Smart Systems for Buildings

Rebeca Baron Kalbermatter

Dissertation presented to the School of Technology and Management of Bragança to obtain the Master's Degree in Industrial Engineering.

Work oriented by:
Prof. Dr. José Luís Sousa de Magalhães Lima
Prof. Dra. Ana Isabel Pereira
Prof. Dra. Ângela Ferreira
Prof. Dr. Guilherme Luiz Moritz

Bragança

2022

Smart Systems for Buildings

Rebeca Baron Kalbermatter

Dissertation presented to the School of Technology and Management of Bragança to obtain the Master's Degree in Industrial Engineering.

Work oriented by:
Prof. Dr. José Luís Sousa de Magalhães Lima
Prof. Dra. Ana Isabel Pereira
Prof. Dra. Ângela Ferreira
Prof. Dr. Guilherme Luiz Moritz

Bragança
2022

Dedication

To my parents, Rolando Leroy Kalbermatter and Miriã Alves Baron Kalbermatter. To my grandmother, Maria Alves Baron (*in memorian*).

"Ainda que a figueira não floreça, nem haja fruto na vide; o produto da oliveira minta, e os campos não produzam mantimento; as ovelhas sejam arrebatadas do aprisco, e nos currais não haja gado, todavia, eu me alegro no Senhor, exulto o Deus da minha salvação." (Hab 3:17,18)

Acknowledgement

First of all, to God, who allowed me to realize all my goals so far through His blessings.

My family, especially my parents, for all the unconditional support they have provided me. Thank you for the love and dedication you have had for me. Without you, the achievement of my goals would not be possible. This small thank you cannot compare to the gratitude I feel for everything you have done for me. And to Luis H. O. Alves, thank you for all the help in all moments, being my inspiration and partner for everything.

To my grandmother, Maria Alves Baron (*in memorian*), who passed away during my time in Portugal, leaving sweet memories and a lot of missing me. Thank you to be my inspiration in life.

To my supervisors Dr. José Lima, Dra. Ana Isabel Pereira, Dra. Ângela Ferreira, and professors Thadeu Brito and João Braun, for all the orientation and patience for the development of this project. Also, the availability and orientations by the co-supervisor of this project, Dr. Guilherme Moritz.

To my course friends, Matheus Guerra, Mateus Rocha, Eduardo Rebello, and Lucas Ritzmann, thank you for making college fun and being my support in times of difficulty.

To the friendships, I made in Portugal, especially in São Roque. Each one of you was fundamental for this experience to be even better. My heartfelt thanks to both sides of the wall in this story, especially to Alexandre Carelli for becoming a true brother during this period. And to my new housemates, thank you for the unconditional support you have provided me to accomplish this project.

To my family and friends who always support me, whether near or far.

To everyone who participated directly and indirectly in the realization of this project.

Abstract

Life in society has initiated a search for comfort and security in social centers. This quest has generated revolutions within the knowledge about the technologies involved, making the environment automated and integrated. Along with this increase, ecological concerns have also arisen, which have been involved since the conception of intelligent structures, remaining throughout their use. Based on these two pillars, the present study aims to cover three central systems inside the apartments of the Apolo Building (Bragança, Portugal). The monitoring of the energy and water consumption and waste disposal systems are integrated by storing the data in a single database in InfluxDB. The data is collected through the sensors and transmitted via Wi-Fi, allowing real-time monitoring through the Grafana application. This data collection makes it possible to track the resident's behavior through the applied machine learning algorithm. In addition, the collected data should be visible to the resident, bringing together ecological solutions that minimize the expenditure of the building's resources. Through continuous observation of the data, it will be possible to analyze whether there have been changes in the resident's behavior through the data presented. Besides dealing only with technical data, this analysis can bring advantages in the search for better ecological awareness of each resident. There is no point in seeking efficient solutions for the building if each resident does not has own actions do not aim at it.

Keywords: Internet of Things; Smart Buildings; InfluxDB; Machine Learning.

Resumo

A vida em sociedade iniciou uma procura ao conforto e à segurança nos centros sociais. Esta procura gerou revoluções dentro do conhecimento sobre as tecnologias envolvidas, tornando o meio ambiente automatizado e integrado. Juntamente com este aumento, surgiram também preocupações ecológicas, que têm estado envolvidas desde a concepção de estruturas inteligentes, permanecendo ao longo do tempo da sua utilização. Baseado nestes dois pilares, o presente estudo pretende abranger três sistemas centrais no interior dos apartamentos do Edifício Apolo (Bragança, Portugal). A monitorização dos sistemas de consumo de energia, de água e descarte de lixo são integrados através do armazenamento dos dados numa mesma base de dados no InfluxDB. Os dados são coletados através dos sensores e transmitidos via Wi-Fi, permitindo a monitorização em tempo real através da aplicação Grafana. A partir da aquisição desses dados é possível fazer o rastreamento do comportamento do residente através do algoritmo de machine learning aplicado. Além disso, os dados adquiridos devem ser visíveis ao próprio morador, trazendo junto soluções ecológicas que minimizam os gastos dos recursos do edifício. Através da observação contínua dos dados, será possível analisar se houve mudanças de comportamento do morador mediante os dados apresentados. Essa análise, além de tratar apenas dos dados técnicos, pode trazer vantagens na busca por uma melhor consciência ecológica de cada morador. Não só adianta buscar soluções de eficiência ao prédio, se cada morador não tiver sua percepção e atos também voltados a isso.

Palavras-chave: Internet das Coisas; Edifícios Inteligentes; InfluxDB; Machine Learning.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Buildings have become centers of social spaces, whether for work or residence. In the 1970s, with the dissemination of controlled processes, the spread of systems that allowed controlling the air conditioning of these social centers became feasible. This advance in microcontroller processes has given the rise in the last decade to a significant increase in demand for intelligent buildings, which take advantage of connectivity between the Internet of Things (IoT), sensors, and the cloud to remotely monitor and control various building systems. Efforts to make buildings "smart" focus on interoperability to ensure systems control, reduce operational, maintenance, and energy costs, and increase the environment's well-being. Communication and connection between the systems in a building are one of the main characteristics of a smart building, which can do much more than turn the lights on and off. For its built environment, it is possible to extract value from IoT solutions, customizing them according to the client's needs and budget.

Although the technologies applied to smart buildings are not recent, the key to implementing them in this branch is to create connectivity between them, control them with appropriate sensors, and use centralized management systems. Froufe and Soares [1] pointed out eight central systems found in smart buildings: heating, ventilation and air conditioning (HVAC), lighting, energy management, security, telecommunications, fire prevention and combat, vertical transportation, and hydraulic system, being possible to categorize three that act directly in energy reduction and five acting in comfort improvement. Buildings in use consume almost 40% of the world's energy and are responsible for 30% of CO_2 emissions [2], which increases the importance of efficient energy management to reduce CO_2 emissions. Monitoring specific sensor networks makes it possible to adjust the environment to maximize efficiency, count and locate occupants, and provide data on building usage in real-time. It is possible to automate the systems using the information obtained from data collection so that some operations can even be controlled autonomously.

Sensors provide valuable data that can lead to more assertive decision-making, improving space utilization, including savings on daily expenses and equipment maintenance. The allied use of learning and prediction should develop over time, interpreting data from current and past use, allowing occupant choices to be used to create a higher level of comfort and satisfaction.

Therefore, beyond just monitoring the data, it is essential to make it visible to the consumer aiming at changing his critical view on using the building's resources, not just by automatically controlling the environment but by helping the resident take ecologically biased measures. Constant monitoring allied with preventive actions contributes to the best use of resources. Thus, the benefits of a smart building go beyond IoT technologies. Smart buildings' convenience and comfort encourage occupants to be more productive and efficient in their daily tasks [3].

1.1 Objectives

This work aims to developd and analyse a system able to integrate the monitoring procedure and data of three central systems in Apolo Building (Bragança, Portugal): energy, water, and waste disposal. The building, which will provide hotel service, has seven apartments divided into three floors, and also contains a part for pastries on the ground floor. A more detailed description of the pilot apartment that was used for this project will be given in Chapter 3. From this monitoring procedure, the overall systems must be able to predict the consumption generated and make the data visible to the residents themselves, presenting ecological solutions that can be applied in his day-to-day to reduce their ecological footprint.

As specific objectives, the work includes the following aspects:

- Use low-cost sensors to monitor apartment parameters such as temperature, humidity, energy consumption, water consumption, and waste disposal.
- Create a database able to centralize the measured parameters.
- Monitor the data in real time.
- Develop and validate a prediction algorithm by using machine learning.
- Create a specific application to the resident.

1.2 Document Structure

The present work is divided as follows:

- Chapter 1 presents an introduction to the topic of smart buildings, contextualizing the objectives of the work and the structure in which the document is presented.
- Chapter 2 presents the literature review on smart buildings, highlighting definitions associated with this term and its main characteristics. Then, reviews of main previous work done on the systems approached by this study are presented.
- Chapter 3 presents details about the sensors chosen for monitoring and the software tools used for data collection, storage, and communication, followed by the machine learning technique.
- Chapter 4 describes the implementation of the developed system, starting with the data acquisition processes and their particularities, followed by the processing and storage in the InfluxDB database, the monitoring in the Grafana application, and finally, the machine learning application.

- Chapter 5 is dedicated to validating the developed processes, presenting the data collected in the Apolo building, followed by the data treated with the machine learning application and its real-time monitoring.
- Chapter 6 concludes the results obtained and presents proposals for future work.

Chapter 2

State of the Art

This chapter deals with the concepts of intelligent building and the proposals for monitoring a building's electrical power, water, and waste disposal systems, intending to be viable solutions in the economic and sustainable scope. In the first part, a brief review of related literature will be performed, followed by an explanation of the tools used and the rationale for the choice.

2.1 Historical Evolution of Intelligent Buildings

As society advanced in social connivance, there was a need to create spaces that brought more comfort and safety to the individual [4]. The buildings brought this social centralization, where people gathered, making this construction a core of economic activities. Over the years, these spaces have to adapt to the needs that people presented, especially in terms of comfort.

The first centralized control equipment appeared in the 1960s, and was used mainly in room air conditioning equipment. From the 70s on, with the greater dissemination of microcontrollers, there was an expansion in the control processes, which allowed advances in automation and supervision. The oil crisis, in the mid-1970s, contributed to making the aspects related to energy conservation and rationalization more practical, in a way that strongly contributed to the implementation of increasingly controlled and supervised systems.

As early as the 1980s, with the advancement of the requirements for comfort and security in places, telecommunication services, and flexibility in workplaces, three pillars emerged for the intelligent building system: automation, telecommunication, and computational systems.

In 1986, the organization *Intelligent Building Institute* (IBI) was created in the United States of America. This institution aims to support and promote topics related to Intelligent Buildings. To generalize and in an attempt to centralize, the Institute defined a narrow concept for this concept, being [5]:

> A building that provides a productive and cost-effective environment through optimization of its four basic elements - structure, systems, services and management - and the interrelationship between them. Intelligent buildings help business owners, property managers and occupants to realize their goals in the areas of cost, comfort, convenience, safety, long-term flexibility and marketability.

Later, in 1998, it is also declared by the European Intelligent Building Group that an intelligent building is [5]:

> One that creates an environment which maximizes the effectiveness of the building's occupants, while at the same time enabling efficient management of resources with minimum life-time costs of hardware and facilities.

As highlighted in [5], there are over 30 definitions for the word "intelligent" when associated with buildings. According to [6], this becomes a problem when "new buildings will not be optimally designed to meet the next century".

According with [7], the lack of a specific concept creates openness for a definition according to "degrees" of building automation, being them:

> A determined building as having a basic intelligence (25% automation of the systems), moderate intelligence (50% automation of the systems), or sophisticated intelligence (over 80% automation of the systems).

The fact is that a building is a long-term investment. A building has a lifespan of 50 to 100 years today, depending on the materials and installations [8]. Therefore, investing in factors that make the environment more pleasant to live in is an investment in own quality of life. In [9], a list is made of the main systems to be monitored in a building for human comfort standards, listed under temperature, luminosity, hearing level, and air quality (i.e., levels of carbon dioxide (CO_2), carbon monoxide (CO), airborne mold and mildew, among others).

Thus, an intelligent building will not necessarily be the most automated one. However, the ones with the most integration between the systems are the building and cabling infrastructure, automation of the systems and their integrated control, management and maintenance. In this context, it can be concluded that the concept of the intelligent building is not limited to the definition of controlling the building through a remote computer, but as an environment capable of offering comfort and productivity with a value for money between the systems and services used, for the longest time possible [10].

2.2 Energy System

Energy consumption is one of the primary factors of concern regarding energy efficiency in smart buildings. It is necessary to have an insight into how and where energy consumption occurs and to manage it accordingly. As early as 2012, buildings were consuming 40% of the total electricity consumption in the European Union, with this consumption increasing yearly [11].

Energy efficiency, in particular, has become a specific target when associated with intelligent buildings, mainly due to costs and ecological care [12]. It is a concern that must be taken into account through the stages of design, construction, and use of the building. The optimization must go through the monitoring and control of the process in order to obtain increasingly efficient buildings, understand the data presented, and apply, if necessary, improvements in the problematic points. In Portugal, for this specific issue, there is legislation in force through the Decree-Law 101-D/2020 [1], which establishes requirements for energy efficiency improvements in buildings.

In [13] a system is developed using the ESP8266 microcontroller and the EmonCMS data cloud to monitor temperature, relative humidity, and air quality data using a retractable solar panel as the power source for the house. The data collected by the sensors is used to control the devices in the house to maintain comfort standards with the least possible energy expenditure.

When dealing with the monitoring process in buildings that have already been constructed, the measurement of electrical energy can be done without the need to modify the construction. This may be implemented by using passive sensors, which can measure the voltage and/or electrical current in the building wiring, through the Hall Effect.

Sensors based on this effect are able to measure direct and alternating currents without the need for intervention between the monitoring circuit and the power circuit. Their output voltage is proportional to the magnetic field in which the sensor is subject to [14]. Examples of equipments based in this principle are, for instance, IoTaWatt [15] and Shelly EM [16]. The function of this equipment will be better explained in Chapter 3. In summary, these devices use clamp-type terminals and have the same principle of operation, differing in the capacity of the number of readings being performed, with the IoTaWatt being able to measure up to 14 distinct circuits and the second one measuring only two simultaneously.

2.3 Water System

Water use efficiency becomes important when it turned into an ecological concern in the smart building industry. The immediate concern is not only its availability but also how it has been managed [17]. According to the United Nations World Water Development Report (UNWWDR), in 2015, the rate of water demand growth already exceeded twice the rate of population growth [17].

