Project – Indoor activity detection using gas sensors

Luiz Miranda
18/03/2024

Indoor air quality (IAQ)

> 80% Indoors

In homes

Daily activities

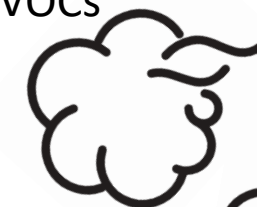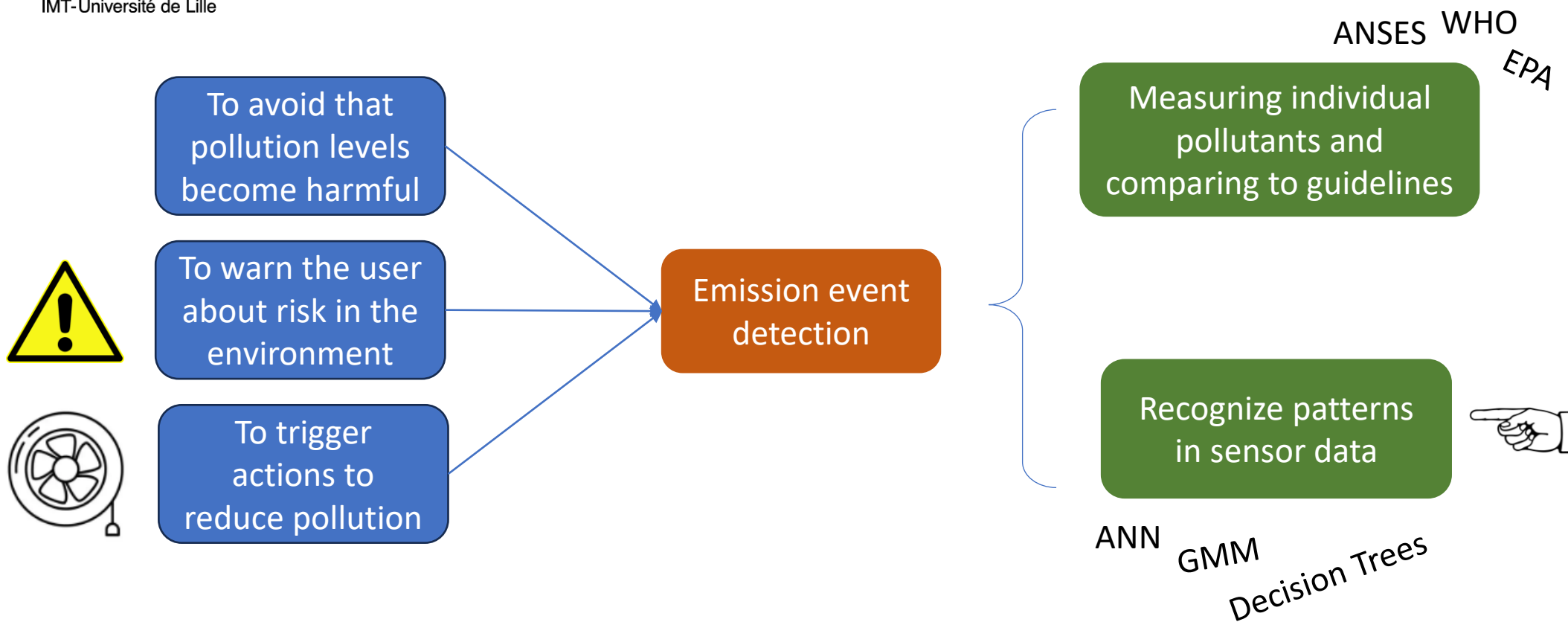VOCs

PM

Nausea
Asthma
Allergies
Irritations
Sick building Syndrome (SBS)

2

To avoid that pollution levels become harmful

To warn the user about risk in the environment

To trigger actions to reduce pollution

Emission event detection

Measuring individual pollutants and comparing to guidelines

ANSES WHO EPA

Recognize patterns in sensor data

ANN GMM Decision Trees

21 unique models of MOX

13 m² (46 m³) room

To simulated a set of usual indoor activities we proposed the following list:

