ANOMALY DETECTION IN MANUFACTURING SYSTEMS BASED ON DIGITAL TWIN AND MACHINE LEARNING

JOURNÉES DE LA SAGIP – CT IMS² (INTELLIGENT MANUFACTURING & SERVICES SYSTEMS)

Farah Abdoune

Nantes Université

2nd year PhD student

Olivier Cardin, Maroua Nouiri, Pierre Castagna

DIGITAL TWIN PARADIGM

- (*Liu et al., 2018*) describe DT as "a **living model** that **continually adapts** to **change** in the environment or operation using **real-time data** and can **forecast** the future of the corresponding **physical asset's behavior**"
- According to (*Kaur et al., 2020*), a digital twin's main feature is its ability to **correctly replicate** physical space using a combination of **physics-based** and **data-driven models**
- DT can check the consistency of **monitored data**, perform data mining to **detect** existing and **forecast** upcoming problems, using **Artificial Intelligence** to support effective **decisions**. (*Asimov et al., 2018*).

Liu, Zheng, Norbert Meyendorf, et Nezih Mrad. 2018. « The role of data fusion in predictive maintenance using digital twin ». AIP Conference Proceedings 1949

Kaur, Maninder Jeet, Ved P. Mishra, et Piyush Maheshwari. 2020. « The Convergence of Digital Twin, IoT, and Machine Learning: Transforming Data into Action ». In Digital Twin Technologies and Smart Cities

Asimov, R., S Chernoshey, Ingmar Kruse, et Vital Asipovich. 2018. DIGITAL TWIN IN THE ANALYSIS OF A BIG DATA

DIGITAL TWIN SERVICES

Considering the nine services listed by (*Tao et al., 2018*), they can be grouped in these categories:

- Real-time state **monitoring;**
- Energy consumption analysis;
- Failure analysis and prediction, and maintenance strategy;
- Intelligent optimization and update;
- User behavior analysis and user operation guide;
- Product virtual maintenance and product virtual operations;

Tao, Fei, Jiangfeng Cheng, Qinglin Qi, Meng Zhang, He Zhang, et Fangyuan Sui. 2018. « Digital Twin-Driven Product Design, Manufacturing and Service with Big Data ». The International Journal of Advanced Manufacturing Technology 94 (9-12)

MONITORING DEFINITION FROM PHM POV

- According to (*Saez et al. 2020*), Process monitoring aims to :
  - Demonstrate the **current status** of the system;
  - Show **undesired or unallowed deviations** from the acceptable behavior➔ **Fault/ anomaly !!**
- Necessity to detect and diagnose these faults ➔ **potential premature failure**
- **Fault/ anomaly detection and diagnosis scheme:**
  - **Detection**: Detect malfunctions in real-time
  - **Isolation**: Isolate the component to find the root cause
  - **Identification**: Estimate the size and type of the fault

Saez, Miguel A., Francisco P. Maturana, Kira Barton, et Dawn M. Tilbury. 2020. « Context-Sensitive Modeling and Analysis of Cyber-Physical Manufacturing Systems for Anomaly Detection and Diagnosis ». IEEE Transactions on Automation Science and Engineering 17

DIFFERENT APPROACHES OF ANOMALY DETECTION

Model based

Verify the consistency between observations and expected system behavior

i.e. Observers, Kalman filter, parity equations…etc.

Data driven

Learn the system behavior pattern based on historical data

i.e. ANN, PCA, SVM, RF…etc.

Knowledge based

The assessment of online monitored data according to a set of rules specified by expert knowledge

i.e. Expert system, Fuzzy logic

LITERATURE REVIEW

Aims of study

- Summarize the current research and application of DT in Fault detection and diagnosis
- Evaluate the limitation of DT in FDD
- Provide the potential ideas of future study

Sources

- Academic databases
  - ScienceDirect
  - IEEE Xplore
  - MDPI
  - Web Of Science
  - Google Scholar

Search criteria

- Keyword : « Digital twin » & « fault diagnosis» | « fault detection » | « anomaly detection»
- Retrieval scope : Literature title and Abstract
- Languages of publication: English
- Date of publication: from 2018 to 2022

Documents Types

- Peer-viewed journal article
- Conference paper

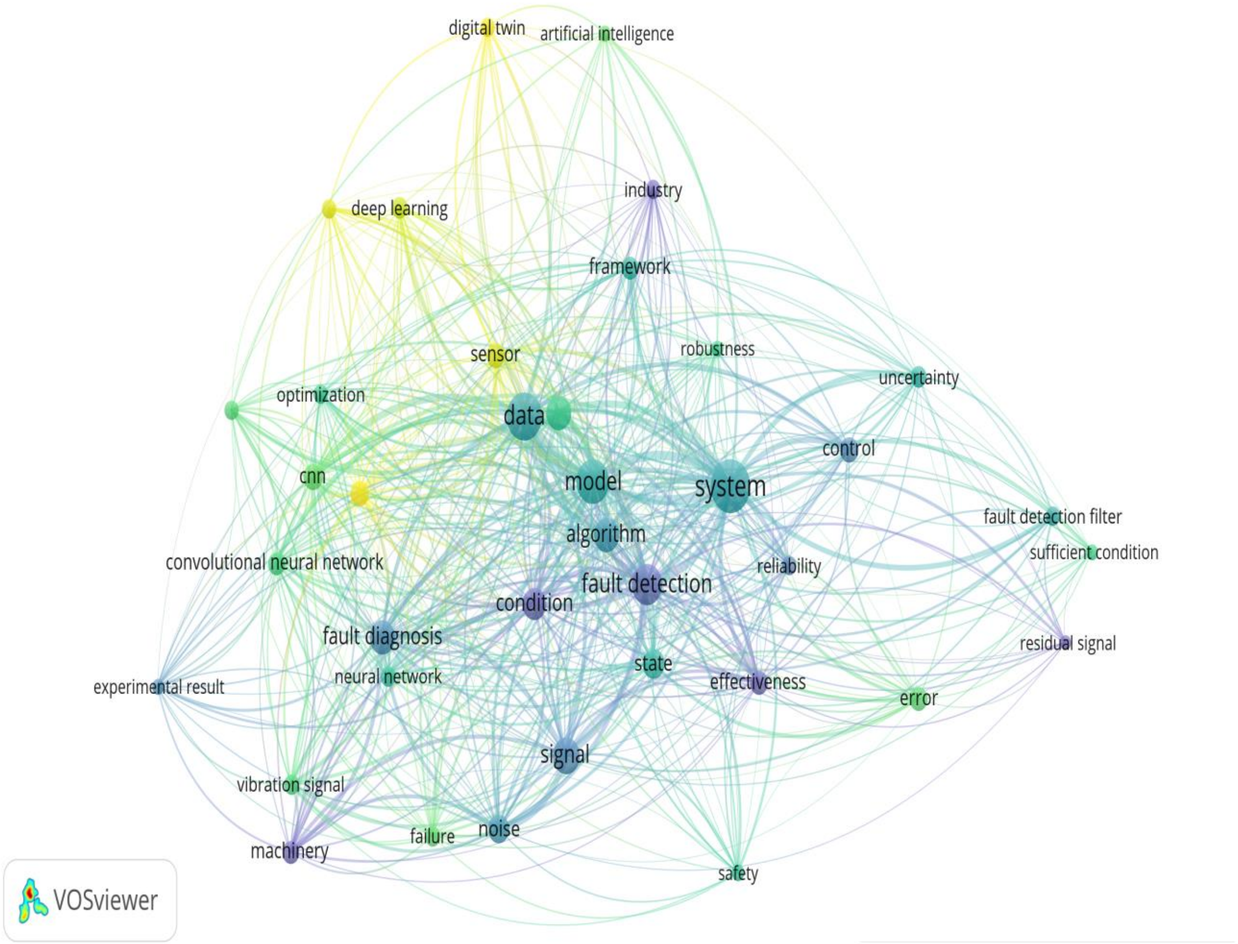
Screening method

- Review the title and abstract of the literature on Rayyan QRCI

Classification

The literature is classified by :

- Research area
- Literature type
- Application purposes
- DT features
- FDD approaches



LITERATURE REVIEW RESULTS

Overall findings

DT can be used to:

1) Generate synthetic data
2) Track the machinery degradation process
3) Monitor production process
4) Improve the decision-making process in maintenance
5) Support production issues after abnormality: rescheduling, energy-saving…etc.

