production on automated assembly lines.

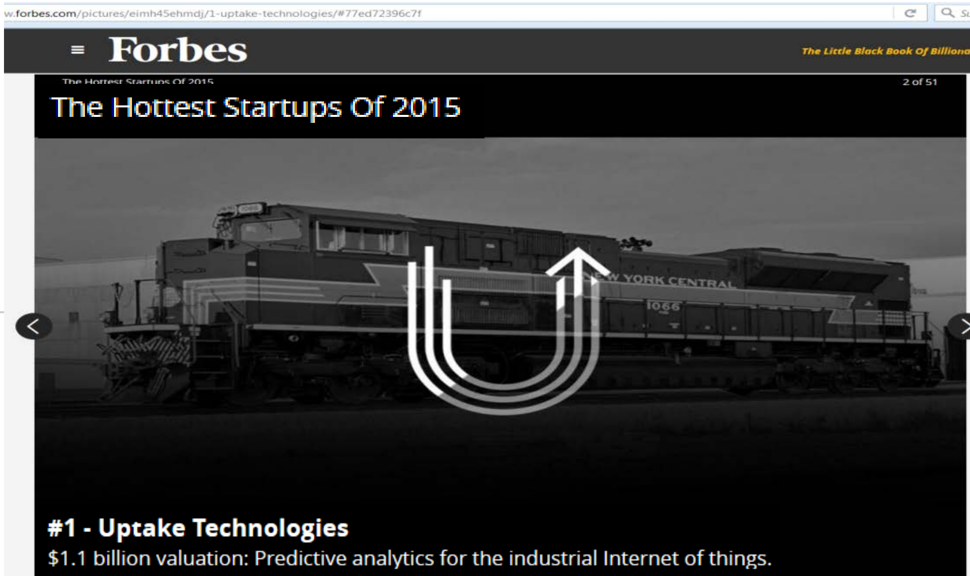
- Goals:
  - Show through examples that the application of machine learning in manufacturing can lead to increased productivity and decreased production costs.
  - Gain insight in machine learning challenges

ECSEL Joint Undertaking
Electronic Components and Systems for European Leadership

Motivation

- $400 billion worth semiconductor industry
- 73.9 million vehicles produced worldwide in 2015
- Increased volume and complexity of production data



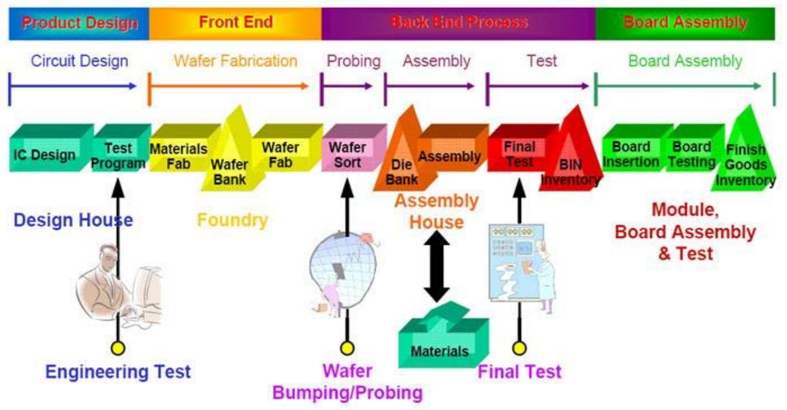


- Machine Learning applications:
  - Virtual Metrology
  - Fault Detection
  - Predictive Maintenance
  - Root Cause Analysis

https://www.semiconportal.com/en/industryinfo/industry-at-a-glance.html

Semiconductor manufacturing example

- Highly complex process
  - Production is based on wafers; wafers are organized in lots
  - Hundreds (even thousands) production steps
    - Etching
    - Lithography
    - Chemical Vapor Deposition (CVD) …
  - The quality of the process is assessed by measuring one or more parameters on wafer (for CVD it is the thickness of the deposited layer)
  - Common practice to save time/money is to test only one wafer from a lot thus resulting in incomplete measurement/test data -> hard to find faults in non-assessed wafers



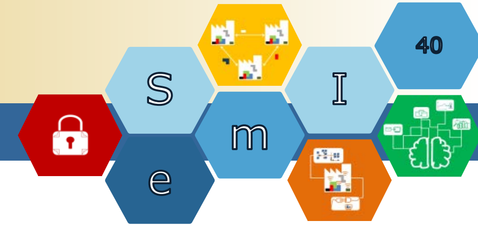
http://www.li-sion.com.tw/lision/upfile/2012210164428.jpg

Solution in Virtual Metrology :