- Static Activity (Working on the computer)
- Cooking / Eating (Eggs)
- Humid bathroom (Drying towels)
- Cleaning (vacuum and volatile products)
- Low intensity exercise
- Scented candles
- Aeration
- Manual work (hot glue / wood cutting)

activity	Started	Ended	Comments
Saber	14/11/2022 15:55	14/11/2022 16:42	
Aera	14/11/2022 16:42	14/11/2022 16:53	
Nett	14/11/2022 17:00	14/11/2022 17:15	floor: javel + vaseille
Asp	14/11/2022 17:15	14/11/2022 17:25	
AS1	15/11/2022 9:30	15/11/2022 10:00	
Aera	15/11/2022 10:00	15/11/2022 10:10	
Saber	15/11/2022 10:15	15/11/2022 10:45	
Asp	15/11/2022 11:10	15/11/2022 11:20	
Bougie	15/11/2022 11:20	15/11/2022 11:50	
SdB	15/11/2022 14:30	15/11/2022 15:30	
AS1	15/11/2022 15:34	15/11/2022 16:04	
Saber	15/11/2022 16:10	15/11/2022 16:45	
Bougie	15/11/2022 16:45	15/11/2022 17:00	
Aera	15/11/2022 17:18	15/11/2022 17:28	
BricoP	16/11/2022 10:00	16/11/2022 10:15	Hot glue
Saber	16/11/2022 10:35	16/11/2022 11:05	
Saber	16/11/2022 11:20	16/11/2022 11:40	
Asp	16/11/2022 14:35	16/11/2022 14:45	
AS1	16/11/2022 14:45	16/11/2022 15:15	
Aera	16/11/2022 15:15	16/11/2022 15:25	