[1] Available in https://files.dre.pt/1s/2020/12/23701/0002100045.pdf

In order to make consumption more conscious, the work developed by Teixeira [18] aimed to develop a certification and labeling system for water efficiency in order to make knowledge of such available to consumers. One of the studies presented observed that the application of water efficiency measures incurs into 63.9% reduction in water consumption.

To make already constructed buildings more efficient, the problem of modifications being more limited opens up. To have less interference with these changes in already installed plumbing installations, there are ultrasonic meters on the market today. This system has been developed rapidly due to its accuracy and convenience in use. This type of measurement was started in 1954 by researcher Lynnworth, who used the ultrasonic sensor to measure flow in plastic pipes. Later, in 1957, it was suggested to change to metal pipes in order to obtain a better propagation along an oblique path. In 1964, in Japan, tests were made in larger diameter and steel pipes, and from 1970 this technology began to be imported to the United States. This technology has had a considerable advance in the last 40 years because until then, there were no accurate means of measuring transit times in the magnitude of microseconds [19]. A significant barrier that stands in the way of using this type of sensor is that ultrasonic sensors are sometimes much more expensive when compared to other means in IoT.

For the study in question, the building's water system was complete, but the water metering devices had not yet been installed. This made it possible to open up other options for measuring water flow, like automatic water metering.

The water metering devices from the company Contimetria [20] is a water meter with a pulse emitter. This equipment would allow a reading through a microcontroller of the pulses emitted, calculating the water flow rate used in each apartment.

The Brazilian company Basic Sanitation Company of the State of São Paulo (SABESP), responsible for water supply, collection, and treatment of sewage in the region of São Paulo, has been implementing a project for intelligent water meters that allow the measured data to be sent directly to a database, without the need of reading by a company employee [21]. This equipment is still being tested. In the same way, the American company Sensus [22] provided the water clock that is an intelligent meter capable of sending data automatically through its interface, with a high-performance battery lasting up to 15 years. The data can be accessed by the consumer and assists in detecting leaks in the system.

Although these are very advanced solutions, they have the negative factor of cost, which can discourage users. Another way to perform the measurement is through a flow sensor, which does not depend on the hydrometer, and can be installed anywhere in the hydraulic system. As the goal is to measure the total consumption, the sensor must be installed right after the water meter. This sensor is a turbine-type sensor and works by transforming the rotations caused by the internal flow into pulses at its output.

2.4 Waste System

In order to make cities greener, Smart Waste Management (SWM) is growing in developing smart cities. Through Internet of Things (IoT) technology, the system covers the monitoring of containers and the planning of collection (even involving the route to be taken). Waste management can be segmented into three stages [23]:

(i) planning and implementation of waste collection;

(ii) transportation of the waste to the specific locations;

(iii) recycling process and preparation for reuse.

In the research done by Cerchecci et. al [24], an IoT architecture is proposed that optimizes waste management through a system in charge of measuring the capacity of waste cans and simultaneously transmitting the data to a data collection center. The system architecture is based on *Lora* communication, composed by a microcontroller, ultrasonic sensor, and a Lora module for transmission. It has been observed that data transmission in an urban area can reach 1.1 to 3km.

Bakhshi and Ahmed [25] propose a solution based on IoT and data analytics, and composed of a Wi-Fi IoT communicator, *Raspberry Pi* board and ultrasonic sensors attached to the container. This way, the real-time status of the dumpster and data analysis techniques determine the collection planning. It was observed that the proposed project reduced collection time by up to 18%, optimizing waste policies in the long term.

In this context, the present work has as a proposal to develop a platform that will be able to measure the weight of the garbage drum in order to optimize the waste disposal and collection process inside the apartments. For this, a load cell will be used in each garbage drum, respecting the European waste separation system: one for glass, one for plastic and metal, another for paper and cardboard, and finally, another for non-recyclable waste.

2.5 Internet of Things

The idea "Internet of Things" (IoT) emerged in the late 1990s, and in 1999 Kevin Ashton launched the term as the technology became known [26]. IoT enables inanimate objects to connect, transmit, and perform functions in various ways. According to Bandyopadhyay and Sen [27], the IoT system is a network of physical objects and virtual devices capable of communicating. Therefore, the use of the internet of things brings multidisciplinary connectivity between the professionals acting in the project. Whether to build the hardware part or to plan the software, it is necessary to identify the surrounding parts of each one and the objects to be used to build quality systems and security [28].

Its growing evolution, and consequently, the increase of complexity that involves this technology, brings challenges, defined by Ma [29], as heterogeneity, availability of devices, large amounts of data, and security and privacy.

According to Nasar and Kausar [30], the architecture of IoT systems is divided into four layers, which can be seen in Fig 2.1, and are described as:

- Physical Layer: layer that contains the physical objects ("the Things") that generate the data.
- Communication Layer: layer responsible for the bridge between "the Things" and the data analysis and storage layer.

- Middleware Layer: stores the data sent by the "the Things" and analyzes the data in order to extract the necessary information.

- Application Layer: takes advantage of the information extracted by the previous layer and creates intelligent applications.

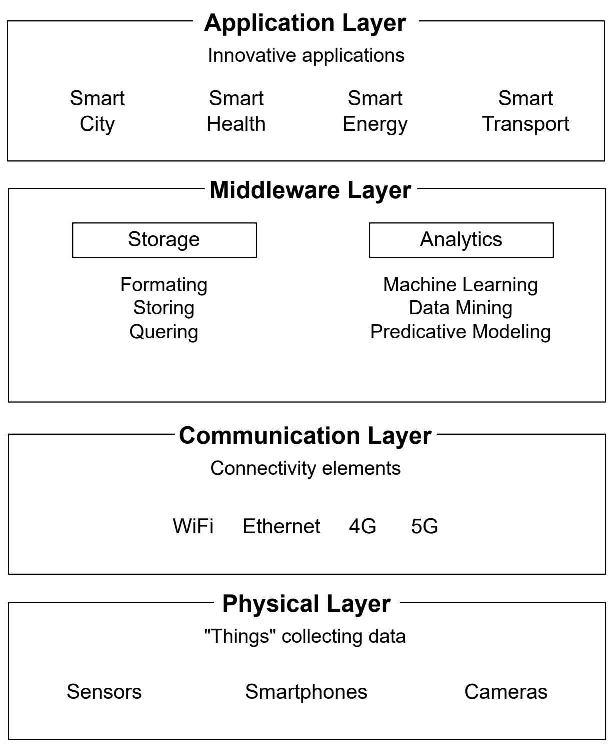

Figure 2.1: Internet of Thing Architecture. Adapted from [30].

In the context of Smart Buildings, the IoT aims to integrate equipment used in everyday life by connecting it to the internet or a database, where it will be treated in a necessary way. In this work, the study interest is related to this communication between sensors and their data analysis, focusing on the physical, communication, and middleware layers of the IoT architecture.

2.6 Time Series Database

The Time Series Database (TSDB) is an optimized database for time series or time-stamped data. This type of database was created precisely to deal with time-stamped measurements so that users can create, update, manipulate and organize time series more efficiently, allowing the visualization of long-term changes [31].

The InfluxDB database is a time series tool developed by InfluxData [32], which can be used open-source or close-source. TICK Stack, the open-source version of the platform, provides various database-related services, which can run in the cloud or on a local node. The closed-source versions offer extra functionality, such as higher availability, scalability, backup and restore, and can run on-premises (InfluxInterprise) or in the cloud (InfluxCloud). For the work developed here, the open-source version of the InfluxDB platform will be used on a local node.

Chapter 3

System Methodology

In the present work, the choice of the sensors to be used is of utmost importance, since there are specific needs in monitoring the intelligent system. This chapter will discuss the components of the designed system and the application of the system in Apolo Building.

3.1 Apolo Building

The Apolo building, located in Bragança, Portugal, underwent a modernization phase to make the building more energy efficient. The building is divided into two parts, the first on the ground floor intended for pastry and the second for residence on the subsequent three floors. During the entire project period, the building was still undergoing renovation, so to make the study and data collection feasible, one of the seven apartments was chosen as a pilot, which served as a meeting room for the building's administration. The apartment, which has the format presented in Figure 3.1, was located on the second floor of the building and was therefore named "Apartment 2" by the administration itself. Also, in Figure 3.1, it is possible to see figuratively how the sensors were allocated in the apartment, except for the YF-B2 sensor, which will be placed at the building entrance when possible.

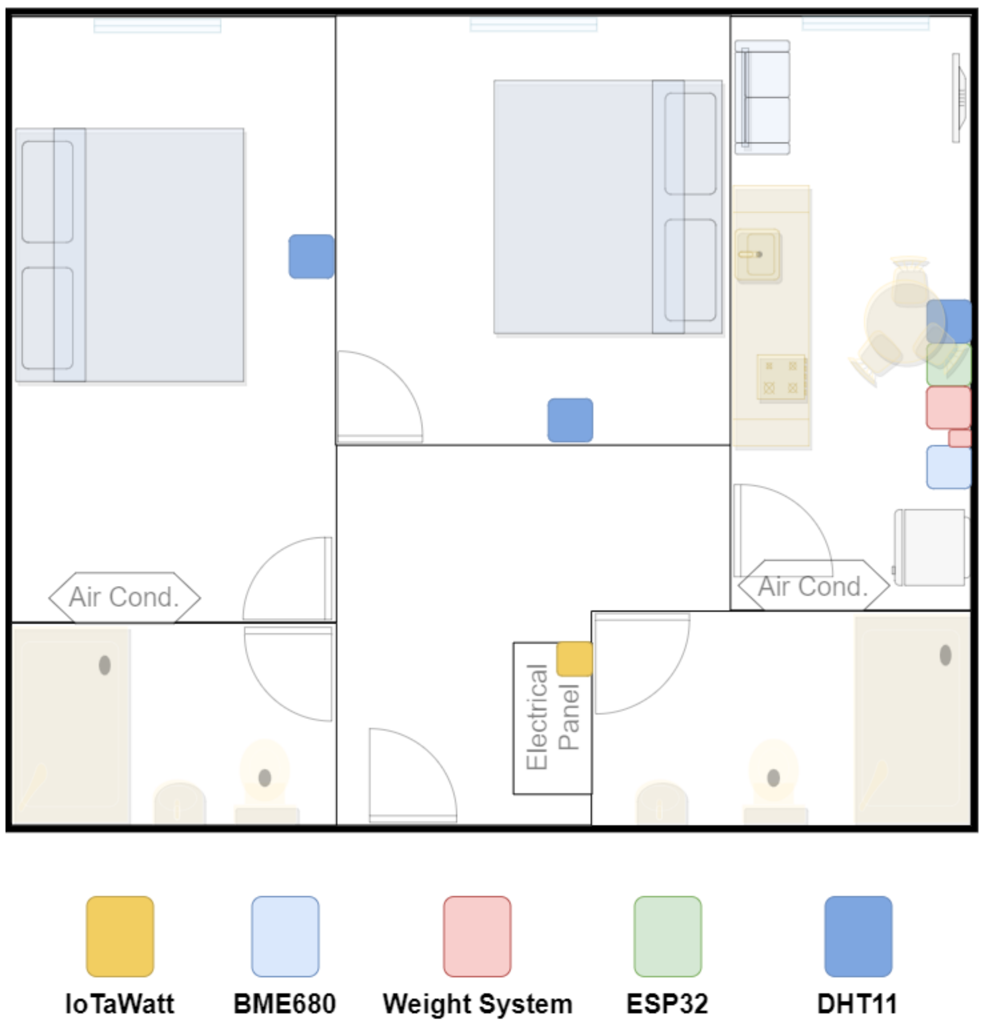

Figure 3.1: Representative sketch of the apartment and location of the sensors.

3.2 Developed Hardware System

The hardware part of the developed project is composed of the sensors, responsible for collecting data, and the microcontroller, responsible for sending the data via Wi-Fi to the database. The diagram presented in Figure 3.2 is a simplified representation of the system architecture, and the operation of the equipment will be discussed throughout this chapter.

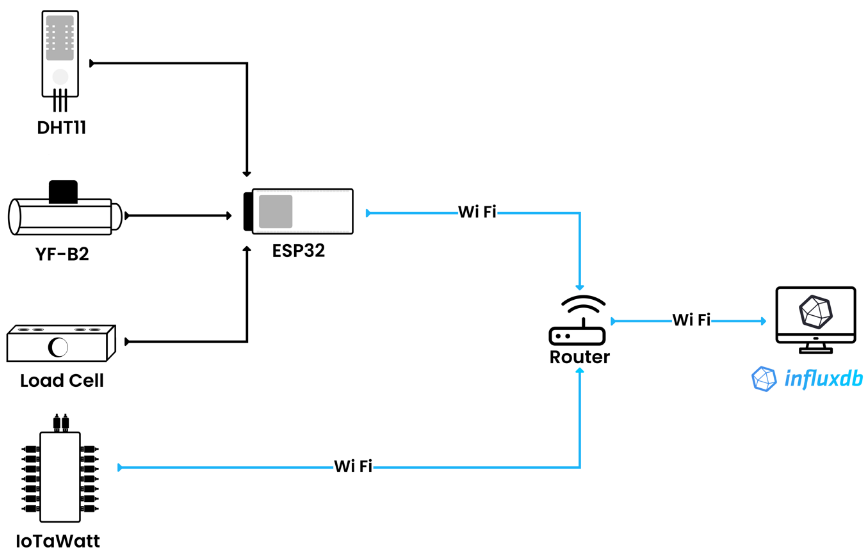

Figure 3.2: Simplified System Architecture.

3.2.1 IoTaWatt and ShellyEM

Based on the Hall Effect operating principle, Shelly and IoTaWatt's devices use a passive sensor to measure the current of a circuit by connecting to one of the insulated wires via a clamp sensor. Figure 3.3 shows the two devices, followed by a brief explanation below.