The hybrid approach is better to overcome the drawbacks of each method ➔ DT is often combined with Data-driven methods

Artificial intelligence methods are widely used, especially Deep learning

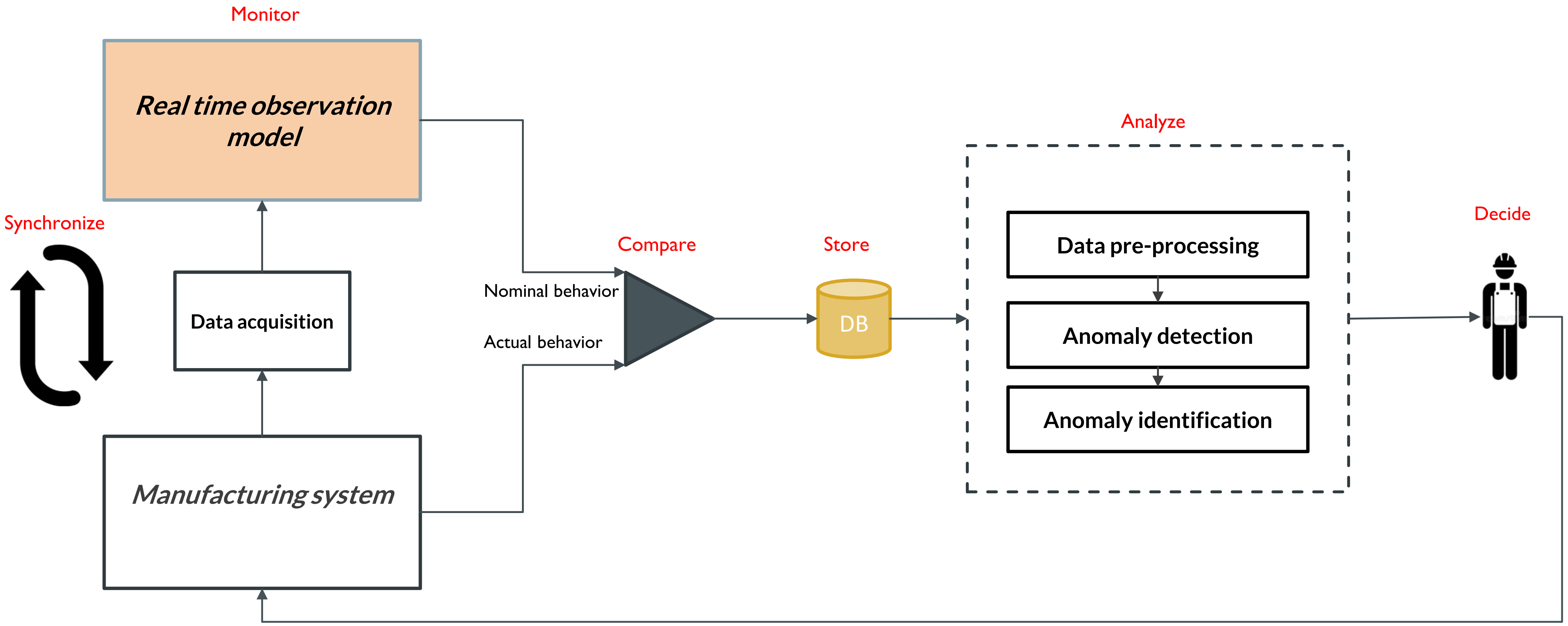
Research gap

Fault detection and diagnosis at the machine or component level ➔ less common on the system level

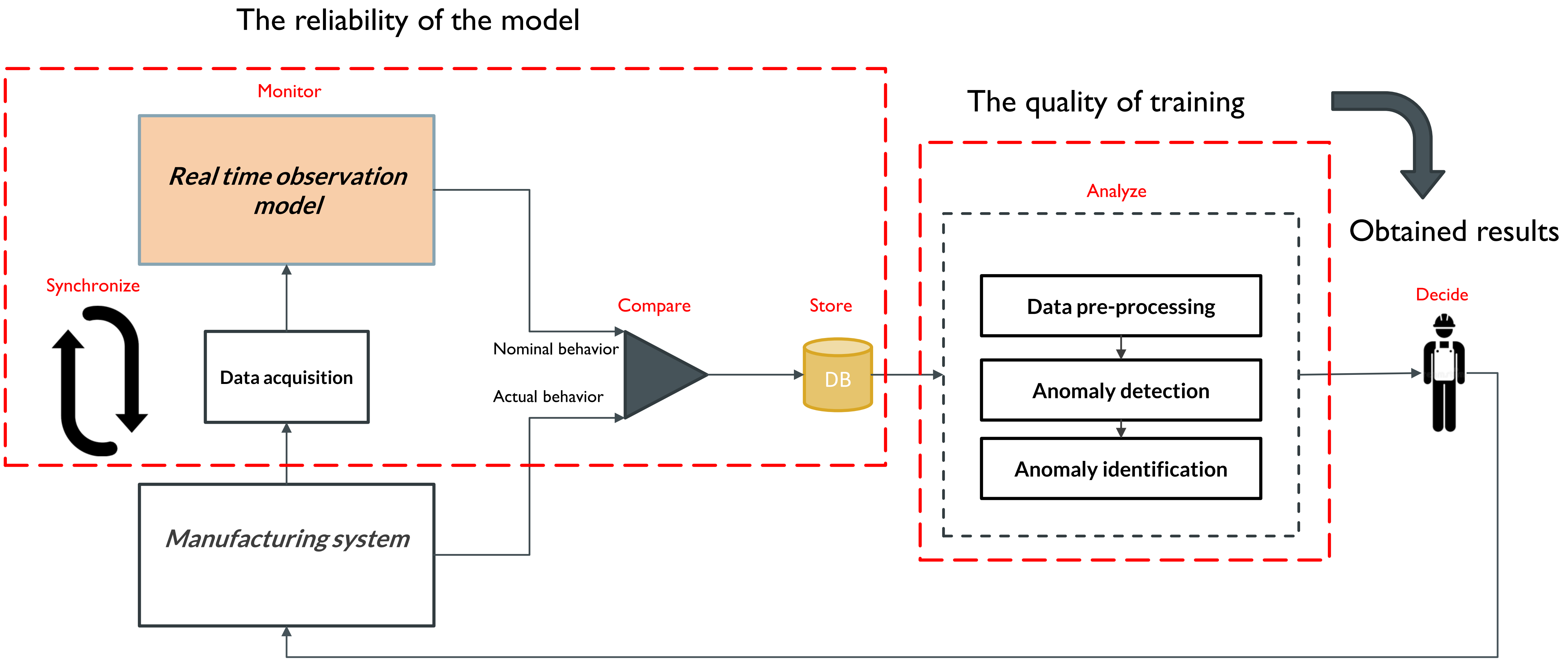
Lack of interest to behavioral anomalies