Labelled activities with recorded duration

4
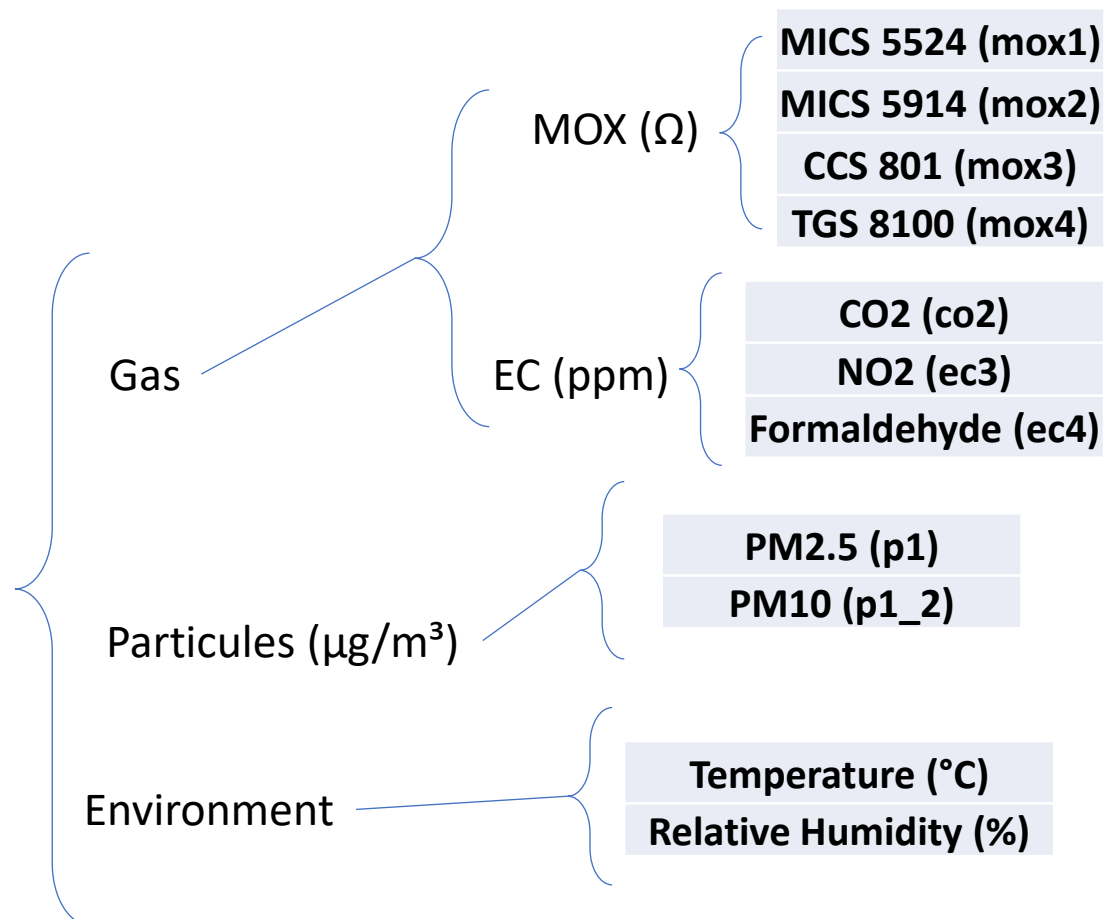
Libellium New

Sampling Time = 4 s

Gas

MOX (V)

| TGS 4161 |
| MICS 2714 |
| TGS 2442 |
| MICS 5524 |
| TGS 2602 |
| TGS 2620 |

Environment

| Temperature (°C) |
| Relative Humidity (%) |

| RH | Temp | 4161 | 2714 | 2442 | 5524 | 2602 | 2620 |

```
21/11/2022 11:11:46    37.904  23.300  0.474  0.331  11.574   0.258  0.227  0.785
21/11/2022 11:11:50    38.036  23.300  0.522  0.310  268.518  0.255  0.371  1.196
21/11/2022 11:11:54    38.057  23.280  0.550  0.296  370.370  0.252  0.431  1.316
21/11/2022 11:11:57    37.968  23.270  0.561  0.291  458.333  0.250  0.451  1.340
21/11/2022 11:12:01    37.868  23.280  0.567  0.289  432.870  0.250  0.456  1.342
21/11/2022 11:12:05    38.023  23.280  0.570  0.287  500.000  0.248  0.458  1.338
```

6

Piano (thick and thin)

Sampling Time = 10 s

Thick

MOX (Ω)

Thick	Thin
TGS 2603	GM 102B
TGS 2620	GM 302B
COM 31	GM 402B
H2SM 31	GM 502B
VOCM 31	GM 602B
TGS 2600	GM 702B
TGS 2602	GM 802B
SB 5100	MICS 5524
SBAQ 106	MICS 2714
	MICS 5914

Thin

element	date	piano_TGS2603I00	piano_TGS2620I00	piano_COM31I00	piano_H2SM31I00	piano_VOCM31I00	piano_TGS2600I00	piano_TGS2602I00	piano_SB5100I00	piano_SBAQ106I00
# Units		-	-	-	-	-	-	-	-	-
# Types		custom	custom	custom	custom	custom	custom	custom	custom	custom
# References		piano_TGS2603I00	piano_TGS2620I00	piano_COM31I00	piano_H2SM31I00	piano_VOCM31I00	piano_TGS2600I00	piano_TGS2602I00	piano_SB5100I00	piano_SBAQ106I00
# Sensor_visibility		visible	visible	visible	visible	visible	visible	visible	visible	visible
IMT_Thick	2022-11-14 00:00:00+0100	300430.1563	30498.78906	1723107.75	2078335.75	1709399.375	48975.01172	76671.95313	46717.46875	48220.27
IMT_Thick	2022-11-14 00:00:10+0100	300656.2813	30509.89063	1720434.25	2070651.375	1707974.75	48972.03906	76559.14063	47053.16016	46395.68
IMT_Thick	2022-11-14 00:00:20+0100	300028.5938	30546.36914	1710234.625	2104660.5	1706473	48990.30078	76548.07813	47389.05859	45363.54
IMT_Thick	2022-11-14 00:00:30+0100	300264.5625	30575.08008	1722924.75	2062511.375	1703642	49029.26953	76548.82813	47584.05078	48003
IMT_Thick	2022-11-14 00:00:40+0100	300044.25	30544.41016	1726786.5	2084093.75	1714282.5	49029.64063	76525.78125	46774.17969	44914.82
IMT_Thick	2022-11-14 00:00:50+0100	300021.7188	30499.17969	1735418.375	2084271.75	1709616.75	49022.44922	76510.82031	47450.03906	46643.48

