2024 SNU Cryptography Seminar

Feb 02, 2024

Industrial AI Technology in Manufacturing

Sunghee Yun

VP, AI Technology & Product Strategy @ Erudio Bio, Inc.

Today

1 Why Manufacturing AI?

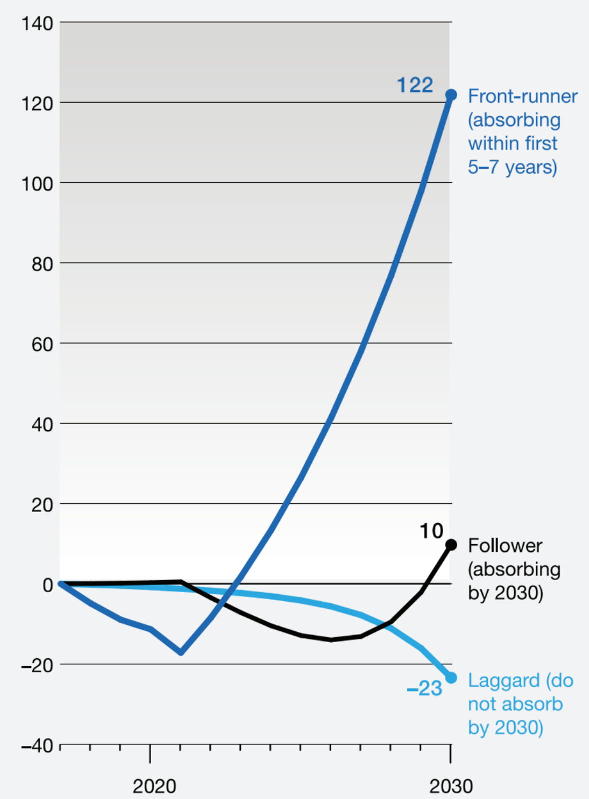
2 Computer vision ML for manufacturing

3 Time-series ML for manufacturing

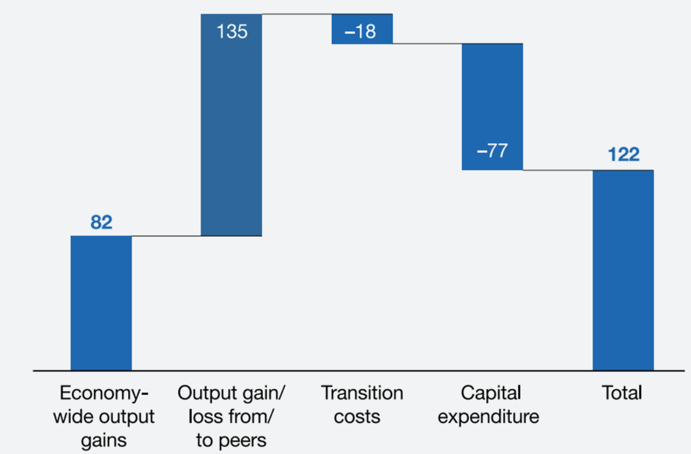
4 Difficulties with time-series ML in manufacturing

Fast AI adoption WILL create **way larger economic gains**

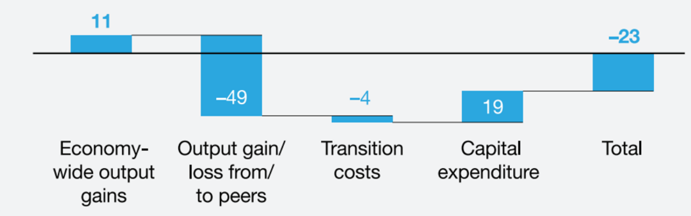
Relative changes in cash flow by AI-adoption cohort, cumulative % change per cohort



Front-runner breakdown, % change per cohort



Laggard breakdown, % change per cohort



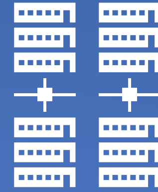
Note: Numbers are simulated figures to provide directional perspectives rather than forecasts.

* Source: McKinsey Global Institute Analysis (2019)