The first, IoTaWatt, can receive up to 14 distinct circuits and store the data locally and can use the client's web server to monitor, configure or control its devices. According to the company's official website [15], despite being an open system for the public, that is, available to everyone, the data is collected and stored privately. This is possible since the equipment has an integrated web server, where it is possible to see the status of the

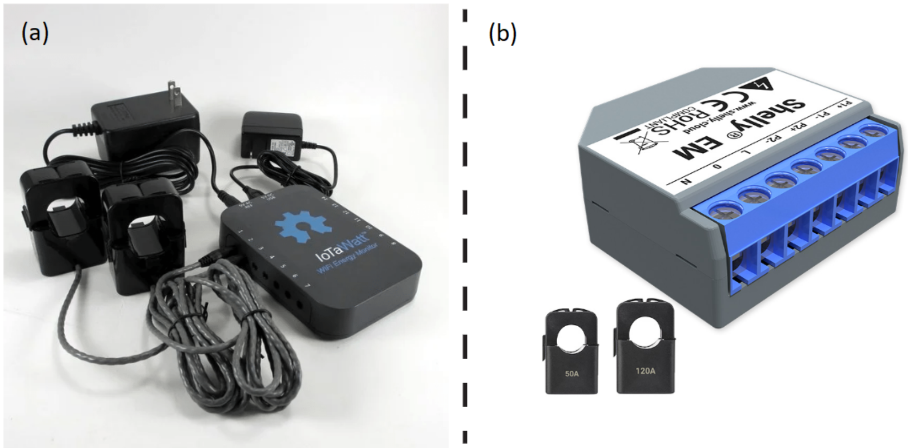

Figure 3.3: (a) IoTaWatt; (b) Shelly EM. Source: [15], and [16].

equipment and manage the system through the computer, tablet, or cell phone. The equipment is capable of monitoring different systems in more than 60 countries, once they split 120V/240V systems, as in the USA, single-phase and three-phase 230V systems as in Europe, as in some Iberian countries with 400V systems. The equipment uses the acquired current and voltage signals to obtain the power measurement. The clamp-type sensor captures the first signal, and the voltage is captured by a wall transformer, a simple step-down transformer, associating the measurement to the electric frequency.

On the other hand, the ShellyEM equipment, despite having a very similar operation, has the limitation of measuring only two circuits at a time. In addition to monitoring, this equipment also allows the interruption of the circuit when consumption reaches a predefined limit. Because of its small dimensions, this equipment is sometimes used in places with limited space.

The choice of these sensors is justified precisely by the ease of carrying out measurements without the need to open the lines to install new equipment. These two models were used in parallel to have a comparison of performance, but the IoTaWatt was chosen to give continuity to the project, for allowing more distinct readings.

3.2.2 Water Flow Sensor

This type of sensor is integrated with a magnetic Hall Effect sensor, responsible for generating an electrical pulse with each revolution. The flow rate through the sensor is calculated by the ratio of the number of pulses generated by the volume of water flowing. This ratio becomes unique for each type of sensor and must be calibrated before installation for better accuracy. The calibration factor is found in the sensor's datasheet, and its programming will be explored in more depth in Chapter 4. The sensor used for the development of this work was the YF-B2, shown in Figure 3.4. As previously referenced, this sensor was chosen for its low cost and ease of operation configuration. A negative point to be highlighted is the fact that it requires intervention in the piping to be installed.

Figure 3.4: Sensor YF-B2. Source: [33].

3.2.3 Mass Sensor

The load cell work as a force transducer that converts the load applied to it into an measurable electrical output. Types of load cells can be characterized by the output signal (pneumatic, hydraulic or electrical) or by the principle of weight detection (bending, shear, tension, among others).

The working principle of this sensor was developed by Samuel Hunter Christie in 1833, but physicist Charles Wheatstone became famous for this assembly when he described the resistor bridge in 1843, becoming known as the Wheatstone Bridge. In 1940 the first resistance wire type strain sensors were developed. However, the application of this technology could only advance after modern electronics.

The deformation sensor used for this project changes the electrical resistance of the signal according to the mechanical deformation applied. This electrical signal is sent through the data acquisition module, HX711, which sends it to the ESP32 microcontroller. An example of the type of load cell used is shown in Figure 3.5.

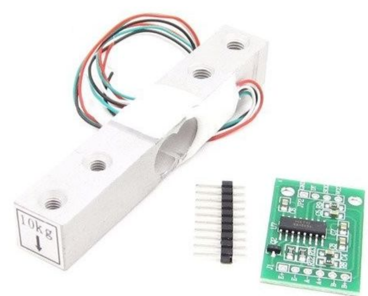

Figure 3.5: Load Cell and HX711 Module. Source: [34].

Since each scale will be assigned to a garbage can by the apartment, the maximum weight was chosen to be 5 kg, since the garbage drums will be small and divided according to recycling classification.

3.2.4 Microcontroller

The variety of possible modules that can be used in IoT applications is quite large. However, most of them present cost, size, and performance problems. Figure 3.6 presents a comparison performed by Maier and Sharp [35] and presents 4 of the main microcontrollers used in IoT projects.

Chip (Module)	ESP32 (ESP-WROOM-32)	ESP8266 (ESP8266-12E)	CC32 (CC3220MODSF)	Xbee (XB2B-WFPS-001)
Details:				
CPU	Tensilica Xtensa LX6 32 bit Dual-Core at 160/240 MHz	Tensilica LX106 32 bit at 80 MHz (up to 160 MHz)	ARM Cortex-M4 at 80 MHz	N/A
SRAM	520 KB	36 KB available	256 KB	N/A
FLASH	2MB (max. 64MB)	4MB (max. 16MB)	1MB (max. 32MB)	N/A
Voltage	2.2V to 3.6V	3.0V to 3.6V	2.3V to 3.6V	3.14V to 3.46V
Operating Current	80 mA average	80 mA average	N/A	N/A
Programmable	Free (C, C++, Lua, etc.)	Free (C, C++, Lua, etc.)	C (SimpleLink SDK)	AT and API commands
Open source	Yes	Yes	No	No
Connectivity:				
Wi-Fi	802.11 b/g/n	802.11 b/g/n	802.11 b/g/n	802.11 b/g/n
Bluetooth®	4.2 BR/EDR + BLE	-	-	-
UART	3	2	2	1
I/O:				
GPIO	32	17	21	10
SPI	4	2	1	1
I2C	2	1	1	-
PWM	8	-	6	-
ADC	18 (12-bit)	1 (10-bit)	4 (12-bit)	4 (12-bit)
DAC	2 (8-bit)	-	-	-
Size	25.5 x 18.0 x 2.8 mm	24.0 x 16.0 x 3.0 mm	20.5 x 17.5 x 2.5 mm	24.0 x 22.0 x 3.0 mm
Prize	£8	£5	£16	£23

Figure 3.6: Comparison of the main IoT devices. Source: [35].

A brief analysis of the table shows that the ESP32 has a great advantage over the other microcontrollers. This microcontroller is an embedded system board produced by Espressiff System, commonly used in IoT systems because it is low-cost, integrating Wi-Fi and Bluetooth in the same system. Its operating range varies from 2.2V to 3.6V, and its power consumption is around 0.3W. The ESP32-WROOM-32D module used in this project integrates the ESP32 microprocessor, making it easy to use on printed circuit boards. The description of the pins is presented in Figure 3.7.

All source codes of the systems used are integrated into the microcontroller, being a bridge between data acquisition and sending to the InfluxDB platform. The connections between the ESP32 and sensors will be described throughout Chapter 4.

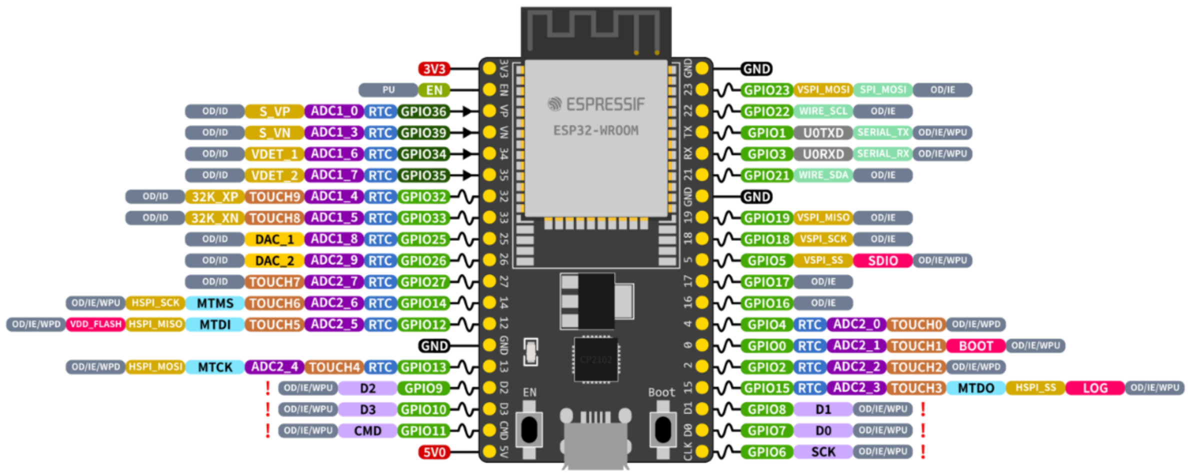

Figure 3.7: Pin description of ESP32. Source: [36].

3.3 Communication and Database

This section focus on the tools used in the software part of the work developed.

3.3.1 ESP32 and Arduino IDE

To program the components involved, the Arduino Integrated Development Environment (IDE) was used. The environment, developed in Java language, accepts programming in C++ based languages [37]. Through this IDE, the communication code between the sensors and the InfluxDB database was developed.

3.3.2 InfluxDB

InfluxDB [32] is a platform that allows recording in an optimized way all state changes already aggregated in the system. Because it accumulates in an optimized way, it is able to store thousands and thousands of data using little memory space. The main use of this platform is for real-time system monitoring, as these systems demand a large amount of reading and writing to the database. InfluxDB can handle this volume without crashing the entire system.

InfluxDB uses the Structured Query Language (SQL), and its schema can be described as follows [38] [39]:

- Database: container for different data time-series, configured through users, inputs, retention policy, and continuous query.
- Series: represents data collection that shares the same measurement, retention policy, and continuous query.
- Measurement: data stored in the same fields.
- Tags: separated in keys and values and used to store metadata.
- Fields: also divided into keys and values, it stores metadata and the current collection, each value being associated with a timestamp.
- Retention Policy: configures how long the data will be kept in the database. By default, it is always started as infinite (data is not deleted).
- Continuous Query: a database query in InfluxQL that executes automatically and regularly. In order to conserve persistent storage, it is possible to aggregate older data.

Since it is an open-source platform, it allows the database to be easily created. The configuration performed on the platform and microcontroller to send the data will be explained in the next chapter.

3.3.3 Grafana

To allow a better visualization of the collected data, the Grafana [40] application becomes an important tool. This platform allows the creation of interactive dashboards, with automatic data updates and even in real time. In addition, it also allows to send alerts to E-mail, SMS, among others.

3.3.4 Node-RED

The Node-RED programming environment, an open-source platform commonly used in real-time data management, helps hardware communicate with other services through visual programming. It provides a browser-based editor that makes it easy to link flows using functions or nodes in the palette that can be deployed at runtime with a single click. In addition to the functions already present in the application, it is possible to install any new libraries that may be needed, since these are available as an open-source project [41].

3.4 Machine Learning

This section presents the technique used for applying machine learning (ML), which is supported Multiple Linear Regression (MLR).

3.4.1 Multiple Linear Regression

The linear regression of ML technique obtains the equation that best fits the variable of interest being studied (dependent variable) and a set of variables (independent variables). The model is called simple linear regression or just linear regression (LR) in the case of having only one independent variable. When more than one independent variables are used, the model becomes a multiple linear regression model. The objective is to measure the value of y (dependent variable) over more than one covariate. The mathematical model for this solution is:

$$y_t = \alpha + \beta_1 X_{t1} + \beta_2 X_{t2} + ... + \beta_k X_{tk} + e_t \tag{3.1}$$

where t is the index of $1, ..., T$, y_t is the dependent variable, X_{tk} the independent variable, α the coefficient intercept when the independent variables are zero, β_k partial regression coefficients and e_t a constant. It is essential to point out that the more independent variables related to each other for applying the MLR model, the more efficient the results [42] [43] [44].

The code for the machine learning application was developed in Python language. For this, the PyCharm IDE was used, which allows programming in this language and code development in JavaScript, SQL, HTML/CSS, among others. The evariablet integrates tools and libraries, such as NumPy and Matplotlib, which allow the developer to work with matrix visualizers and interactive graphics [45].

Chapter 4

Development

This chapter presents the problem formulation, as well as the paths found for its solution. The chapter follows with the hardware architectures used, followed by an explanation of the step-by-step approach to data acquisition, storage, processing and visualization.

4.1 Problem Formulation

Due to the need of making buildings energy efficient and intelligent, aligning their functions and use [46] [47] for minimal effects on the environment, smart building technology has started an essential role in the operation of making a smart building with a close relationship to emission reduction [1] [48]. As one of the goals aimed at ecological awareness, the decision-making during the project was conducted to monitor the collected data and make its consumption accessible to the resident. Figure 4.1 presents the block diagram that summarizes the developed project, highlighting the system's collection, storage, and visualization phases.

4.2 Data Aquisition

The hardware part performed the data acquisition, composed of the ESP32 microcontroller, the DHT11 and YF-B2 sensors, the HX711 load cell and acquisition module, and

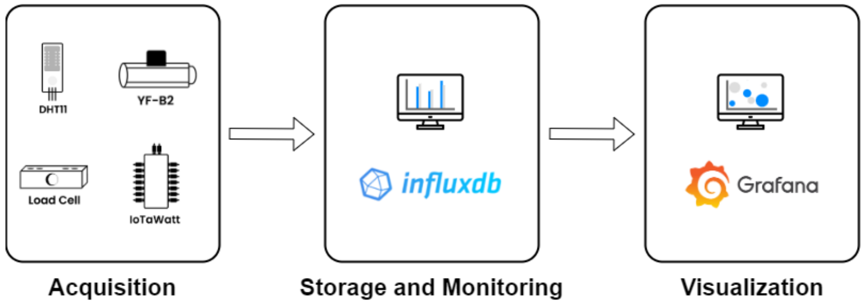

Figure 4.1: Simplified Diagram Architecture.

the IoTaWatt equipment. The choice of these sensors was due to their low cost, as seen in Table 4.1. Although the Shelly EM equipment was cheaper in unit price, it limited monitoring to only two circuits simultaneously, as explained in Chapter 3. Taking on the basis that one of the goals is to have more circuits in monitoring, it was chosen to proceed with only the IoTaWatt equipment. For the study performed, it was developed a printed circuit board (PCB), represented by Figure 4.2, that allows the centralization of the sensors along with the microcontroller stored in a box prototyped in 3D printing. Due to the time limitation for the completion of this project, it was not possible to print the PCB. Besides the sensors mentioned for the project, the PCB also has digital and analog reading ports for inserting new sensors that may be needed in the future, along with a part exclusively for the BME680 sensor.

Table 4.1: Project cost.

Component	**Cost (€)**
ESP32	6,50
DHT11 Sensor	4,50
YF-B2	7,40
Load Cell + HX711 Module	11,40
IoTaWatt (Module + 5 sensors)	201,50

Next, the operation of each sensor will be explored, as well as their connections to the microcontroller and the particularities encountered.