Piano pico

MOX (Ω)
- **BME68x_gas**
- **SGP41_mox0**
- **SGP41_mox1**

Gas

CO2 (ppm)
- **Ens160_tco2**
- **BME68x_tco2**

TVOC (ppm)
- **Ens160_tvoc**
- **BME68x_tvoc**

Environment (BME68x)
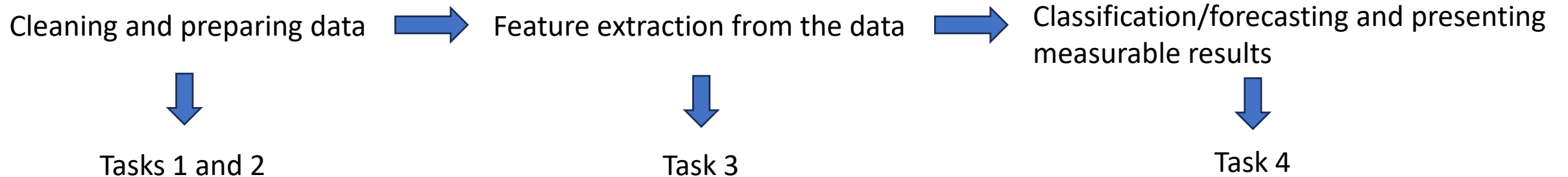- **Temperature (°C)**
- **Relative Humidity (%)**

Sampling Time = 10 s

8
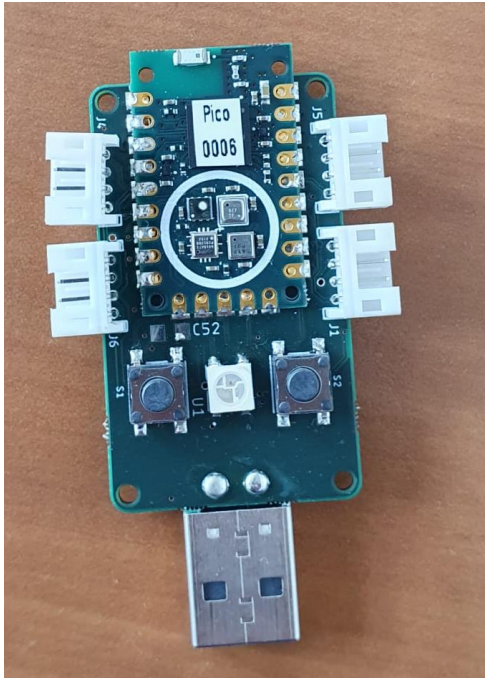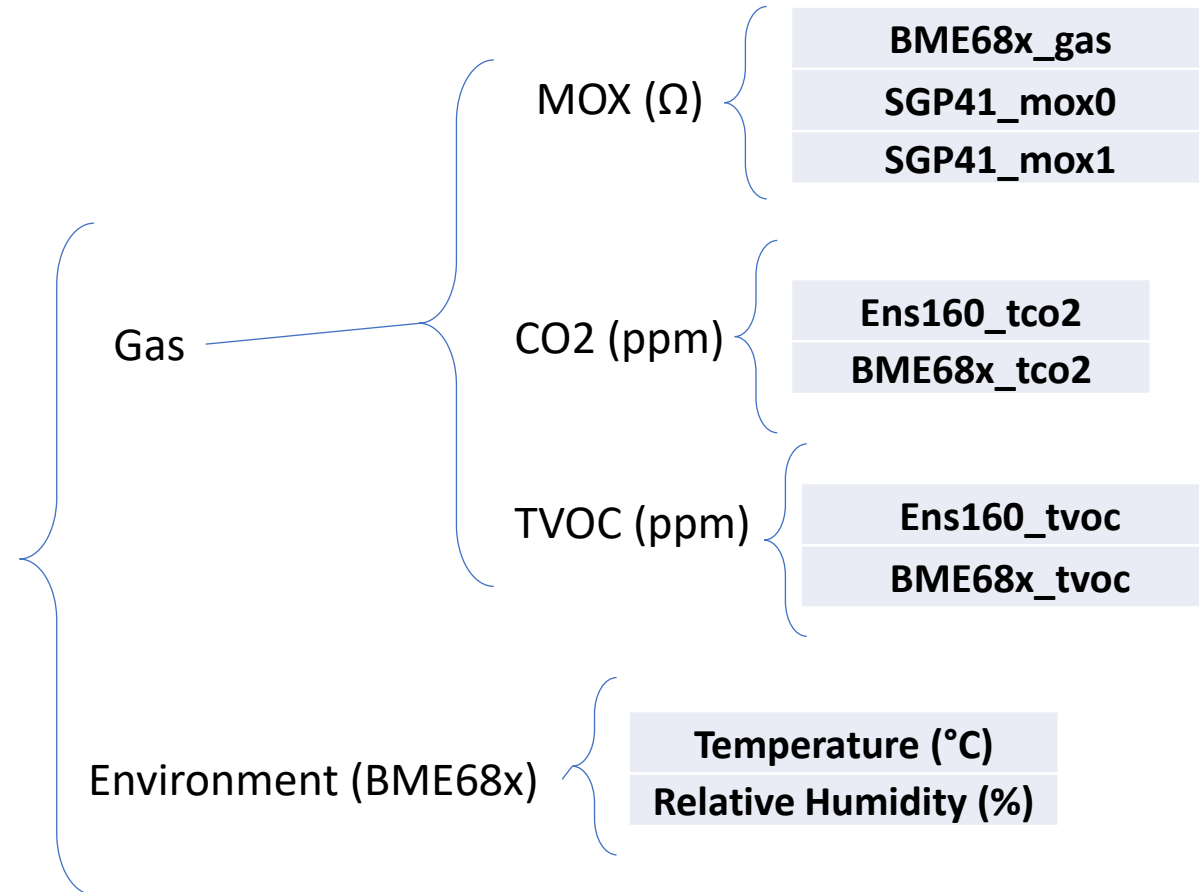
Done in pairs (i.e. Groups of 2 people)!

