The Use of Cognitive Digital Twins on an IoT System for Edge Resilience and Anomaly Detection

Laboratoire
Décision et Information
Pour les Systèmes de Production

Contributors:

- Meriem Smati
- Boubou Thiam Niang
- Jannik Laval

Presented By:

- Meriem Smati

INSA INSTITUT NATIONAL DES SCIENCES APPLIQUÉES LYON

Université Claude Bernard Lyon 1

Université Lumière Lyon 2

Université de Lyon

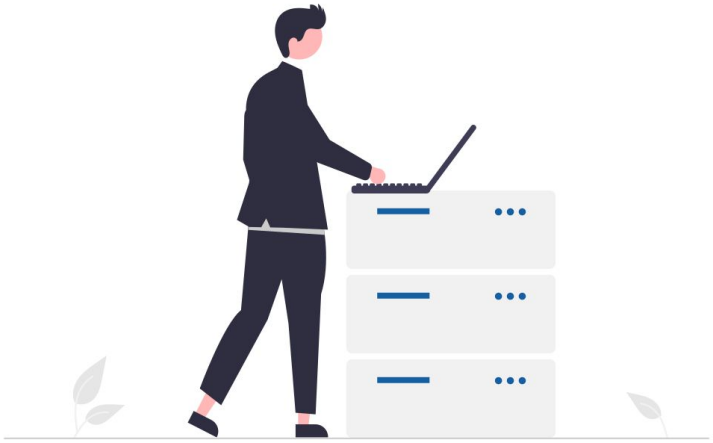
Part I

Laboratoire
Décision et Information
Pour les Systèmes de Production

Section I.1 : Problem Statement, Objectives and Backgrounds

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.1: Problem Statement



- Resilience.
- Maintenance.
- IoT Systems.

How to use Digital Twins for IoT Systems resilience?

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.1: Objectives

- Define the DT's concepts.
- Analyze different articles to create a new improved framework.

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.1: Different Involved Domains

- Internet of Things
- Artificial Intelligence
- Machine Learning
- Deep Learning
- Digital Twins

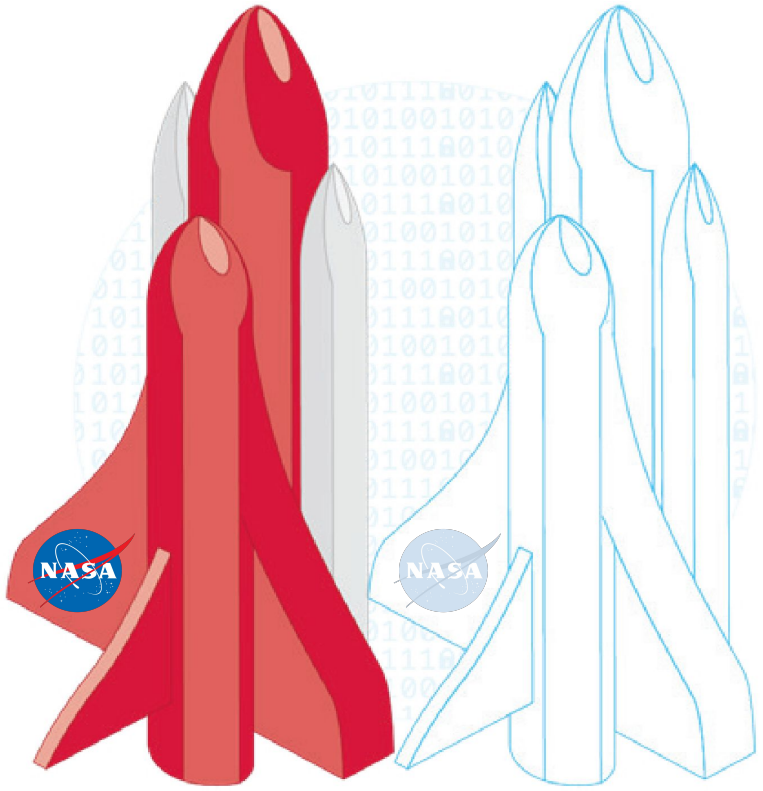
Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.1: Digital Twins Background

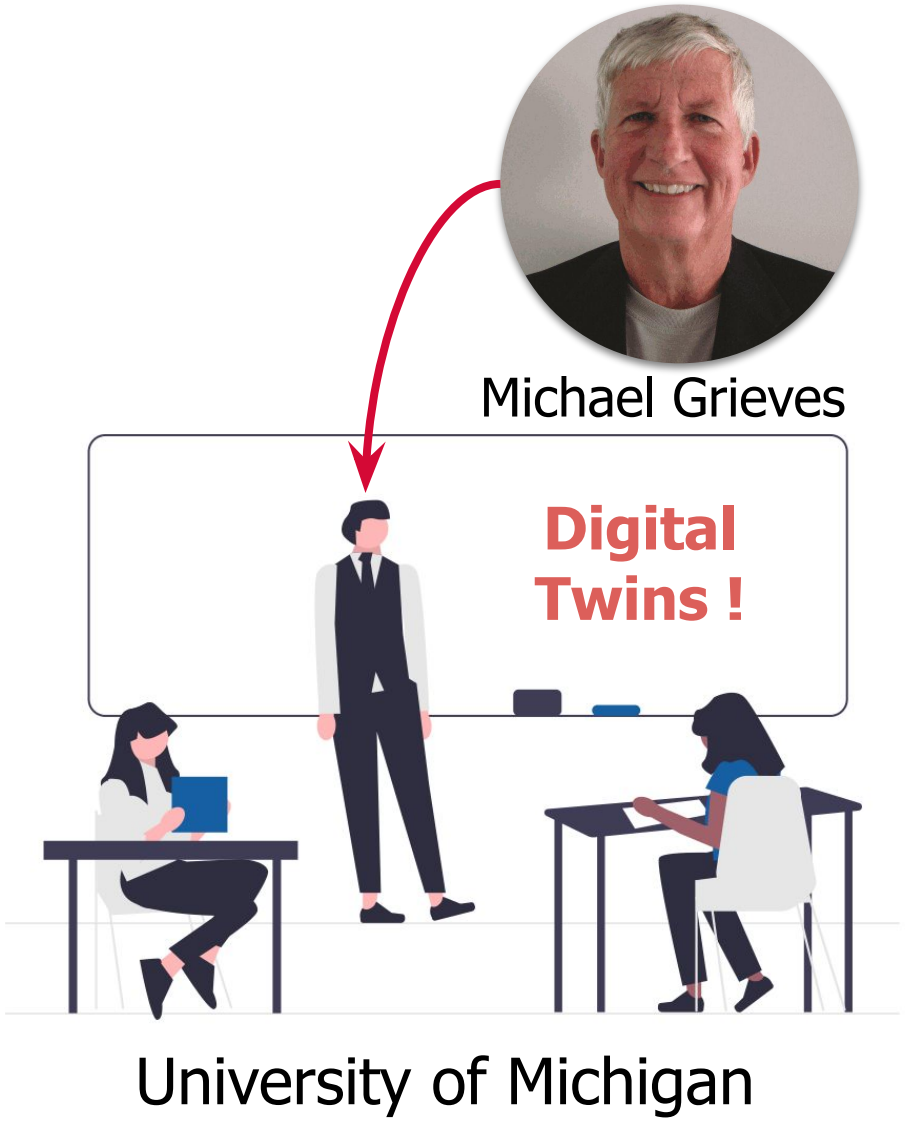
1970

- The concept of a "Twin" has its roots in the NASA Apollo program of the 1970s.

- Create a replica of space vehicles on Earth.



Part I.1: Digital Twins Background

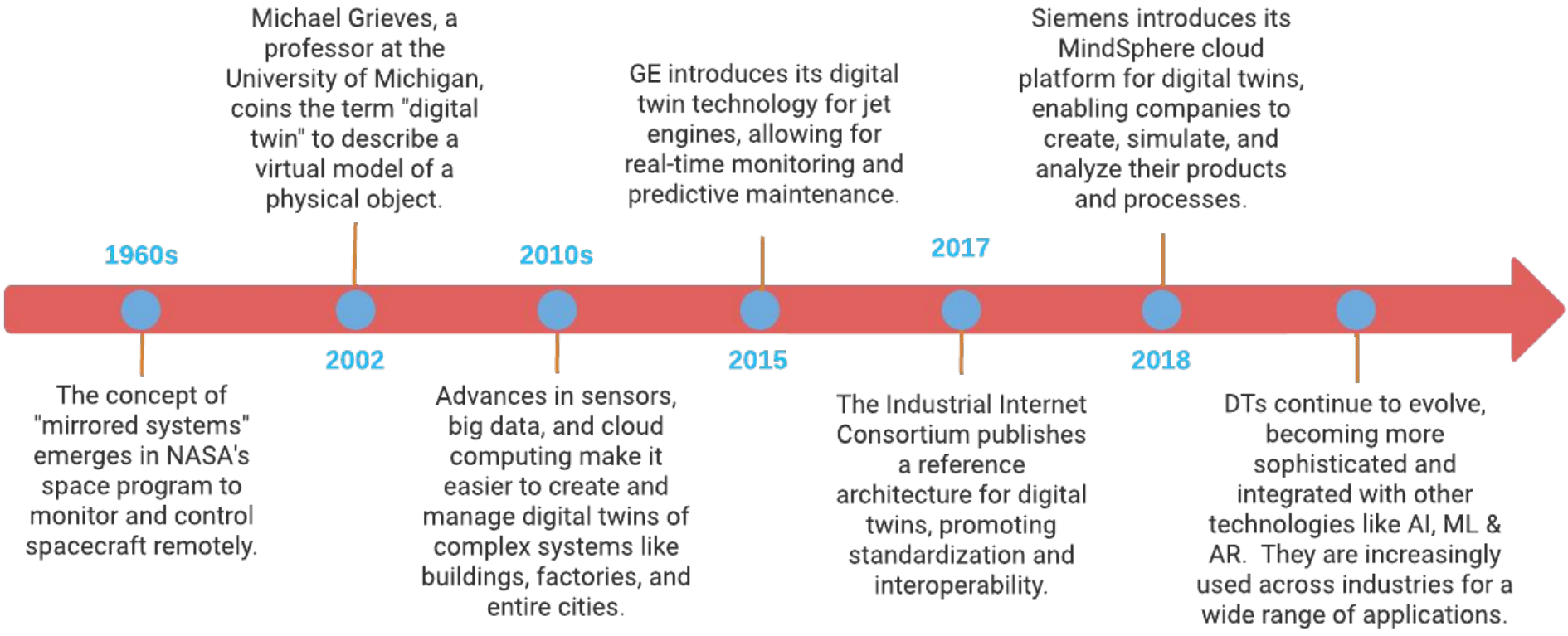


2003

- Michael Grieves proposed the idea of a "Digital Twin" in his Product Life-Cycle Management (PLM) course.

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.1: Digital Twins Background



Laboratoire
Décision et Information
Pour les Systèmes de Production

Section I.2 :
State of the Art

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2: State of The Art

State of the Art

- Digital Twins Concepts
- Digital Twins for resilience

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.1: Digital Twin's Principales

Digital Twin's Definitions

Digital Twins lake a universal definition and highly depend on the use case and the domain of application.

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.1: Digital Twin's Principales

Digital Twin's Definitions

In Aerospace:

A DT is an integrated multiphysics, multiscale, probabilistic simulation of an as-built vehicle that uses the physical models and other relevant information to accurately replicate the life and behavior of its corresponding flying counterpart.

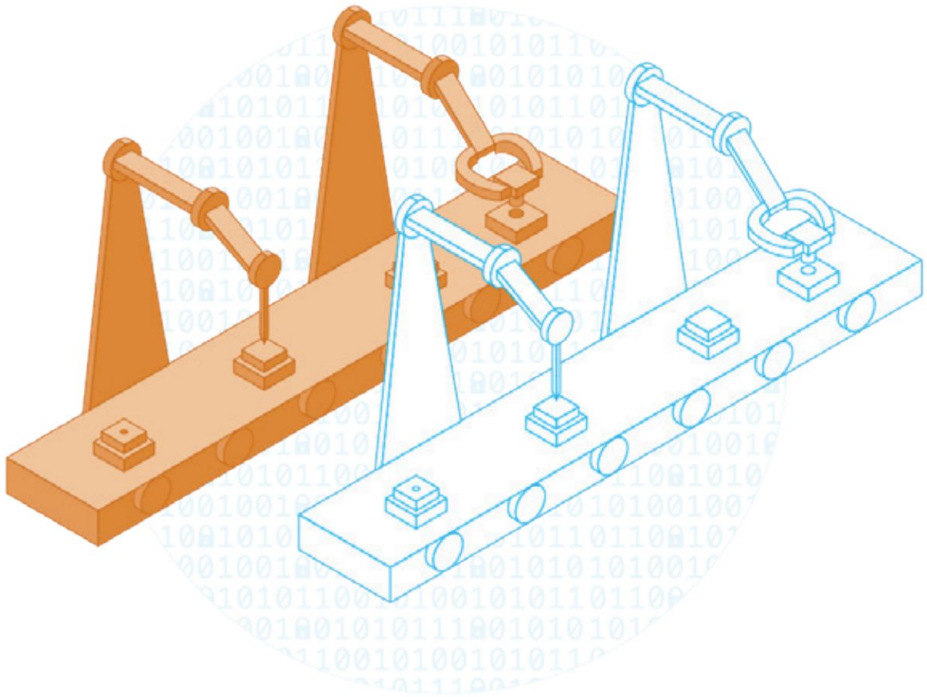
Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.1: Digital Twin's Principales

Digital Twin's Definitions

In the Industry:

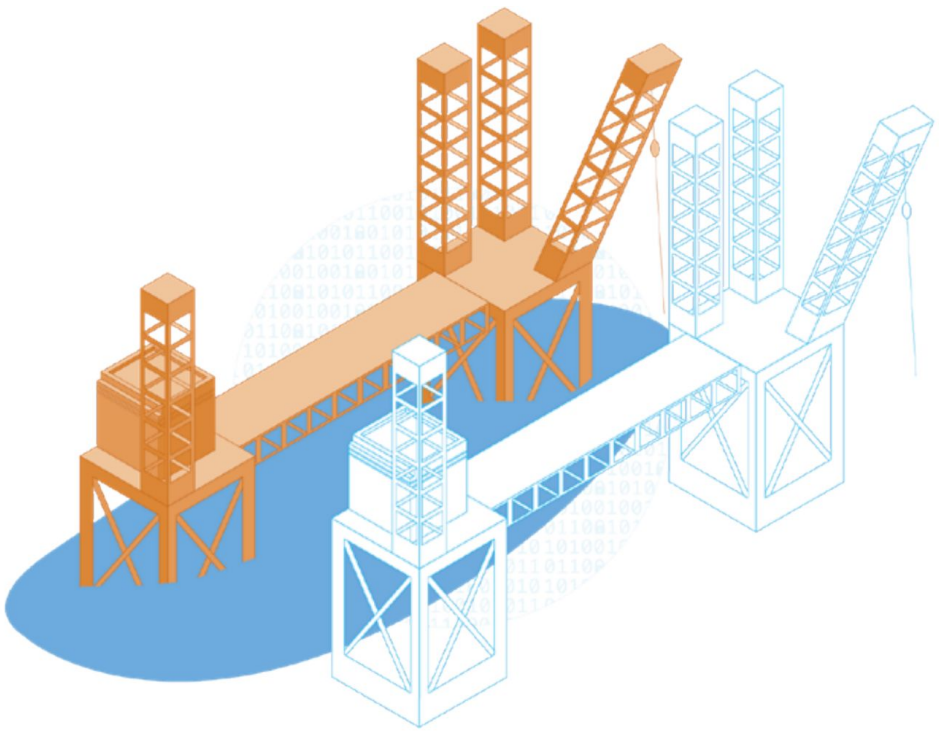
A DT is an evolving digital profile of the historical and current behavior of a physical object that helps optimize business performance. It is based on massive, cumulative, real-time, real-world data measurements across an array of dimensions.



Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.1: Digital Twin's Principales

Digital Twin's Definitions



In Engineering:

A DT is a digital replica of physical assets, processes, and systems that can be used for various purposes, such as simulation, optimization, and monitoring.

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.1: Digital Twin's Principales

Digital Twin's Definitions

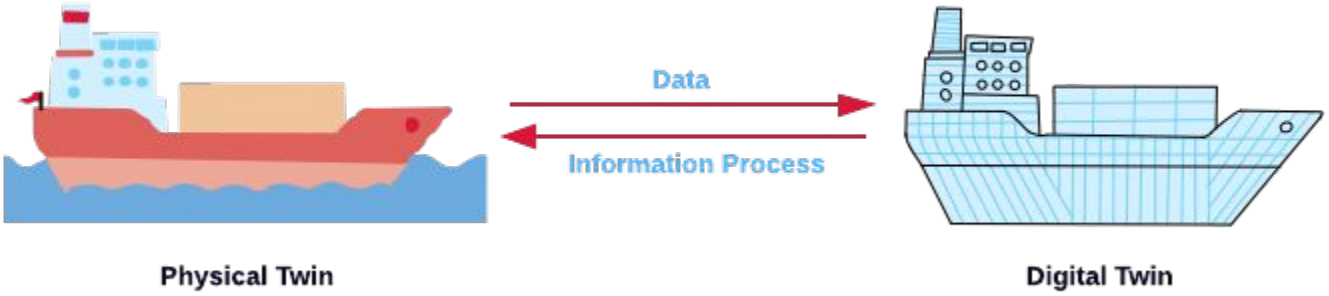
And much more other definitions have been conducted in Healthcare, Manufacturing, Agriculture etc.

Part I.2.1: Digital Twin's Principales

Digital Twin's Definitions

We can conclude the following definition for "Digital Twins in the IoT Domain":

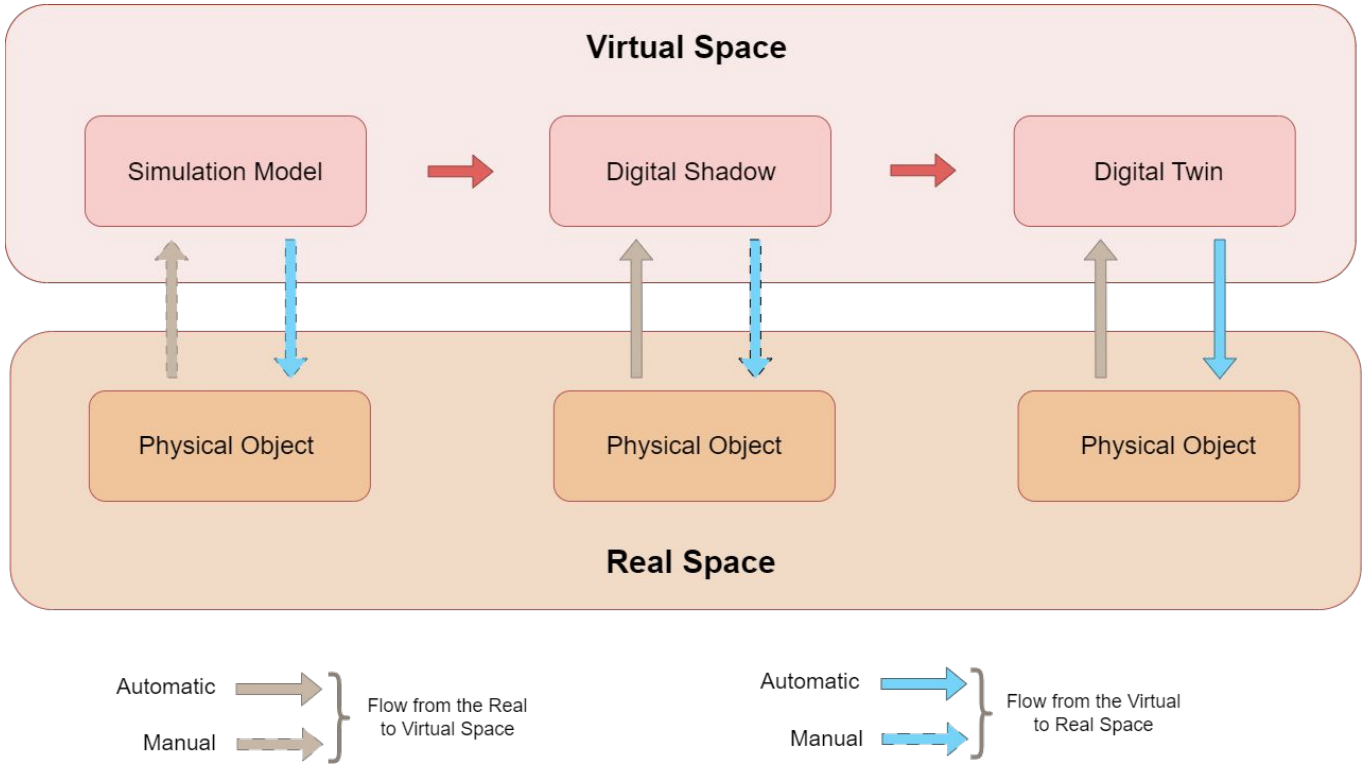
"The combination of virtual machines and computer-based models that enable the simulation, emulation, or mirroring of the behavior and characteristics of a physical entity".



Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.1: Digital Twin's Principales

Digital Twin's Predecessors



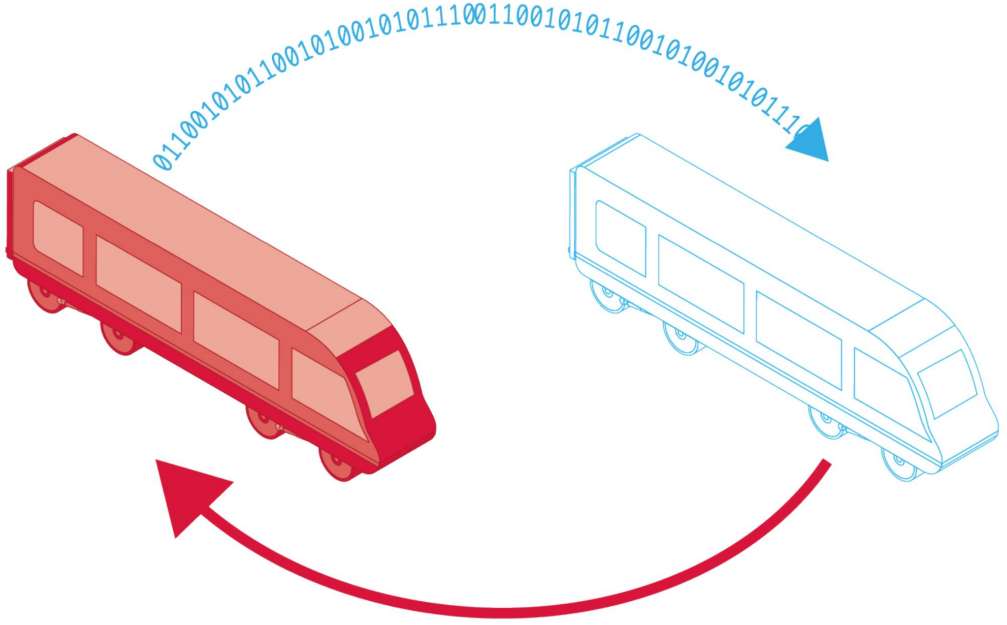
Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.1: Digital Twin's Principales

Digital Twin's Components

Required Components

- Physical Asset
- Digital Asset
- Continuous Bijective Relation



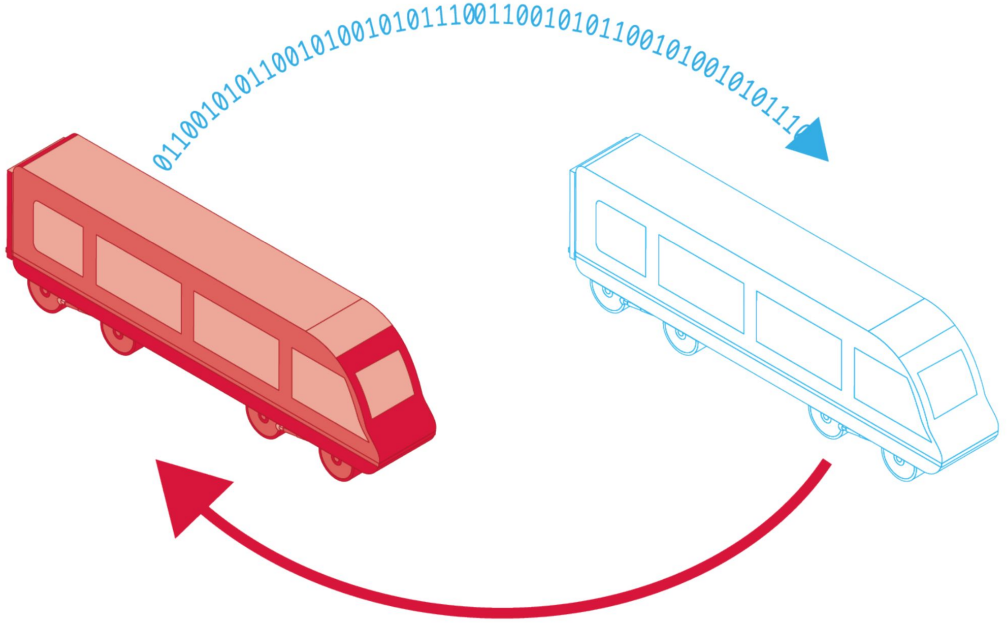
Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.1: Digital Twin's Principales

Digital Twin's Components

Optional Components

- Time Continuous Data
- Time-series data
- Knowledge Database
- Security
- ML
- Evaluation metrics
- IoT
- etc.



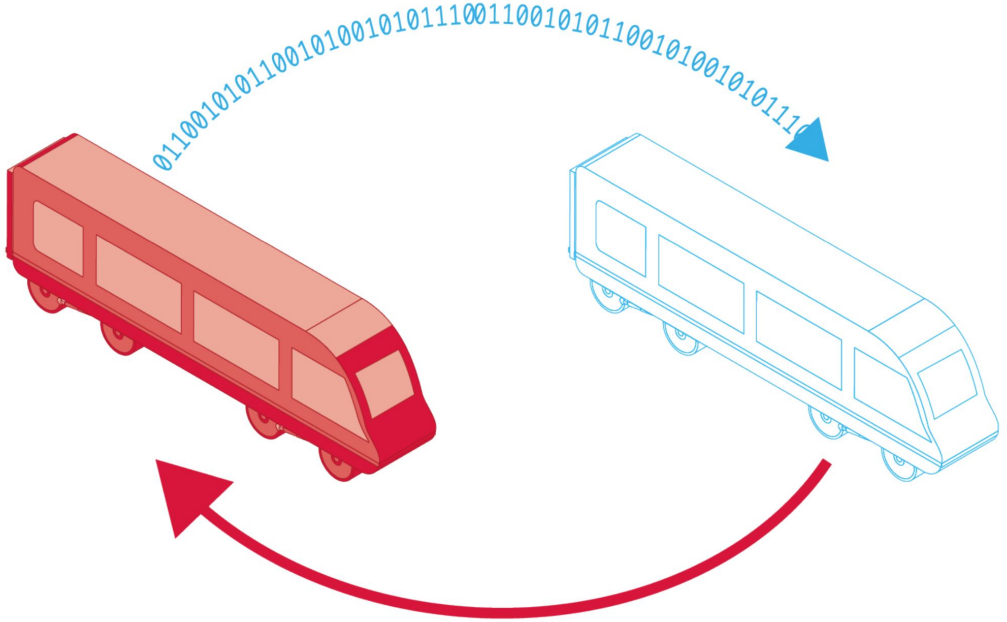
Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.1: Digital Twin's Principales

Digital Twin's Components

Optional Components

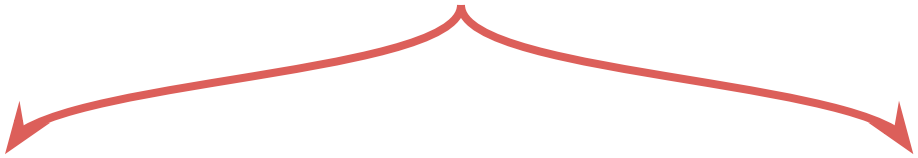
- Time Continuous Data
- Time-series data
- Knowledge Database
- Security
- ML
- Evaluation metrics
- IoT
- etc.



Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.1: Digital Twin's Principales

Digital Twin's Types

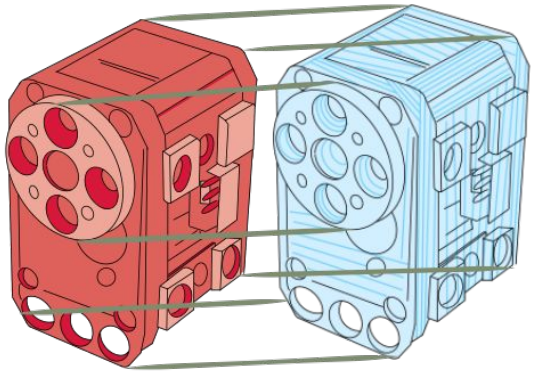


- Level of granularity
- Twin's dynamism

Part I.2.1: Digital Twin's Principales

Digital Twin's Types

Level of granularity



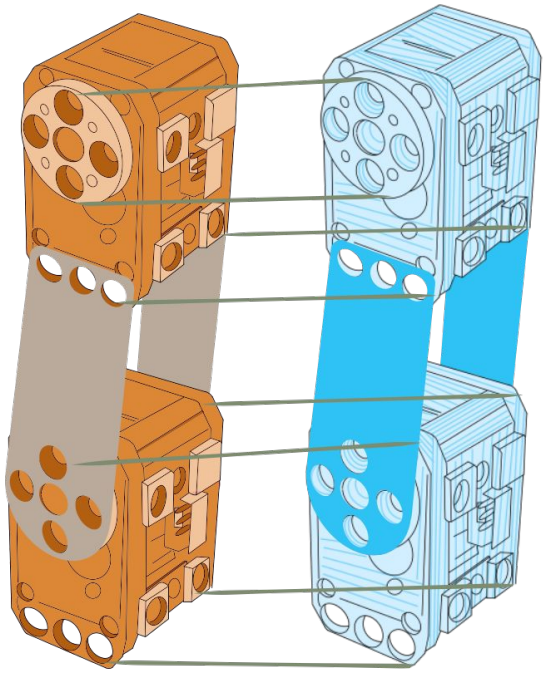
Components Twins:

The basic unit of a DT and the smallest example of a functioning component.