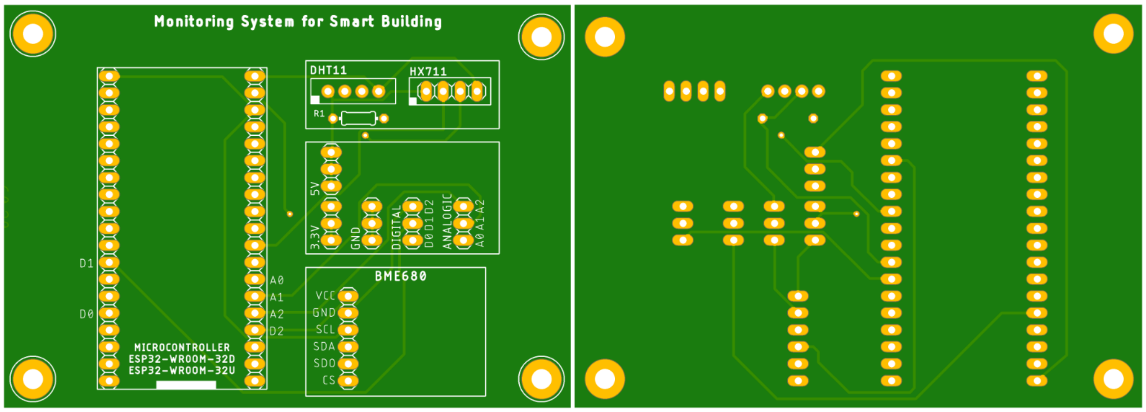

Figure 4.2: Printed Circuit Board

4.2.1 DHT11

The connection of the DHT11 sensor to the microcontroller is shown in Figure 4.3, having its Vcc connected to the ESP32's 5 V power pin, the GND pin connected to the ESP32's GND, and the read pin to the ESP32's GPIO16 pin via a 10 kΩ pull-up resistor. This mount also allows the switch to the DHT22 sensor from the same class of sensors as the DHT11. In order to get a correct reading, it is necessary to access the code through the ESP32 and change the sensor type by uncommenting the line of code for the new sensor.

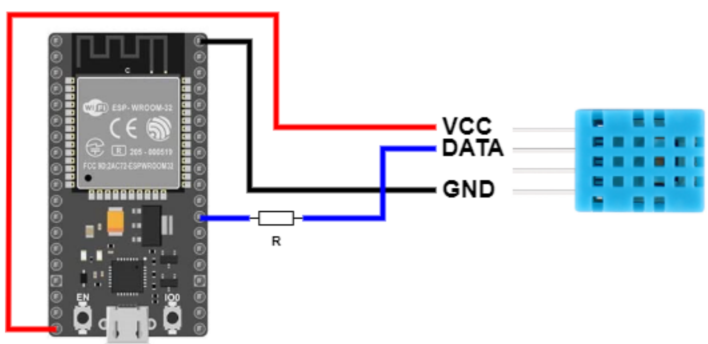

Figure 4.3: Electronic connection between DHT11 and ESP32.

4.2.2 Load Cell and HX711

For the garbage disposal readout, shown in Figure 4.4 is the connection between the data acquisition module, HX711, and the ESP32. The power pin of the module is connected to the 5 V pin of the microcontroller, as well as the ground connected to the same ground of the ESP32. For readout, the DT and SC pin of the HX711 module is connected to the ESP32's GPIO18 and GPIO17 pins, respectively. A 3D printed prototype was created for this system, along with a box for circuit storage, as seen in Figure 4.5.

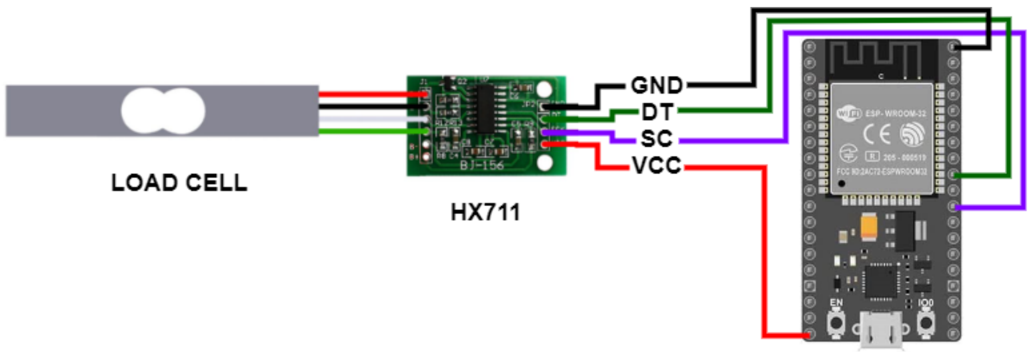

Figure 4.4: Electronic connection between the load cell and ESP32.

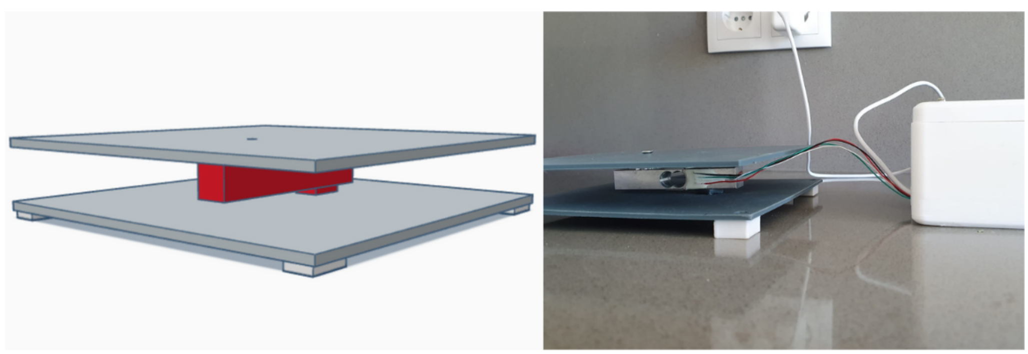

Figure 4.5: 3D modeling and weighting balance system.

4.2.3 IoTaWatt

As described previously, this equipment has an internal microcontroller, which allows sending data directly to the database. For this, it is necessary to configure the address to store the data within the IoTaWatt interface. The entire path for this configuration is provided by the equipment supplier, in [15], and Figure 4.6 shows the configuration made for this project. The IP and location data have been hidden for security reasons.

The equipment was installed in the energy box of the pilot apartment. As there were no residents in the place until this moment, only the air conditioning and general consumption of the apartment were being measured, as can be seen in Figure 4.6.

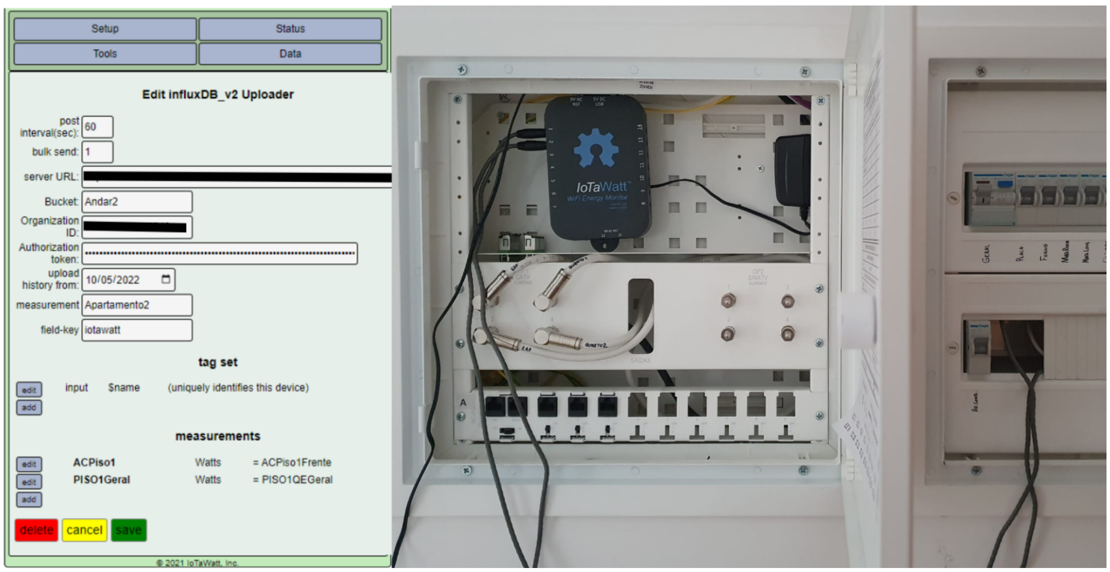

Figure 4.6: IoTaWatt interface and installed equipment.

4.2.4 YF-B2

The water flow reading was planned to be taken by the cold water pipe that enters each apartment near the entrance of the building. For this, the YF-B2 sensor was configured in the form shown by Figure 4.7. This way, the power is supplied by the 5 V pin of the microcontroller, and the GND and the reading pin are connected, respectively, to the GND and GPIO16 pins of the ESP32. An important detail is that an ESP32 has been allocated exclusively for reading the flow sensors. Here it only demonstrated the connection made for the pilot apartment. However, it is expected that in the future, the seven YF-B2 sensors will be read (one for each apartment) through this same microcontroller, with no other classes of sensors connected to it. It is also important to point out that this sensor has a calibration factor in the code, where it is possible to modify it according to what is specified in the datasheet.

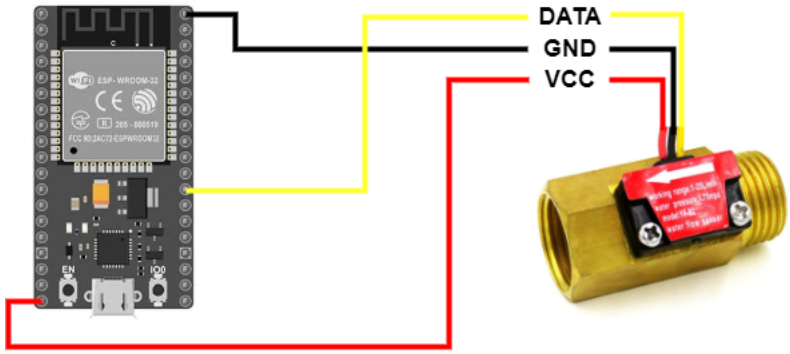

Figure 4.7: Electronic connection between the YF-B2 sensor and ESP32.

Since this sensor can only be installed by a professional, because the need for intervention in the water piping, during the process of this study, it was not possible to accomplish this installation by the administrative decision of the building. For this reason, it is emphasized that both the reading code and the reception of data in the database have already been configured and tested so that as soon as the installation is done, the reading starts.

4.2.5 BME680

First of all, it is necessary to explain why this sensor was not previously mentioned as one of the main ones to be applied in the project. Although it already had an assigned place on the PCB, this sensor was not installed in the Apolo building because it was not the initial goal to monitor the apartment's gases. However, due to the lack of time to have residents in the building, and with this data that would be relevant to the machine learning application, a module was assembled only with the BME680 to be tested at the Research Centre in Digitalization and Intelligent Robotics (CeDRI).

The operation of this sensor, responsible for reading temperature, humidity, altitude, and gas, changes the output resistance according to the gas concentration, with this relationship being inversely proportional. The sensor was connected to the microcontroller using the I^2C communication protocol, which uses two communication wires, SLC and SDA, connected to pins GPIO22 and GPIO21, respectively. The power was supplied through the microcontroller's 3.3 V and GND pins. The Figure 4.8 shows the connection between the sensor and the ESP32.

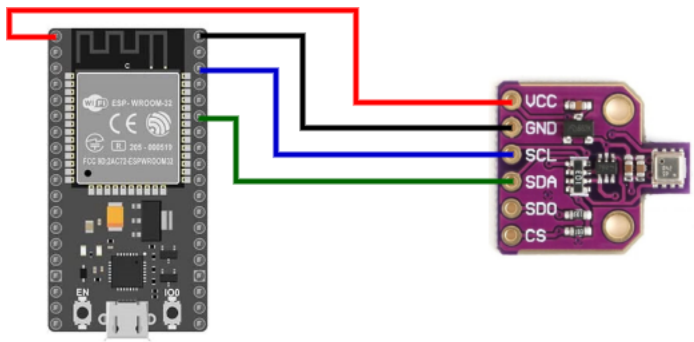

Figure 4.8: Electronic connection between the BME680 sensor and ESP32.

4.3 Data Transfer and Storage

The data transfer process was done using the communication between the microcontroller and the database in InfluxDB through Structured Query Language (SQL) communication and then sent to Grafana for visualization. This process can be summarized by the diagram shown in Figure 4.9. Therefore, for this solution, InfluxDB must be running on local equipment (computer, Raspberry, or similar) to connect to the same Wi-Fi signal as the ESP32. For the progress of the project and due to the lack of Raspberry for sale, a local computer provided by the building administration was used.

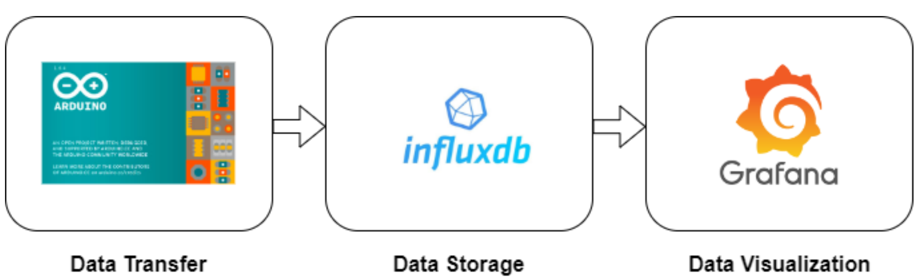

Figure 4.9: Data process diagram.