To pass through the whole of a Data scientist's workflow using real data:

Cleaning and preparing data → Feature extraction from the data → Classification/forecasting and presenting measurable results

Tasks 1 and 2   Task 3   Task 4

9

Module	Temps d'échantillonnage	# de capteurs mox	# de copies	Noms des fichiers
Libelium New	4 s	6	2	mod1.txt / mod2.txt
PODs	10 s	4	3	POD 200085.csv POD 200086.csv POD 200088.csv
Piano Thick	10 s	9	1	IMT_Thick.csv
Piano Thin	10 s	10	1	IMT_Thin.csv
Piano PICO	10 s	3	1	IMT_PICO.csv

Task 1: To group all files with the same name into a single file, organized chronologically and without duplicate samples.

Libelium

Piano

POD

Single file containing data from all sampling periods of a given module
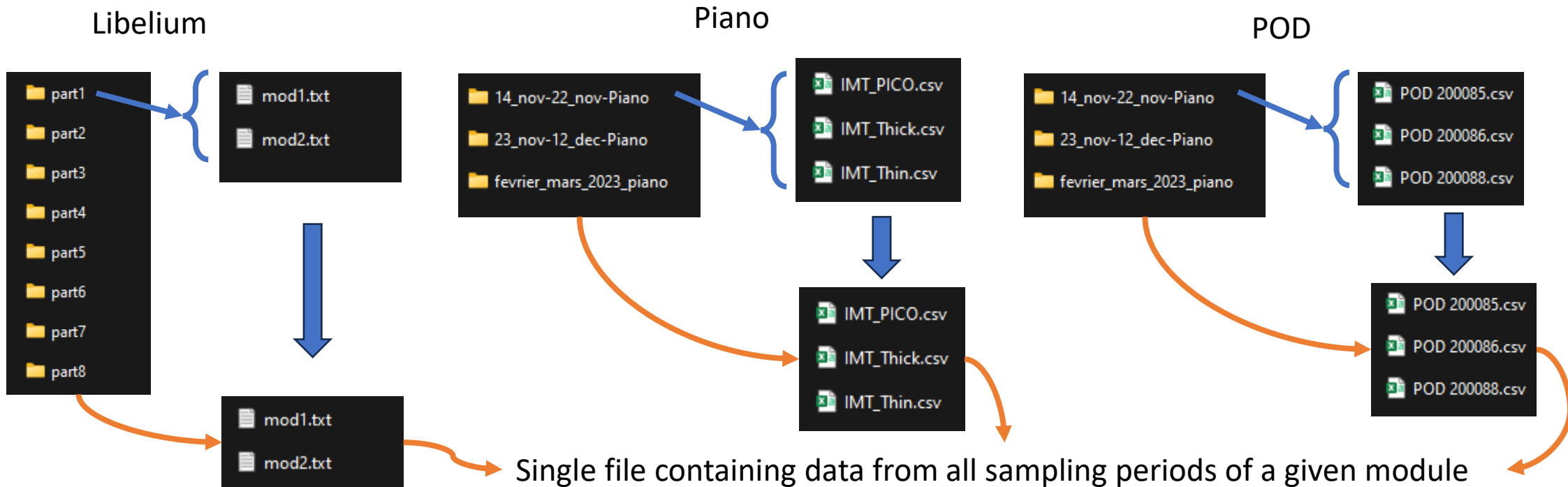
Task 2: To synchronize data from each module using the same sampling time and create a database with all available data in a single .csv file