Part I.2.1: Digital Twin's Principales

Digital Twin's Types

Level of granularity



Assets Twins:

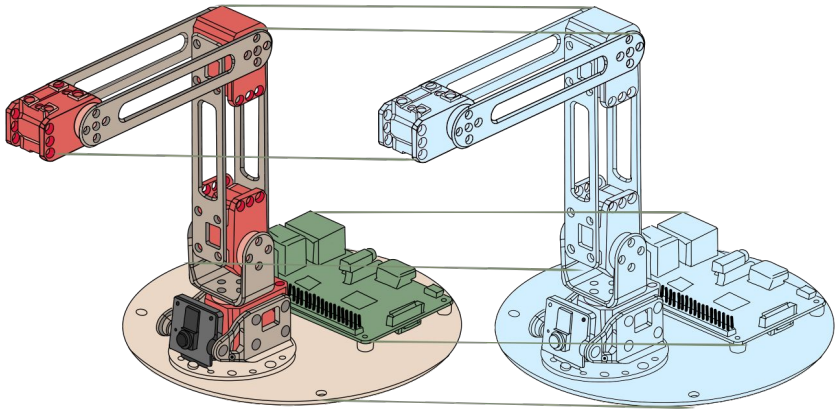
When two or more components work together.

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.1: Digital Twin's Principales

Digital Twin's Types

Level of granularity



System Twins / Unit Twins:

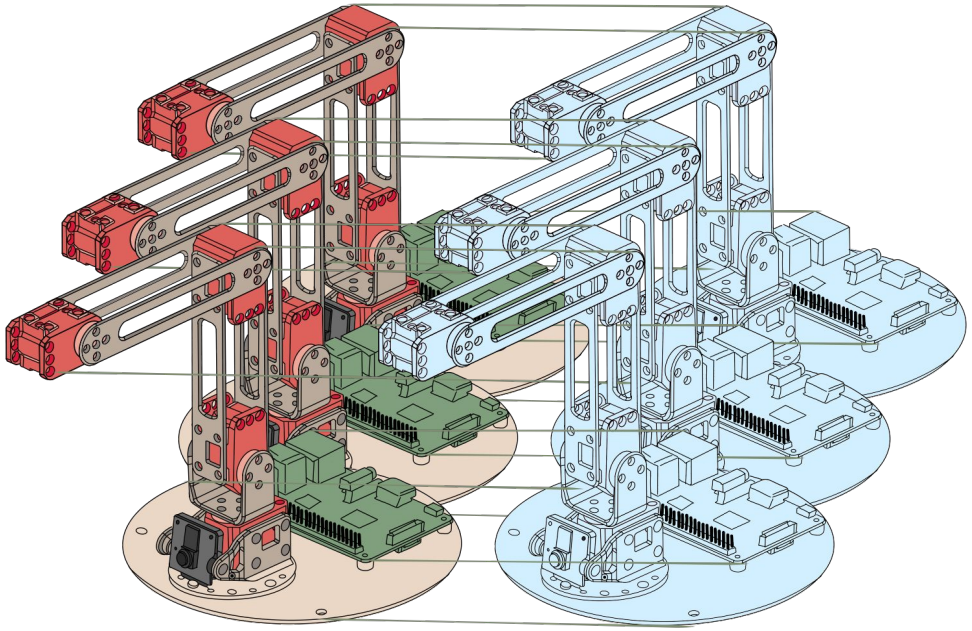
It enables to detect different assets connected to form a whole functioning system.

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.1: Digital Twin's Principales

Digital Twin's Types

Level of granularity



Process Twins:

It is the macro level of magnification. It is the digitalization of entire business processes.

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.1: Digital Twin's Principales

Digital Twin's Types

Twin's Dynamism

a Dynamic DT:

Fed by live data flows from a physical asset. Insights and programmed instructions from the digital twin can then impact the physical twin using real-time control mechanisms.

a Static DT:

Changes periodically as long-term data about a physical asset are added in.

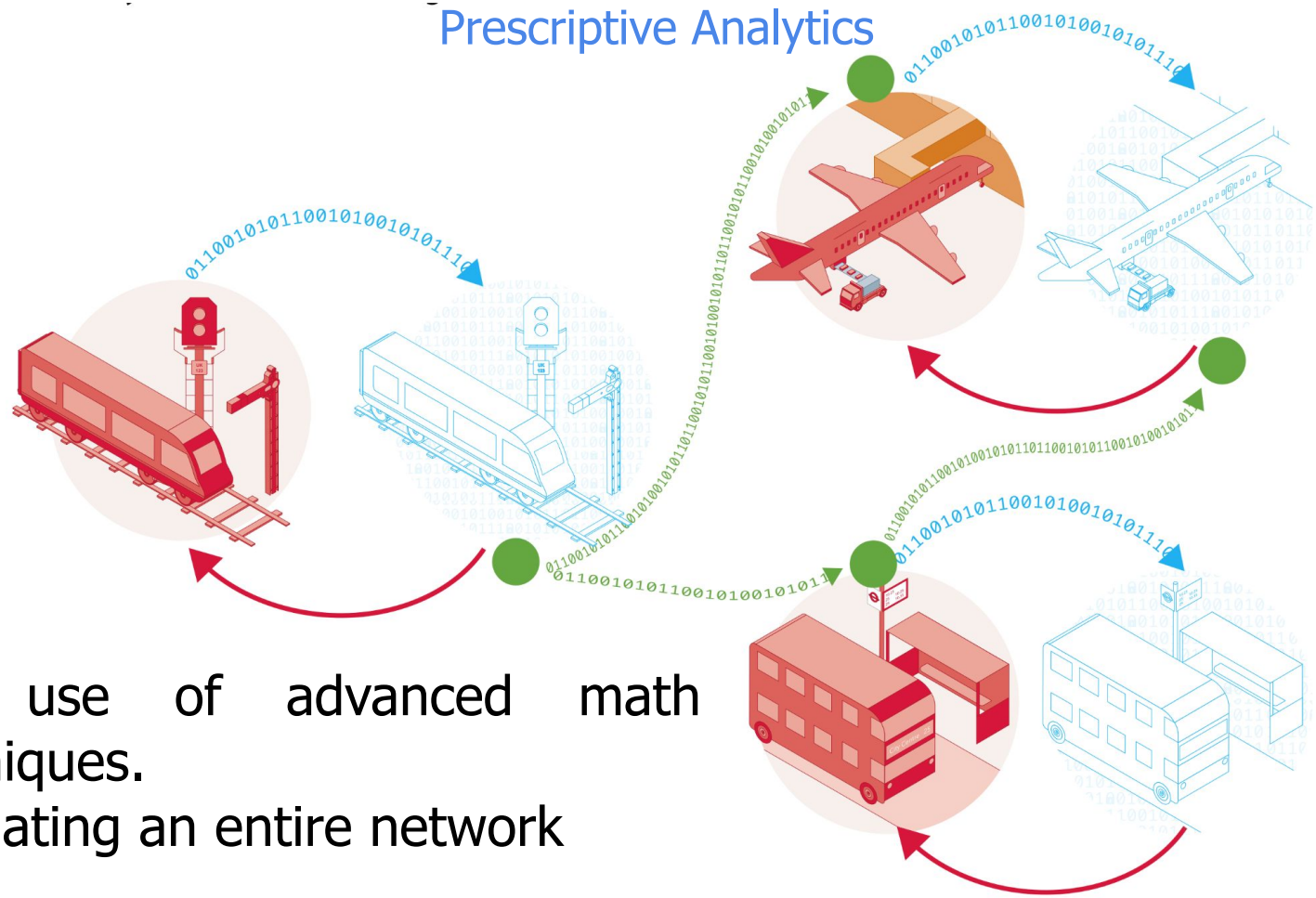
Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.1: Digital Twin's Principales

Digital Twins with ML and DL

Two widely used ML areas in DT

Prescriptive Analytics



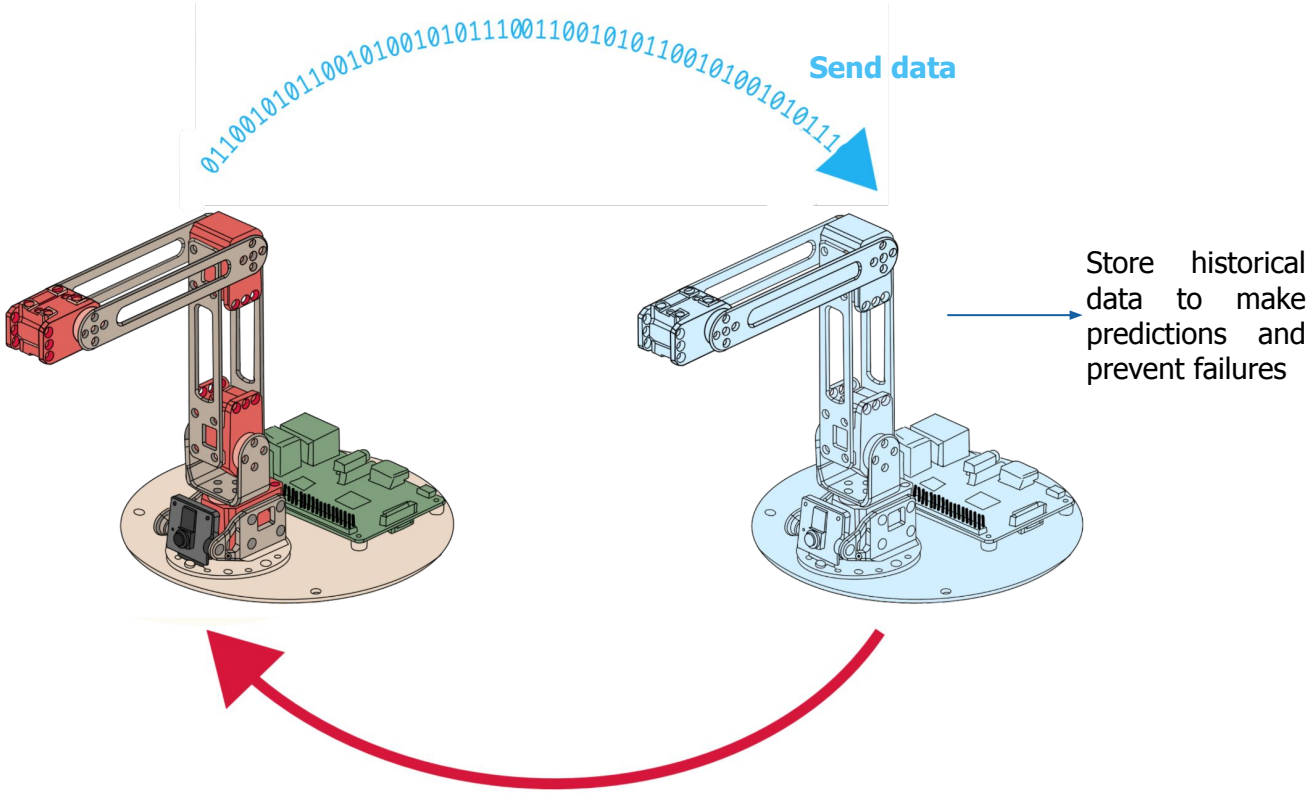
- The use of advanced math techniques.
- Simulating an entire network

Part I.2.1: Digital Twin's Principales

Digital Twins with ML and DL

Two widely used ML areas in DT

Diagnostic and Predictive Analytics



- Diagnose potential problems and predict future behaviors of the system.

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.2: Digital Twins for resilience

Numerous of Articles

In CPSs:	In Production:	In Healthcare:
creation of conceptual framework	Use CDT for detecting failures	Resilience of material
Predict potential failures and recommend actions to prevent them,	Predict potential failures and recommend actions to prevent them,	Predict potential failures

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part I.2.2: Digital Twins for resilience

Contribution

Creation of a CSDT

Cognitive Super-Digital Twin

Part II

Laboratoire
Décision et Information
Pour les Systèmes de Production

Section II.1 : Realization

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.1: Use case

The scenario involves an IoT system aligned with the Physical Twin, featuring a total of four sensors that help detect whether the room is occupied or not:

- A temperature sensor.
- A humidity sensor.
- A light sensor.
- A CO2 sensor.

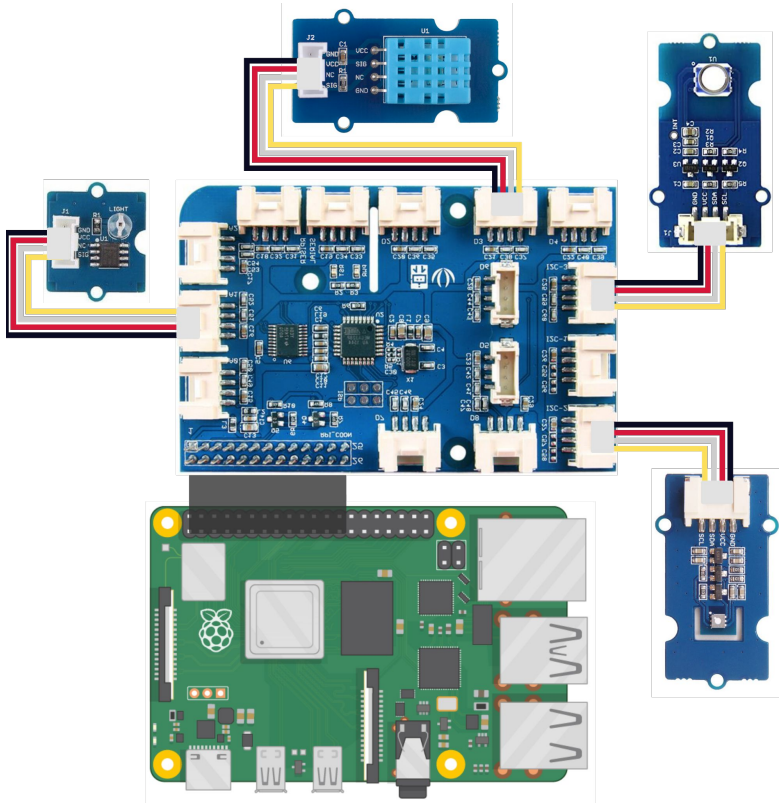
Each Sensor corresponds to a feature in the external dataset (Occupancy Detection Dataset).

The robot poppy ergo jr moves depending on the result of the model and the collected data.

Part II.1.2: Used Hardware and Technologies

First subsystem

- Raspberry Pi 3 Model B.
- GrovePi+ add-on board.
- Grove Temperature & Humidity Sensor (DHT11).
- Grove Barometer (High-Accuracy)
- Grove Light Sensor.
- Grove VOC and eCO2 Gas Sensor (SGP30).