Lack of design methodology, reliability, and verification of the DT

METHODOLOGY

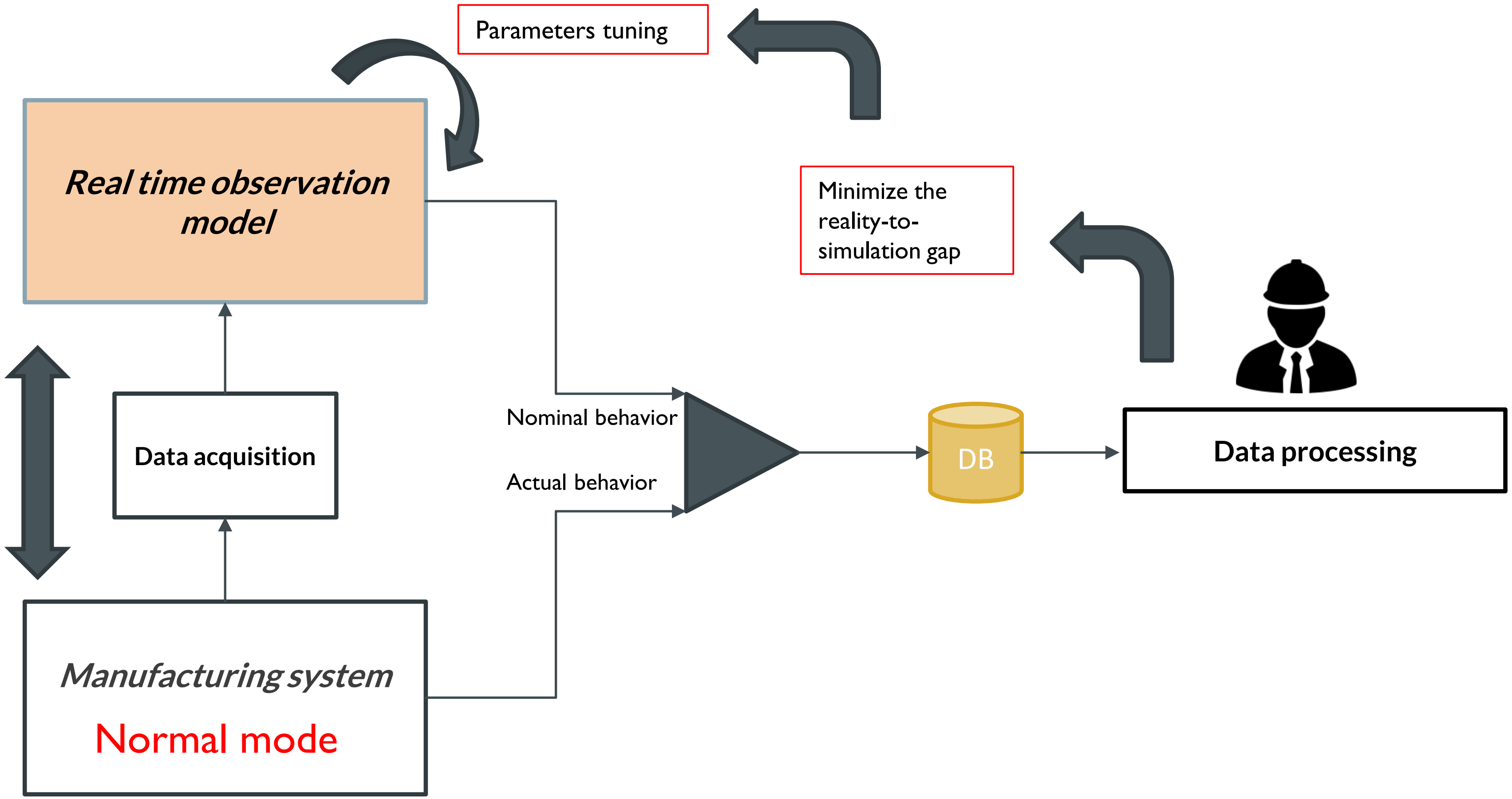


APPLICATION CHALLENGES



RELIABILITY OF THE MODEL

Phase 0: offline

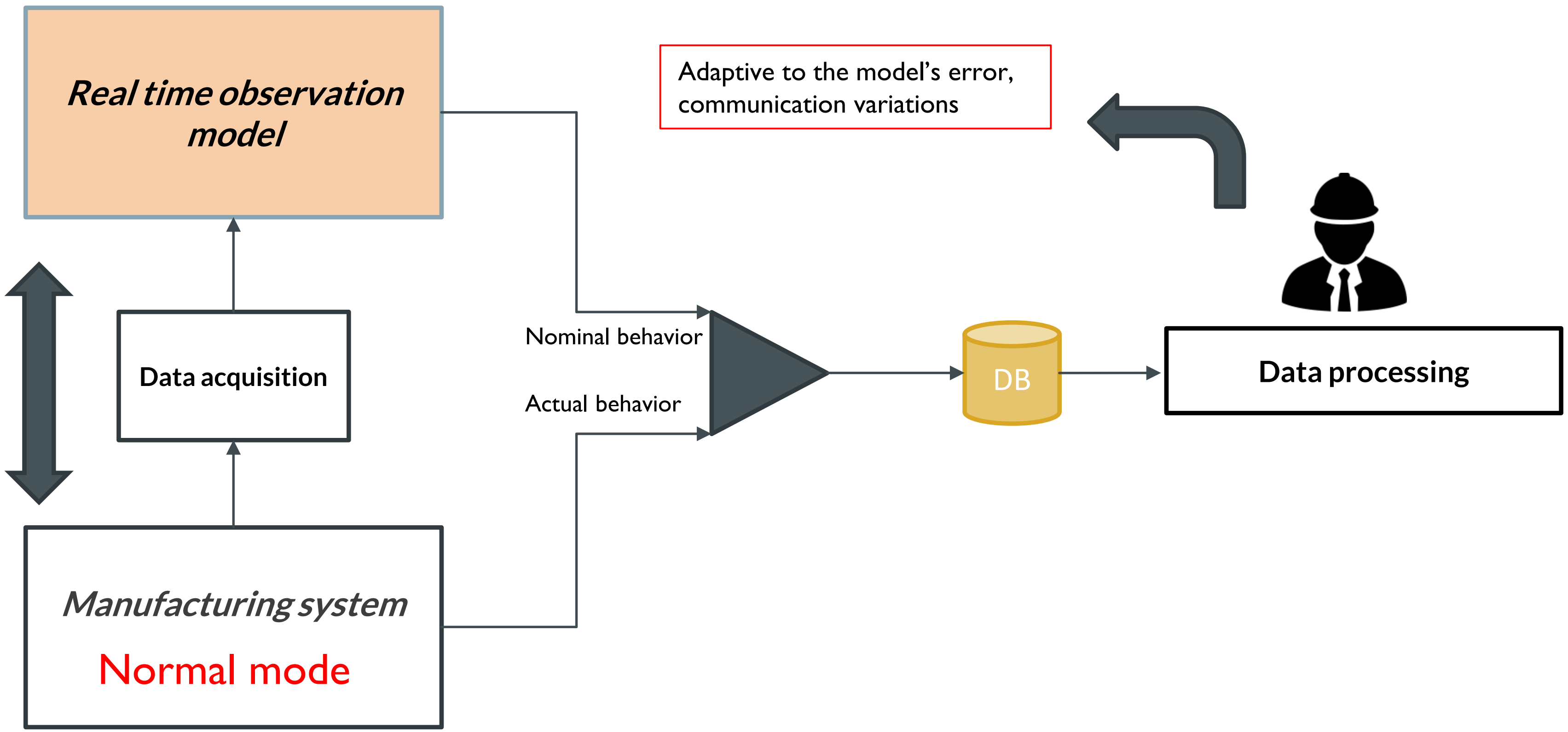


Threshold

- Generate false or missed alarms
- Perfect model

RELIABILITY OF THE MODEL