The code developed in the Arduino IDE implemented connection confirmation parameters to ensure the correct sending of data, which is presented in the flowchart in Figure

4.10. For receiving the data within the InfluxDB platform, the system architecture is represented by Figure 4.11.

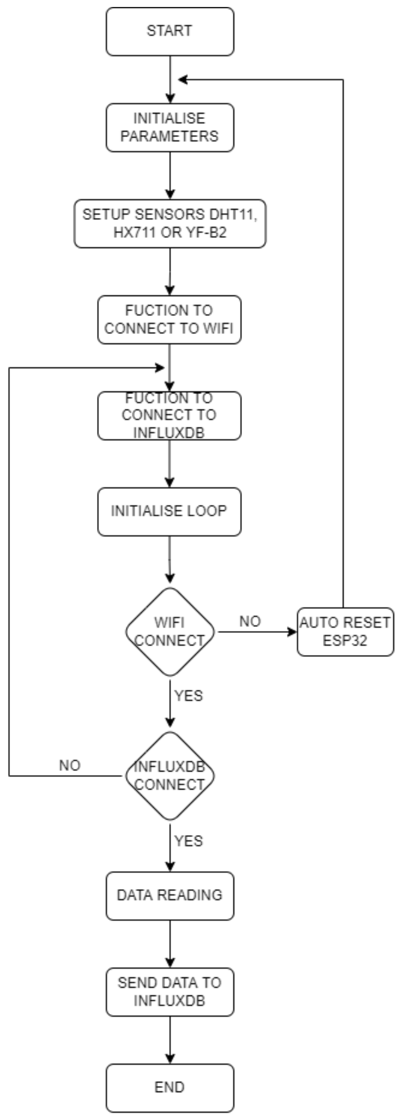

Figure 4.10: Flowchart of the code developed.

4.4 Data Processing

The database generated by the sensor data collection is used to apply the machine learning algorithm. For validation and explanation of the algorithm, only the application made to the temperature variable will be presented here. Initially, a connection is made with

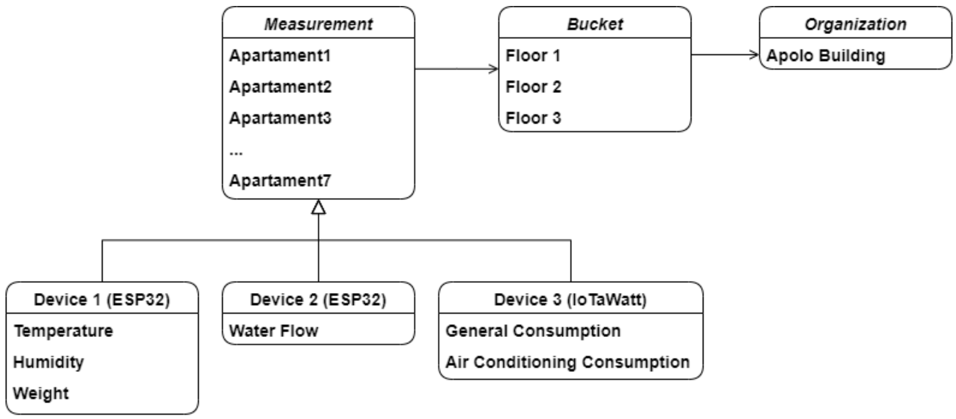

Figure 4.11: Architecture of InfluxDB.

the InfluxDB database to load the data collected over the last seven days. After that, the data is read, and then the dependent variable is selected, which will be used to apply machine learning, as in this example, the temperature. Figure 4.12 shows the step-by-step that the algorithm performs to apply MLR.

The data is collected by the code implemented in the PyCharm IDE, and the obtained result (predicted value) is sent through Node-RED to InfluxDB and displayed by Grafana. Figure 4.13 shows the node used for this application and the results obtained will be presented in Chapter 5.

4.5 Monitoring

Grafana was configured to have separate dashboards for each apartment for better data visualization, updating the data in real-time. Figure 4.14 shows the dashboard of the pilot apartment used in the tests, with data collected from June 20th to 27th, 2022. The temperature sensor and the load cell sensor are allocated in the kitchen and the IoTaWatt is measuring the overall consumption and the consumption of the apartment's air conditioner.

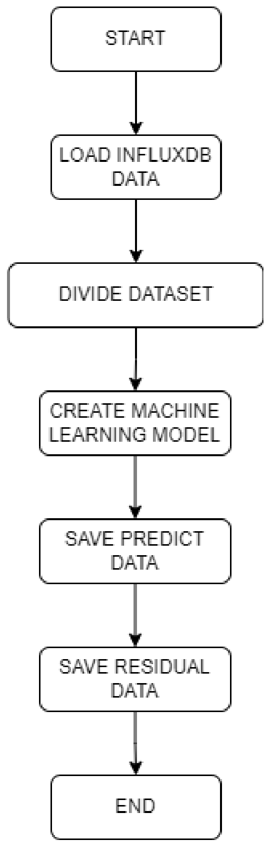

Figure 4.12: Flowchart of Machine Learning Algorithm.

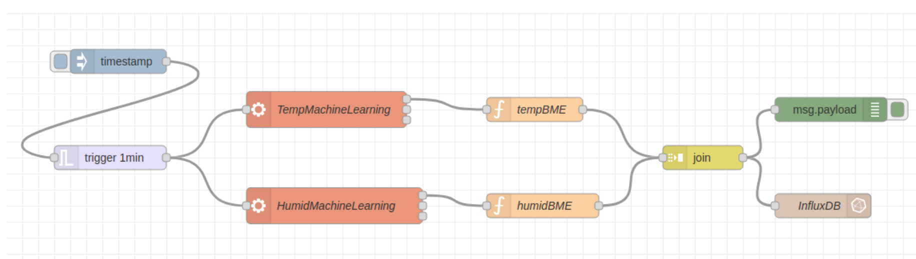

Figure 4.13: NodeRED nodes to storage prediction data in InfluxDB.

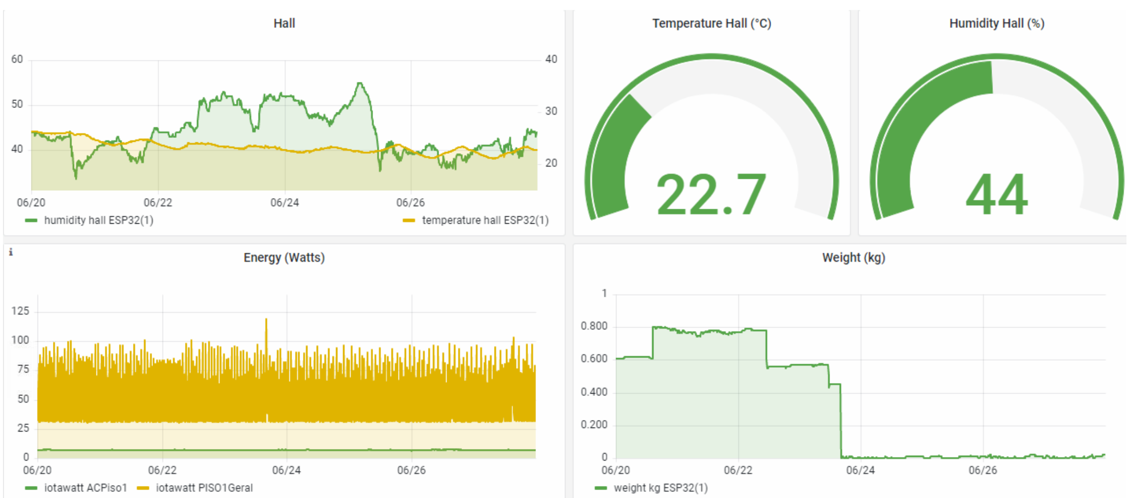

Figure 4.14: Dashboard test in Grafana.

Chapter 5

Results Analysis

This chapter intends to present the results obtained during the project's testing phase, divided into two parts: firstly, with the analysis and results obtained by the sensors in the Apolo building, followed by the analysis of the data in CeDRI laboratory with the results obtained in the ML application. As explained before, the building was in the final stage of reform and had no residents, affecting data collection and machine learning. For this reason, and as described in Section 4.2.5, a module with the BME680 sensor was previously made and left in the CeDRI. The energy parameters and the amount of discarded waste were kept in the building because they are fixed equipment.

5.1 Apolo Building Data

As previously described, the Apolo building database was divided between the floors (three) and the apartments (seven). The apartment chosen as a test for the project, named Apartment 2, has been used as the administration meeting room, allowing to collect data from the energy consumption (general and air conditioning), the waste weight generated and the temperature and humidity local. Thus, it is possible to observe in Figure 5.1 the reception of data from the apartment using InfluxDB 2.0. As mentioned before, the only sensor that could not be installed to collect data was the YF-B2 to measure the water flow. Despite this technical limitation, the database is already configured to receive this data when it is possible to do this. In addition, the database is already configured to receive data from two other DHT11 sensors, which will be inserted in the apartment's rooms, when the renovation is finished. The sensors have been configured to collect and send the data every 1 minute.

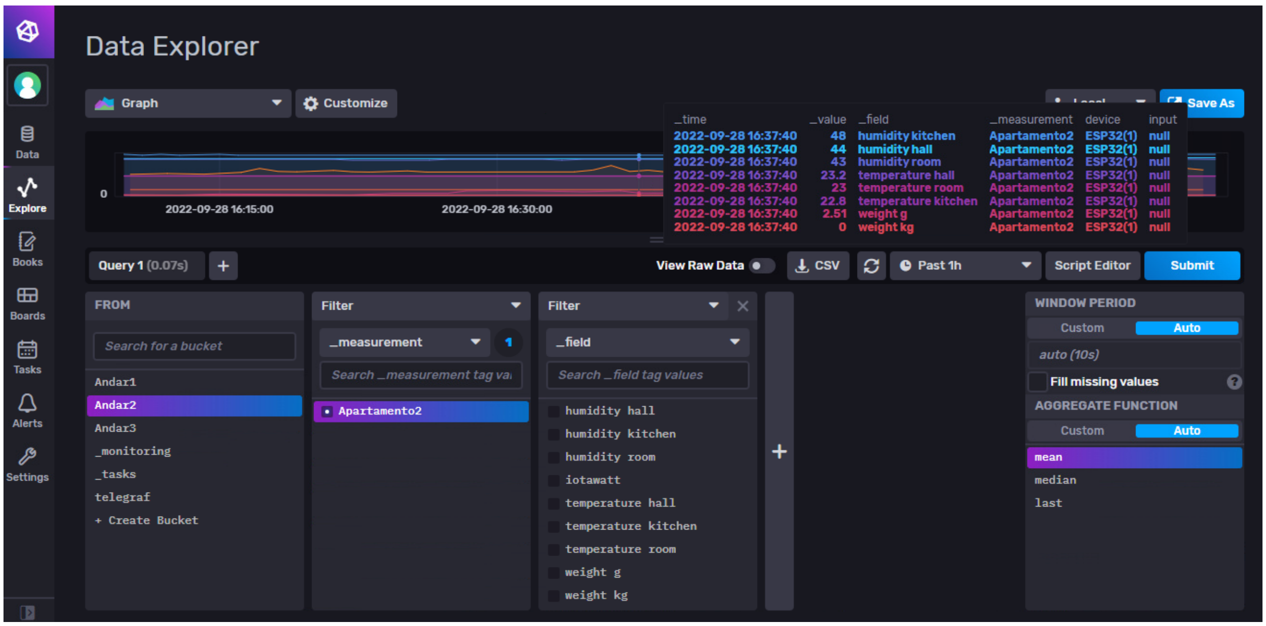

Figure 5.1: InfluxDB platform with the data collected from Apollo.

To better visualize the collected data and with the intention to apply machine learning algorithms, the dashboard for the apartment was created in Grafana. Like InfluxDB, Grafana's dashboards will be divided by apartment. Figure 5.2 shows that the validation is done by receiving the data and a test of the final presentation format to the consumer. It should be noted that the data from the YF-B2 sensor will not be the focus of analysis since it could not be installed, which is why the graph is blank in Figure 5.2.

For the energy consumption, Figure 5.3 presents the data collected by the IoTaWatt equipment, which measured the apartment's consumption, including air conditioning. During the selected period, from June 01st to 30th, there was no need to use the air conditioner, so the consumption for this equipment was 0 W. As for general consumption, it is possible to observe a particular standard behavior. This behavior can be explained by the fact that the apartment has a refrigerator that was constantly on, having its

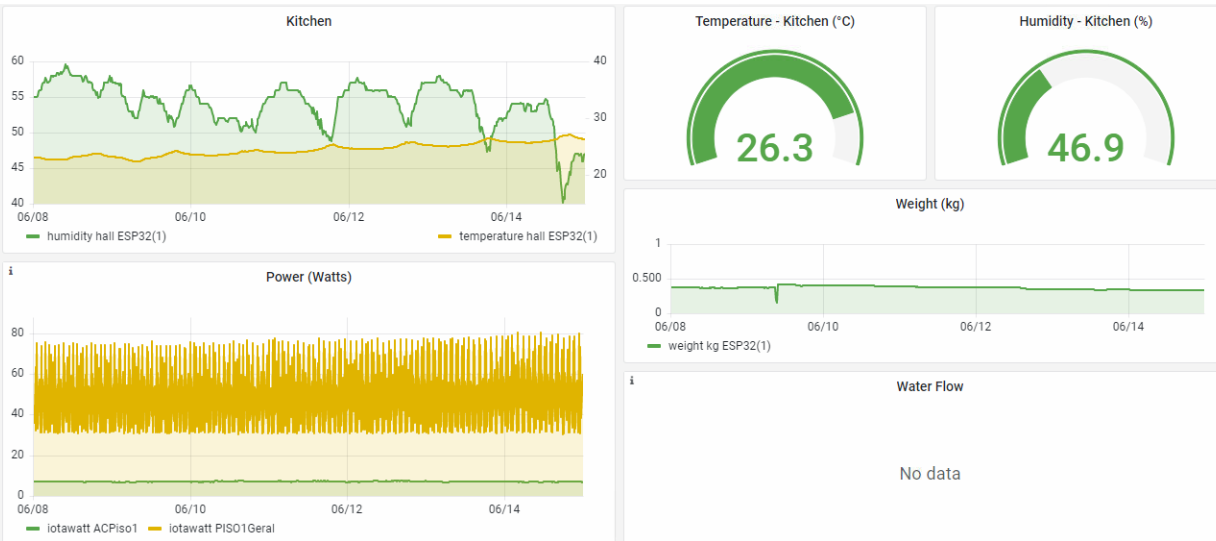

Figure 5.2: Grafana's dashboard used to display data from the pilot apartment.