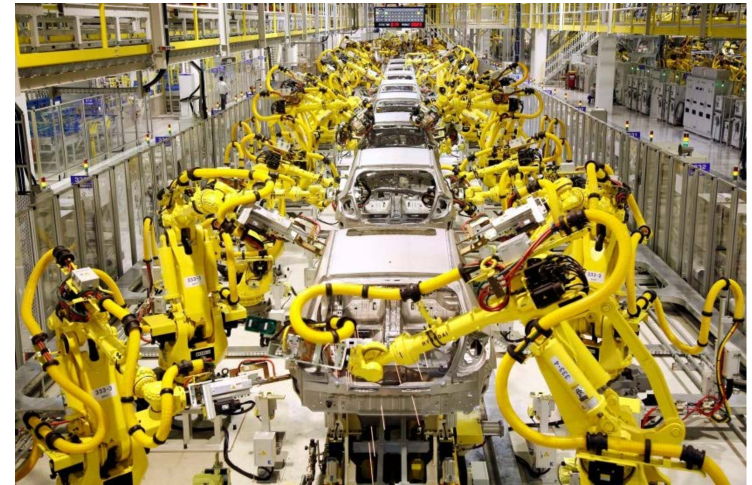
- Exploit the data from production tools for non-assessed wafers (temperature, pressure, …) to estimate the quality of wafer
- This way, at least the estimated quality of the wafer is available
- Example: CVD layer thickness prediction

ECSEL Joint Undertaking
Electronic Components and Systems for European Leadership

Assembly line example (Car-body AL)



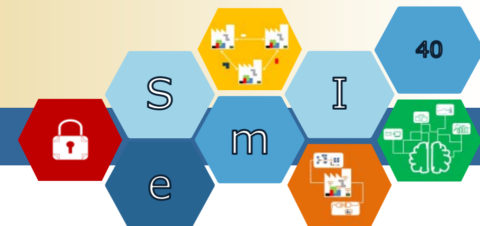
- Complex process
  - Car-body assembly line consists of a large number of consecutive workstations
  - Different quality assessment tools are installed
    - Optical Coordinate Measurement Machines (OCMMs)
    - State of the robotic equipment ...
  - OCMMs allow dimension measurement of all produced car-body assemblies
  - Production data generated after every workstation
  - Data can be used for fault detection, and when faults are detected, root cause analysis can be done



http://www.sme.org/MEMagazine/Article.aspx?id=2147483816&taxid=3440

ECSEL Joint Undertaking
Electronic Components and Systems for European Leadership

Methodology



- Methodology

- Searching for papers published in prominent conferences
  - ICML (International Conference on Machine Learning)
  - KDD (Conference on Data Discovery and Data Mining)
  - NIPS (Conference on Neural Information Processing Systems)

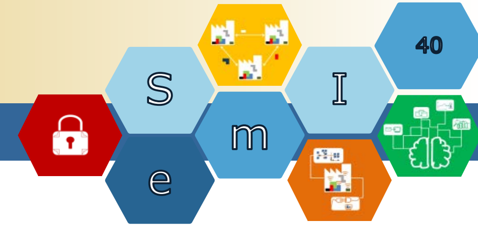
- Searching for articles in journals:
  - JMLR (Journal of Machine Learning)
  - JAIR (Journal of Artificial Intelligence Research)

- Not too many papers of interest found by this method
  - Search for papers of interest in ACM and IEEE libraries
    - still insufficient number of papers describing fault detection in production on assembly lines (with ML)

http://www.sme.org/MEMagazine/Article.aspx?id=2147483816&taxid=3440

ECSEL Joint Undertaking
Electronic Components and Systems for European Leadership

Gained insights (1)



Challenges

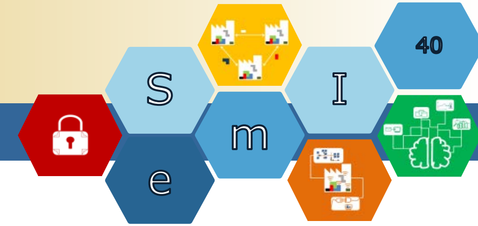
- acquisition of manufacturing data
  - availability of data (lack of data capturing capability, security concerns)
  - quality of data (metadata available?, outliers, redundancy)

- high dimensionality of manufacturing data
  - ML algorithms are able *to handle high-dimensionality* of the data
    - some ML algorithms are designed to work well with such data (e.g. SVM)
    - otherwise dimensionality reduction (e.g. PCA)

- insufficient transparency in manufacturing process
  - the major advantage of ML algorithms is *to discover formerly unknown (implicit) knowledge and to identify implicit relations in data sets*
  - extracted patterns (knowledge) can be used by process owners as a decision making support or for automatic system improvement

http://www.sme.org/MEMagazine/Article.aspx?id=2147483816&taxid=3440

Gained insights (2)