Phase 0: offline

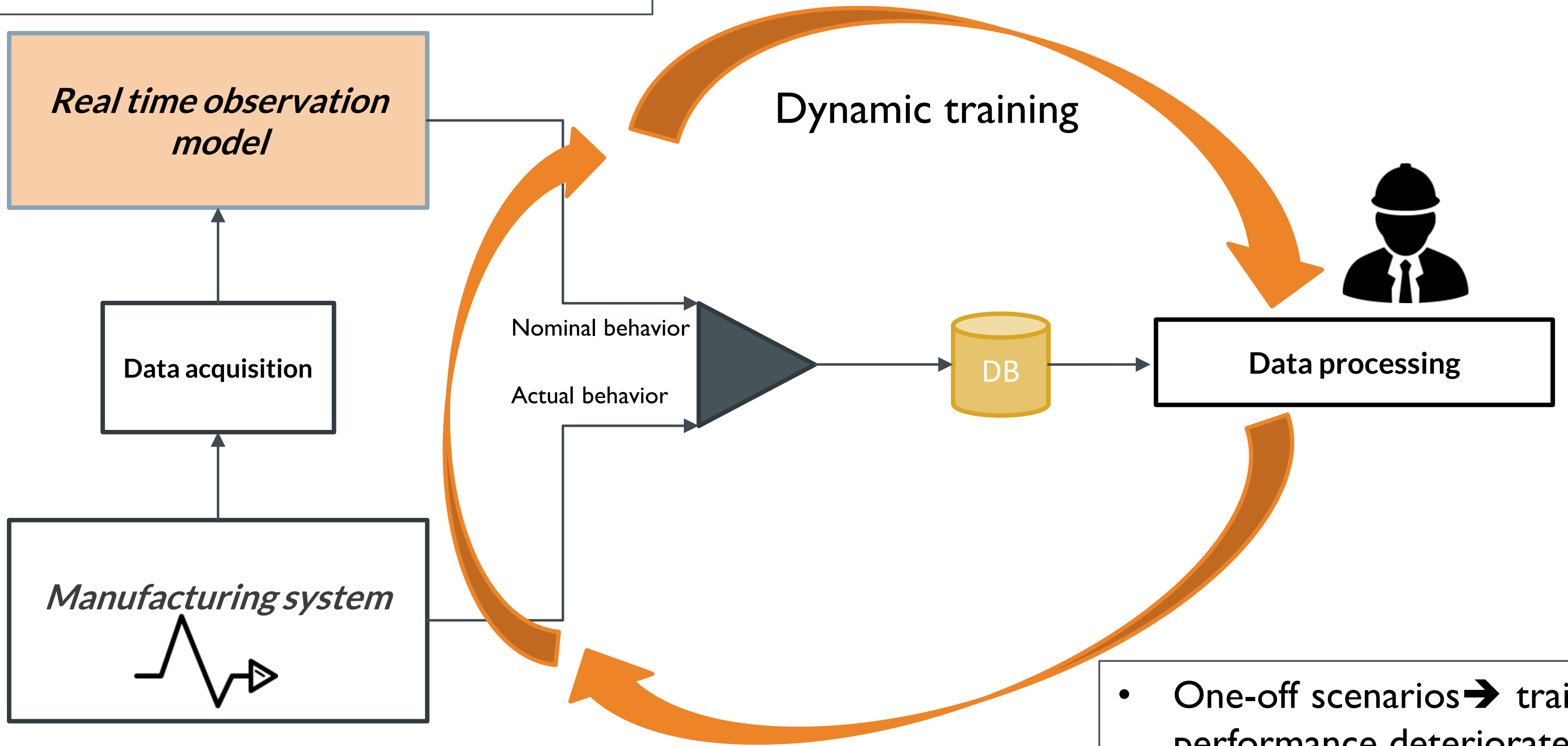


Machine learning

- Learn normal data pattern/ distribution
- Aknowledge « purposely false » behavior

Phase 1: online

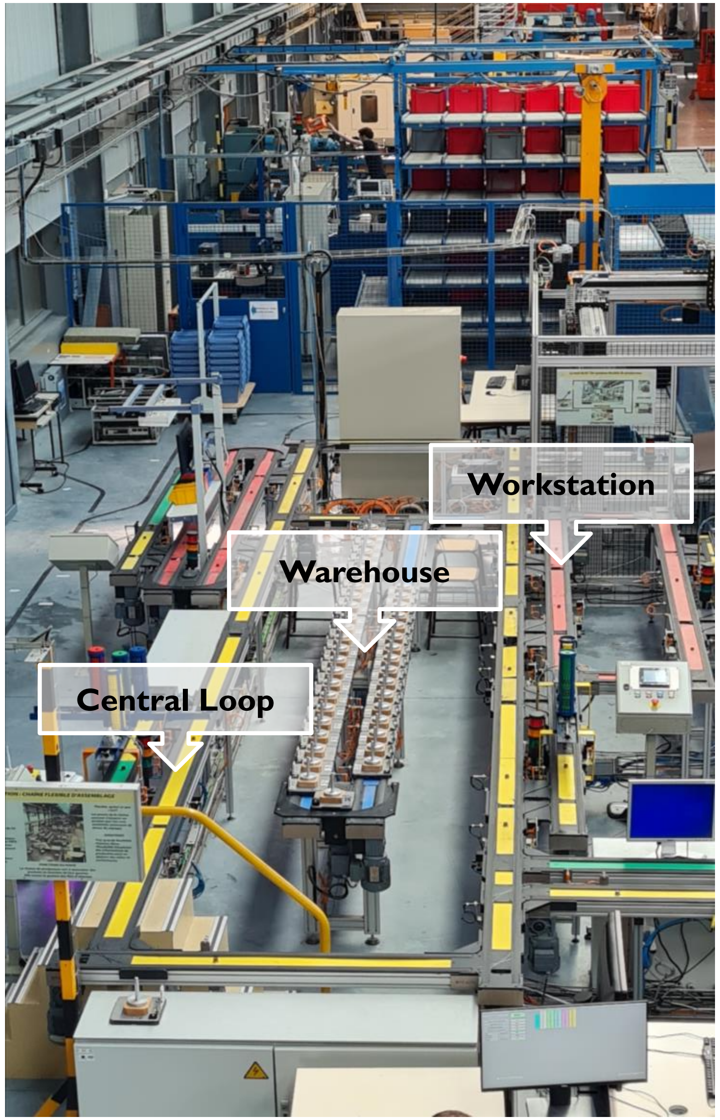
The rigidity of the virtual model ➔ restricted ability to reproduce new patterns that may be detected in real data.