merge(mod1.txt mod2.txt | IMT_PICO.csv IMT_Thick.csv IMT_Thin.csv | POD 200085.csv POD 200086.csv POD 200088.csv) ➡ syncedDatabase.csv
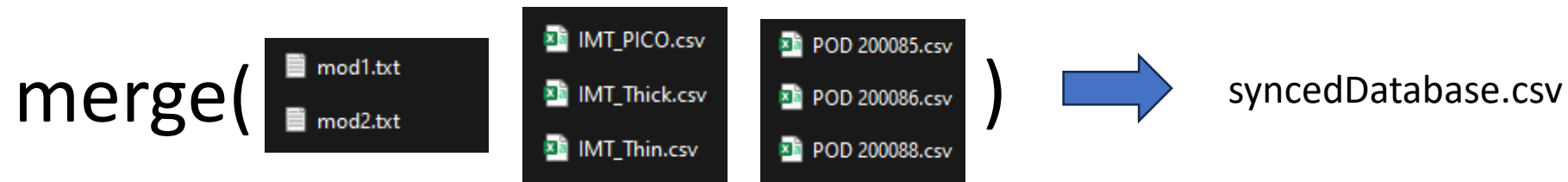
- The individual files have become a single file with all the data present in all the individual files;
- The data are all with the same sampling time (10 s);
- Samples from the "Libellium new" system that are between samples from the other systems must be averaged and put into the same instant of time as the other systems;
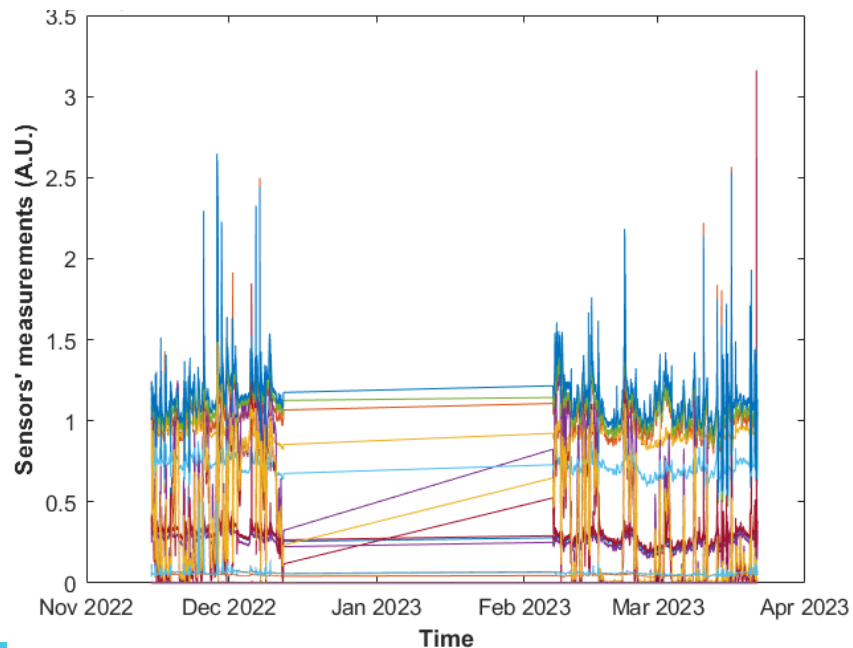- Module columns marked with "aqi" or "iaq" should be deleted;