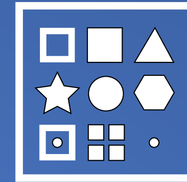
Data Characteristics

Virtuous (or Vicious) Cycle

Data-centric AI



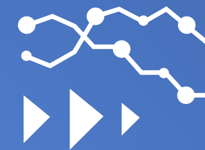
Volume



Variety

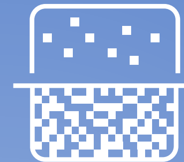
Velocity

FatData



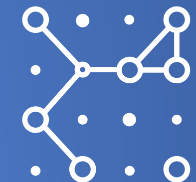
Shift/drift

Imbalance



Quality

Nonlinearity

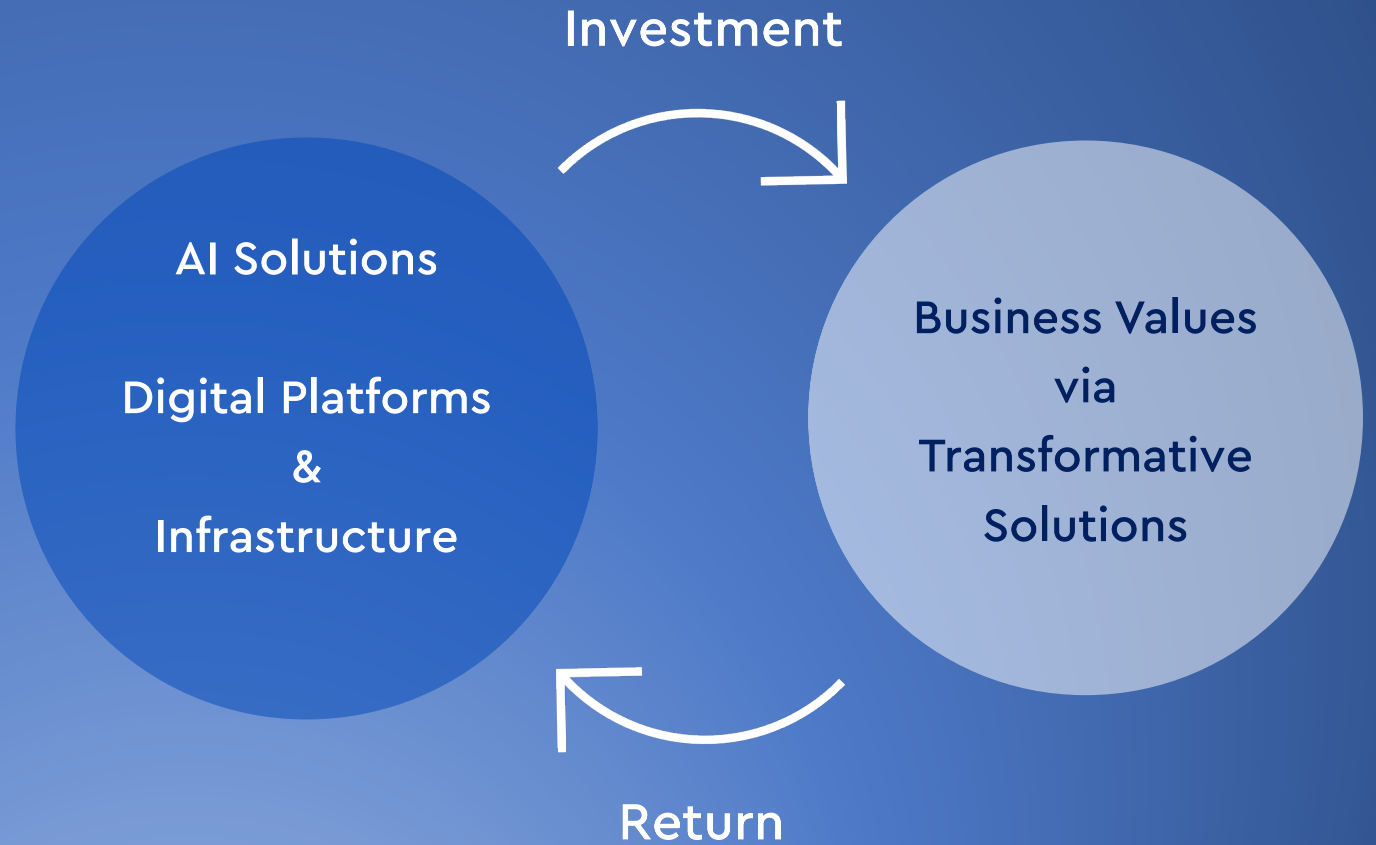


Complexity

Data Characteristics

Virtuous (or Vicious) Cycle

Data-centric AI



"We need 1,000 models for 1,000 problems" – Andrew Ng

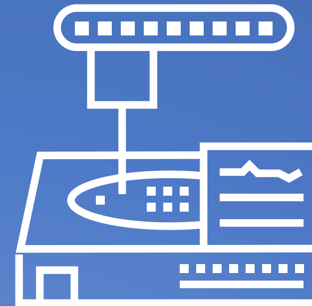
Data-centric AI

Discipline of systematically engineering the data used to build an AI system

Data Characteristics

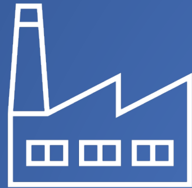
Virtuous (or Vicious) Cycle

Data-centric AI



Every company or sector has its own problems

Our initial focus for 10x changes*

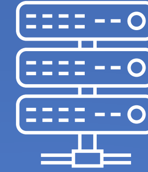


Semiconductor Fab

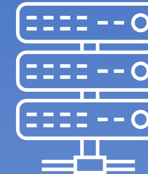
A modern mega fab has …

- ~1,000 process equipement
- ~100 metrology equipment
- ~1,000 wafers per day, per equipment
- ~1,000 processes per wafer
- 3-6-month cycle time

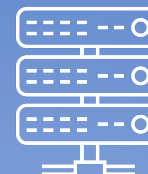
Servers and Systems



Equipment Sensor Data
(~1M parameters, ~1Tb/day)



Metrology Image Data
(~1M images, ~10 Tb/day)



Manufacturing Execution Data
(~10M events, ~10 Gb/day)

Why Semiconductor?

Data availability from advanced digitalization

Diverse and sophisticated processes, ideal for expanding to new customers & sectors

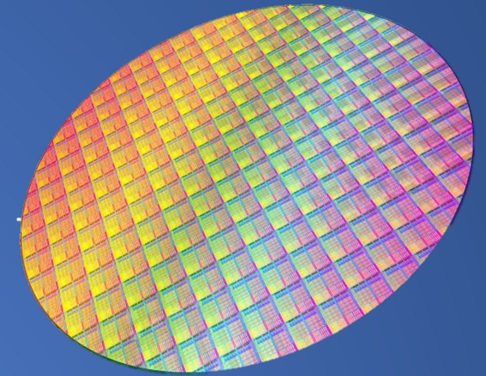
Huge impact even within the sector itself

Computer vision and time-series ML in Manufacturing

lots of image data to measure and inspect

Scanning electron microscope (SEM) images, transmission electron microscope (TEM) images, etc.