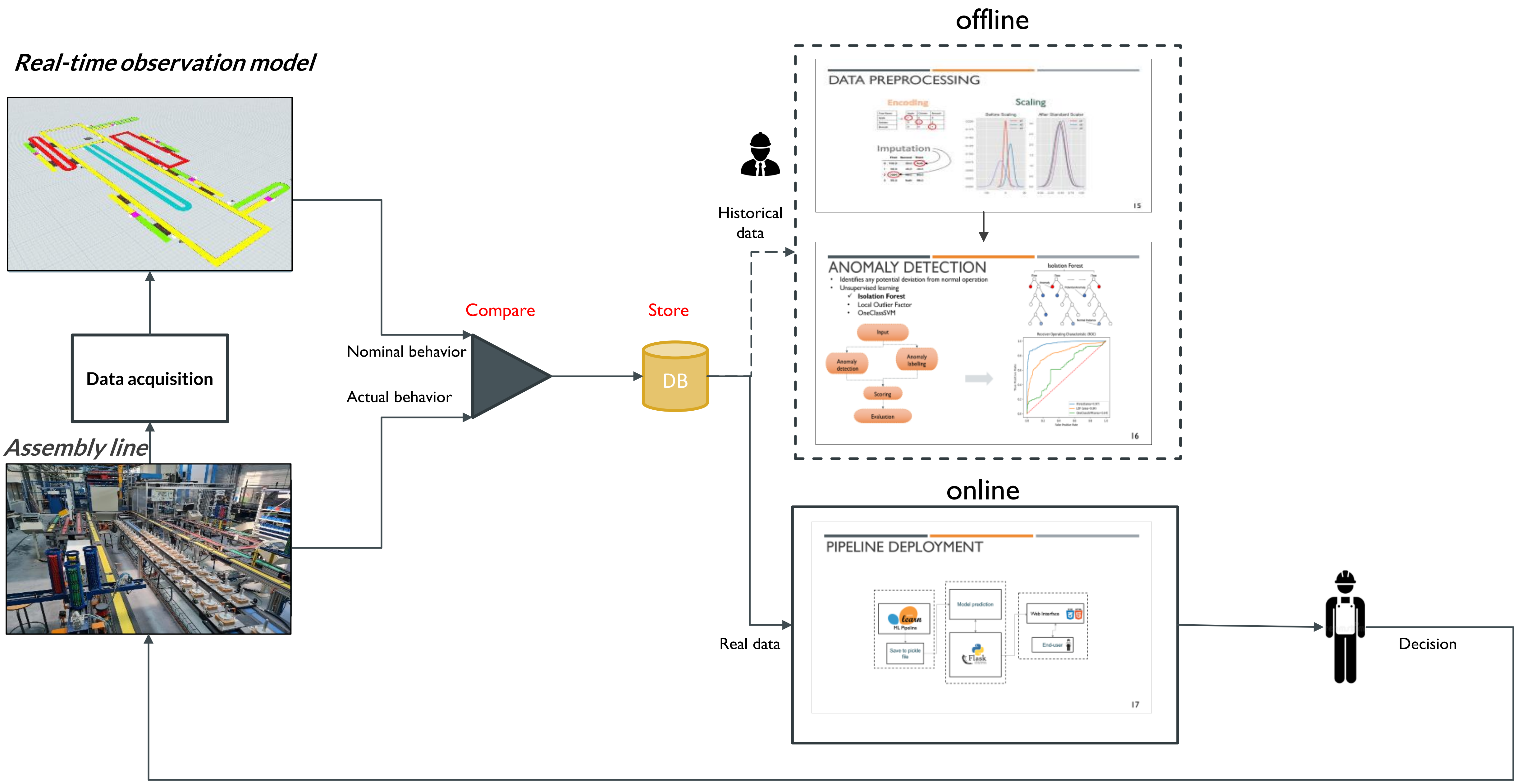
- Dynamic, non-linear, stochastic, complex behavior of the system
- Variations in external conditions

- One-off scenarios➔ trained ad-hoc, refreshed when its performance deteriorates
- Batch training ➔constantly refresh the model based on the latest train.
- Online training➔ train and learn continuously online.

ASSEMBLY LINE

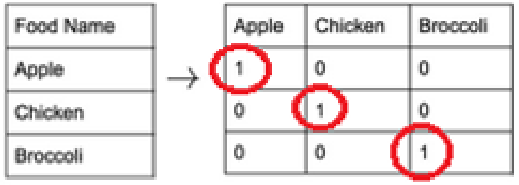


METHODOLOGY



DATA PREPROCESSING

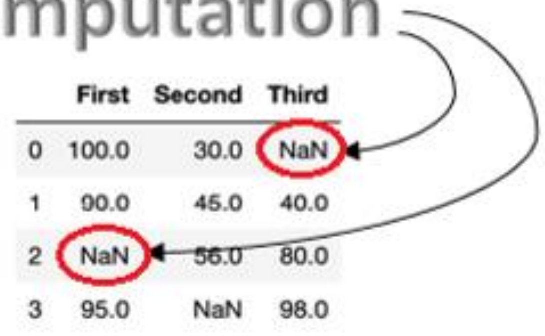
Encoding



Food Name
Apple
Chicken
Broccoli

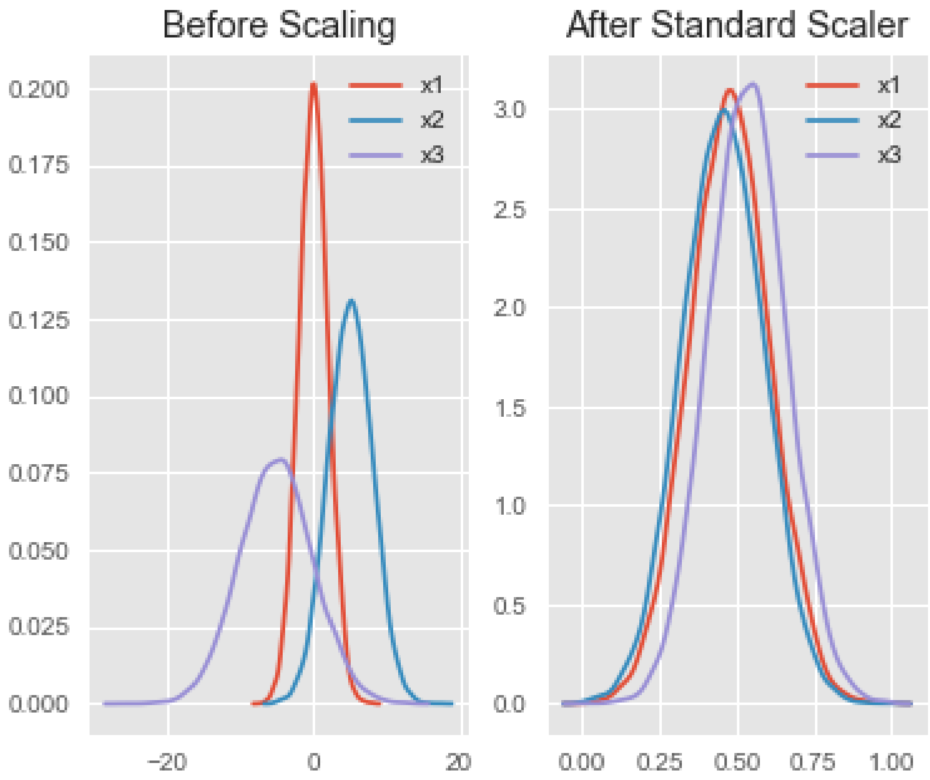
Apple	Chicken	Broccoli
1	0	0
0	1	0
0	0	1

Imputation



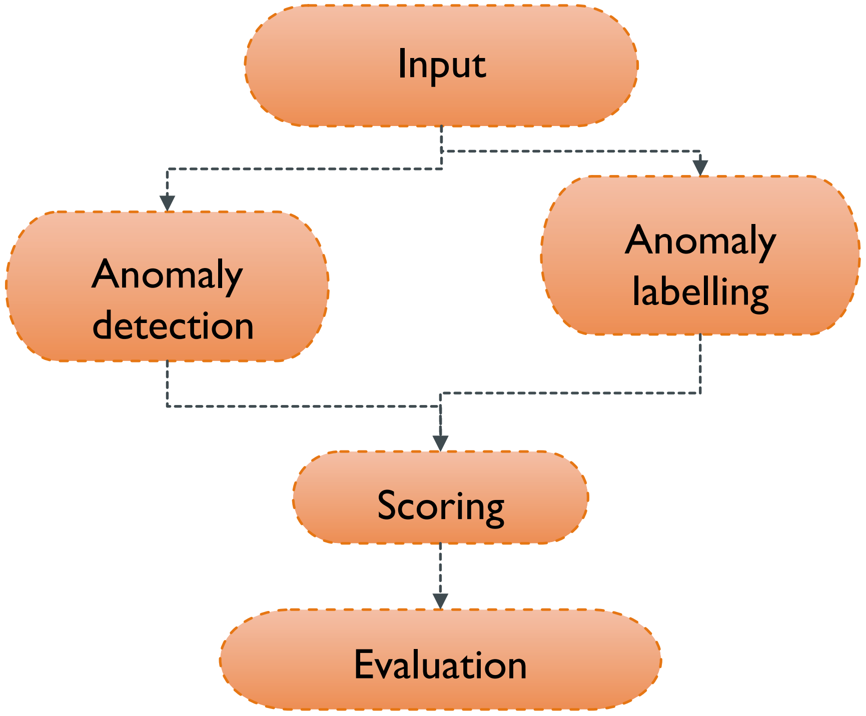
	First	Second	Third
0	100.0	30.0	NaN
1	90.0	45.0	40.0
2	NaN	56.0	80.0
3	95.0	NaN	98.0