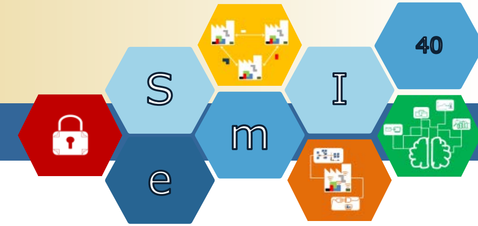


- data preprocessing (critical impact on the result)
  - normalizing, balancing, filtering of data
  - replacing missing values
- selection of ML algorithm based on
  - available data (labeled?, unlabeled?, expert-knowledge available?)
  - general applicability of algorithm (high-dimensionality problem)
  - previous application of ML algorithm on similar problems
- Interpretation of results
  - format, visualisation of the result
  - parameters, settings of used ML algorithm
  - data including its preprocessing

http://www.sme.org/MEMagazine/Article.aspx?id=2147483816&taxid=3440

ECSEL Joint Undertaking
Electronic Components and Systems for European Leadership

Conclusion

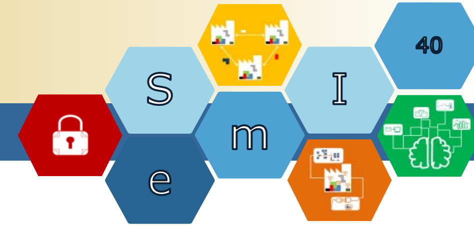


- Plenty of papers on machine learning in semiconductor manufacturing
- Lack of papers on machine learning in production on assembly lines (assembly line load optimization excluded)
- Most of the reviewed papers provide almost none concrete information about the data set used
- In general, it is hard to obtain real production data
  - mainly because of the security issues
  - sometimes the data capturing ability on production machines is not available

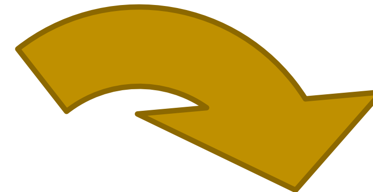
http://www.sme.org/MEMagazine/Article.aspx?id=2147483816&taxid=3440

ECSEL Joint Undertaking
Electronic Components and Systems for European Leadership

UseCase Understanding (Example)



Production Line

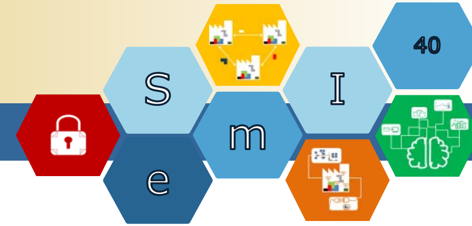


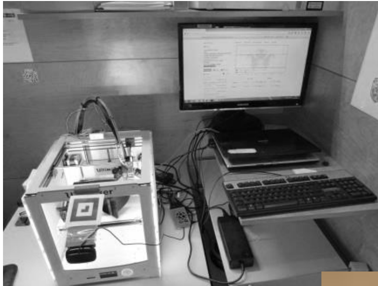
Generated data

adc_1	adc_2	adc_3	adc_4	adc_5	adc_6	adc_7	adc_8
-0,10291539	-0,90486263	-0,69228403	-0,97892402	-0,84420413	-1,57017771	2235,95995	28,6999994
-0,12437945	-0,83609674	-0,72347857	-0,97555038	-0,83693934	-1,57137108	2239,86995	28,6799994
-0,12218427	-0,78831028	-0,75398499	-0,97868304	-0,83648529	-1,56015342	2226,08995	28,7999994
-0,11194005	-0,81838078	-0,72990097	-0,97699622	-0,833761	-1,5594374	2238,83995	28,6899994
-0,11706216	-0,8288705	-0,64778594	-0,97747817	-0,83080967	-1,5513225	2197,88995	29,0599994

https://en.wikipedia.org/wiki/Ultimaker

„Machine Learning Lab“

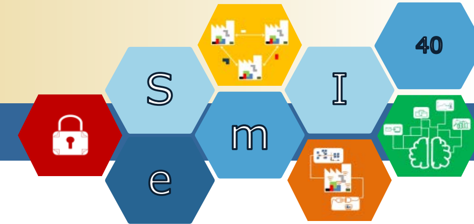




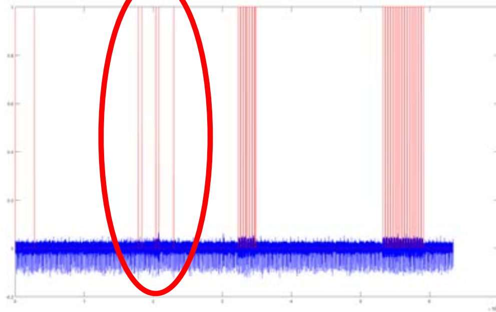
Machine Learning – Simplified production example