11

Task 3: To find the average signature of each activity. By signature, we mean the time series (with all sensors) that represents each activity on average
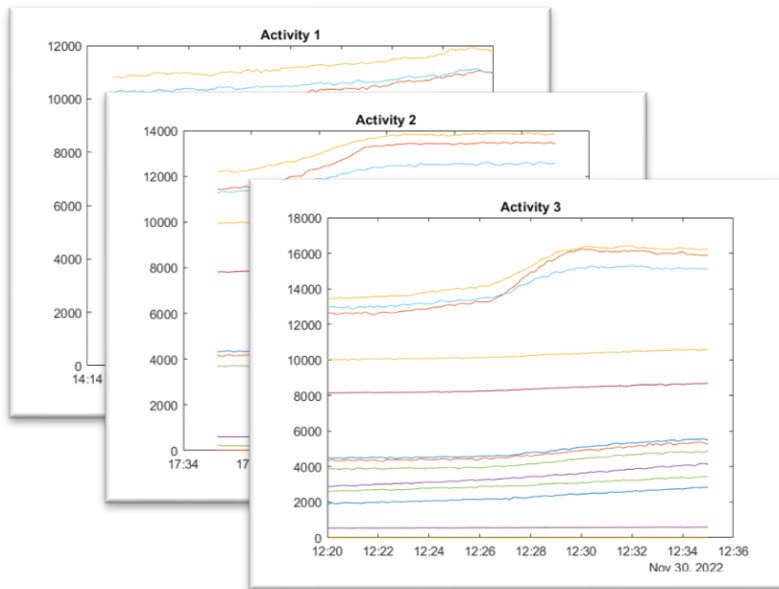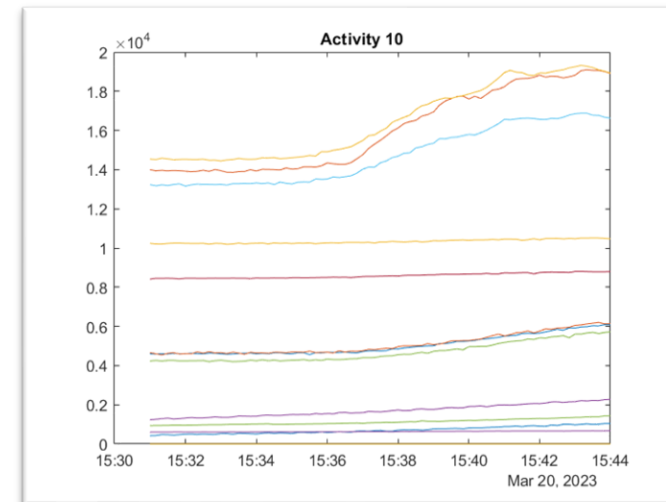

syncedDatabase.csv
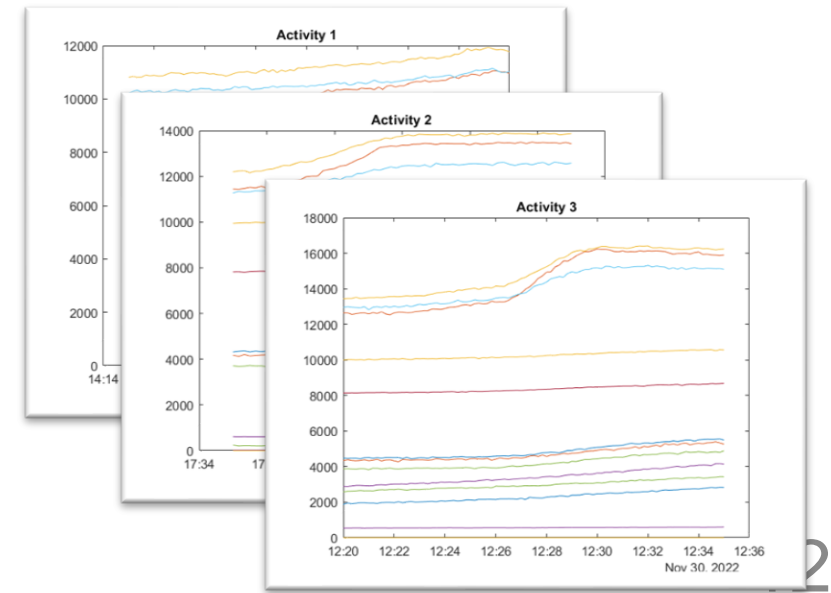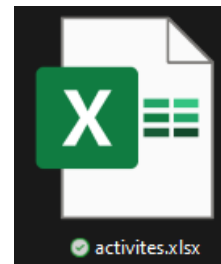
Task 3: To find the average signature of each activity. By signature, we mean the time series (with all sensors) that represents each activity on average