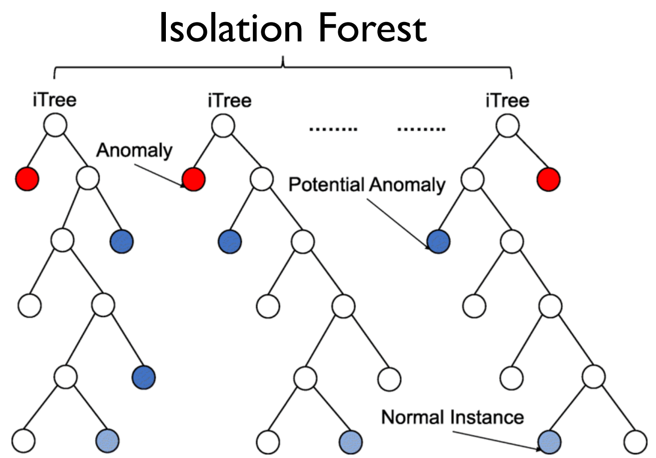
Scaling

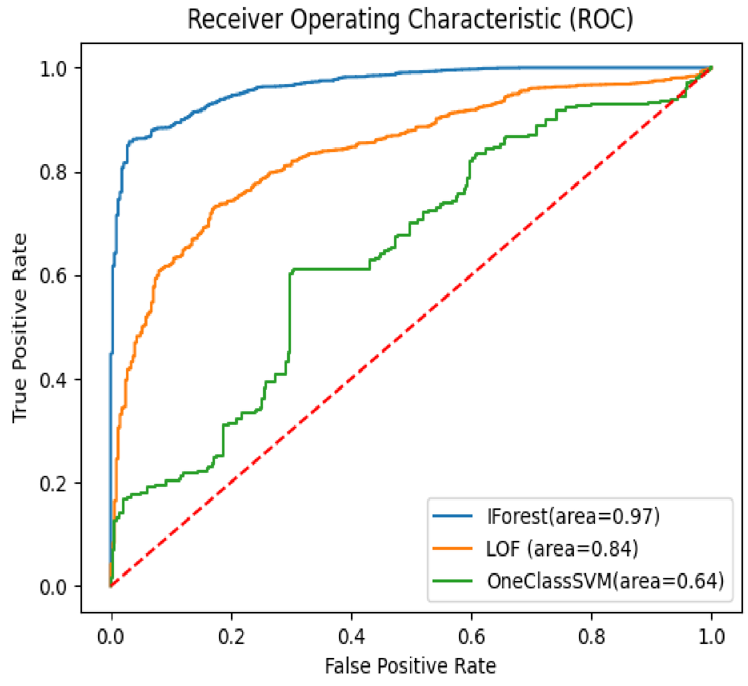


ANOMALY DETECTION

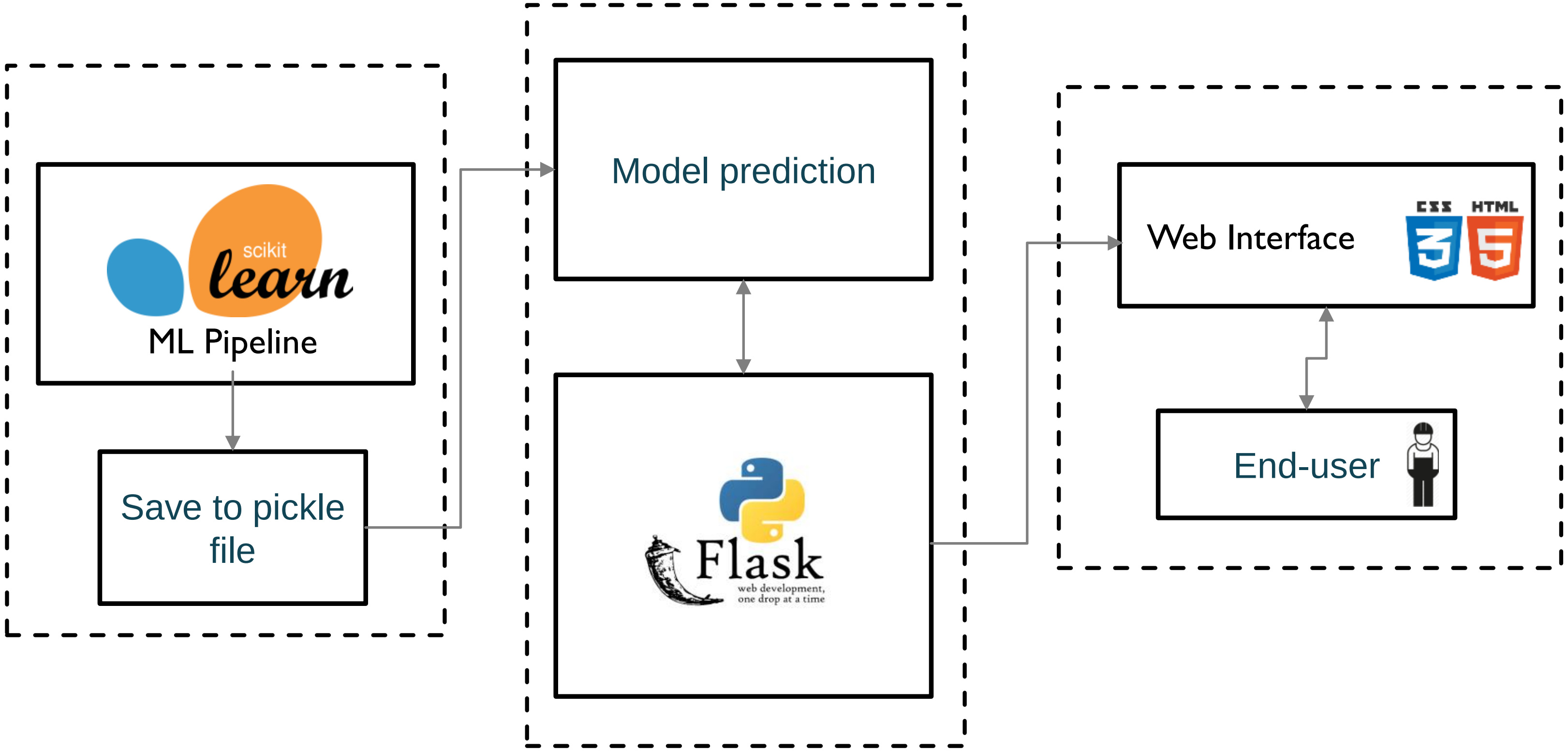
- Identifies any potential deviation from normal operation
- Unsupervised learning
  - ✓ **Isolation Forest**
  - Local Outlier Factor
  - OneClassSVM