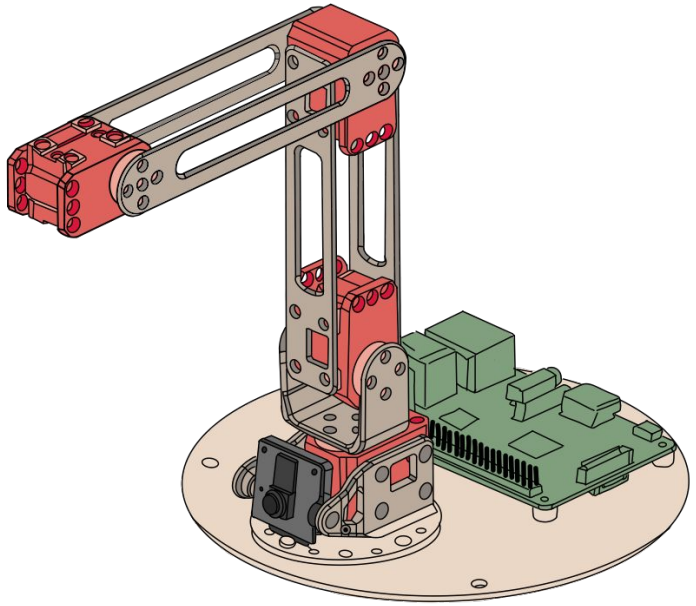


Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.2: Used Hardware and Technologies

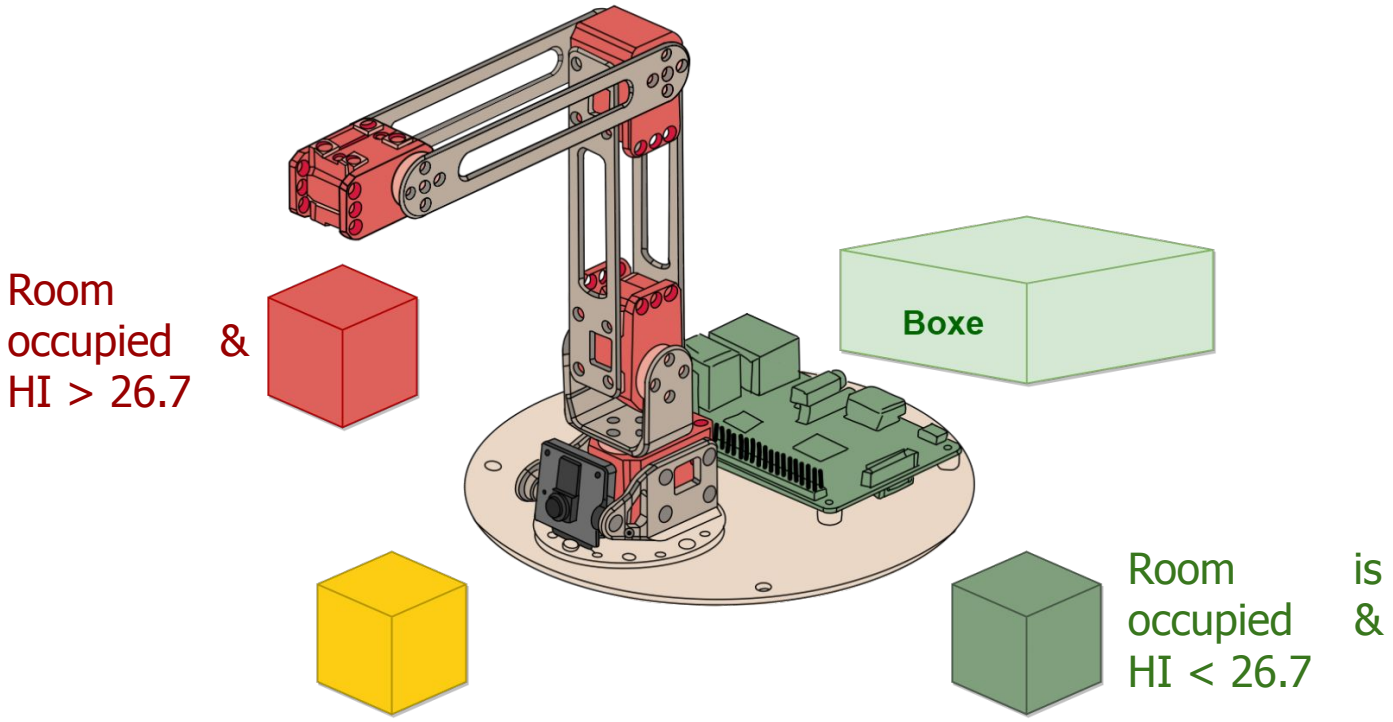
Second Subsystem - Poppy Ergo Jr

- Raspberry Pi 3 Model B.
- Six motors.
- 3D printed parts.



Part II.1.2: Used Hardware and Technologies

Second Subsystem - Poppy Ergo Jr



Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.2: Used Hardware and Technologies

Second Subsystem - Poppy Ergo Jr

$$HI = c_1 + c_2 \cdot T + c_3 \cdot RH + c_4 \cdot T \cdot RH + c_5 \cdot T^2 + c_6 \cdot RH^2 + c_7 \cdot T^2 \cdot RH + c_8 \cdot T \cdot RH^2 + c_9 \cdot T^2 \cdot RH^2$$

where :

HI is the Heat Index

T is the temperature in Celsius

RH is the relative humidity in percentage

$c_1, c_2, \ldots, c_9$ are the coefficients specific to the equation

$$c_1 = -8.78469475556$$

$$c_2 = 1.61139411$$

$$c_3 = 2.33854883889$$

$$c_4 = -0.14611605$$

$$c_5 = -0.012308094$$

$$c_6 = -0.0164248277778$$

$$c_7 = 0.002211732$$

$$c_8 = 0.00072546$$

$$c_9 = -0.000003582$$

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.2: Used Hardware and Technologies

RabbitMQ - MQTT

- It is an extension to RabbitMQ that enables support for the MQTT protocol. MQTT is a lightweight messaging protocol designed for efficient communication between devices or client applications in constrained or unreliable networks.

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.2: Used Hardware and Technologies

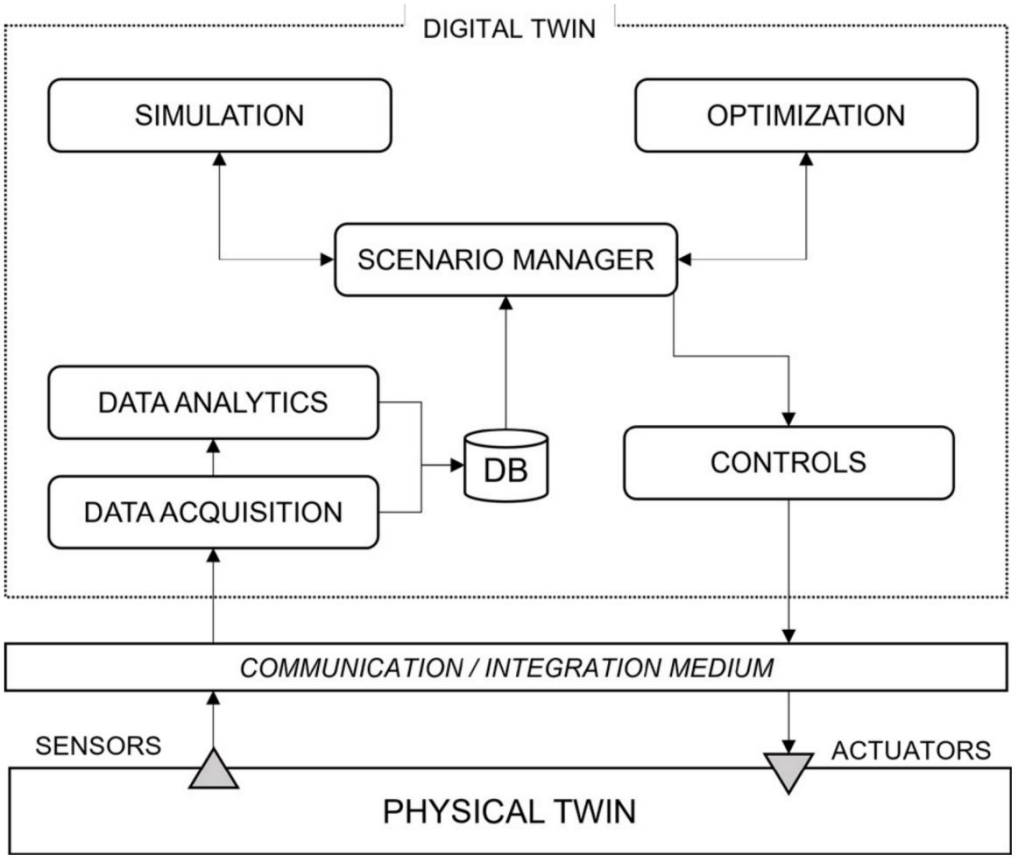
InfluxDB

- Open-source time series database written in Go programming language for storing and retrieving time series data.

Laboratoire
Décision et Information
Pour les Systèmes de Production

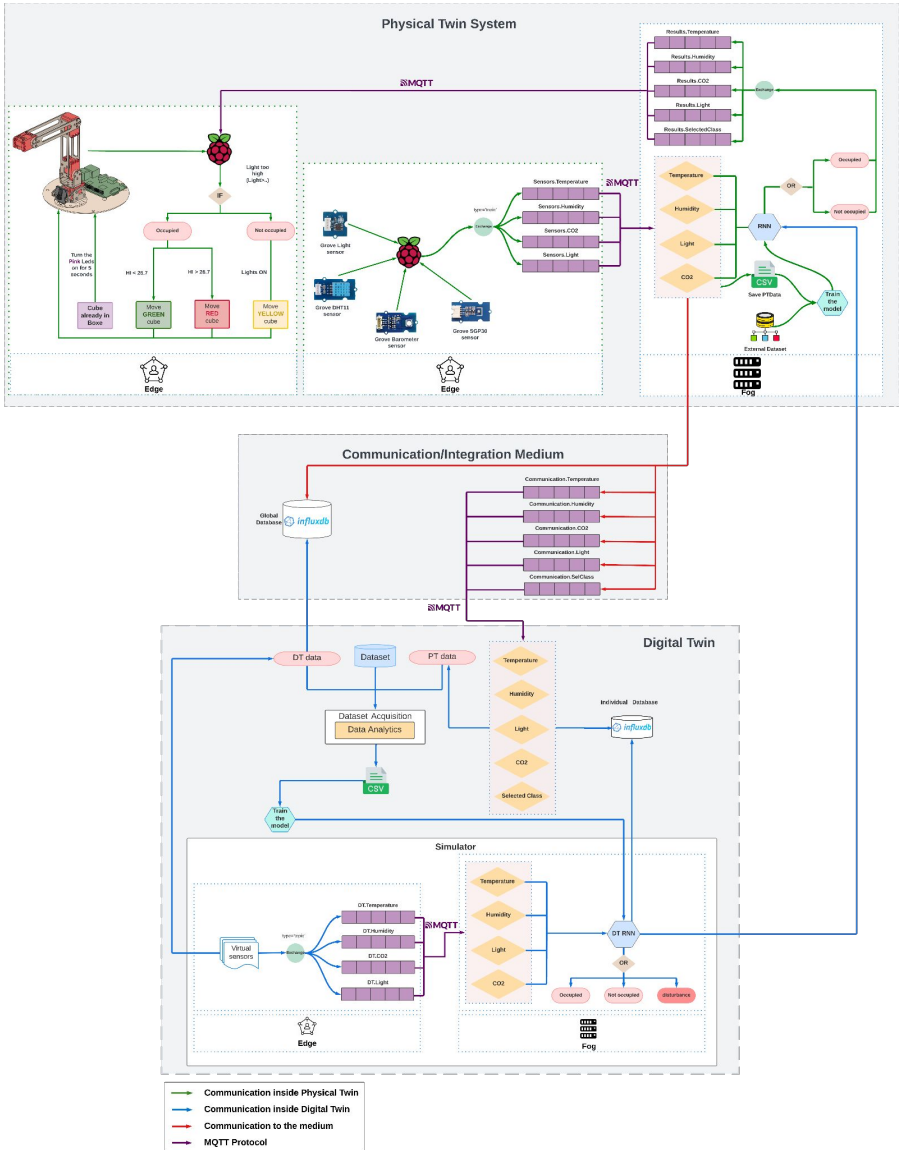
Part II.1.3: General Architecture

Inspired from:



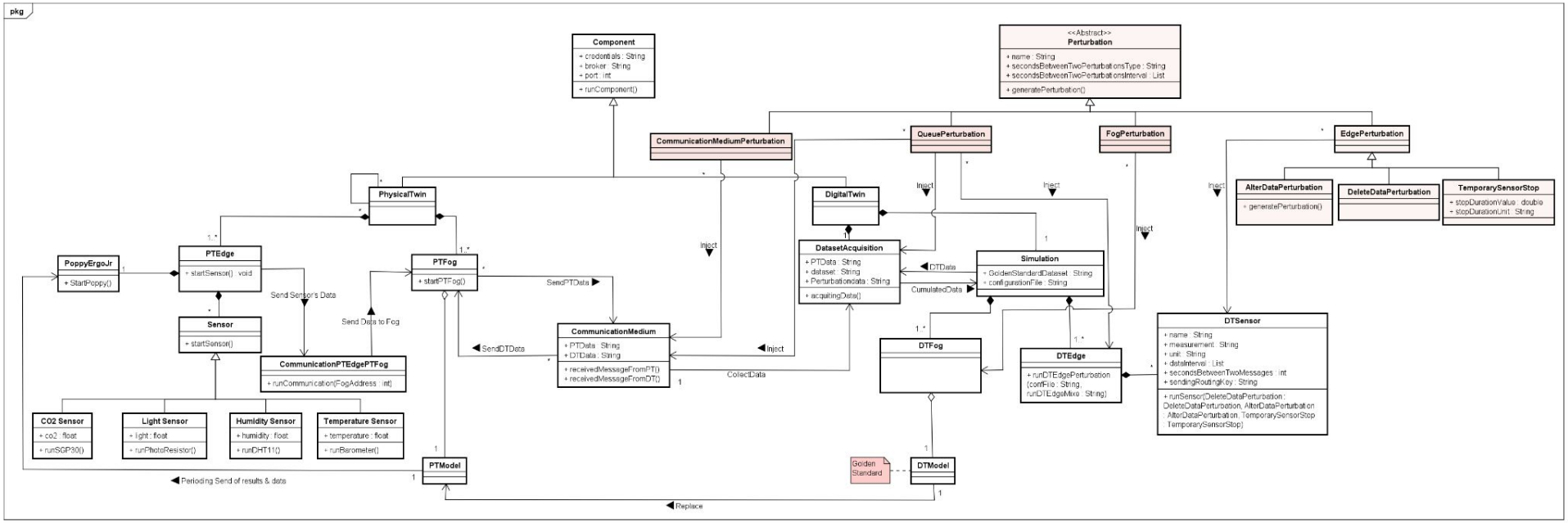
Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.3: General Architecture



Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.3: Class Diagram



Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.4: Sources of Data

The Digital Twin has Three sources of data:

- External Dataset (Occupancy Detection Dataset).
- Physical Twin's Data.
- Disturbed Data.