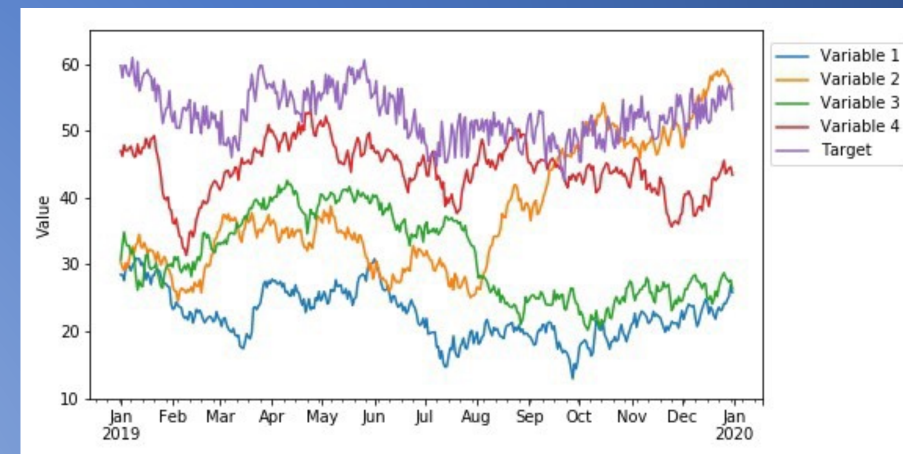
→ Image pattern classification, image anomaly detection, defect inspection, etc.



(almost) All the data coming from manufacturing are time-series data

Equipment sensor data, process times, material measurement, *etc*.

→ time-series (TS) regression / prediction/estimation, TS anomaly detection, etc.



Computer Vision ML for manufacturing

Metrology
Measurement of critical features

Inspection
Anomaly detection, localization and classification

Image courtesy of ASML

Scanning Electron Microscope*

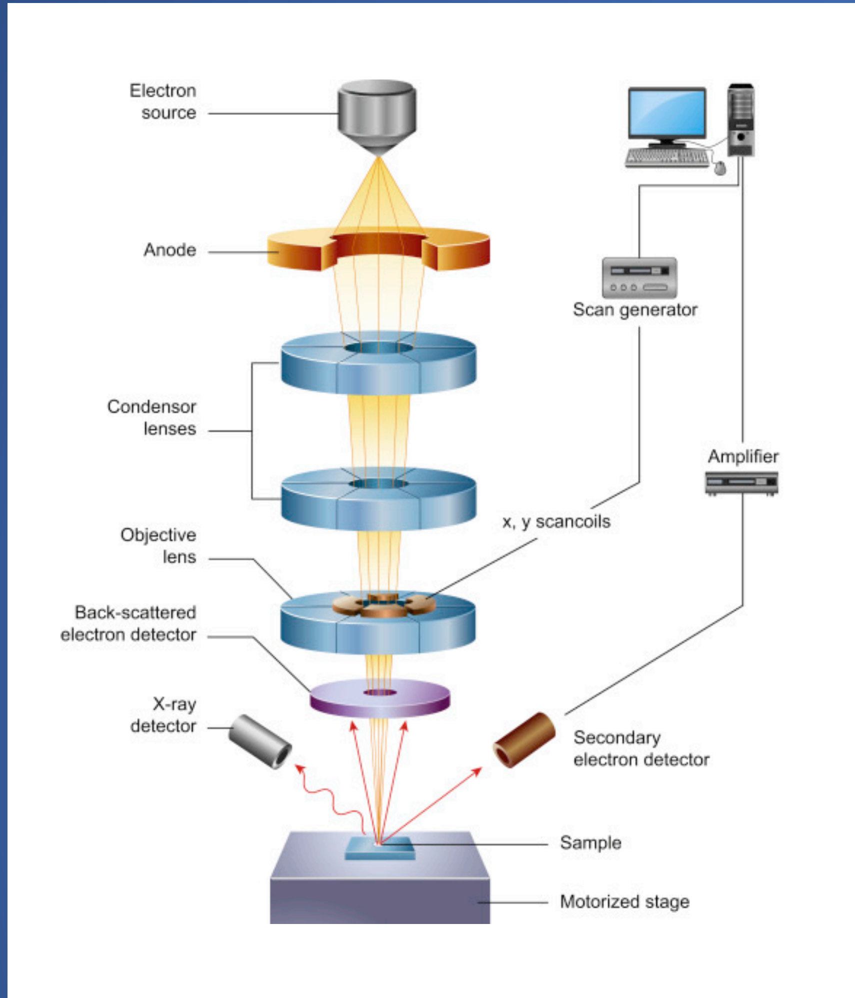
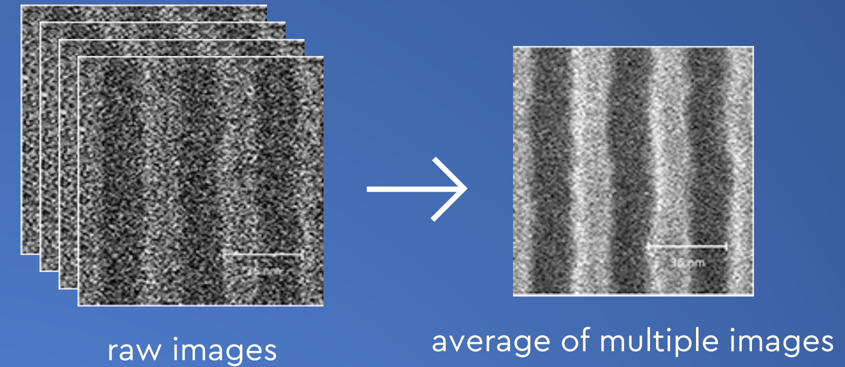


Image courtesy of https://www.sciencedirect.com/science/article/pii/B978008100040300002X