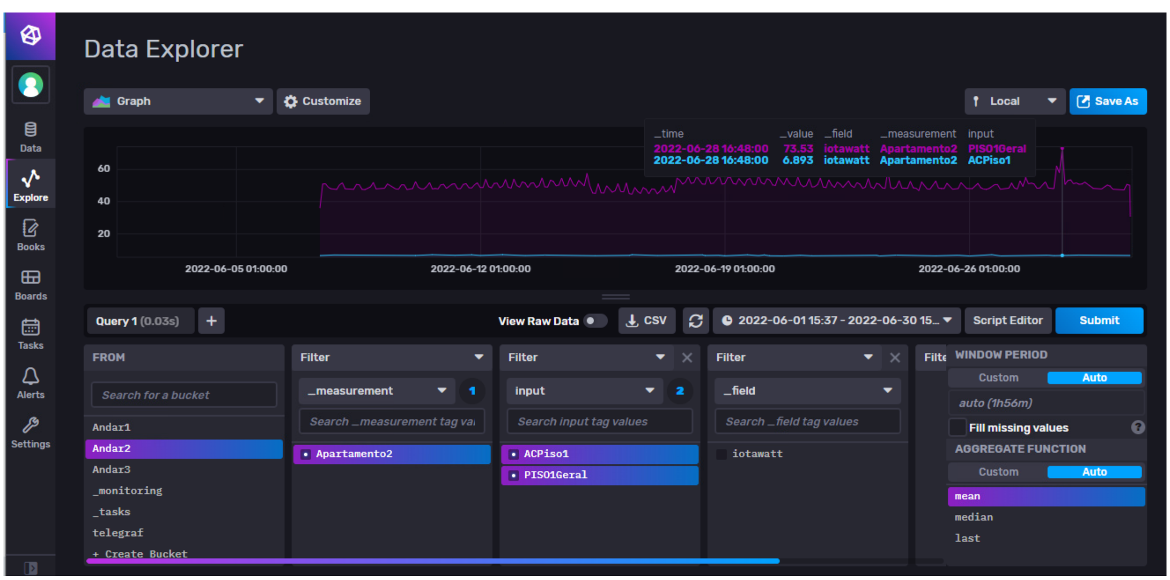

Figure 5.3: Data collected by IoTaWatt stored in InfluxDB.

refrigeration cycle captured by IoTaWatt.

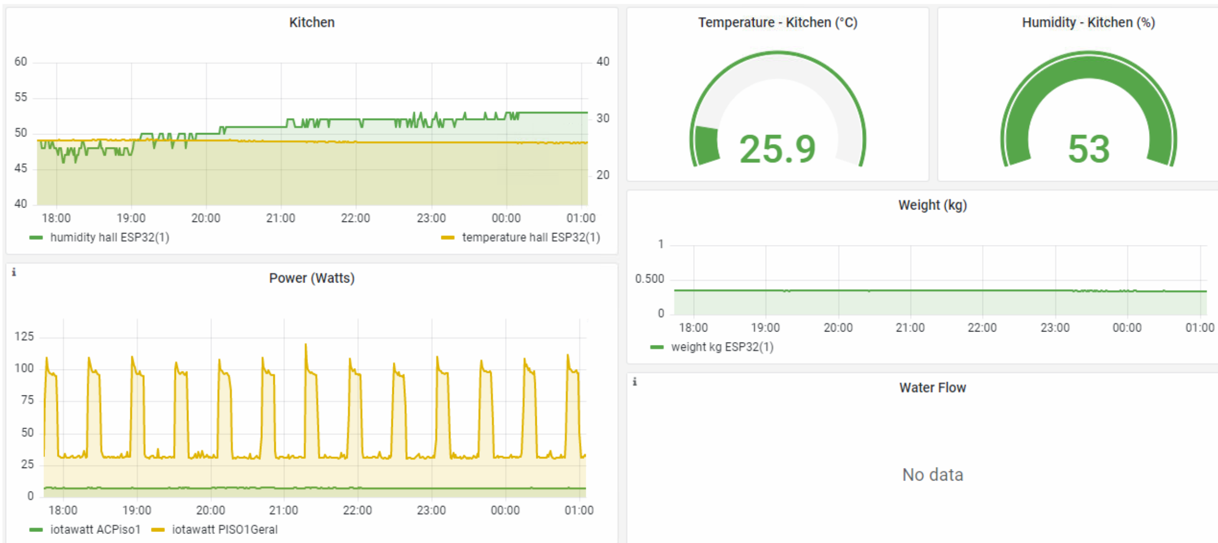

Figure 5.4: Grafana's Dashboard amplified.

For better visualization, Figure 5.4 shows consumption on day 13, where it is possible to see the intervals of approximately 30 minutes between the switching on-off the refrigeration. This equipment was installed in February when it was already configured for data collection. Due to a communication problem between IoTaWatt and the database, there was a break in the data collection, being restored only in May, when the error was corrected. However, it is possible to access the data through the equipment's interface, which stores it in a cloud. Thus, it is possible to notice the data collection from March 1st to 10th, when the air conditioning had to be turned on, and consequently, the increase in the apartment's overall energy consumption for approximate peaks up to 1.4 kW. The graph was generated in the IoTaWatt application itself and shown in Figure 5.5.

The temperature and humidity sensor (DHT11) was also validated through its data, as presented in Figure 5.6. As reported in the histogram of the same figure, it is noted that the temperature, presented in purple, for one week did not vary too much, averaging between 23°C and 27°C. The humidity, shown in blue, had more variation, showing higher values at night.

As for the garbage system, Figure 5.7 shows the data collection performed from June

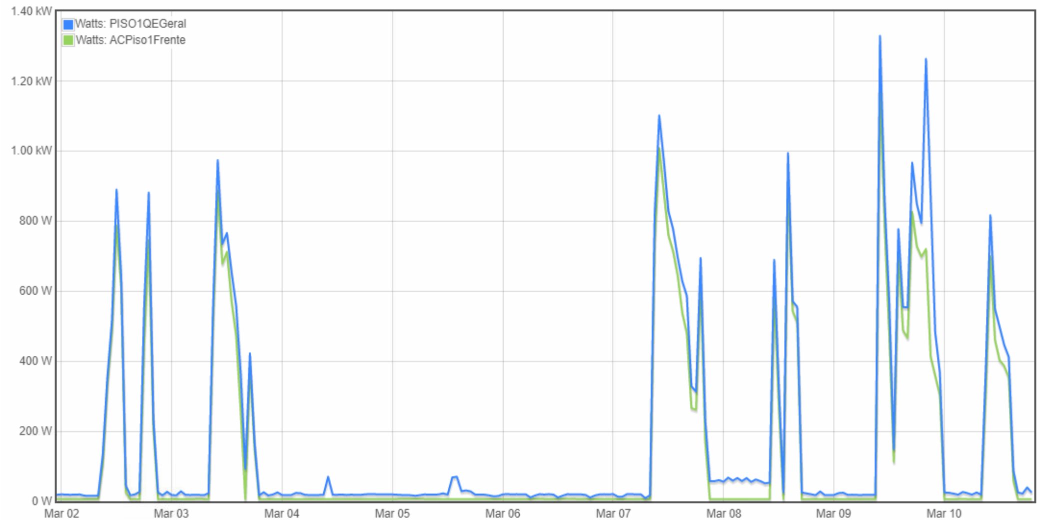

Figure 5.5: IoTaWatt interface for viewing consumption data.

Figure 5.6: Humidity and temperature of the DHT11 sensor for one week.

7th to June 30th. It is possible to see the use of the developed system but was little used, the value did not exceed 1 kg. The garbage disposal presented a certain instability of reading, having an average value of the absolute error of $+-10$ g.

Figure 5.7: Weight of the discarded garbage (in grams) collected by the load cell.

5.2 Data Analysis

As mentioned in the previous chapters, the Apolo apartments were not completed in time to finish this project and do the tests with actual daily consumption data. As this problem could occur and make the tests possible with machine learning, it was explained in Section 4.2.5 that a board with only the BME680 sensor was tested at CeDRI, a research laboratory of the Polytechnic Institute of Bragança. This sensor was responsible for collecting the temperature, humidity, gas level, and pressure data. The module sends the data every 2 minutes, where is possible to see the data obtained for example in Figure 5.8.

Figure 5.8 presents the measurements taken from January 01th to 31th, 2022, which can be seen as a standard behavior in temperature and humidity. During this period in Bragança was winter, where external temperatures ranged from -5°C to 18°C [49]. It is possible to see that inside the laboratory, the temperature varied within a range of 16°C to 25.1°C, with the humidity varying between 18.3% to 49.8%. It is interesting to note that at the lowest humidity of 18.3%, the temperature was at its highest, at 24°C. This inverse relationship between the two variables is also notable during weekends when the temperature tends to fall, and humidity rises. Using this period for analysis, the data for

Figure 5.8: Data collected by the BME680 sensor at CeDRI.

temperature and humidity are obtained, described in Figure 5.9.

	TEMPERATURE (°C)	HUMIDITY (%)	GAS (kΩ)	PRESSURE (hPa)
MÁX	25,1	49,8	222,0	954,0
MIN	16,0	18,3	63,9	845,0
AVERAGE	20,1	31,2	155,2	947,6

Figure 5.9: Maximum, minimum and average for the variables.

Analyzing the data, it is possible also to note the changes that occurred during the day, with variations in the working and departure times of the researchers and the difference between weekdays and weekends. Filtering to analyze these differences, one can observe from Figure 5.10 these nuances of temperature and humidity. Figure 5.10 highlights the 15th and 16th of January, which were weekends, where it is possible to notice that there is still the presence of people in the laboratory on Saturday. However, due to the reduced number of people, the temperature did not rise as much as on previous days, reducing even more on the 16th (Sunday). As of Monday, the 17th, it is already noticeable that the temperatures return to have significant variation when the researchers arrive. Also, from Figure 5.10, the inverse relationship of humidity with the presence of people in the laboratory is highlighted. While the temperature tends to rise in this situation, the humidity remains lower than during the weekend, despite its significant variance, when fewer people are in the lab.

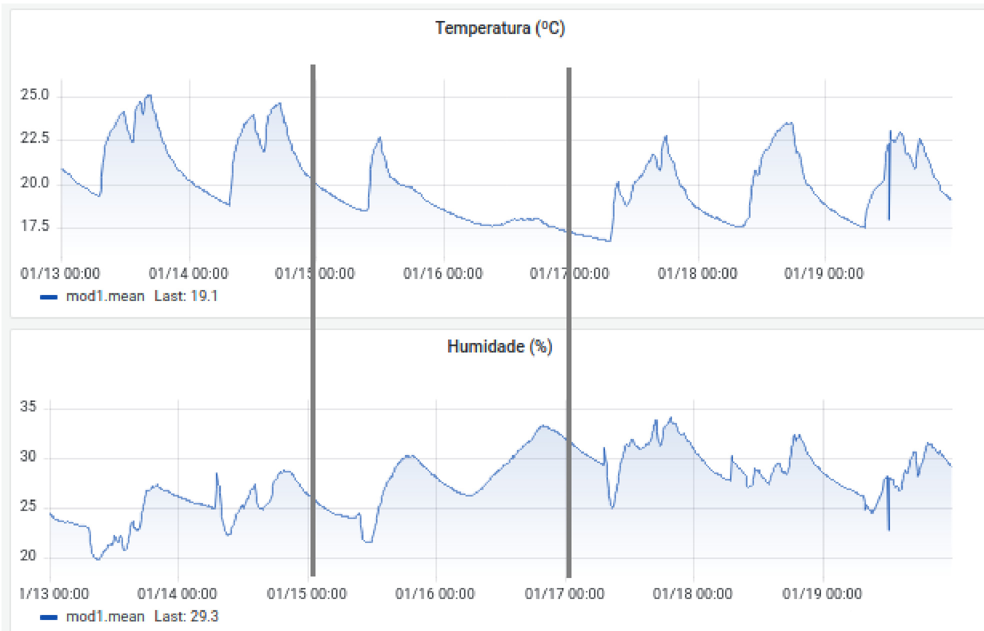

Figure 5.10: Behavior pattern of the data during the weekend.

Another possible analysis with this sensor is the quality of the air present in the laboratory. The BME680 sensor has its output resistance inversely proportional to the amount of gas in the environment, as explained in Section 4.2.5. Therefore, the data presented on the graph is that the higher the sampled rate, the less amount of gas in the environment, and vice versa. Figure 5.11 shows the concentration of gas from the same period in January. Through it, we can observe the same relationship commented previously between the days of the week and the weekend, in this case, the weekends are highlighted in gray.

For an even closer look, Figure 5.12 shows the data collected between January 17 and 22. Through it, it is possible to observe the variation during the day, highlighting the laboratory's start, exit, and lunch times (on average), being 9hr00, 19hr00, and 12hr00. Note that this behavior extends to temperature, humidity, and gas concentration.

As an example of this same behavior analysis for these variables, Figure 5.13 presents the data collected during March 08-15, 2022. During this analysis period, it is pretty noticeable the difference in the data during the weekday and weekend periods has also

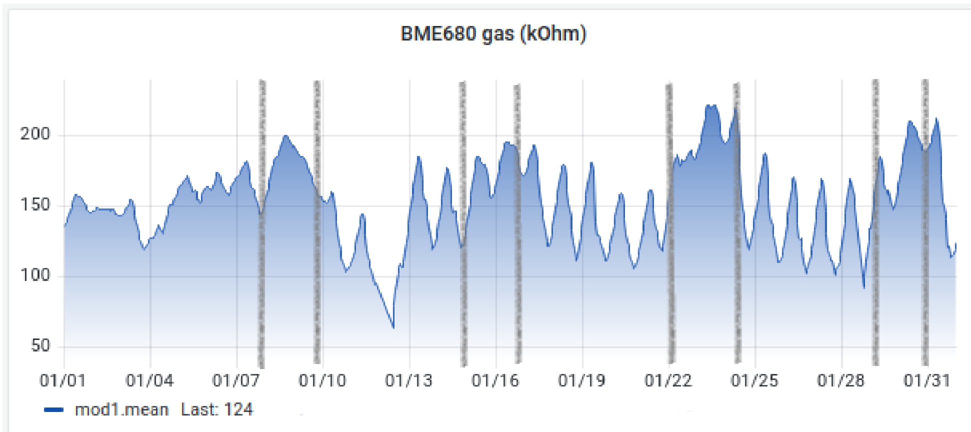

Figure 5.11: Behavior pattern of the gas concentration by the BME680 sensor.