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.4: Sources of Data

External Dataset

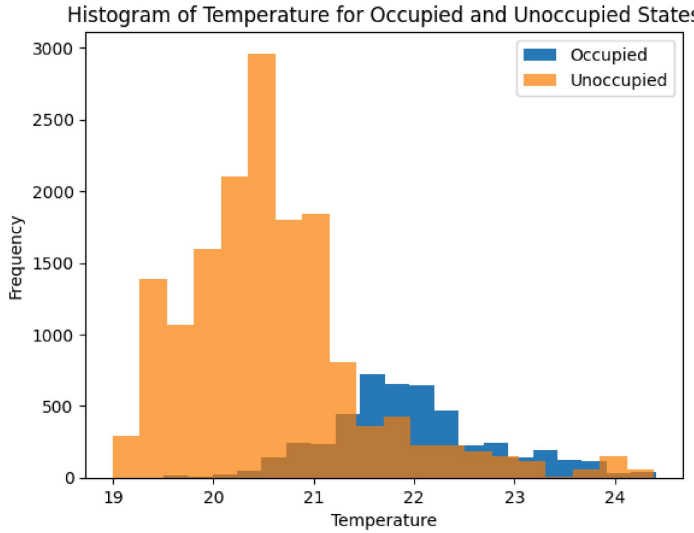
Date	Temperature	Humidity	Light	CO2	Occupancy
2015-02-02 14:19:00	23.7	26.272	585.2	749.2	1
2015-02-02 14:19:59	23.718	26.29	578.4	760.4	1
2015-02-02 14:21:00	23.73	26.23	572.66	769.66	1
2015-02-02 14:22:00	23.7225	26.125	493.75	774.75	1
2015-02-02 14:23:00	23.754	26.2	488.6	779	1

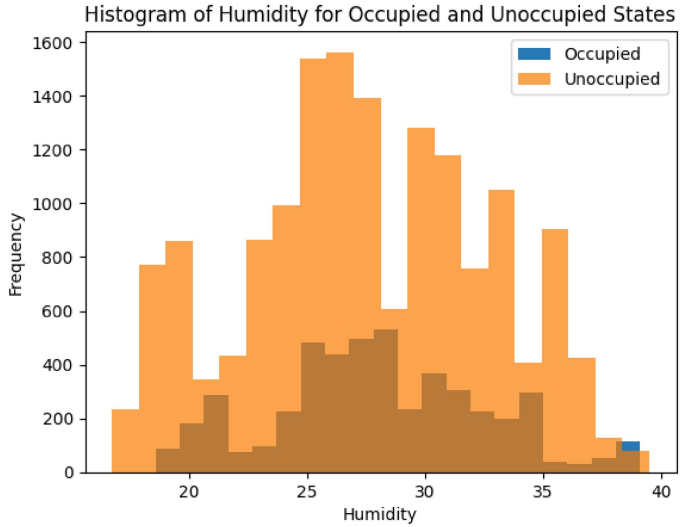
University of California Irvine - Machine Learning Repository

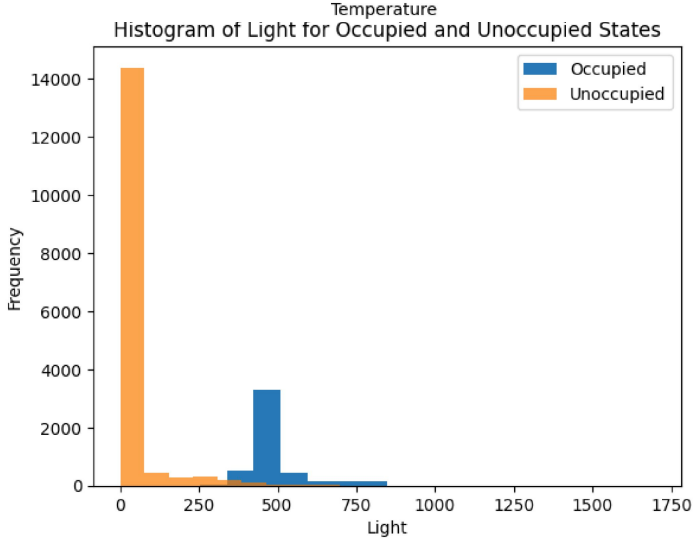
Laboratoire
Décision et Information
Pour les Systèmes de Production

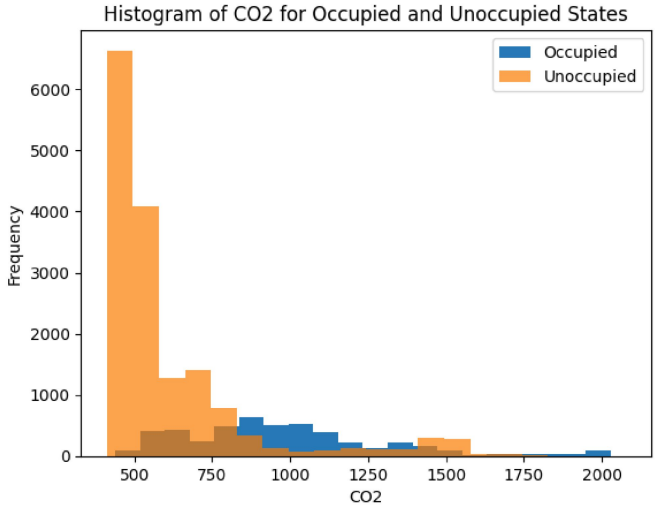
Part II.1.4: Sources of Data

External Dataset









Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.4: Sources of Data

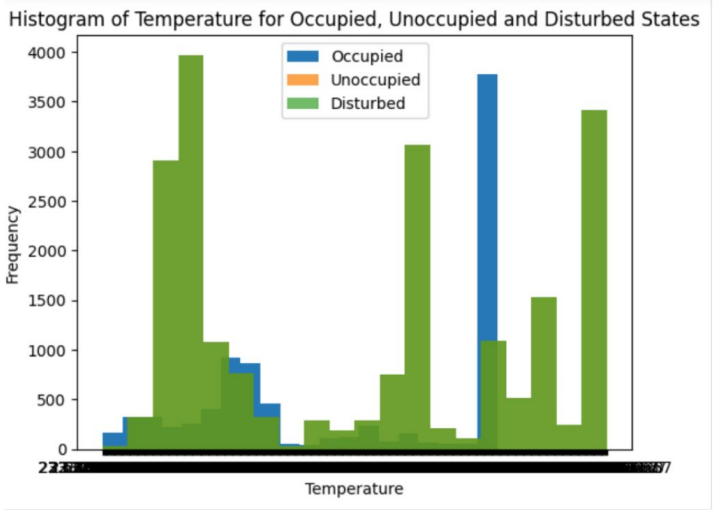
Physical Twin's Dataset

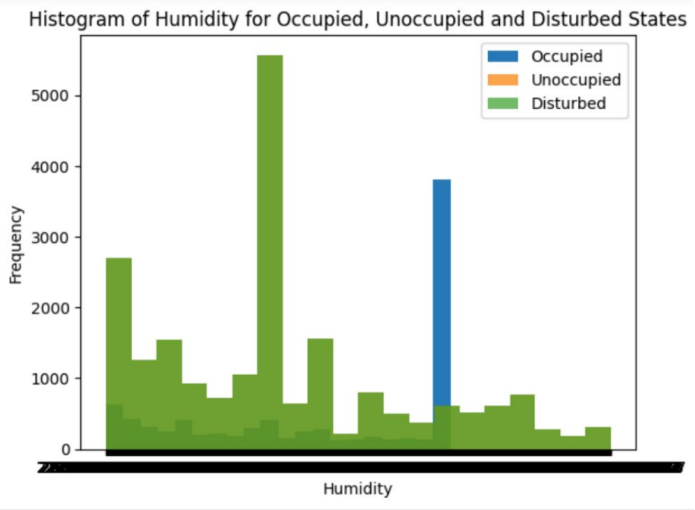
Date	Temperature	Humidity	Light	CO2	Occupancy
2023-06-09 15:33:00	26.12	17.0	766.0	809.0	1
2023-06-09 15:33:33	26.18	17.2	766.0	808.4	1
2023-06-09 15:34:03	25.73	18.23	572.66	769.66	1
2023-06-09 15:34:25	25.74	18.125	600.75	774.75	1
2023-06-09 15:35:06	25.754	17.2	568.6	779	1

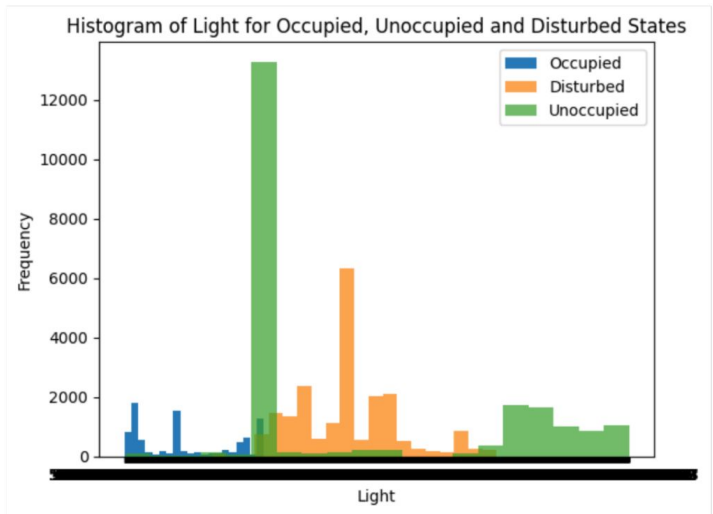
Laboratoire
Décision et Information
Pour les Systèmes de Production

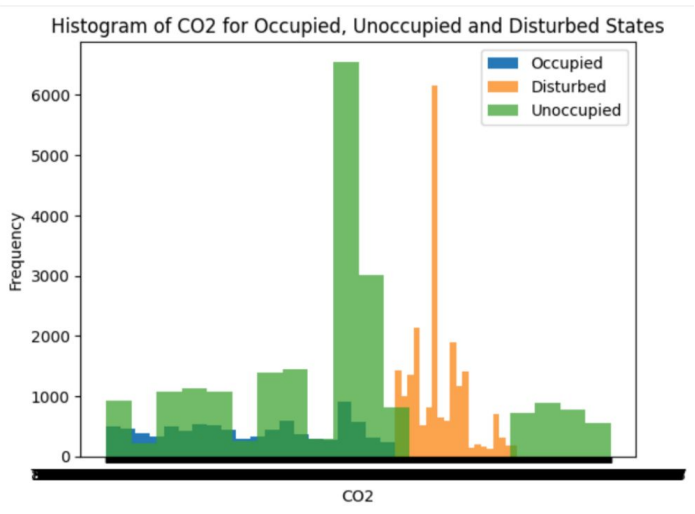
Part II.1.4: Sources of Data

Physical Twin's Dataset









Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.4: Sources of Data

Disturbed Data


```
{
        "Sensor name":"Temperature 1",
        "Measurement":"Temperature",
        "topic":"DT/temperature",
        "Unit":"°C",
        "Data type":"double",
        "DTData":[
           {
              "Occupancy": 0,
              "Data Interval":[19.1,24.39]
           },
           {
              "Occupancy": 1,
              "Data Interval":[20.29,26]
           },
           {
              "Occupancy": 2,
              "Data Interval":[
                 [-100,10],
                 [50,1000]
              ]
           }
        ]
     }
```

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.4: Selecting a Model for the Physical Twin

Train ML & DL models on the External dataset.

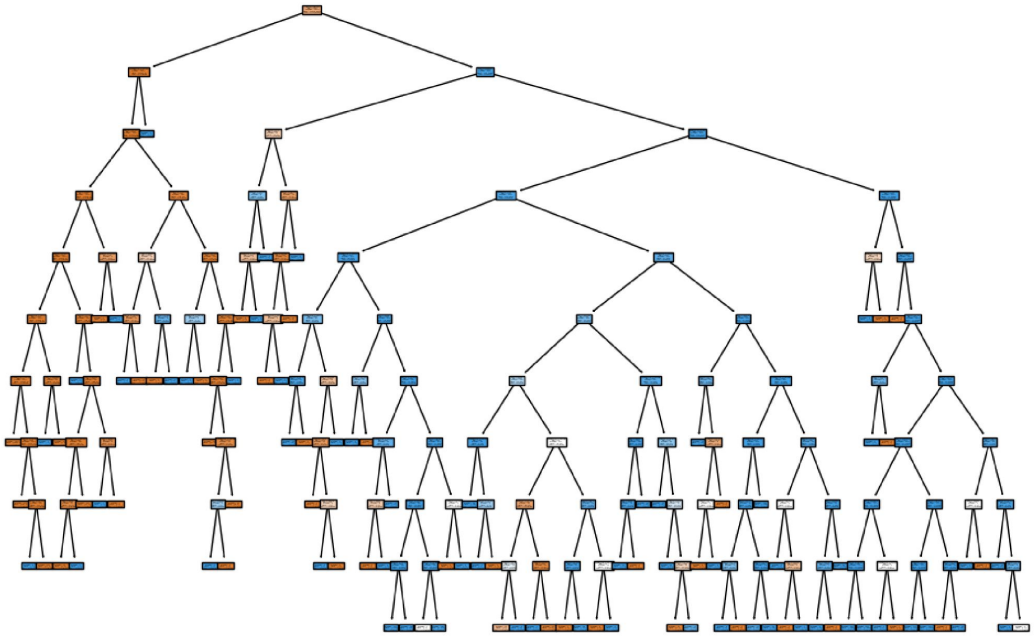
- Decision Trees.
- Random Forests.
- KNN.
- Naive Bayes.
- RNN.
- MLP.

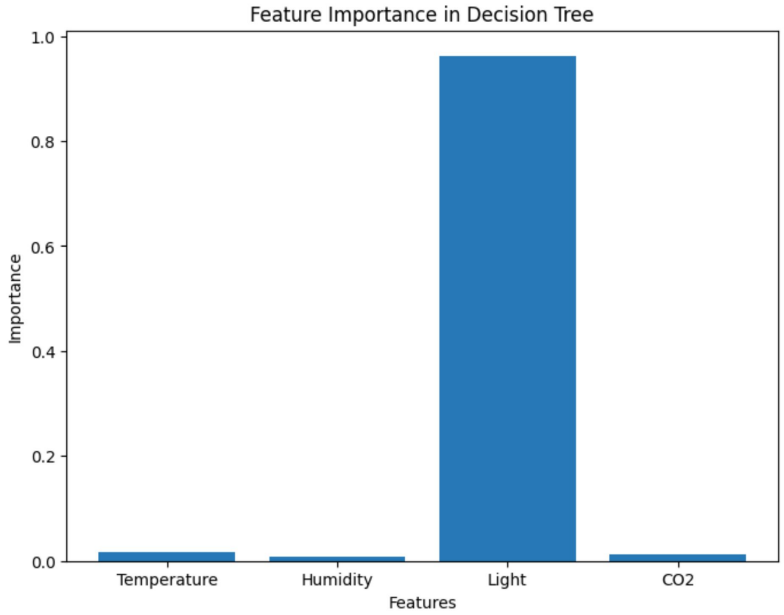
The training is done either with GridSearch or RandomSearch

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.4: Selecting a Model for the Physical Twin

Physical Twin - Decision Tree



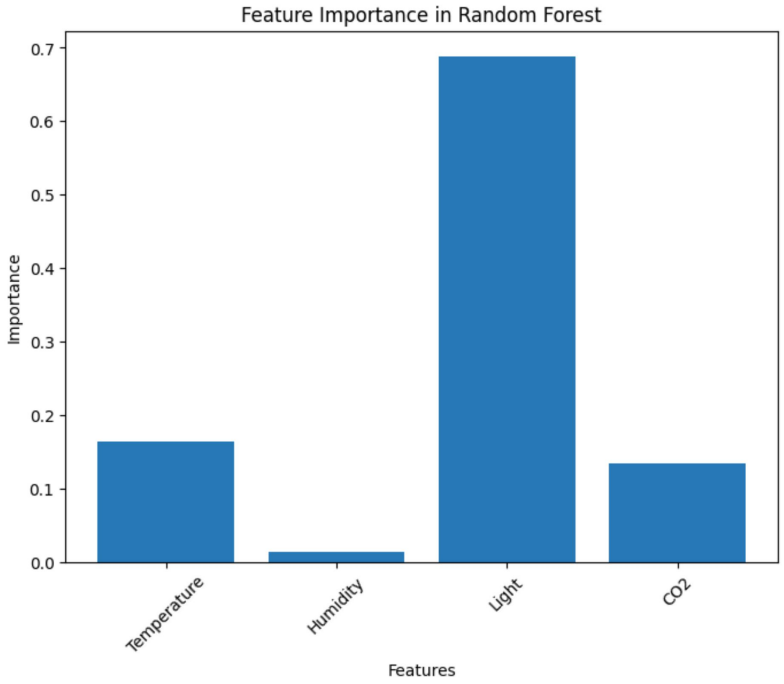


- Accuracy: 99.14%
- Precision: 97.73%
- Recall: 98.47%
- F1_Score: 98.10%

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.4: Selecting a Model for the Physical Twin