Shot Noise Image courtesy of https://en.wikipedia.org/wiki/Shot_noise

Image restoration

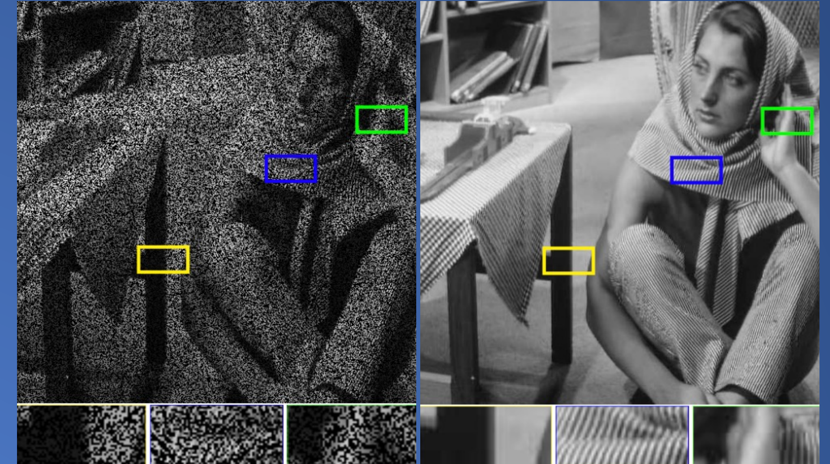
Inverse problem of image corruption

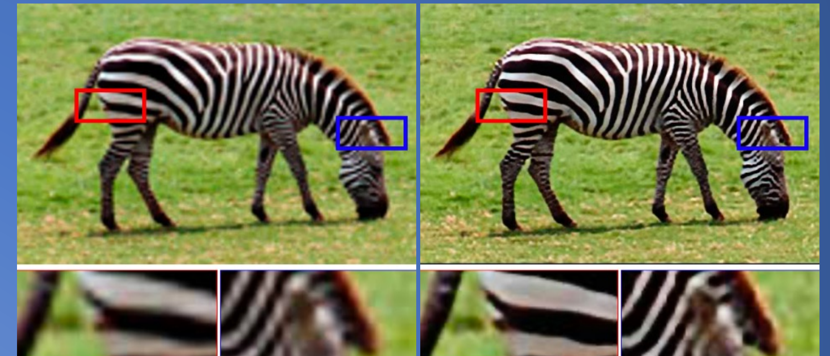
$x = f(y) + n$

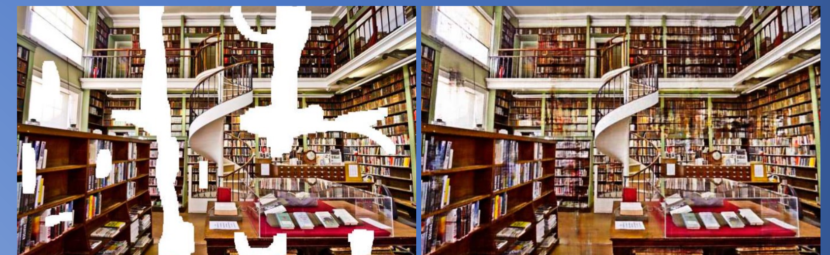
- y: clean image
- x: corrupted image
- n: noise

f(·) and corresponding solutions

- Noising: Identity function → Denoising
- Downsampling → Super-resolution
- Missing pixels → Inpainting

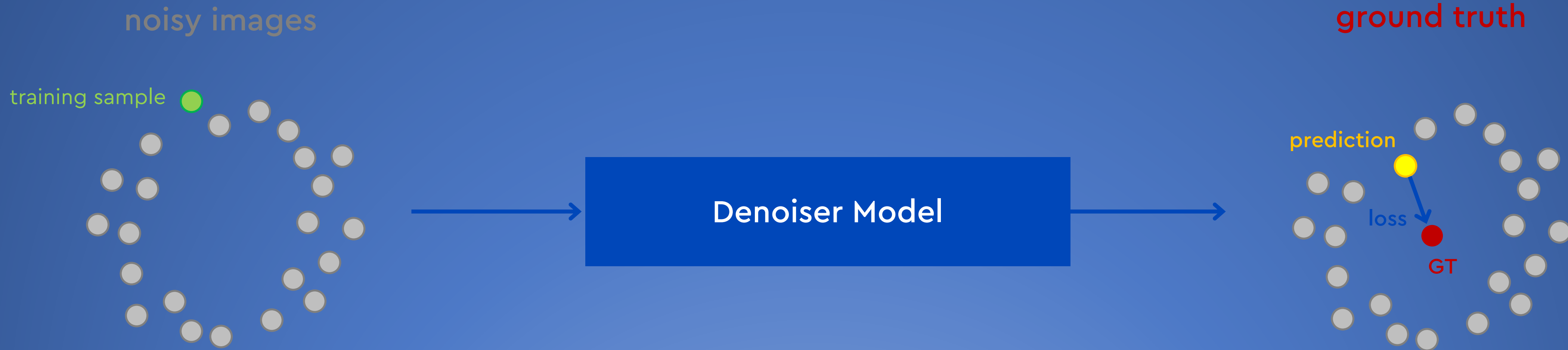






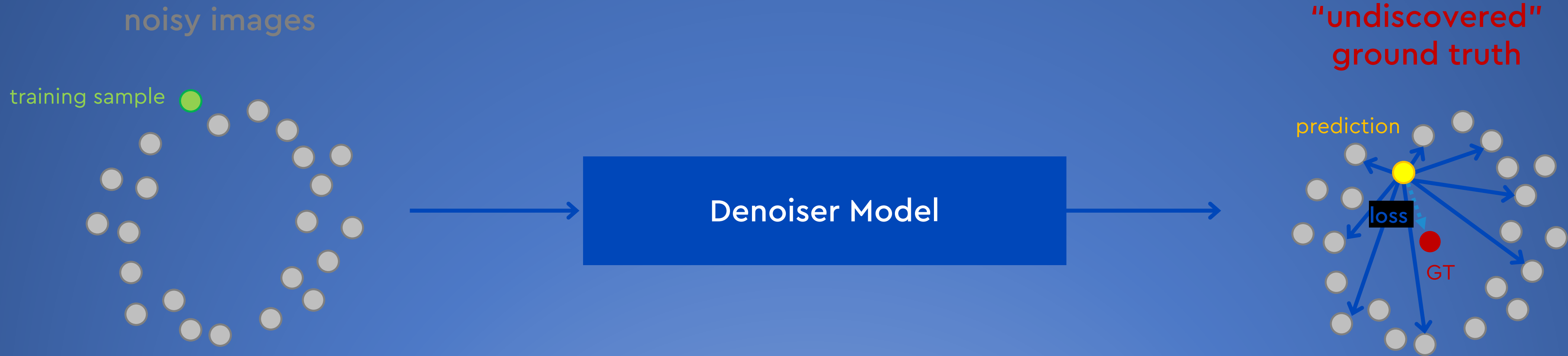
Dmitry Ulyanov et al. Deep Image Prior. CVPR, 2018

Supervised image denoising



However, it is not possible to acquire ground-truth images from SEM device, in practice.