Mean( ) =

The length of the segmented activities may be different, you must find a solution to average them anyways.
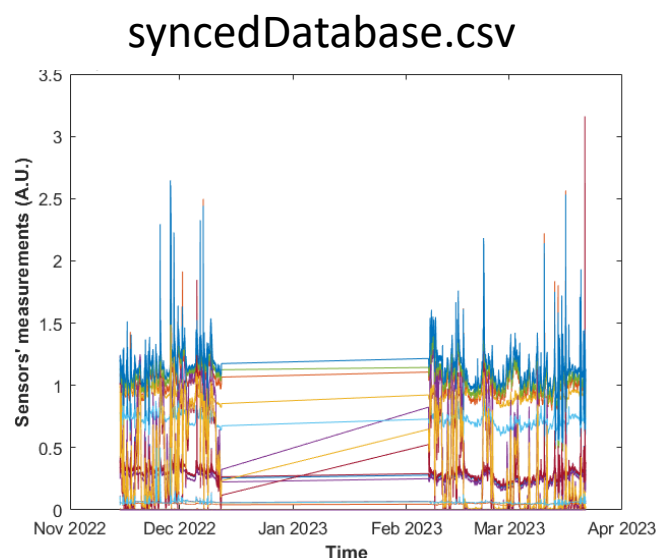
13

Task 4: propose a methodology for recognizing activities using data in the database. All proposed techniques must be justified. You must also execute the proposed methodology and measure its activity recognition performance.
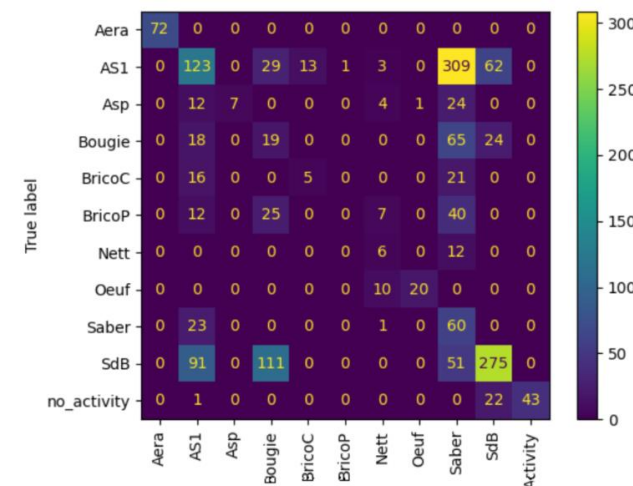
LSTM   CNN   K-NN

syncedDatabase.csv

Proposed methodology

Figure 8.1: Acurracy = 0.385

14

- You must produce a report **(per pair), between 2 and 5 pages long**, in which you describe the solutions you have developed for each TP task.

- The report should be clear and well-explained, with tables and figures if necessary. Beyond the explanations, you must make the codes **(in Python)** used available on your GitHub page (or repository of your preference), the link to which must be in the report.

- The grade will be composed of the quality of the report and the success of each task. The assessment of task success will be made with the code you add to the repository. The code must be clear and well commented to make the evaluation possible.

- The report must be consistent with the codes, i.e. the solutions described in the report must be the same as those implemented in the codes.

- Full description of the components of grade are in the Support pdf.

⚠️ **In case of unsuccessful execution of tasks 1 or 2, a database will be provided so the subsequent tasks may be executed. The points relative to the related task are then lost.**

Email: luiz.miranda@imt-nord-europe.fr

Luiz Miranda
18/03/2024