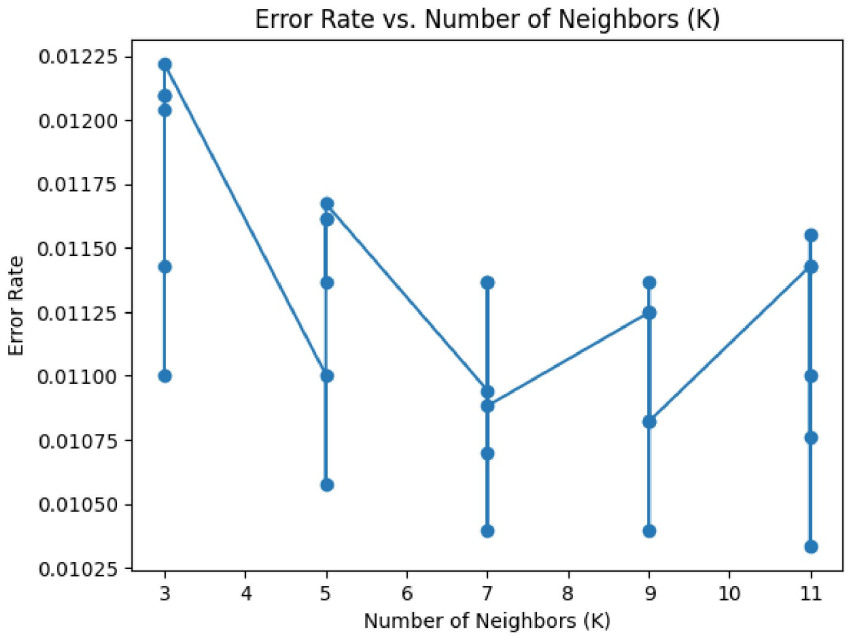
Physical Twin - Random Forest

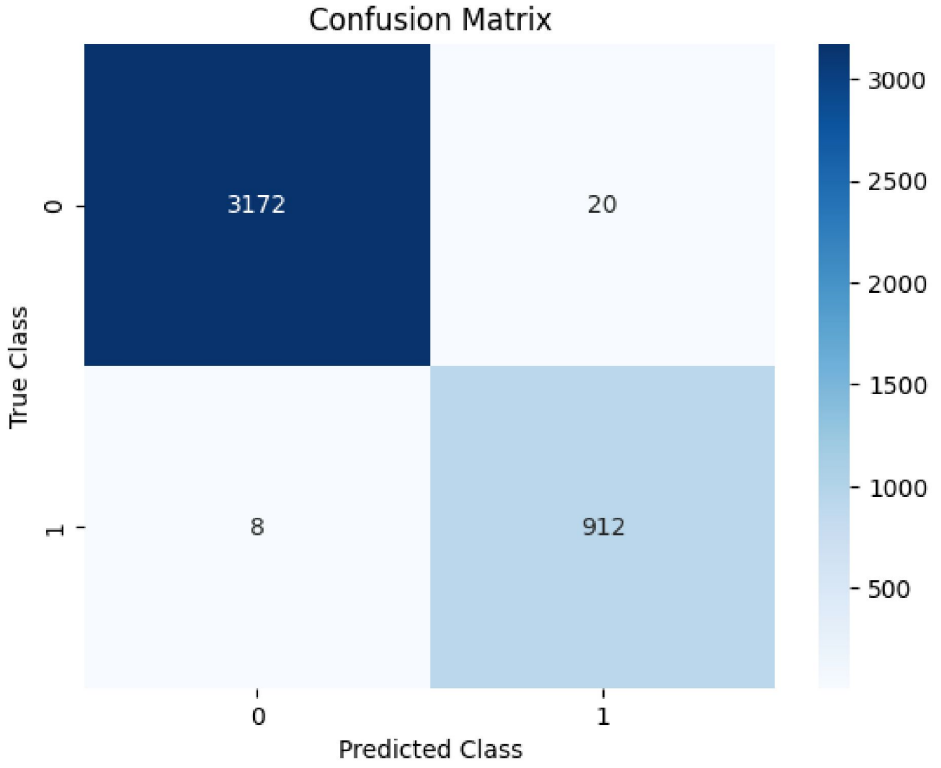


- Accuracy: 99.34%
- Precision: 97.85%
- Recall: 98.23%
- F1_Score: 98.54%

Part II.1.4: Selecting a Model for the Physical Twin

Physical Twin - KNN



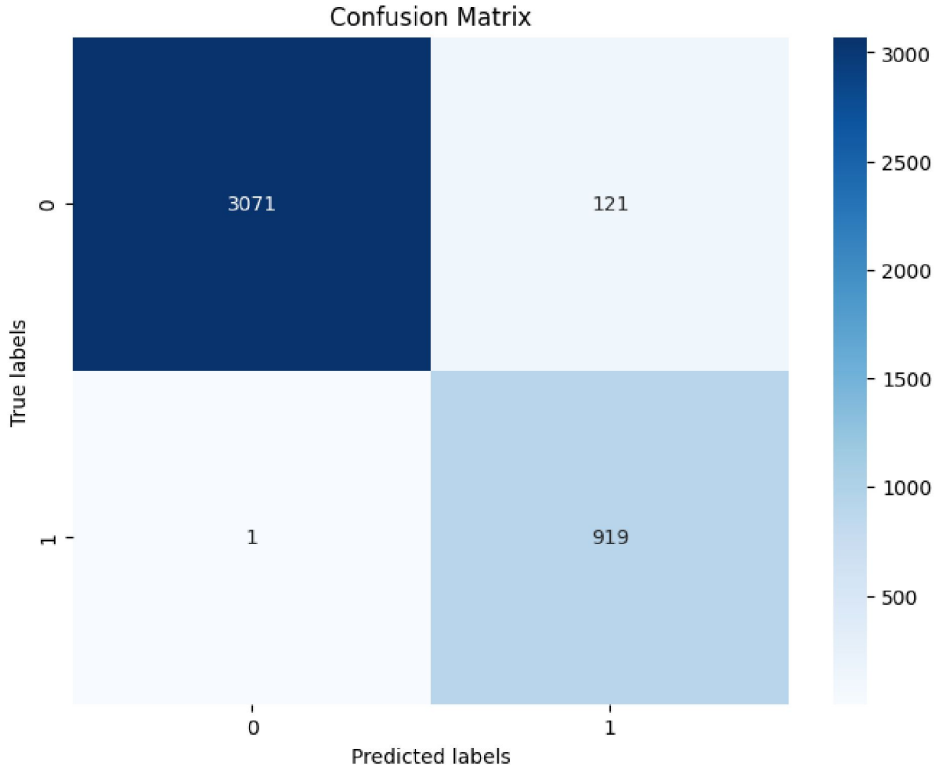


- Accuracy: 99.31%
- Precision: 97.85%
- Recall: 99.13%
- F1_Score: 98.48%

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.4: Selecting a Model for the Physical Twin

Physical Twin - Naive Bayes

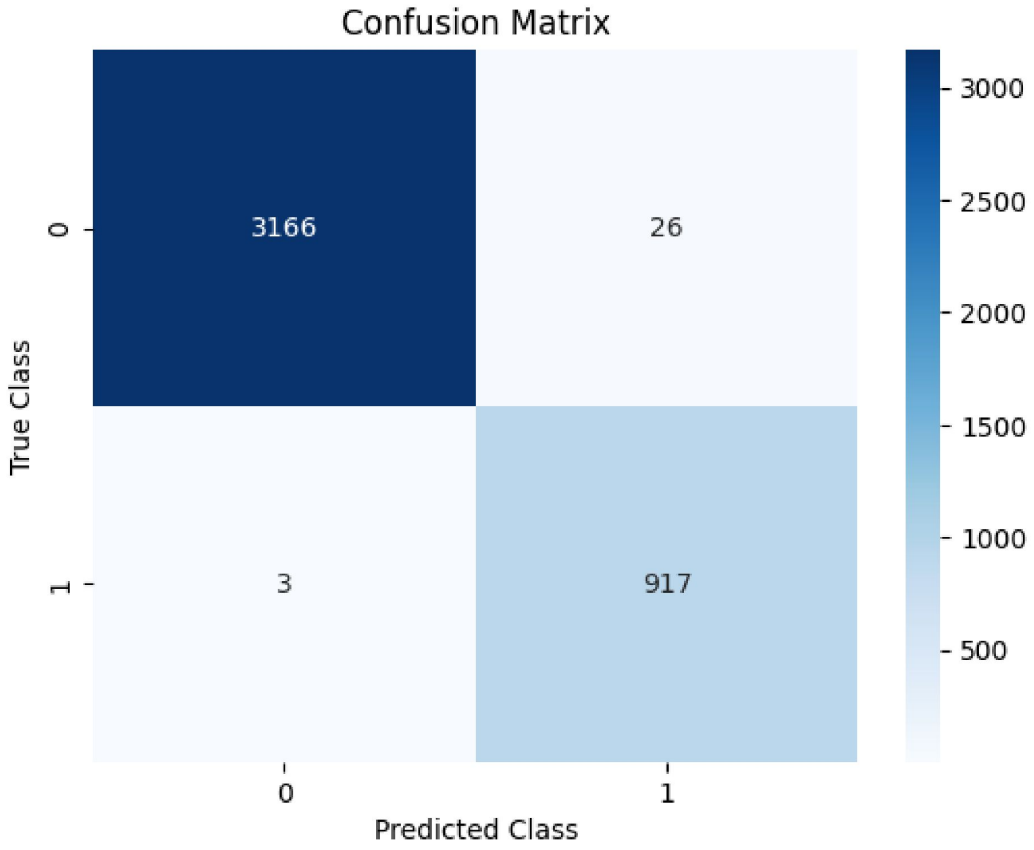


- Accuracy: 97.03%
- Precision: 98.36%
- Recall: 99.89%
- F1_Score: 93.77%

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.4: Selecting a Model for the Physical Twin

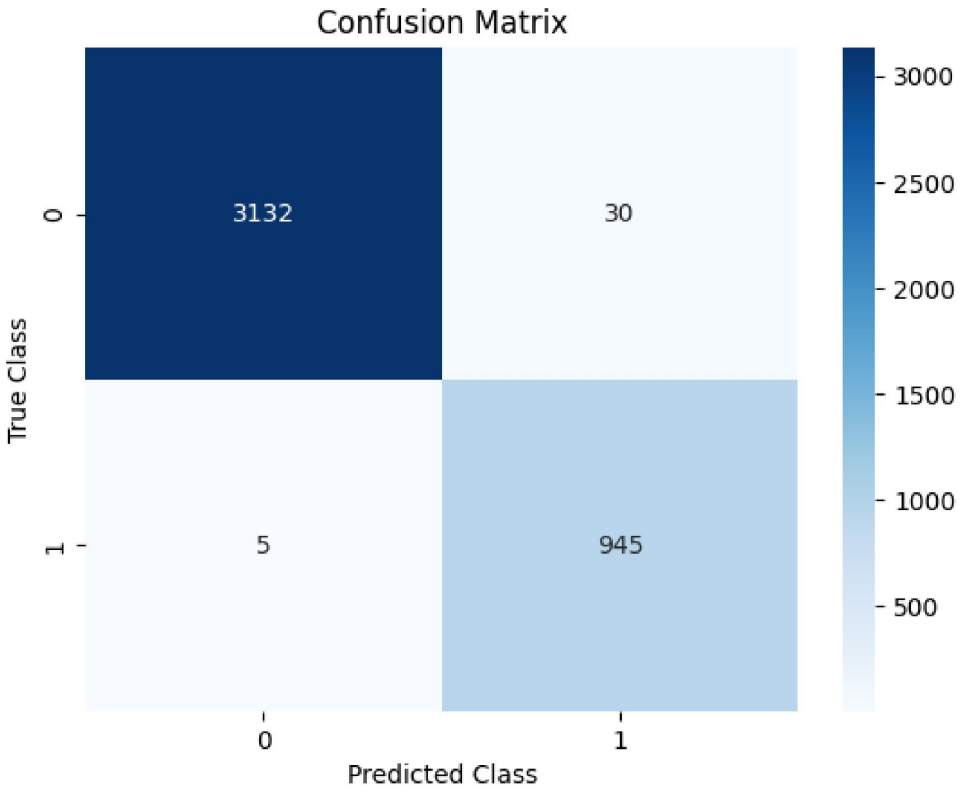
Physical Twin - MLP



- Accuracy: 97.27%
- Precision: 97.34%
- Recall: 99.46%
- F1_Score: 98.39%

Part II.1.4: Selecting a Model for the Physical Twin

Physical Twin - RNN

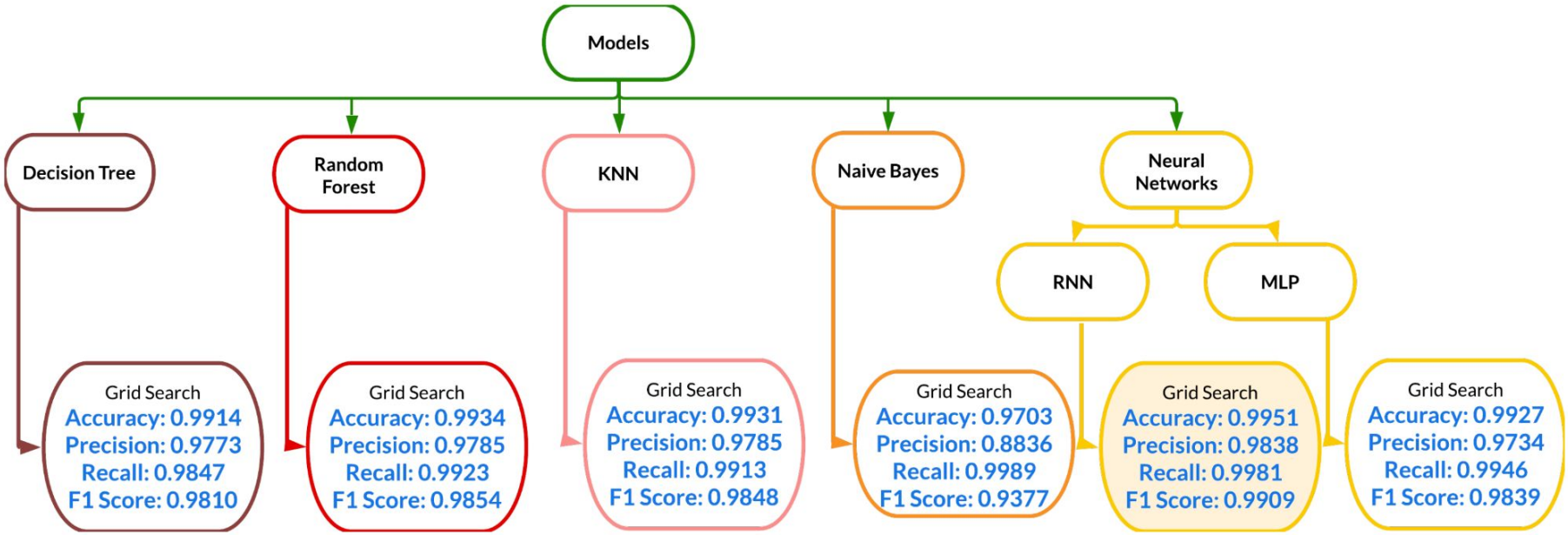


- Accuracy: 99.51%
- Precision: 97.38%
- Recall: 99.81%
- F1_Score: 98.09%

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.4: Selecting a Model for the Physical Twin