Blind denoising without ground truth



If the mean of the noise is zero, the average of the gradients that model takes is same with the gradient to the ground truth

J. Lehtinen, et al. Noise2Noise: Learning Image Restoration without Clean Data. ICML, 2018.

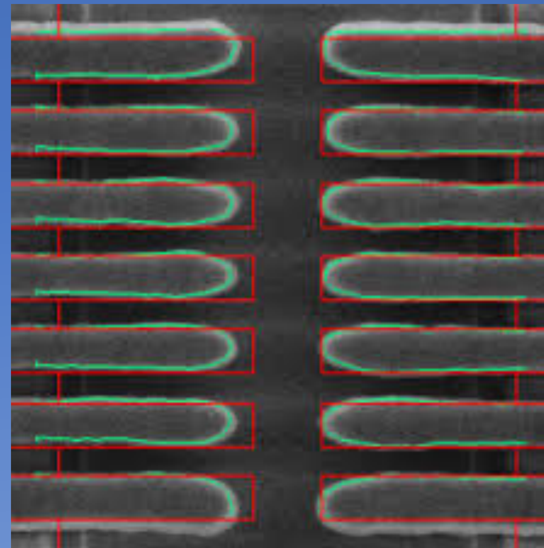
Metrology based on segmentation and pattern recognition

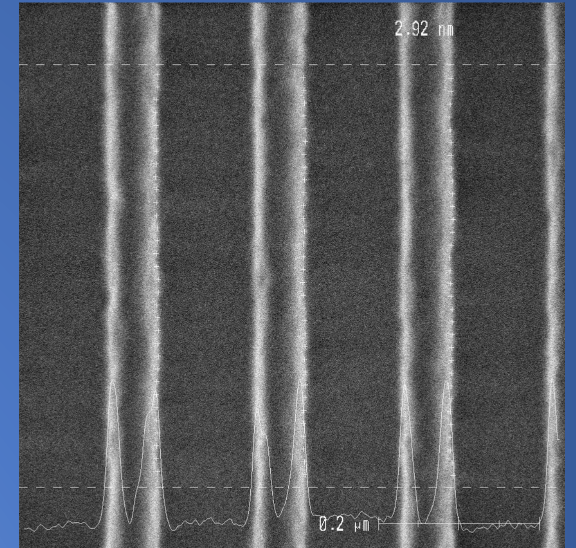
Investment

Automatic measurement of critical dimensions

Approaches

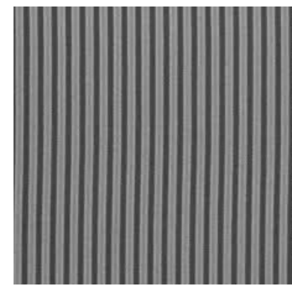
- Unsupervised texture segmentation
- Repetitive pattern recognition

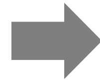


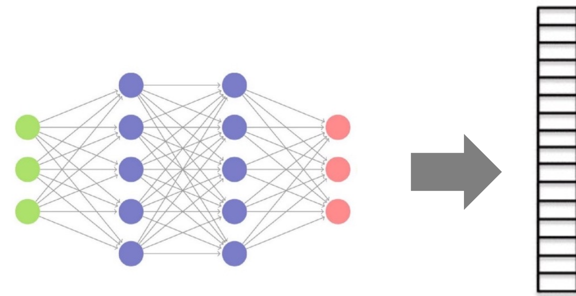


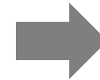
<0.1 nm measurement precision guaranteed

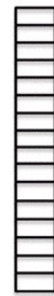
Anomaly detection in unsupervised learning*


input image

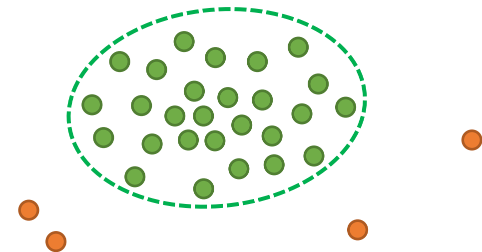



pretrained NN

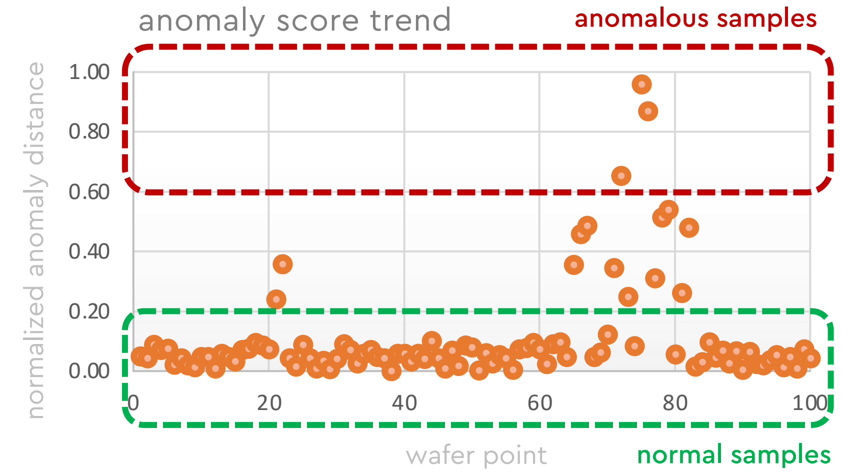



feature vector

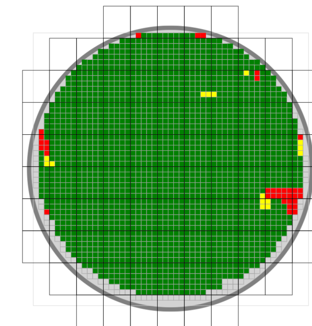
distance to the
normal vector cluster in kNN