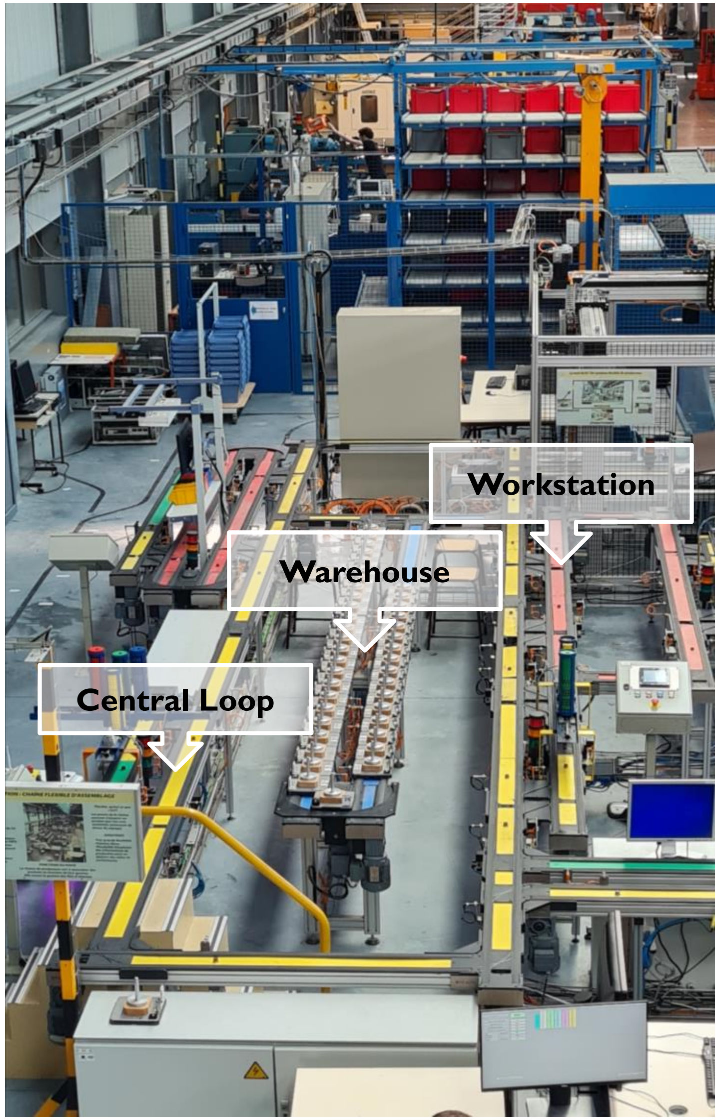


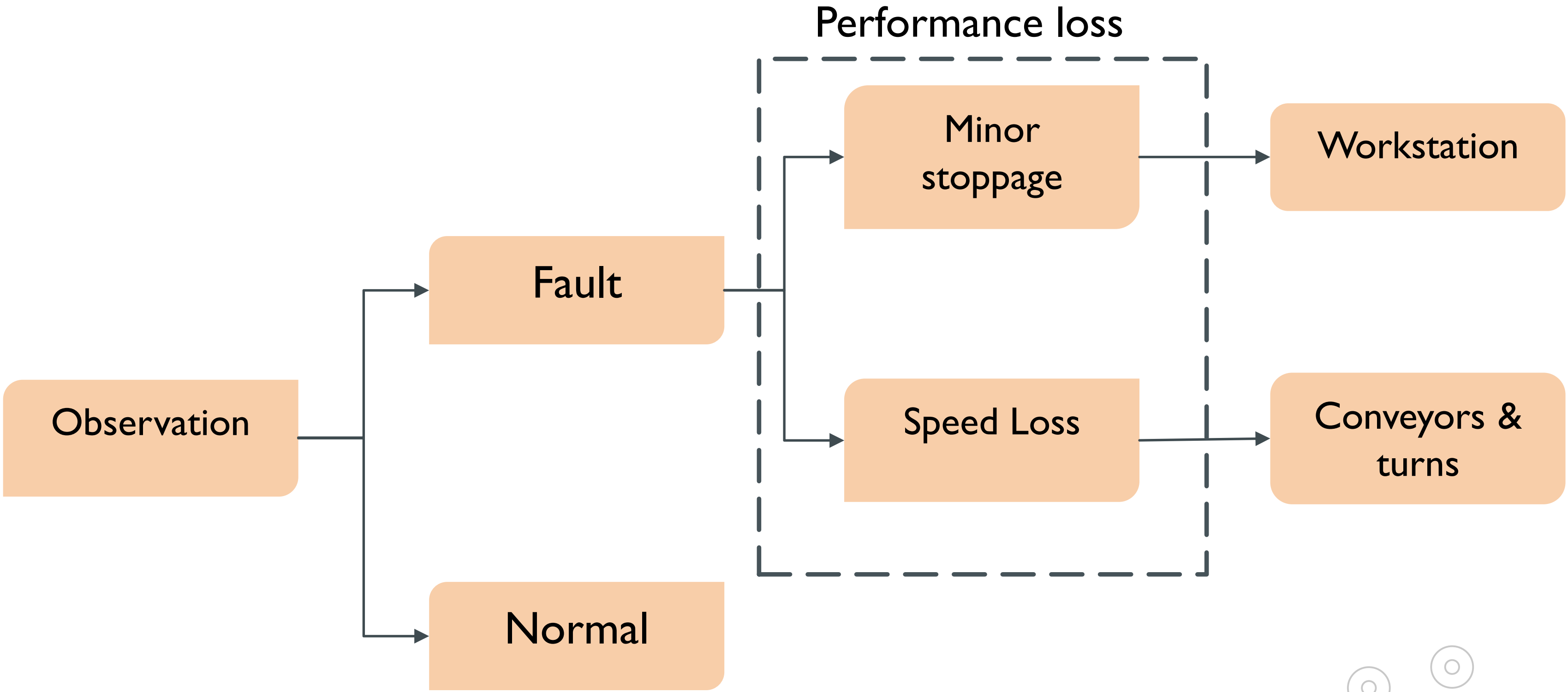


PIPELINE DEPLOYMENT



EXPERIMENTAL SETUP

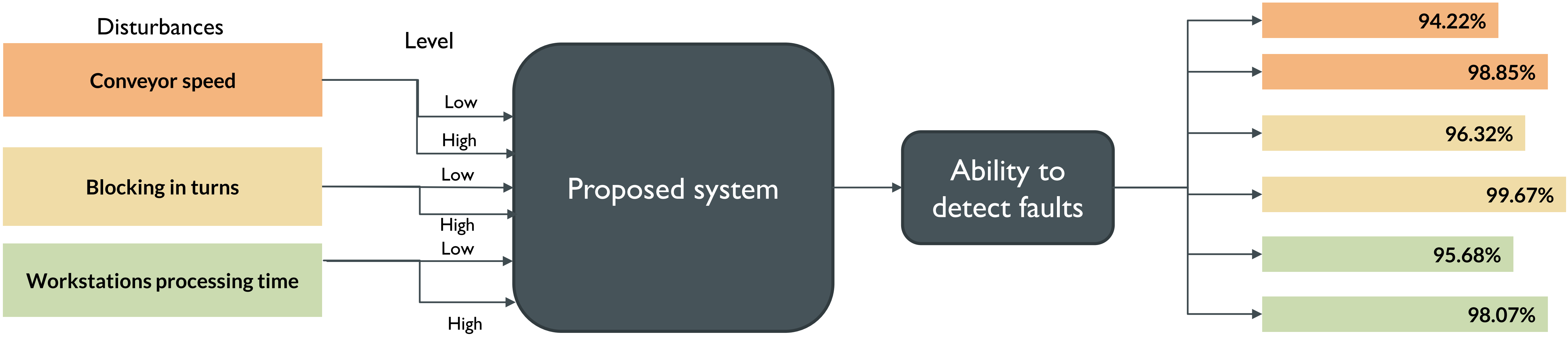


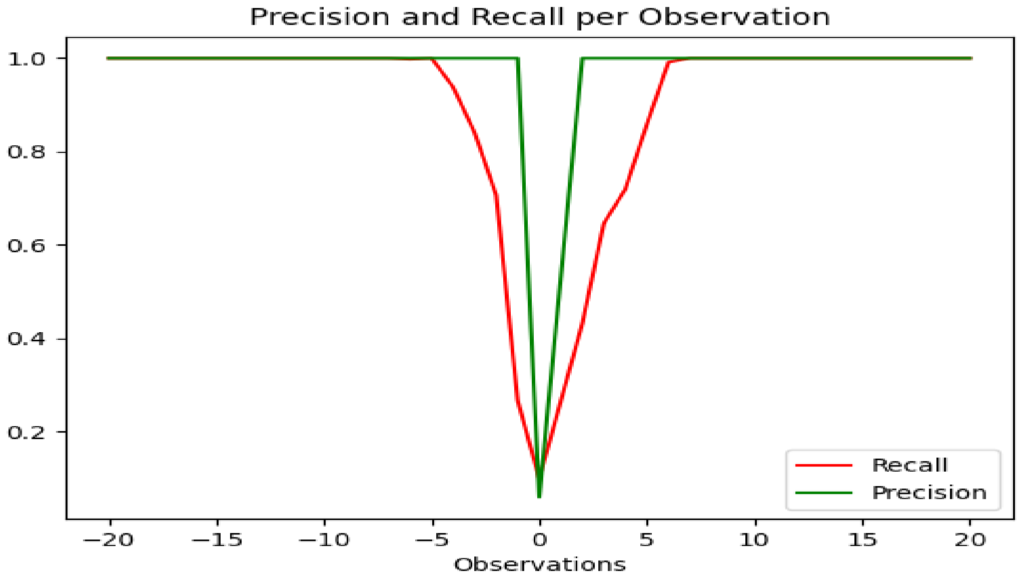


Objective

Apply a series of different type and level of severity of disturbances and observe the ability of the system to detect faults

PRELIMINARY STUDIES



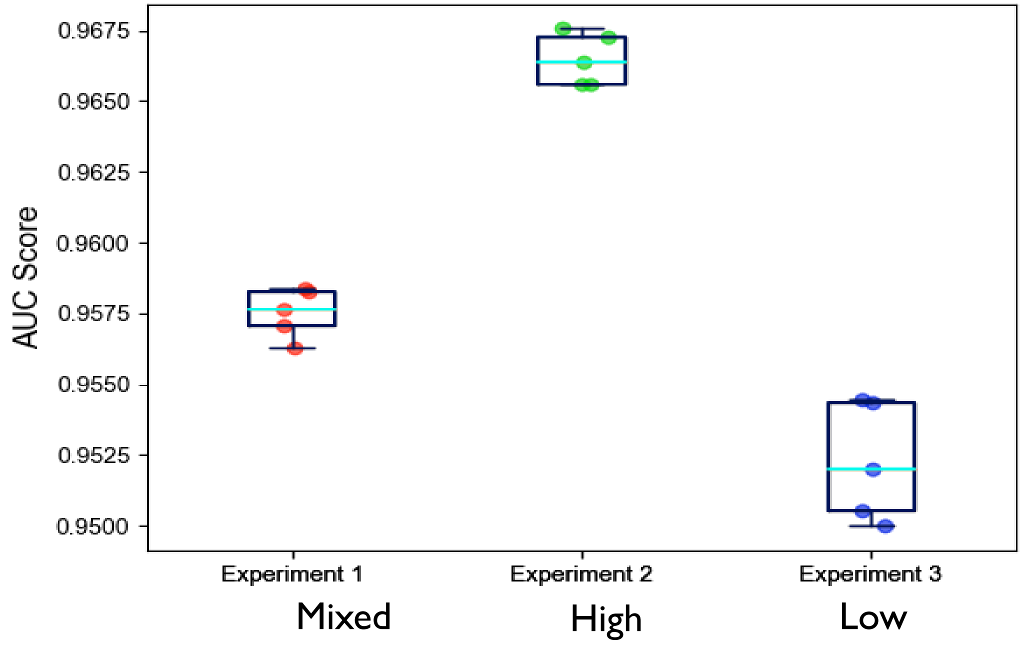


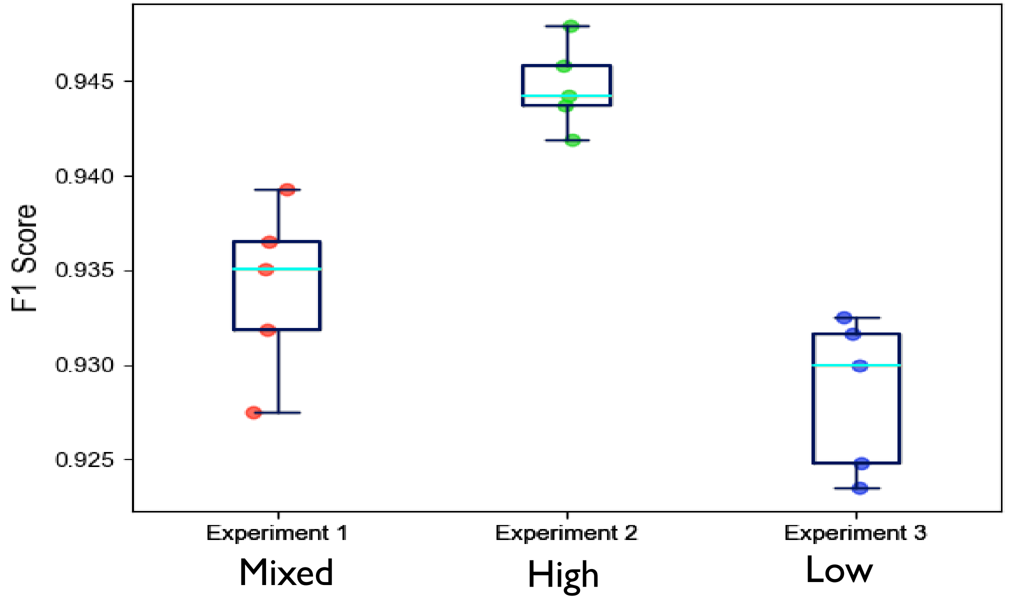
- False Negatives ∈ [-5;5]
- False positives ∈ [0;1]
- Low distrubances : [-5;5]
- High distrubances: elsewhere

VALIDATION

- 3 experiments :
  - Low disturbances
  - High disturbances
  - Mixed disturbances

- Simulation horizon per replica =4 hours
- N replicas = 5 per experiment
- Various faults at time
- Type of faults
  - Minor stoppages ==> 6 workstations
  - Speed loss==> conveyors





CONCLUSION

- The combination of methods (Model-based and data-driven) as a hybrid method to monitor, detect and diagnose behavioral anomalies in Manufacturing systems
- Reduce performance losses (Minor stoppages and speed losses) through the used method
- Experimental demonstration of the proposed framework
- Reliability of the DT model to assess the nominal behavior of the physical system
- Subtle challenge of training an accurate AI model
- Online training/updating of the model is considered in future work

THANK YOU FOR YOUR ATTENTION !

DIGITAL TWIN PARADIGM

- (Liu et al., 2018) describe DT as "a **living model** that **continually adapts** to **change** in the environment or operation using **real-time data** and can **forecast** the future of the corresponding **physical asset's behavior**"
- According to (Kaur et al., 2020), a digital twin's main feature is its ability to **correctly replicate** physical space using a combination of **physics-based** and **data-driven models**
- DT can check the consistency of **monitored data**, perform data mining to **detect** existing and **forecast** upcoming problems, using **Artificial Intelligence** to support effective **decisions**. (Asimov et al., 2018).