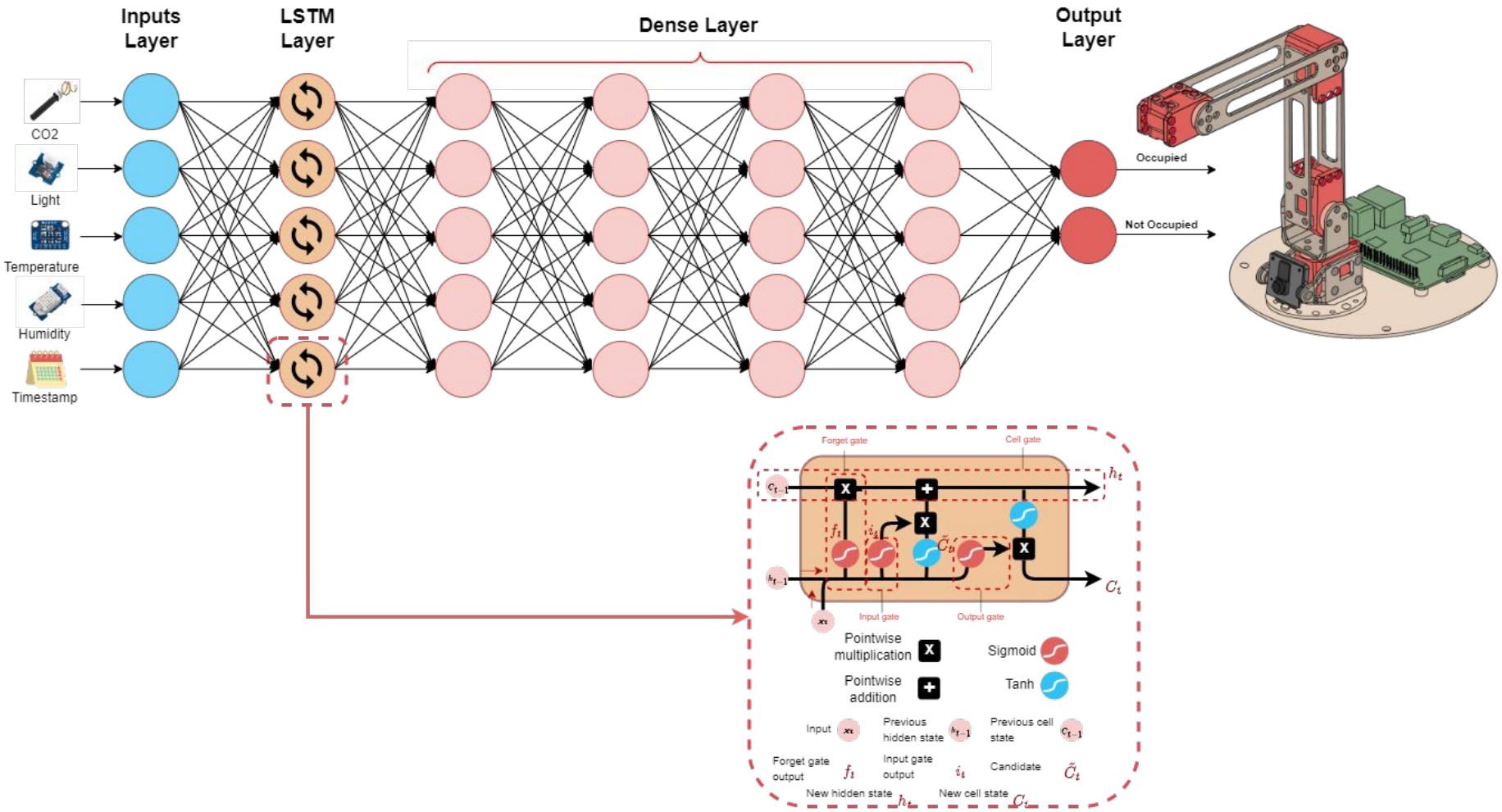
Physical Twin's Models



Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.1.4: Selecting a Model for the Physical Twin

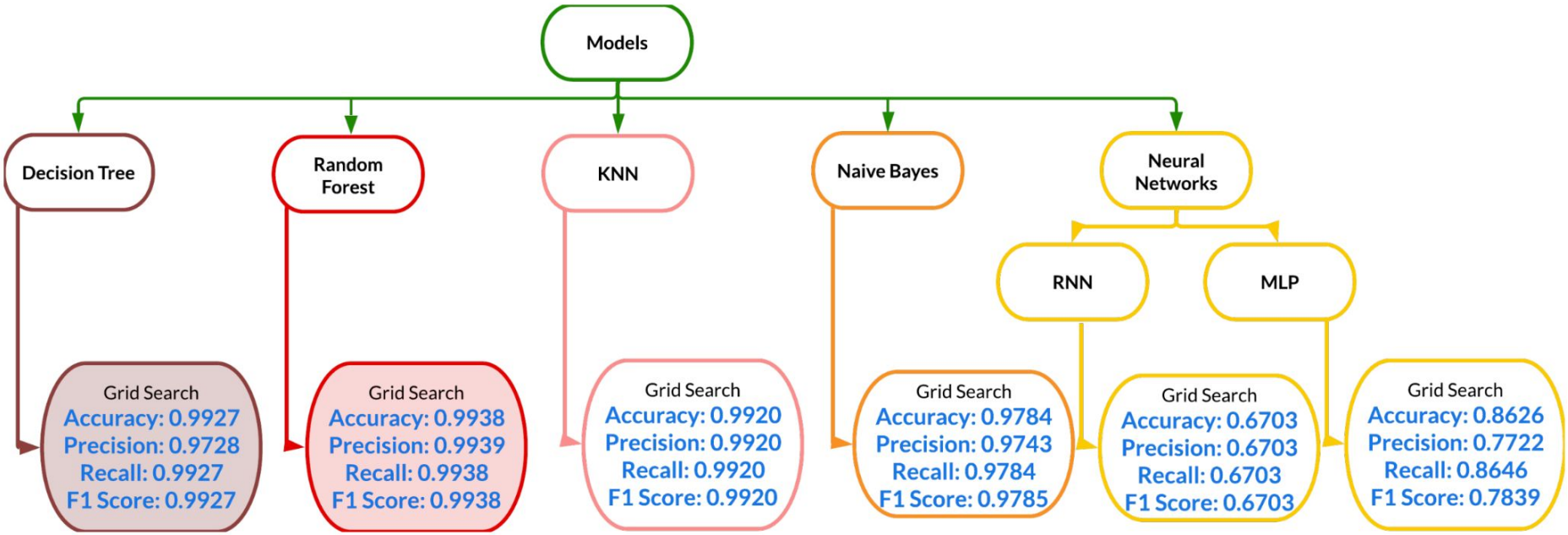
Physical Twin's Selected Model



Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.4: Selecting a Model for the Digital Twin

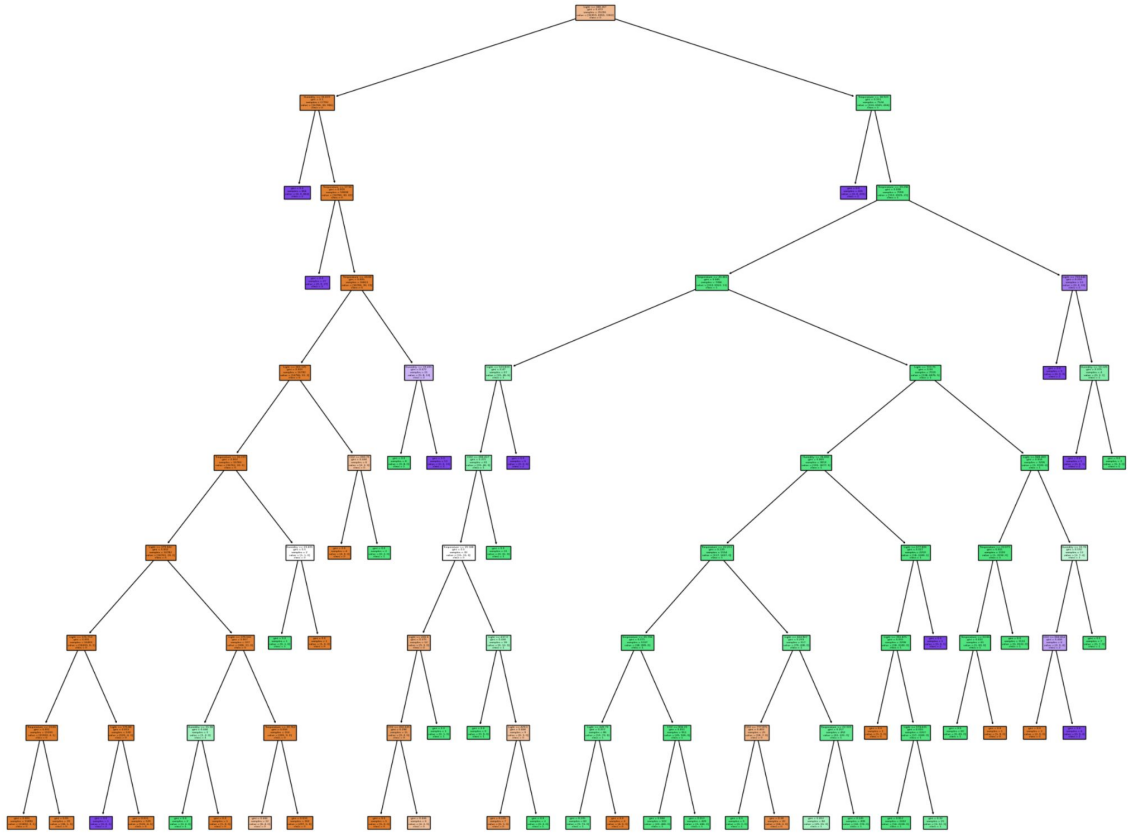
Digital Twin's Models

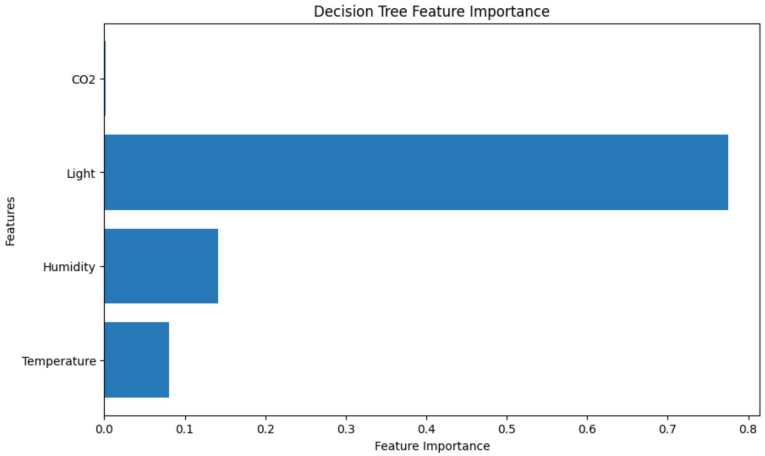


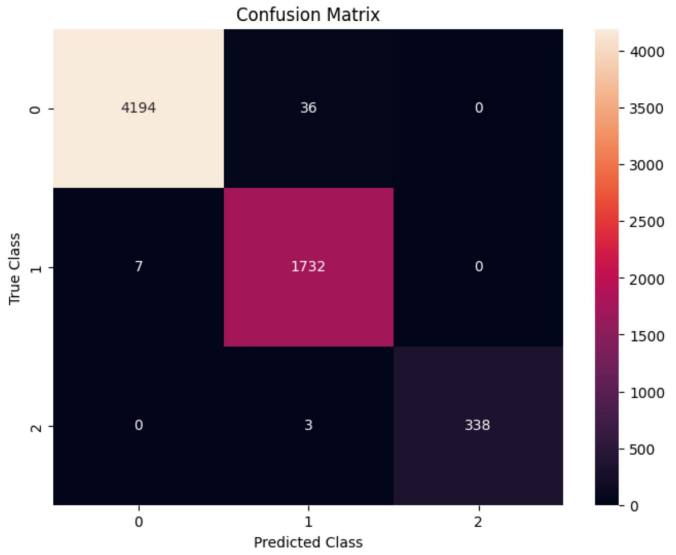
Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.4: Selecting a Model for the Digital Twin