precomputed in
unsupervised learning

vectors of normal images
vectors of anomalous images



anomaly score
wafer map



AI-accelerated metrology system

SEM images
(grayscale)

Image Enhancement

Enhanced images

CD Measurement

CDs

Repetitive Pattern Recognition

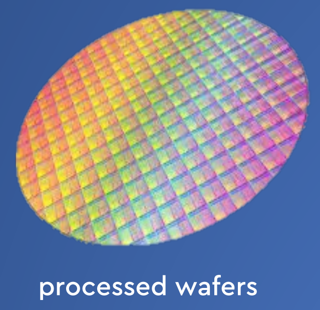
Segmentation

Anomaly Detection

anormalous

alarm

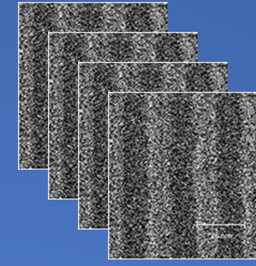
Automatic measurement for semiconductor manufacturing



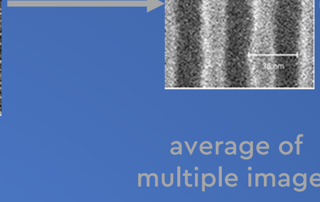
processed wafers

Scanning Electron Microscope (SEM)

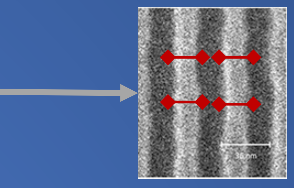
image capture



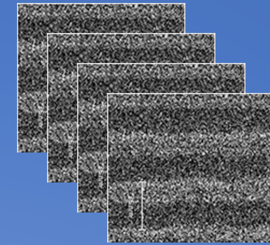
raw images

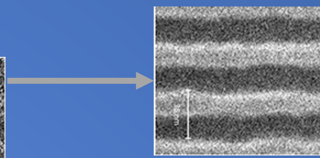


average of multiple images



measure





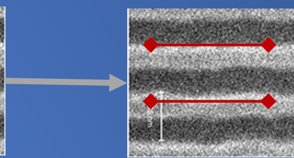
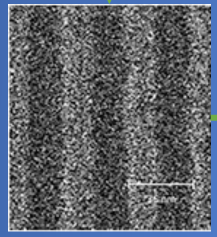
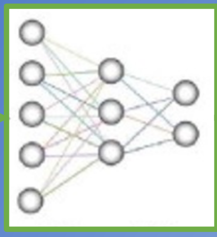