Detection of vibrations induced on printing plate of the 3D printer

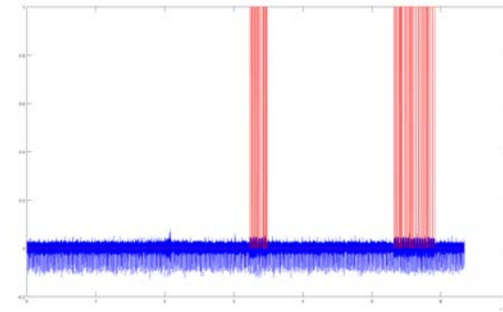
Ergebnisse – 3D Drucker



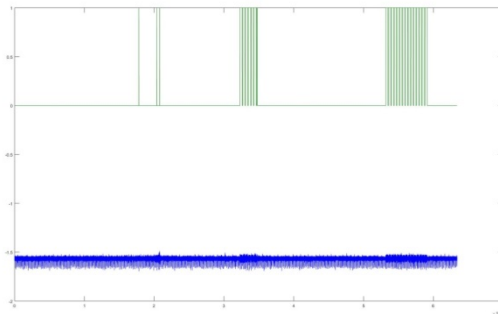
C4.5 (J48)



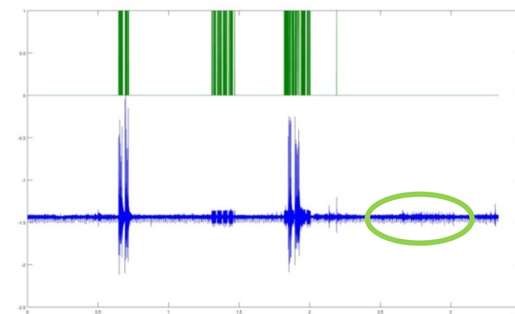
Nearest neighbour



Random Forest



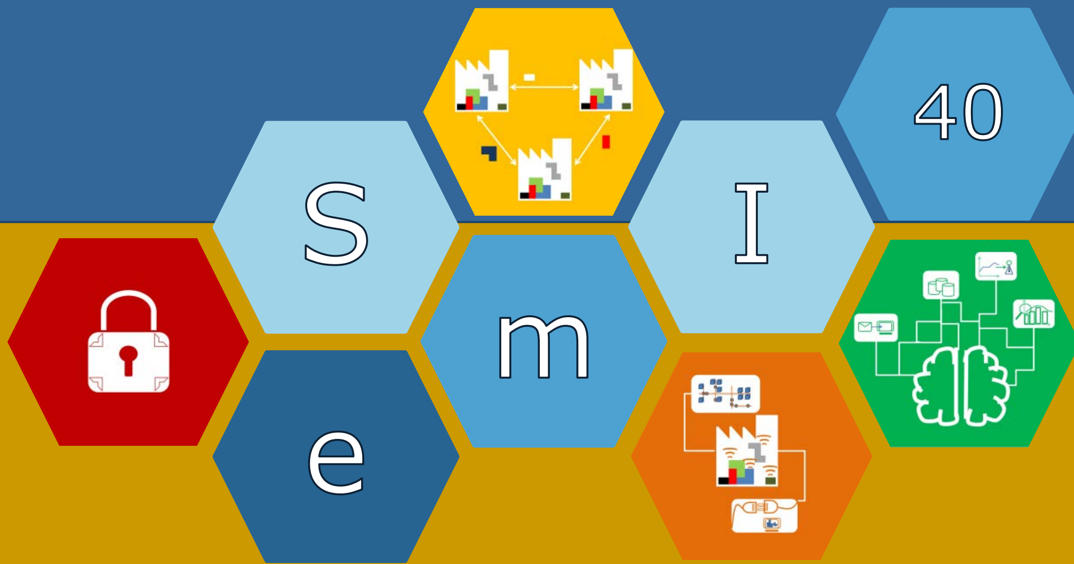
K Means Clustering



- Classification vs. Clustering
- False-positives / False-negatives

ECSEL Joint Undertaking
Electronic Components and Systems for European Leadership

Thank you!



Contact:
DI Darko Stanisavljevic
Junior Researcher at Virtual Vehicle Research Center, Graz
darko.stanisavljevic@v2c2.at

The project SemI40 is co-funded by grants from Austria, Germany, Italy, France, Portugal and the ENIAC Joint Undertaking and is coordinated by Infineon Technologies Austria AG