Physical Twin - Decision Tree



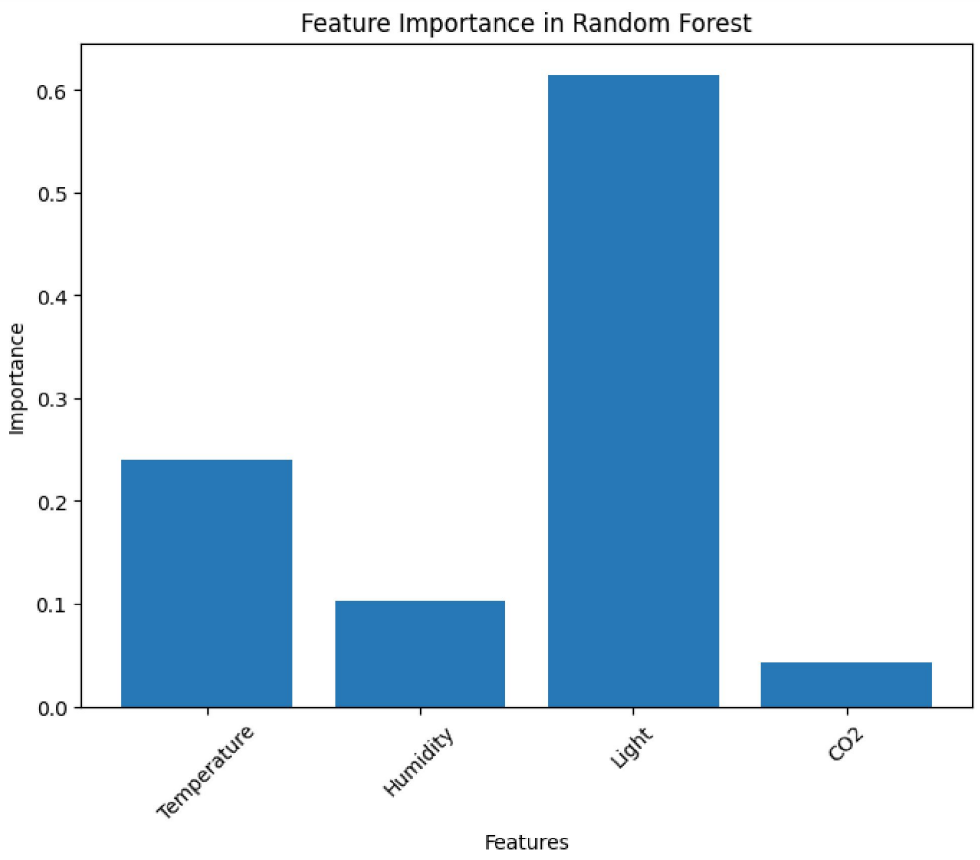




Laboratoire
Décision et Information
Pour les Systèmes de Production

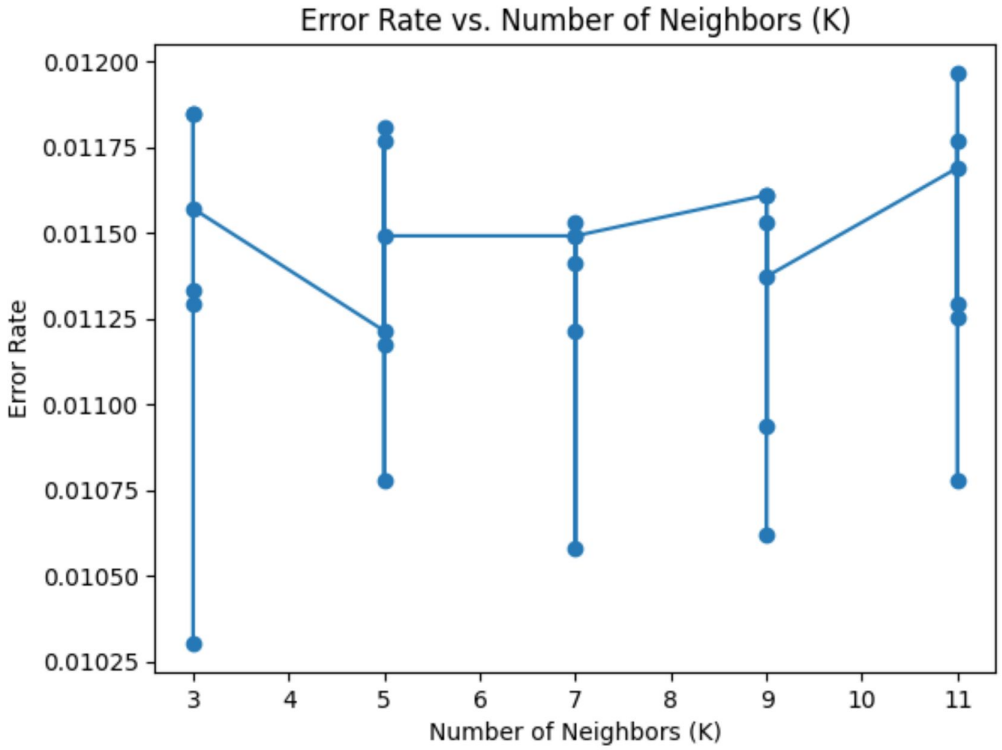
Part II.4: Selecting a Model for the Digital Twin

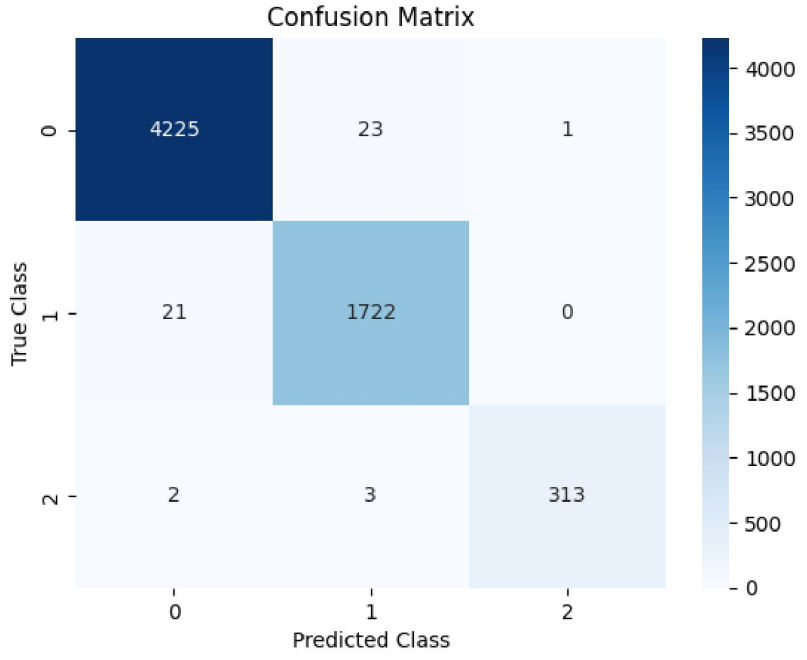
Physical Twin - Random Forest



Part II.4: Selecting a Model for the Digital Twin

Physical Twin - KNN





Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.4: Selecting a Model for the Digital Twin

Decision Trees perform better on tabular data.

Laboratoire
Décision et Information
Pour les Systèmes de Production

Section II.2 : Implementation

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.2: Implementation

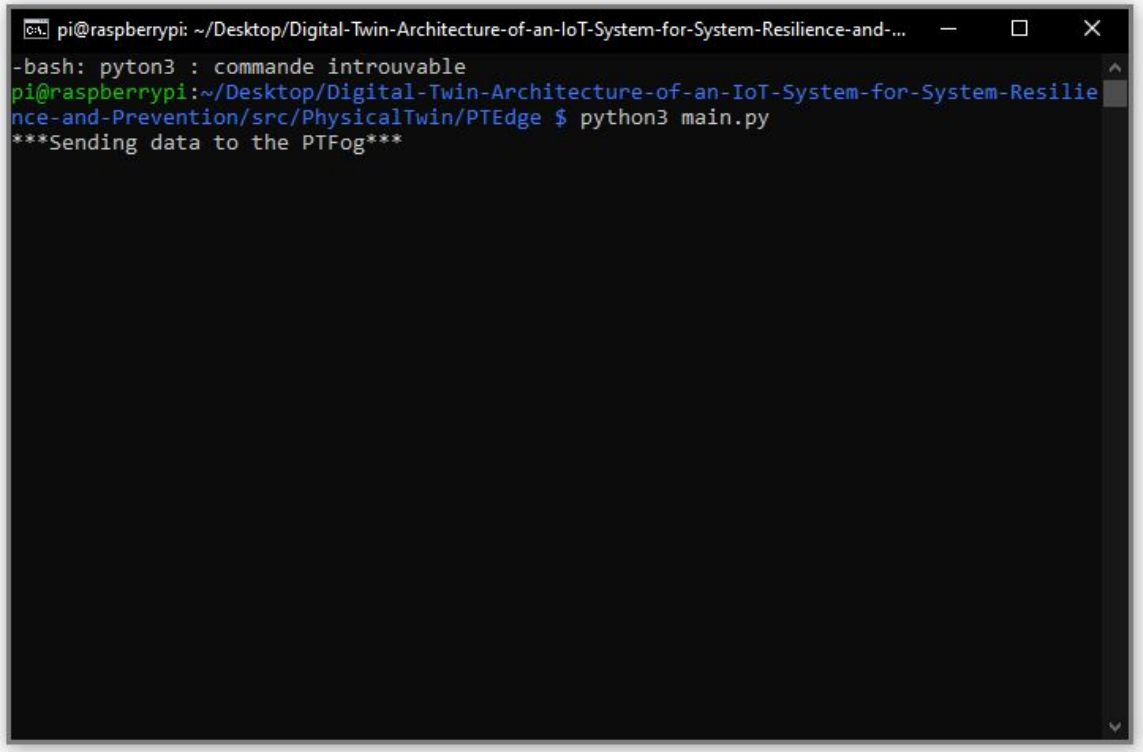
Starting the Database

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.2.1: Implementation

Sub-System 1 - Raspberry Pi

Physical Twin's Edge



- Collecting the data and send it to its corresponding Fog.

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.2.1: Implementation

Sub-System 1 - Raspberry Pi

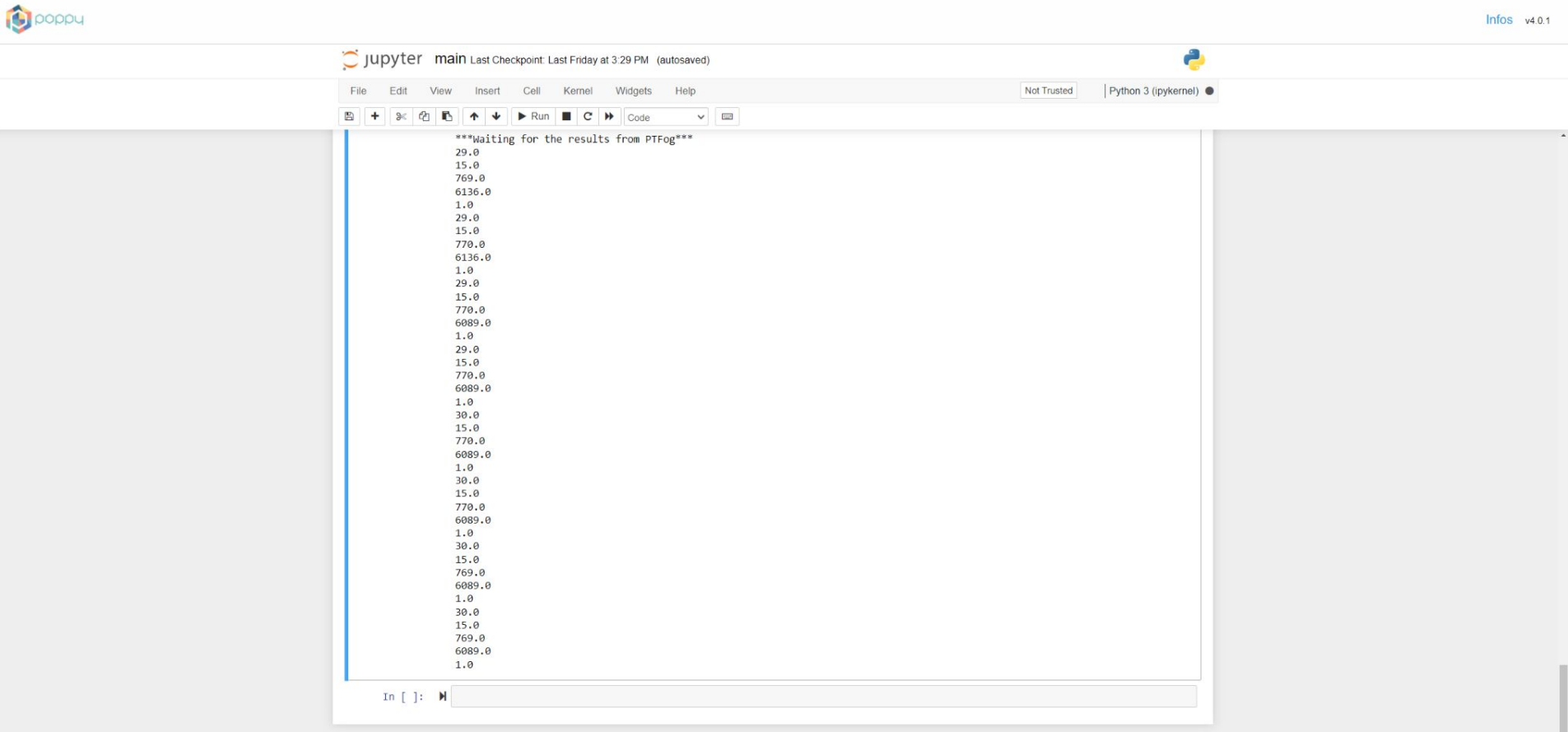
Physical Twin's Fog

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL                                   Python
re names, but DecisionTreeClassifier was fitted with feature names
  warnings.warn(
1
round:  1
C:\Users\msmat\AppData\Local\Programs\Python\Python39\lib\site-packages\sklearn\base.py:439: UserWarning: X does not have valid featu
re names, but DecisionTreeClassifier was fitted with feature names
  warnings.warn(
1
round:  0
***Sending results to the second Edge Poppy Ergo Jr***
Send data to Poppy Ergo Jr
C:\Users\msmat\AppData\Local\Programs\Python\Python39\lib\site-packages\sklearn\base.py:439: UserWarning: X does not have valid featu
re names, but DecisionTreeClassifier was fitted with feature names
  warnings.warn(
1
round:  10
C:\Users\msmat\AppData\Local\Programs\Python\Python39\lib\site-packages\sklearn\base.py:439: UserWarning: X does not have valid featu
re names, but DecisionTreeClassifier was fitted with feature names
  warnings.warn(
1
Ln 5, Col 12  Spaces: 4  UTF-8  CRLF  Python  3.9.13 64-bit  Go Live
```

Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.2.1: Implementation

Sub-System 2 - Poppy Ergo Jr



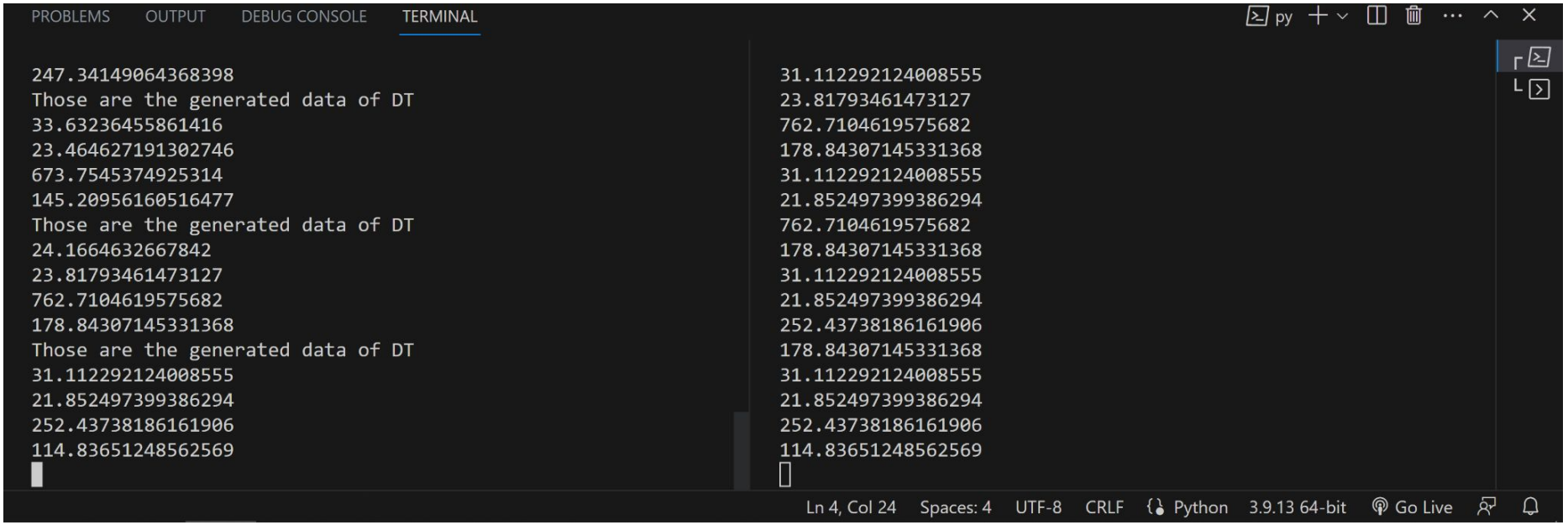
Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.2.1: Implementation

Digital Twin

DT Edge

DT Fog



Laboratoire
Décision et Information
Pour les Systèmes de Production

Part II.2.52: Prospective Endeavors

- Change the type of the used DT from a static twin to a dynamic one where data and responses are done in real-time.
- Change the level of granularity from a system twin to a process twin.
- Develop the use case and include reinforcement learning.
- Find a solution to include different environments in the use case to not have the problem of different distributed data.
- Add a cloud layer to upgrade the architecture to an Edge/Fog/Cloud and deal with big data.

Laboratoire
Décision et Information
Pour les Systèmes de Production

Thank you!