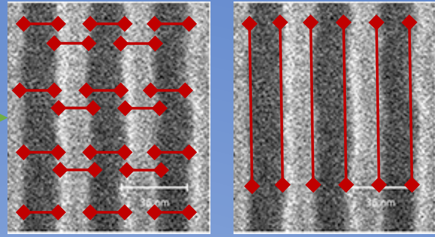
image capture



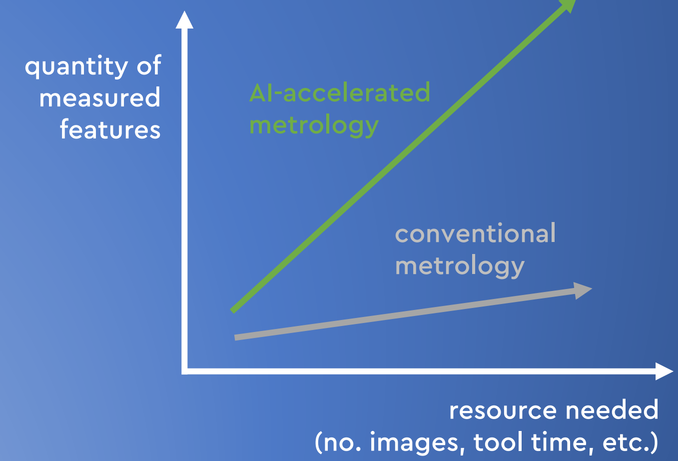
raw image



AI metrology algorithms



automatic measurement in en hanced image

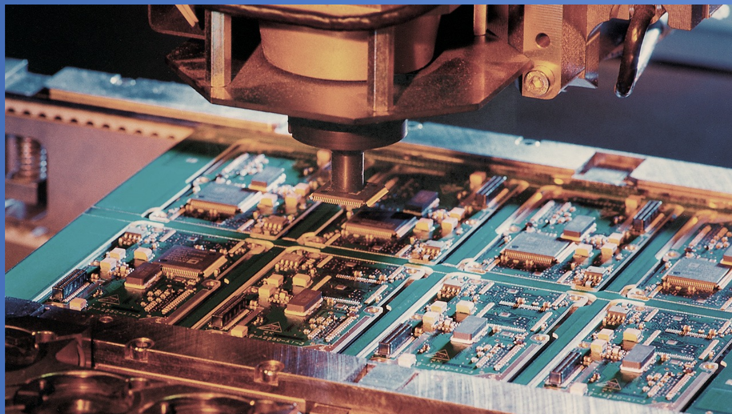


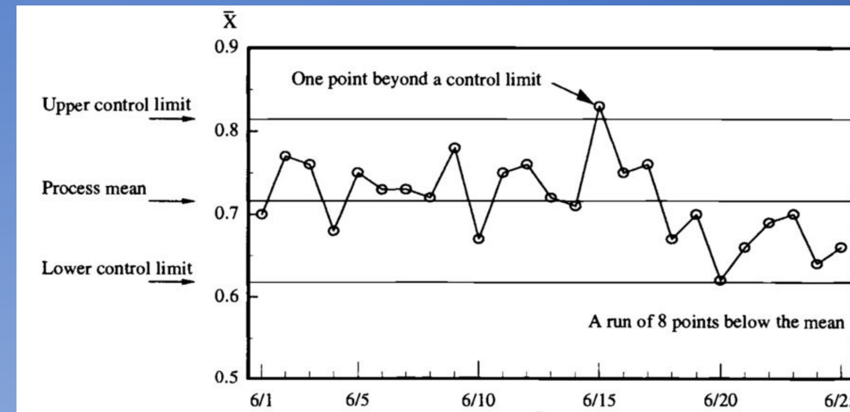
Time-series ML for manufacturing

Why time-series ML?

manufacturing application is about one of the followings:

- prediction of time-series values - virtual metrology, yield prediction
- classification of time-series values - equipment anomaly alarm generation
- anomaly detection on time-series data - root cause analysis, yield analysis





Time-series regression/prediction/estimation

- virtual metrology

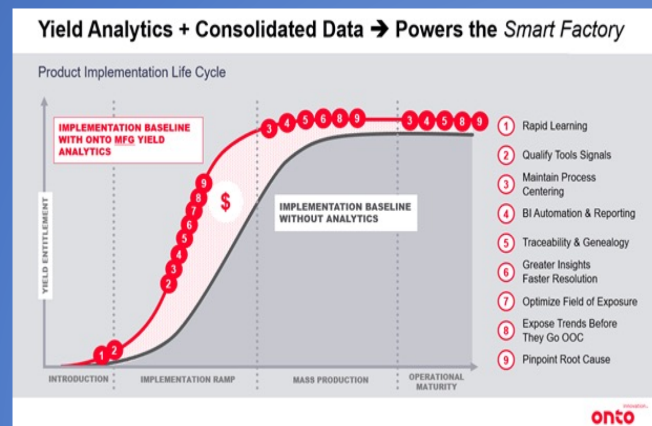
  - *measure unmeasured processed materials using equipment sensor signals*

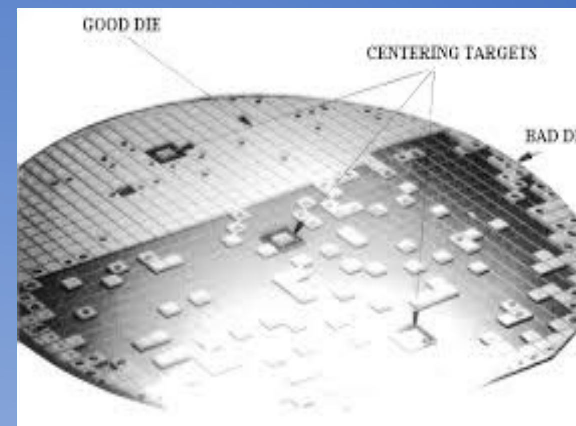
  - *save investment on measurement equipment, downstream applications such as process control, statistical process control, yield improvement*

- yield prediction

  - *predict yield (# working dies / # total dies*

  - *better product quality and larger profit, business impact*

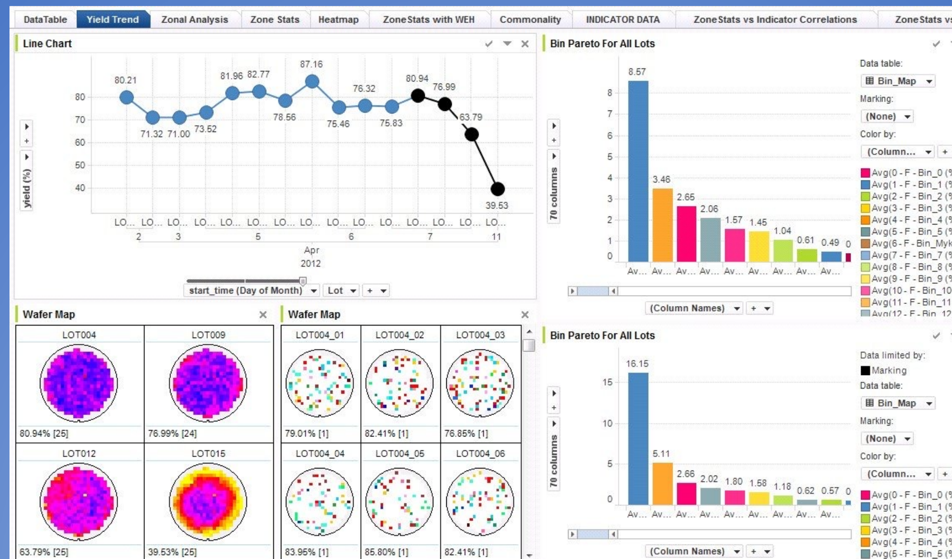




Root cause analysis using time-series anomaly detection*

- equipment alarm root cause analysis
  - *when alarm goes off, find responsible equipment and root causes*
  - *reduce equipment downtime, make process engineers' lives easier*

- yield analysis
  - *find responsible equipment and root causes for yield drop*
  - *a few % yield improvement brings profit increase of tens of millions of dollars!*