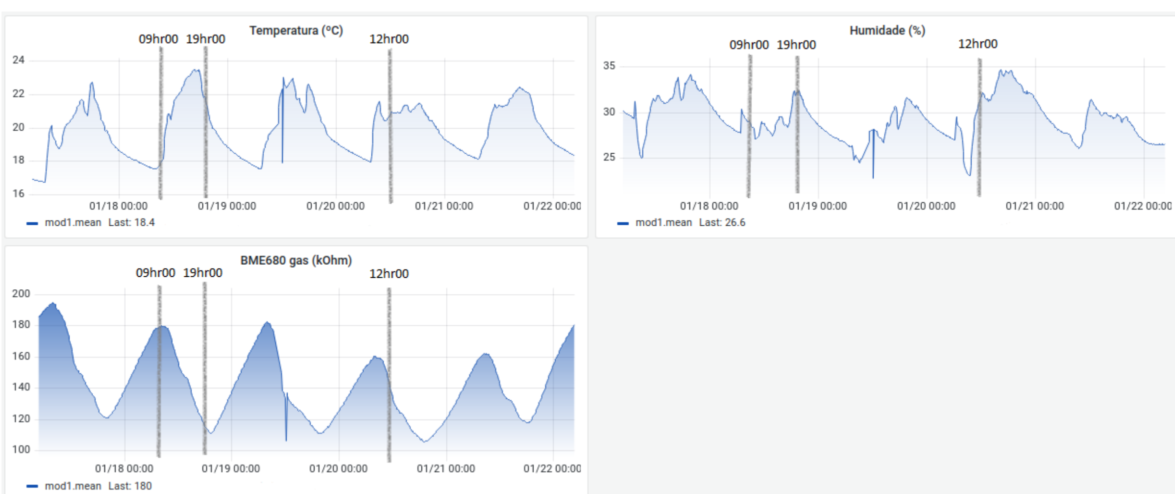

Figure 5.12: Behavior pattern of the variables during the day.

been highlighted in gray. This March period is only to prove the concept of the behavior of the variables at different times of the year since, in this period, the outside temperatures were a little milder. It is possible to note from the graph of the humidity its rise between days 10th and 11th, where there was precipitation [49]. All these results and analyses prove the stable operation of the module, with very coherent and critical results about the state of the monitored environment.

Figure 5.13: Temperature, humidity, and gas by the BME680 sensor during a week.

As described in Section 3.4.1, the machine learning model applied was multiple linear regression in order to predict each variable used in the building. However, as it is not possible to use directly in the building, the algorithm was used to predict the variables present in the CeDRI, such as temperature, humidity, pressure, and gas concentration. This type of regression allows the variable to be predicted with several independent variables. That is, to predict the temperature, the remaining variables will be used, such as time, humidity, pressure, and gas concentration. This will happen for each predicted variable, isolating it as a dependent variable and using the others as independent variables. In this case, it was used all the parameters collected by the sensor BME680. In addition, data from another sensor present in the laboratory to measure the quality of the air was also used. This sensor, of the class MQ135, was used with the objective to compare the data between BME680 and MQ135.

For analysis and testing of the code, test periods of 30, 15, and 7 days were chosen between January 1-31, 10-25, and 13-20. These periods of days were chosen to analyze the accuracy of the prediction algorithm in the variety of the amount of data and different data paths for the regression and were divided between training and testing as shown in Figure 5.14, separating approximately 15% of the data for tests. To improve the efficiency of the prediction, the data had been treated to remove values that were outside the standard deviation. The training and tests were done for all the variables, but here will show the results found in the test for the temperature.

CASE	DAYS of JANUARY	TRAINING	TEST
A	1-31	1-25	26-31
B	10-25	10-20	21-25
C	13-20	13-17	18-20

Figure 5.14: Division between training and testing days.

As previously analyzed, the month of January in Bragança is very cold, and consequently, the heater was used to keep the environment comfortable. These factors considered “external” had an influence on the MLR results, presented in Figure 5.15, amplifying the results of residuals in the second graph. In case “A”, the test model obtained as result a precision of 60.5%, defined by the coefficient determination R^2, with a Mean Absolute Error (MAE) for the training period of 1.02°C and for the test period of 1.05°C, where represents only 5% of the average real temperature of this period. The range of residuals averaged between -2.7°C and 4.4°C, where the residuals are the difference between the real and predicted value. For better visualization, the test period was separated by a green line, representing the date from the 26th day. The results that will be shown from the MLR algorithm were exported directly from PyCharm, and as a result, the X-axis of the table is a function of the sample instead of the date.

Following the algorithm analysis, one can observe from Figure 5.16 the results obtained when applying MLR for the training in case “B”. This application brought, as a result, an MAE of only 0.9°C to the training period and of 1.9°C for the test period, with a precision of 71.3%, and residuals between -2.7°C to 4.0°C. The result is better than before presented, which enables to move forward for the analysis in case “C”.

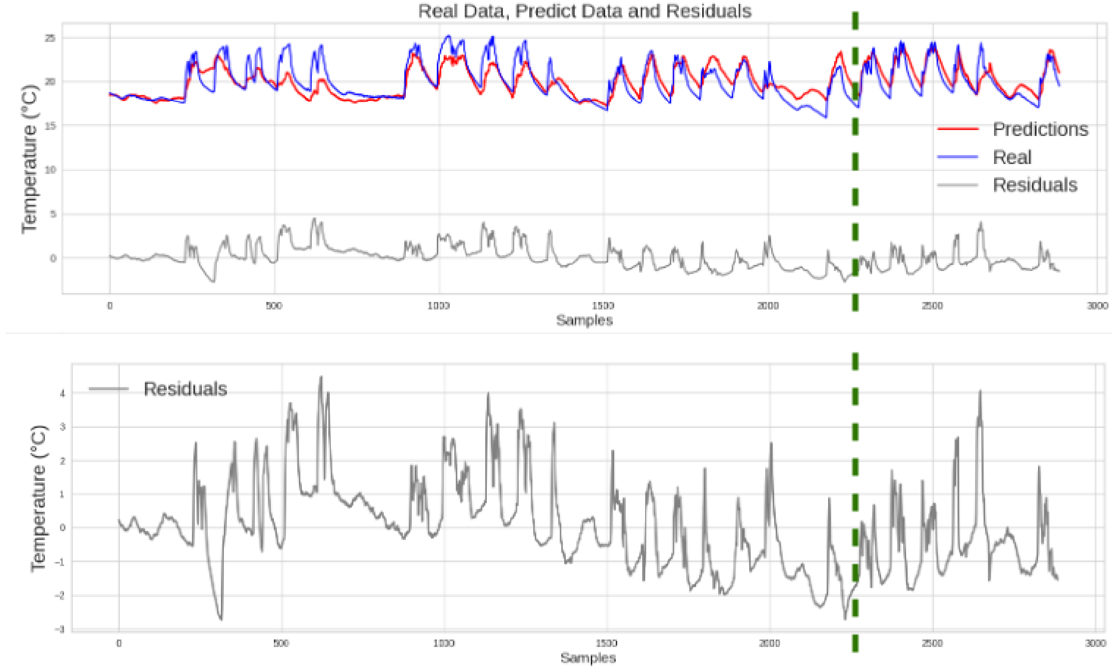

Figure 5.15: Result of the temperature prediction for January by MLR in case “A”.

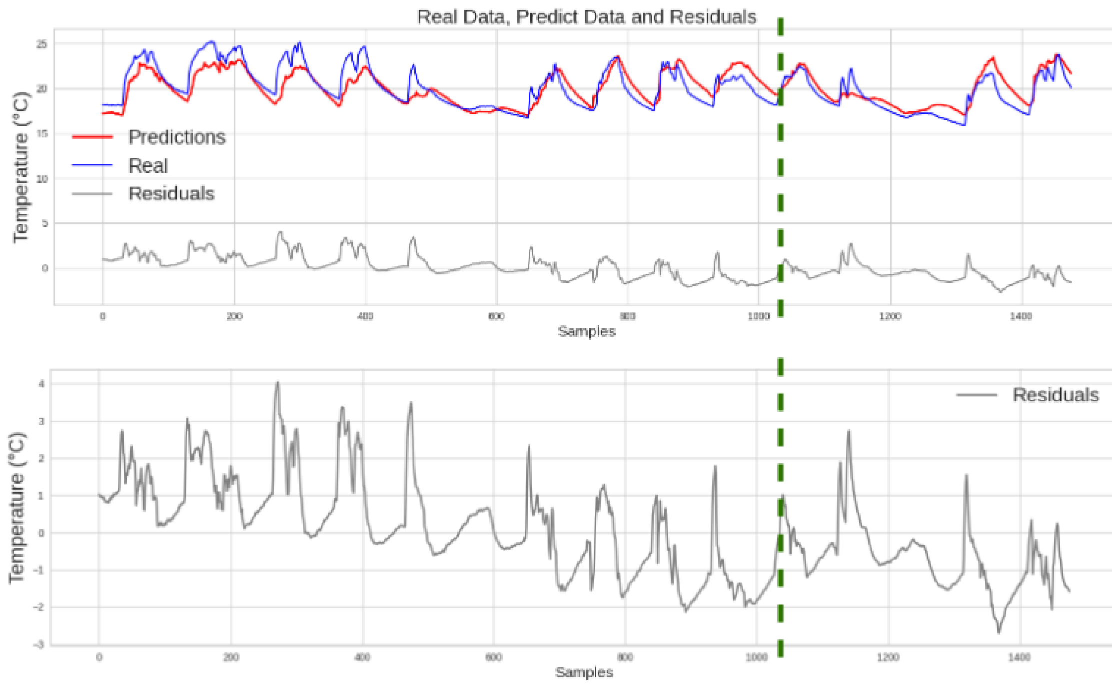

Figure 5.16: Result of the temperature prediction for January by MLR in case “B”.

In continuation, Figure 5.17 shows the results obtained by applying the MLR for the seven days of regression, in case “C”, between January 13th and 20th. This application presents the better results among these three cases, with a range of residues between -2.3°C to 2.5°C, an MAE of 0.4°C and 1.0°C, respectively, to training and test periods, and precision of 91.8%. For the temperature case presented for tests, it is possible to conclude that for better results from the MLR algorithm, case “C” is the most suitable, using five days for training. The correlation indices between the variables used for this prediction are presented in Table 5.1.

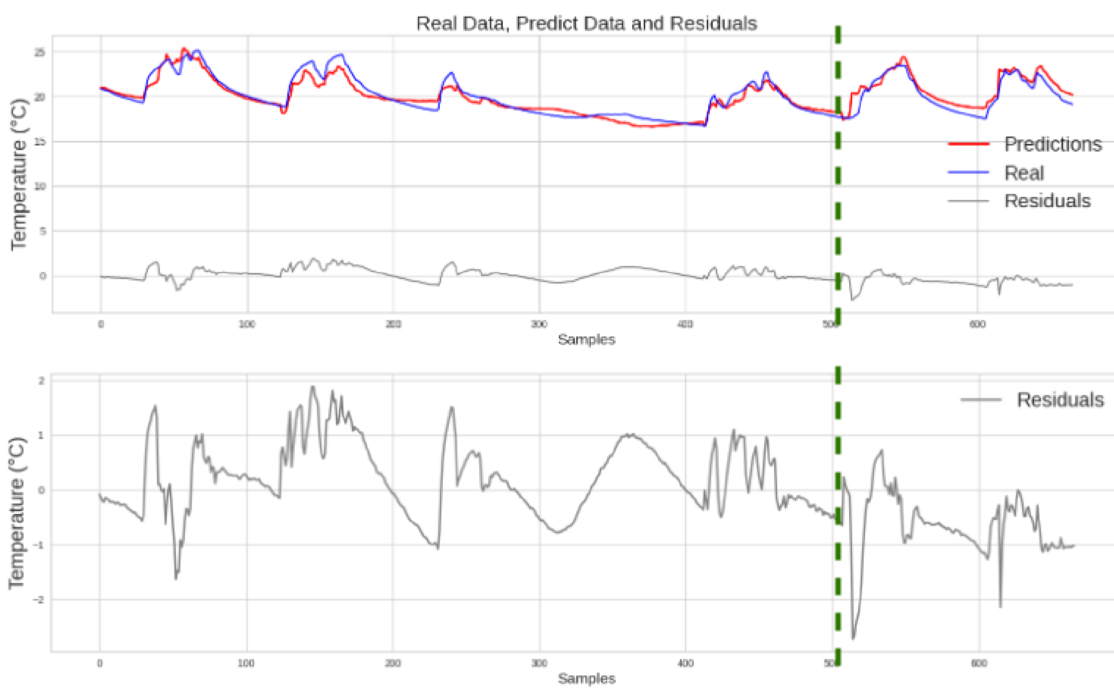

Figure 5.17: Result of the temperature prediction for January by MLR in case “C”.

Table 5.1: Correlation indices between the variables predict.

CORRELATION	Gas Concentration BME680	Humidity	Gas concentration MQ135	Pressure	Temperature
Gas Concentration BME680	1.00	-0.22	-0.79	0.03	-0.58
Humidity	-0.22	1.00	0.35	0.22	-0.37
Gas Concentration MQ135	-0.79	0.35	1.00	0.00	0.66
Pressure	0.03	0.22	0.00	1.00	-0.17
Temperature	-0.58	*-0.37*	0.66	-0.17	1.00

Using this same range for training, the MLR was applied to the other variables collected by BME680, such as humidity, gas concentration and pressure. The results are summarized in Table 5.2. It is possible to see that this algorithm has a good result for the humidity, a median result for gas concentration and for the pressure presents the worst result. The correlation indices between the variables explain this. It is possible to see in Table 5.1 that the variable "pressure" presents the lowest correlation indices and consequently will have the lowest prediction accuracy since it depends practically on zero of the other variables. To improve the MLR algorithm for this variable, it would be necessary to insert other arguments that have a more significant relationship, such as altitude.

Table 5.2: Prediction results for humidity, gas concentration and pressure data.

RESULTS	Humidity	Gas Concentration	Pressure
Accuracy (R^2)	86,90%	58,70%	6,50%
MAE Training	0,99%	12,1kΩ	1,4hPa
MAE Test	1,96%	13,5kΩ	0,86hPa
Residuals	-4,5% to 2,8%	-30,2kΩ to 30,2kΩ	-4,7hPa to 3,4hPa

These results show the efficiency of the MLR algorithm when applied for five training days. It is possible to see that improvements will need to do to have better accuracy in all the parameters. The goal is to use this same multiple linear regression in the various parameters measured for the Apolo building, seeking an efficient relationship between the data and the accuracy of the prediction. These results can be used in the energy efficiency of the building, predicting excessive consumption that can be dealt with in the right way. In Chapter 6, possible future work will be addressed to make this application more trustworthy.

Moreover, real-time monitoring of the prediction provided by MLR was done using Node-RED, presenting the data directly in Grafana along with the data collected by the sensors. Due to technical problems and data loss, this monitoring will not be shown in full here, presenting only a preview of the last measurement performed, on October 2022, in Figure 5.18. This graph will be improved when the data is sufficient to prove the efficiency of the real monitoring with the algorithm.