Difficulties with Time-series ML in manufacturing

Data challenges

- covariate shift & concept drift

  $p(x(t_k), x(t_{k-1}), \ldots)$ *changes over time*

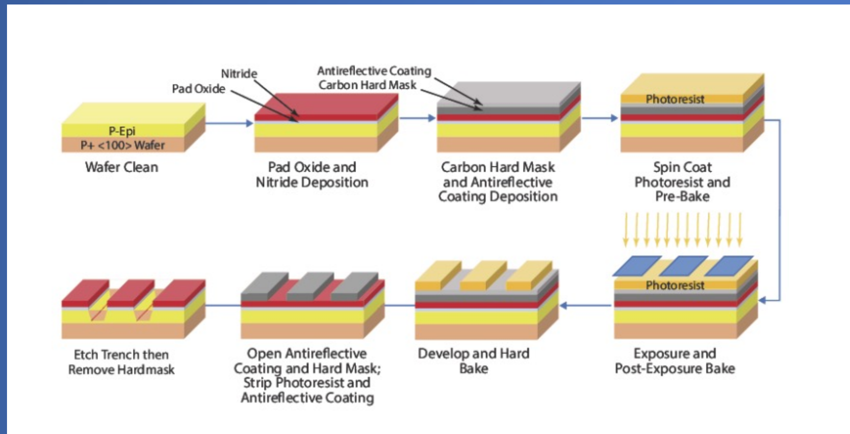
  $p(y(t_k) \mid x(t_k), x(t_{k-1}), \ldots, y(t_{k-1}), y(t_{k-2}), \ldots)$ *changes over time*

- fat data, *i.e.*, # features way larger than # data
- poor data quality; missing values, anomalies, wrong formats
- huge volume of data to process
- multi-modality - different types of data

Domain knowledge and fully home-grown algorithms

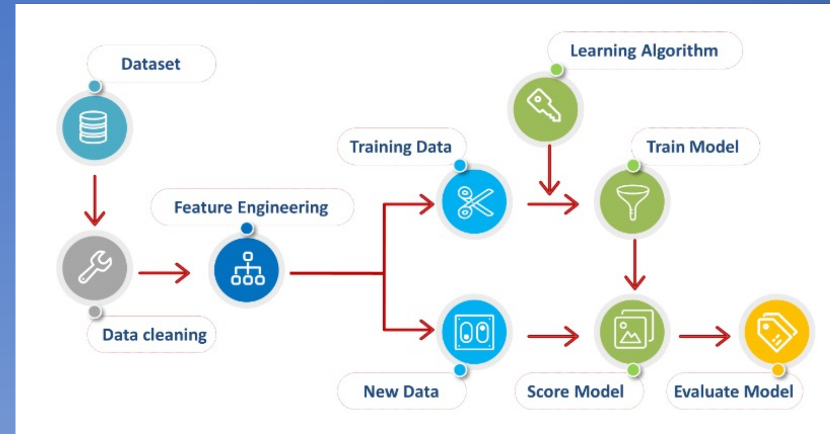
in most cases,
domain knowledge is critical!

close collaboration with customers required



off-the-shelf algorithms
not working!

developing fully customized algorithms needed



Virtual Metrology (VM)

What is VM?

in many cases, we cannot measure all processed materials

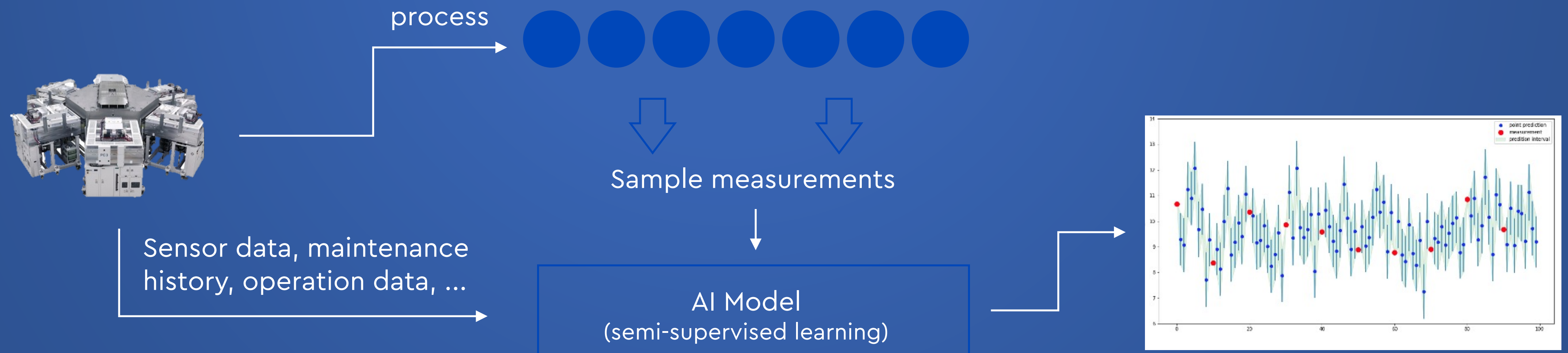
- measurement equipment is too expensive
- measuring every materials makes production slow inducing low throughput

thus, we do sampling (with very low sampling rate)

- in semiconductor manufacturing line, average sampling rate is less than 5%

PROBLEM

- predict the measurement of unmeasured material using indirect signals
- sensor data, maintenance history, operation data, . . .



Business Impact made by VM

To the best of our knowledge

- no organization has even been successful with VM

Gauss Labs VM

- uses online learning to cope with data drift/shift
- RMSE comparable to measurement equipment precision
- also predicts uncertainty of predictions - providing prediction reliability information

VM implications

- measuring ALL wafers – equivalent to investing on 100x measurement equipment
- enables optimal re-allocation of limited measurement resources

Conclusion

supervised and unsupervised ML everywhere in industrial AI applications

lots of challenges

- data challenge, domain knowledge required, need for customizing algorithms

huge changes potentially made via various applications

Impacts

- Tens of Millions of dollars by 1% yield increase
- 100x measurement equipment save by VM