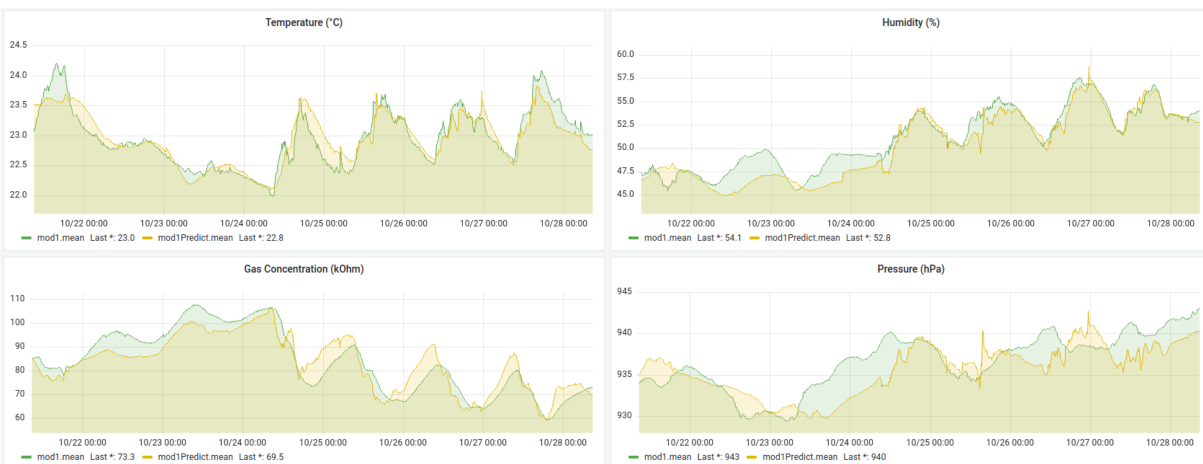

Figure 5.18: Dashboard for real-time data visualization.

Chapter 6

Conclusion and Future Work

This work presented solutions for monitoring the Apolo building, which has developed to become an intelligent building through its restructuring. Through real-time monitoring of the data in the building, energy and waste efficiency was proposed to make the inhabitant's consumption more conscious. The data validation in the building was done, although it was impossible to test it in normal use to the lack of time for the end of the renovation. The project had to undergo some adaptations during the realization due to technical problems due to the renovation. Despite this, the data obtained and the system allocated there present favorable results, allowing its expansion to other apartments. The developed PCB also allows the easy integration of new sensors that may be needed in the future, increasing the quality of monitoring even more. The algorithm presented for the prediction of variables was proposed in order to, in the future, have enough data to predict sustainable solutions for the use of the apartment using additional tools, like gamification. Although the data from CeDRI was used to validate the machine learning algorithm, the principle of applying and sending the data will be the same in Apolo. From the point of view of the efficiency of the MLR prediction, it can be seen that for the predicted variable, the regression showed better results in the 7-day range, using five days for training. Despite promising results for the temperature and humidity variables, it is perceptible that there is a need to improve the relationship between the data to have a better prediction efficiency for the pressure and gas concentration in the environment. When it is possible to perform the application of this project entirely in the building, it is expected to perform better tests to increase the prediction efficiency of the algorithm. Due to the problems in data storage in the Apolo building, the project did not have enough time to fulfill its objective of creating a specific application for the inhabitant, remaining only the visualization through Grafana. In addition, as a result of this thesis project, papers have been published to expose the partial conclusions along the development [50]–[52], and participation as co-author for [53].

6.1 Future Works

As the next steps, it is suggested to continue the project by increasing its complexity and improving the results obtained:

- Include new sensors in the Apollo module to improve the amount and variety of data collected. With the increased parameters, it will also be possible to improve the multiple linear regression algorithm.

- Improve data visualization in Grafana, creating a central dashboard for building management with the primary data collected and predicted.

- Improve the multiple linear regression algorithm, testing adverse situations that may occur in the apartment and with the real data.

- Insert the data collected by the IoTaWatt equipment and the waste system in the MLR algorithm.

- Create a specific application for the resident that can be shown through the personal smartphone or on the interaction screen in the apartment, for example. This application brings solutions that the resident can use for apartment efficiency and resource utilization.

- Through the application to the resident use resources such as gamification to create new perspectives on conscious consumption, going beyond just showing the data in a static and tiresome way. The use of gamification can bring advantages for changing consumer behavior.

Bibliography

[1] M. M. Froufe and C. A. P. Soares, “Os sistemas que autam nos edifícios iinteligentes e sustentáveis,” *Congresso de Construção Cívil, Universidade de Brasília*, 2020.

[2] P. H. Shaikh, N. B. M. Nor, P. Nallagownden, I. Elamvazuthi, and T. Ibrahim, “A review on optimized control systems for building energy and comfort management of smart sustainable buildings,” *Renewable and Sustainable Energy Reviews*, vol. 34, pp. 409–429, 2014, ISSN: 1364-0321. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S1364032114001889`.

[3] D. Clements-Croome, “Creative and productive workplaces: A review,” *Intelligent Buildings International*, vol. 7, no. 4, pp. 164–183, 2015. DOI: `10.1080/17508975.2015.1019698`. eprint: `https://doi.org/10.1080/17508975.2015.1019698`. [Online]. Available: `https://doi.org/10.1080/17508975.2015.1019698`.

[4] I. P. de Pádua, “Caracterização de edifícios inteligentes - um caso exemplo,” M.S. thesis, Pontifícia Universidade Católica de Minas Gerais, Brasil, 2006.

[5] M. Wigginton and J. Harris, *Intelligent skins*. Routledge, 2013.

[6] A. T. So, A. C. Wong, *et al.*, “A new definition of intelligent buildings for asia,” *Facilities*, 1999.

[7] L. C. Braga, “Estudo de aspectos de eficiência energética de edificações com uma abordagem de automação predial,” 2007.

[8] C. C. Pet, *Você sabe qual é a vida útil de um edifício?* Jun 2018. [Online]. Available: `https://civilizacaoengenheira.wordpress.com/2018/06/25/voce-sabe-qual-e-a-vida-util-de-um-edificio/`. Accessed: 2022-06-14.

[9] W. Fincher and M. Boduch, "Standards of human comfort: Relative and absolute," in The University of Texas at Austin, 2009. [Online]. Available: `https://repositories.lib.utexas.edu/handle/2152/13980`.

[10] R. Nunes and C. Sêrro, "Edifícios inteligentes: Conceitos e serviços," May 2022.

[11] E. Union, *Energy efficient buildings*, [Online]. Available: `https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficient-buildings_en`. Accessed: 2022-06-05.

[12] R. W. D. Martins, "Otimização preditiva de operação energética de edificios inteligentes," M.S. thesis, Universidade do Porto, Portugal, 2016.

[13] M. Al-Kuwari, A. Ramadan, Y. Ismael, L. Al-Sughair, A. Gastli, and M. Benammar, "Smart-home automation using iot-based sensing and monitoring platform," in *2018 IEEE 12th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG 2018)*, 2018, pp. 1–6. DOI: `10.1109/CPE.2018.8372548`.

[14] G. M. Smith, *Como medir a corrente usando sensores e transdutores de corrente*, 2021. [Online] Available: `https://dewesoft.com/br/aquisicao-de-dados/como-medir-corrente-usando-sensores-de-corrente`. Accessed: 2022-04-26.

[15] IoTaWatt, *The best kept energy monitoring secret.* [Online]. Available: `https://iotawatt.com/`. Accessed: 2022-04-20.

[16] Shelly, [Online]. Available: `https://shelly.cloud/`. Accessed: 2022-04-20.

[17] R. Connor, *The United Nations world water development report 2015: water for a sustainable world*. UNESCO publishing, 2015, vol. 1.

[18] R. M. M. Teixeira, "Uso sustentável da água: Avaliação de benefícios adotando eficiência hídrica em edifícios," M.S. thesis, Universidade de Coimbra, Portugal, 2015.

[19] M. M. Campos, "Medidor de vazão ultrassônico de multi-trajetórias.," M.S. thesis, Universidade Federal de Itajubá, Brasil, 2017.

[20] Contimetria, [Online]. Available: `https://www.contimetra.com/PagIA/Produtos.aspx?capitulo=237&catalogo=70#CD`. Accessed: 2022-05-13.

[21] C. de Saneamento Básico do Estado de São Paulo, *Medição do consumo com leitura remota*, [Online]. Available: `https://site.sabesp.com.br/site/interna/Default.aspx?secaoId=732`. Accessed: 2022-05-13.

[22] Sensus, *Iperl® (international) water meters*, [Online]. Available: `https://sensus.com/emea/products/iperl-international/`. Accessed: 2022-05-13.

[23] L. G. de Almeida and J. F. Borin, "Plataforma inteligente de coleta de resíduos baseda em internet das coisas e LPWAN," [Online]. Available: `https://smartcampus.prefeitura.unicamp.br/pub/artigos_relatorios/Lahis-Plataforma_Inteligente_de_Coleta_de_Residuos_baseda_em_Internet_das_Coisas_e_LPWAN.pdf`.

[24] M. Cerchecci, F. Luti, A. Mecocci, S. Parrino, G. Peruzzi, and A. Pozzebon, "A low power iot sensor node architecture for waste management within smart cities context," *Sensors*, vol. 18, no. 4, 2018, ISSN: 1424-8220. DOI: `10.3390/s18041282`. [Online]. Available: `https://www.mdpi.com/1424-8220/18/4/1282`.

[25] T. Bakhshi and M. Ahmed, "Iot-enabled smart city waste management using machine learning analytics," in *2018 2nd international conference on energy conservation and efficiency (ICECE)*, IEEE, 2018, pp. 66–71.

[26] K. Ashton *et al.*, "That "internet of things" thing," *RFID journal*, vol. 22, no. 7, pp. 97–114, 2009.

[27] D. Bandyopadhyay and J. Sen, "Internet of things: Applications and challenges in technology and standardization," *Wireless personal communications*, vol. 58, no. 1, pp. 49–69, 2011.

[28] A. P. F. M. Mascarenhas, S. M. Fernandes, F. D. Freitas, *et al.*, "Desenvolvimento de produtos iot," *Brazilian Journal of Development*, vol. 7, no. 1, pp. 4711–4724, 2021.

[29] H.-D. Ma, "Internet of things: Objectives and scientific challenges," *Journal of Computer science and Technology*, vol. 26, no. 6, pp. 919–924, 2011.

[30] M. Nasar and M. A. Kausar, "Suitability of influxdb database for iot applications," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 10, pp. 1850–1857, 2019.

[31] M. Giacobbe, C. Chaouch, M. Scarpa, and A. Puliafito, "An implementation of influxdb for monitoring and analytics in distributed iot environments," in *International conference on the Sciences of Electronics, Technologies of Information and Telecommunications*, Springer, 2018, pp. 155–162.

[32] I. Data, *Open source time series database*, [Online]. Available: `https://www.influxdata.com/`. Accessed: 2022-05-23.

[33] M. Electronics, [Online]. Available: `https://pt.mouser.com/`. Accessed: 2022-04-20.

[34] Solectro, [Online]. Available: `https://solectroshop.com/pt/`. Accessed: 2022-04-20.

[35] A. Maier, A. Sharp, and Y. Vagapov, "Comparative analysis and practical implementation of the esp32 microcontroller module for the internet of things," in *2017 Internet Technologies and Applications (ITA)*, IEEE, 2017, pp. 143–148.

[36] Espressif, [Online]. Available: `https://www.espressif.com/`. Accessed: 2022-04-15.

[37] Arduino, [Online]. Available: `https://www.arduino.cc/`. Accessed: 2022-04-15.

[38] S. N. Z. Naqvi, S. Yfantidou, and E. Zimányi, “Time series databases and influxdb,” *Studienarbeit, Université Libre de Bruxelles*, vol. 12, 2017.

[39] L. Berbakov, M. Batić, and N. Tomašević, “Mobile application for energy management in smart buildings,” in *International Conference on Information Society and Techology*, 2019.

[40] G. Labs, *The open observability platform*, [Online]. Available: `https://grafana.com/`. Accessed: 2022-05-23.

[41] N. Havard, S. McGrath, C. Flanagan, and C. MacNamee, “Smart building based on internet of things technology,” in *2018 12th International conference on sensing technology (ICST)*, IEEE, 2018, pp. 278–281.

[42] A. C. Coelho and J. V. Cunha, “Regressão linear múltipla,” *Análise multivariada: para os cursos de administração, ciências contábeis e economia. São Paulo: Atlas*, pp. 131–231, 2007.

[43] P. Online, *O que é regressão linear múltipla?* [Online]. Available: `https://psicometriaonline.com.br/o-que-e-regressao-linear-multipla/`. Accessed: 2022-08-26.

[44] G. K. Uyanık and N. Güler, “A study on multiple linear regression analysis,” *Procedia-Social and Behavioral Sciences*, vol. 106, pp. 234–240, 2013.

[45] J. Brains, *PyCharm*, [Online]. Available: `https://www.jetbrains.com/pycharm/`. Accessed: 2022-09-20.

[46] S. V. Bhusari, “Smart building integration,” *Consulting Specifying Engineer*, vol. 51, no. 4, pp. 34–39, 2014.

[47] D. Belani, A. H. Makwana, J. Pitroda, and C. M. Vyas, “Intelligent building new era of todays world,” *Proceedings of the Trends and Challenges of Civil Engineering in Todays Transforming World, Surat, India*, pp. 1–16, 2014.

[48] S. Tadokoro, Q.-S. Jia, Q. Zhao, *et al.*, “Smart building technology [TC Spotlight],” *IEEE Robotics & Automation Magazine*, vol. 21, no. 2, pp. 18–20, 2014.

[49] I. P. do Mar e da Atmosfera, [Online]. Available: `https://www.ipma.pt/pt/index.html`. Accessed: 2022-09-23.

[50] R. B. Kalbermatter, T. Brito, J. Braun, *et al.*, "Smart systems for monitoring buildings - an iot application," *II International Conference of Optimization, Learning, Algorithms and Applications.*, Submitted for publication.

[51] R. B. Kalbermatter, T. Brito, J. Braun, A. I. Pereira, and J. Lima, "Monitoring system for smart buildings - apolo building case," *Symposium of Applied Science for Young Researchers*, Submitted for publication.

[52] R. B. Kalbermatter, T. Brito, J. Braun, A. I. Pereira, Â. P. Ferreira, and J. Lima, "Smart buildings – a case study in braganza," *VII Ibero-American Congress on Entrepreneurship, Energy, Environment and Technology*, [Online]. Available: `http://cieemat.ipb.pt/files/CIEEMAT2022_Book_of_Abstracts.pdf`.

[53] J. Lima, R. B. Kalbermatter, J. Braun, T. Brito, G. Berger, and P. Costa, "A realistic simulation environment as a teaching aid in educational robotics," *13th Workshop of Robotics in Education*, Submitted for publication.